

基于景观形状指数抽样的 GlobeLand30 (2010) 中国区精度验证

王宇^{1*}, 张景雄², 刘迪², 杨文静¹, 张旺乐¹

1. 武汉大学遥感信息工程学院, 武汉 430079; 2. 武汉大学测绘学院, 武汉 430079

摘要: 大尺度土地覆盖数据是地球表层过程研究、生态系统评估、环境建模等科学研究的重要基础。研究现有数据集的特点, 评价它们的精度, 对数据使用者及生产新的数据集都具有指导意义。本研究为顾及地表覆盖空间异质性, 实现稀少地类样本量的均衡分配, 采用了基于景观指数 (LSI) 的抽样方法。在中国区域抽取 1,000 个验证点, 利用 Google Earth 高分历史影像判读验证样本, 评价了 GlobeLand30(2010) 中国区数据集的精度。结果表明 GlobeLand30(2010) 中国区产品的总体精度为 77.90%, 说明中国区 GlobeLand30(2010) 数据和真实地表覆盖数据之间的一致性较好, 能较为准确地反映中国地表覆盖的分布状况。

关键词: GlobeLand30(2010); 中国区; 抽样; 精度验证

DOI: 10.3974/geodp.2018.01.06

1 前言

大尺度土地覆盖数据是地球表层过程研究、生态系统评估、环境建模、可持续发展规划等研究的重要基础。2010 基准年的全球 30 m 分辨率地表覆盖数据集 GlobeLand30(2010) 是国家基础地理信息中心领衔的团队自 2009–2014 年研发的世界首套全球 30 m 地表覆盖数据产品, 2014 年 9 月起向全世界开放共享^[1]。针对全球尺度 30 m 影像地类提取的难点, GlobeLand30 采用了基于像元、对象和知识 3 个层次的分类策略, 根据地表覆盖 10 大类型的光谱、形态特征等特点, 综合现有的分类技术, 设计和制订了稳健的遥感信息提取方法, 依次提取水体、湿地、冰雪、人造地表、耕地、裸地、林灌草和苔原等单一地类, 继而进行数据集成^[2]。

为了验证该数据集的精度, 同济大学、中科院、中国农科院、中国林科院等单位采用二级抽样检验方法, 对 GlobeLand30(2010) 数据的总体分类精度进行了第三方精度评价^[3]。第一级为图幅抽样, 以“图幅”为抽样单元。第二级为要素抽样, 以“图幅内空间分类要素”为抽样单元。通过计算每一级抽样单元需要抽取的样本量, 结合空间变异性分析合理布设样本, 以相当数量的样本来检验地表覆盖分类的准确性。全球共抽取 80 个图幅, 154,586

收稿日期: 2018-01-18; 修订日期: 2018-03-12; 出版日期: 2018-03-25

基金项目: 国家自然科学基金 (41471375)

*通讯作者: 王宇 B-5868-2018, 武汉大学, wangyuchn@whu.edu.cn

数据引用方式: [1] 王宇, 张景雄, 刘迪等. 基于景观形状指数抽样的 GlobeLand30 (2010) 中国区精度验证[J]. 全球变化数据学报, 2018, 2(1): 29–34. DOI: 10.3974/geodp.2018.01.06.

[2] 王宇, 张景雄, 刘迪等. 景观指数抽样土地覆盖验证点(中国)与 GlobeLand30 (2010) 数据产品比对数据集[DB/OL]. 全球变化科学研究数据出版系统, 2018. DOI: 10.3974/geodb.2018.01.13.V1.

个检验样本，得出 GlobeLand30（2010）数据的总体精度为 $80.33\% \pm 0.2\%$ ^[4]。

目前，针对中国区域的 GlobeLand30 产品验证的研究相对较少，但是有一些研究值得学习。杨永可等以中国为研究区，为与其他几套全球现有的地表覆盖数据集进行统一对比，将 GlobeLand30（2010）产品升尺度到 300 m 的尺度，然后人工选取了 1,063 个样本点进行验证，抽样结果表明 GlobeLand30（2010）精度最高，中国区整体精度为 82.39% ^[5]。陆苗等人对比了中国区 2010 基准年的五套全球耕地数据集的精度，结果表明 GlobeLand30（2010）耕地类别精度要优于其它四种，总体精度达到 79.61% ^[6]。显然，不同的抽样设计得出不同的验证结果。本研究旨在采用基于景观形状指数（Landshape shape index, LSI）的样本点布设方法，对 GlobeLand30（2010）中国区产品进行采样效率较高的精度验证，以期得到聚焦面积占比小但却易误分的异质地域的精度验证结果。

