



文献 DOI:

10.11922/csdata.2018.0097.zh

数据 DOI:

10.11922/sciencedb.716

文献分类: 地球科学

收稿日期: 2018-12-21

开放同评: 2019-03-18

录用日期: 2019-09-28

发表日期: 2019-11-29

1. 河南大学环境与规划学院, 河南开封 475004

2. 黄河中下游数字地理技术教育部重点实验室, 河南开封 475004

摘要: 利用 86 景 Landsat 8 卫星 OLI 云量小于 10% 的卫星影像, 以及高时间分辨率的高分一号 WFV 宽视角多光谱相机、无人机影像、Google Earth 影像为基础数据, 根据河南省的地势地貌和土地覆被特征, 选取 4 个覆盖研究区所有土地覆被类型的样方进行全面野外调查, 并获得 1518 个样点, 参考样点和样区的土地覆被特征用 ArcMap 10.3 软件目视解译样区 (到二级类型)。采用专家论证法评价和修正样区解译结果。在样区解译工作的基础上, 再次野外调查并解译河南省剩余区域土地覆被类型, 最终获取野外样点共 4883 个。最后采用全球随机正六边形网格 (12 级) 的方法选取 1895 个地面验证点评价解译精度