2 数据集元数据简介

GlobeLand30（2010）中国区精度验证数据集（ValPlotChinaGlobeLand30(2010)）^[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 景观指数抽样土地覆盖验证点与 GlobeLand30（2010）数据产品比对数据集元数据简表

条 目	描 述
数据集名称	景观指数抽样土地覆盖验证点（中国）与 GlobeLand30（2010）数据产品比对数据集
数据集短名	Samples_LSI_CHN
作者信息	王宇 B-5868-2018, 武汉大学, wangyuchn@whu.edu.cn 张景雄 B-6210-2018, 武汉大学, jxzhang@whu.edu.cn 刘迪 B-6783-2018, 武汉大学, alliu0815@whu.edu.cn 杨文静 B-6720-2018, 武汉大学, yangwj@whu.edu.cn 张旺乐 B-6084-2018, 武汉大学, zhangwl@whu.edu.cn
地理区域	地理范围 3°52'N-53°33'N, 73°40'E-135°2'30"E, 行政范围：中国
数据年代	2010 年
空间分辨率	30 m
数据格式	.rar, .kmz
数据量	88.4 MB（压缩后）
数据集组成	数据集由 3 部分组成： （1）Excel 表格数据, 包括 1,000 个验证点初始 GlobeLand30 的地表覆盖类型、基于 Google Earth 在 30 m × 30 m 尺度和 300 m × 300 m 尺度下的判定类别、Google Earth 图片以及验证点经纬度坐标 （2）.shp 文件, 包括验证点初始类型、判读类别以及经纬度坐标 （3）.kmz 文件, Google Earth 格式的验证点信息
基金项目	国家自然科学基金项目（41471375）
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 ^[8]

3 数据研发方法

不同的抽样设计得出不同的验证结果。本研究为顾及地表覆盖空间异质性，采用了基于景观指数的抽样方法，在中国区域抽取验证点，利用 Google Earth 高分历史影像判读验证样本，继而评价 GlobeLand30（2010）中国区数据集的精度。

3.1 验证点布设

众所周知，地表覆盖检验样本的数量与空间分布直接影响着其数据精度评价的科学性与准确性。对于大范围地表覆盖验证，人们以往主要根据验证成本、面积或专家主观经验确定地表覆盖区域或类型的样本数量，并通常采用整群随机抽样、分层随机抽样和系统抽样，布设样本空间位置。但这些传统方法忽略了大范围地表覆盖所具有的空间异质性，难以保证区域样本量、地类间样本量分配和样本空间分布的合理性。为解决该问题，陈斐提出了一种顾及地表覆盖空间异质性的样本抽样方法，分别计算验证区域、区内各地类和抽样格网三个层级（尺度）的景观形状指数，推导出基于区域级景观形状指数及面积的区域样本量，建立顾及地类级景观形状指数的样本量分配公式，进而提出了基于格网级景观形状指数曲线的样点抽取方法。实践证明，该方法的精度评价结果更接近真实精度，尤其是稀少类精度评价结果的准确度显著提高^[9]。

统计发现，中国区各地类分布较不均匀，空间异质性很大，各地类具体覆盖比例见表2。可以看出耕地、森林、草地、裸地四种地类比较接近，其面积累计占中国区陆地面积整体的94.07%，而灌木地、湿地、水体、人造地表、冰川和永久积雪面积比例很低，约占中国区陆地面积的6%。如果按照传统的分层随机抽样，稀少地类所能获得的验证点比例必然很低，从有限样本量中分配到的样本量较少，所得的精度指标不如非稀少类的可靠。因此，本研究采用国家基础地理信息中心提供的顾及地表覆盖空间异质性的样本点布设方案（以下简称三级LSI），进行GlobeLand30（2010）中国区的验证，验证点具体分布如图1、表2所示。

由表2可知，基于三级LSI的样本点布设方案，较好地顾及了中国区稀少类（灌木地、湿地、水体、人造地表、冰川和永久积雪）的样本分配，避免了稀少类样本数量过少的问题，实现了稀少类样本量的均衡分配，进而取得更接近实际情况的精度评价结果。需要指出的是，中国区域内苔原（寒带环境下由地衣、苔藓、多年生耐寒草本植物和灌木植被覆盖的土地）面积很少，鲜有分布，因此，苔原类不参与中国区验证。

3.2 验证点判读

验证点判读主要利用 Google Earth 的高分辨率影像作为参考数据源，时间优先设定为2010年生长季。判读过程主要包括：建立基本判读单元、判读标准，利用地学知识（形状、大小、模式、阴影、色调、纹理、位置、时间序列、空间序列与海拔高度等）综合判读，对所有样本点进行标注，并记录相关信息。由于GlobeLand30数据集的空间分辨率是30 m，因此，基本判读单元是与以样本点为中心的30 m × 30 m的正方形。考虑到基本判读单元内部出现没有优势类型，或者判读单元只是孤立一点等情况，在30 m正方形外围，增加了一个边长300 m的正方形，以便辅助判读，基本判读单元样图如图2所示。

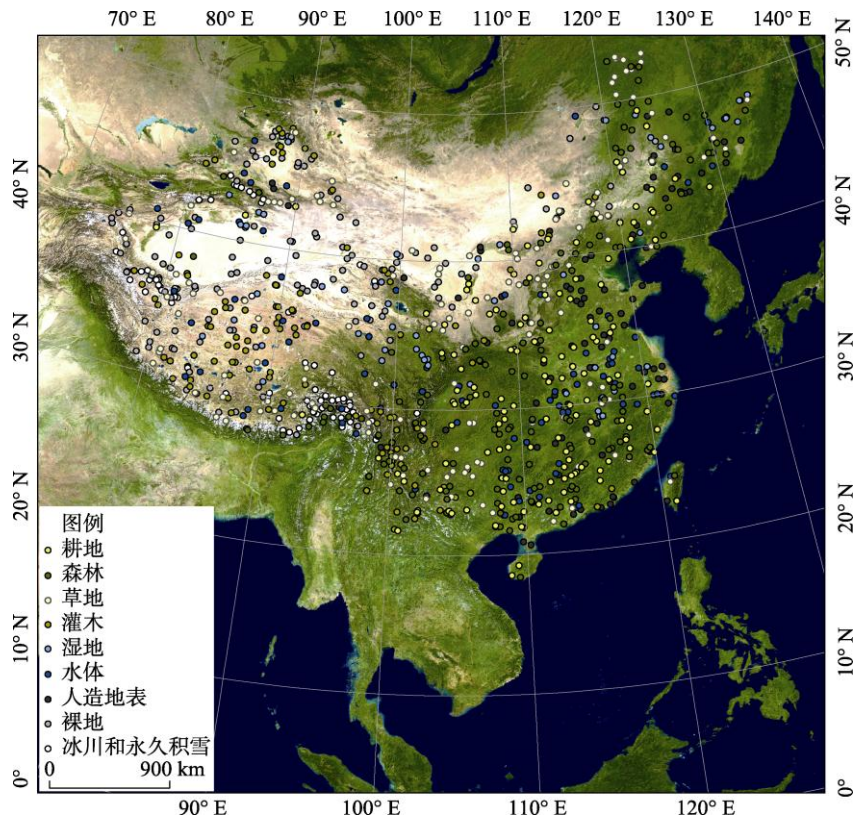


图 1 研究区范围与 GlobeLand30（2010）验证样本分布情况

表 2 GlobeLand30（2010）中国区验证点布设

地类	全国地类面积 占比（%）	布设样 本量（个）
10-耕地	21.34	135
20-森林	22.23	142
30-草地	29.60	146
40-灌木地	1.05	116
50-湿地	0.43	62
60-水体	1.56	96
80-人造地表	1.81	100
90-裸地	20.90	124
100-冰川和 永久积雪	1.08	79
总计	100	1,000

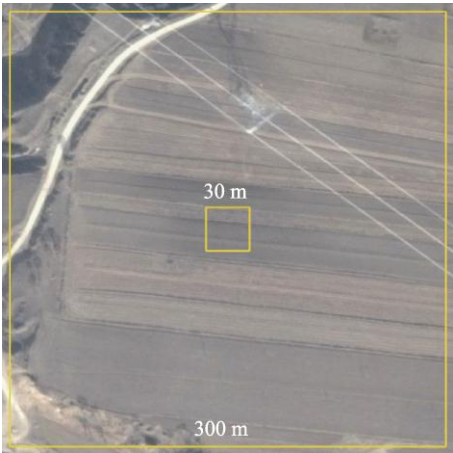


图 2 基本判读单元

4 数据结果

地表覆盖产品的精度评估方法主要是通过数据产品的抽样验证，得到分类结果和真实结果（参考数据结果）的交叉制表，即混淆矩阵，进而计算用户精度（错分率）、生产者精度（漏分率）、总体精度和 Kappa 系数等精度指标^[10]。混淆矩阵也称误差矩阵或应变表，

是一个用于表示分为某一类别的像元个数与实际检验为该类别数的比较阵列。通常，阵列中的列代表参考数据，行代表由遥感数据分类得到的类别数据。混淆矩阵有像元数和百分比表示两种。中国区 1,000 个验证点的混淆矩阵结果如表 3 所示。

表 3 中国区样本点的混淆矩阵

类型	耕地	森林	草地	灌木地	湿地	水体	人造地表	裸地	冰川和永久积雪	小计
耕地	106	15	2	4	0	2	6	0	0	135
森林	8	108	14	8	0	0	2	1	1	142
草地	6	13	102	12	2	2	4	5	0	146
灌木地	2	0	19	88	1	0	0	6	0	116
湿地	1	1	1	0	55	3	1	0	0	62
水体	6	1	6	4	4	72	1	1	1	96
人造地表	14	5	3	6	0	3	67	2	0	100
裸地	1	1	7	3	0	1	0	111	0	124
冰川和永久积雪	0	1	0	0	0	0	0	8	70	79
小计	144	145	154	125	62	83	81	134	72	1,000

混淆矩阵是评估专题图精度的基础，通常利用混淆矩阵计算以下四种精度评价的描述因子：

- （1）总体精度：整个地表覆盖产品的综合描述，总体精度可以利用被正确分类的像元总和与总样本量的比值来表示。
- （2）用户精度：是指地表实际类别的像元个数与地图分类结果的像元个数的比值，即，混淆矩阵中正确分到某类的像元总数（对角线值）与分类器将整个影像的像元分为该类的像元总数（混淆矩阵中该类行的总和）的比值。
- （3）生产者精度：又称制图精度，是指分类器将整个影像的像元正确分为某类的像元数（对角线值）与该类真实参考总数（混淆矩阵中该类列的总和）的比值。
- （4）Kappa 系数：一种用于精确度评估的离散多变量技术，主要用于统计确定误差矩阵是否与另一种误差矩阵有显著差异。Kappa 系数的范围是从 0 到 1，其值越大表示数据结果的可靠性越高。

5 讨论和总结

通过抽样、布样、判读、检验等过程，建立了样本的混淆矩阵，在此基础上计算每个地类的制图精度和用户精度，如表 4 所示。

从生产者精度来看，冰川和永久积雪、湿地、水体、裸地和人造地表精度较高，均高于 80%，而森林、耕地、灌木地和草地则相对较差。原因在于林、灌、草和耕地容易混淆，尤其在异质区，这几类地物在 30 m 的影像上难以分辨。

表 4 各类地表类型的生产者精度和用户精度

类型	生产者精度（%）	用户精度（%）
耕地	73.61	78.52
森林	74.48	76.06
草地	66.23	69.86
灌木地	70.40	75.86
湿地	88.71	88.71
水体	86.75	75.00
人造地表	82.72	67.00
裸地	82.84	89.52
冰川和永久积雪	97.22	88.61

就用户精度而言,除了草地和人造覆盖,其它类别都普遍高于75%,说明数据集所标注的地类类别正确率较高,而在人造地表信息存在一定的高估。GlobeLand30(2010)中国区基准年数据产品的总体精度77.90%,Kappa系数为0.75,而考虑面积加权^[11]的总体精度为77.57%,表明GlobeLand30(2010)中国区数据和真实地表覆盖之间的一致性具有可用价值。与常见的分层随机等方法相比,基于三级LSI的样本点布设方法分配给稀少地类的样本量有所增加,即在空间异质区的样本点密度会高于同质区的样本点密度,会导致基于该方法估算的精度低于基于未考虑空间异质性抽样布设方法的精度评估结果。

整体来说,中国区30 m分辨率的地表覆盖信息产品反映出中国地表覆盖的分布状况,为研究人类活动对地球表面的影响提供了可用的数据。今后仍需在多空间尺度,比如由整体到局部的精度验证等方面进一步努力,可以进一步提高数据的精度。

作者分工:张景雄教授对数据集的开发做了总体设计;王宇、刘迪采集和判读了验证点;杨文静设计了模型和算法;张旺乐做了数据分析和验证;王宇撰写了数据论文等。

致谢:感谢中国国家地理基础数据中心提供GlobeLand30(2010)的数据产品,感谢陈军教授、武昊老师、郑新燕老师提供的关于基于LSI抽样方法的指导以及中国区1,000个验证点的布设信息,笔者在此深表谢忱。

参考文献

- [1] Chen, J., Ban, Y., Li, S. China: open access to earth land-cover map [J]. *Nature*, 2014, 514(7523): 434. DOI: 10.1038/514434c.
- [2] 曹鑫, 陈学泓, 张宏伟等. 全球30 m空间分辨率耕地遥感制图研究[J]. *中国科学:地球科学*, 2016, 46(11): 1426–1435. DOI: 10.1360/N072015-00291.
- [3] Tong, X., Wang, Z., Xie, H., *et al.* Designing a two-rank acceptance sampling plan for quality inspection of geospatial data products [J]. *Computers & Geosciences*, 2011, 37(10): 1570–1583. DOI: 10.1016/j.cageo.2011.02.006.
- [4] Chen, J., Chen, J., Liao, A., *et al.* Global land cover mapping at 30 m resolution: A POK-based operational approach [J]. *Isprs Journal of Photogrammetry & Remote Sensing*, 2015, 103: 7–27. DOI: 10.1016/j.isprsjprs.2014.09.002.
- [5] Yang, Y., Xiao, P., Feng, X., *et al.* Accuracy assessment of seven global land cover datasets over China [J]. *Isprs Journal of Photogrammetry & Remote Sensing*, 2017, 125: 156–173. DOI: 10.1016/j.isprsjprs.2017.01.016.
- [6] Lu, M., Wu, W., Zhang, L., *et al.* A comparative analysis of five global cropland datasets in China [J]. *Science China Earth Sciences*, 2016, 59(12): 2307–2317. DOI: 10.1007/s11430-016-5327-3.
- [7] 王宇, 张景雄, 刘迪等. 景观指数抽样土地覆盖验证点(中国)与GlobeLand30(2010)数据产品比对数据集[DB/OL]. 全球变化科学研究数据出版系统, 2018. DOI:10.3974/geodb.2018.01.13.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]. DOI:10.3974/dp.policy.2014.05 (2017年更新).
- [9] Chen, F., Chen, J., Wu, H., *et al.* A landscape shape index-based sampling approach for land cover accuracy assessment [J]. *Science China Earth Sciences*, 2016, 59(12): 2263–2274. DOI: 10.1007/s11430-015-5280-5.
- [10] Congalton, R.G. A review of assessing the accuracy of classifications of remotely sensed data [J]. *Remote Sensing of Environment*, 1991, 37(1): 35–46. DOI: 10.1016/0034-4257(91)90048-B.
- [11] Olofsson, P., Foody, G. M., Stehman, S.V., *et al.* Making better use of accuracy data in land change studies: estimating accuracy and area and quantifying uncertainty using stratified estimation [J]. *Remote Sensing of Environment*, 2013, 129(2): 122–131. DOI: 10.1016/j.rse.2012.10.031.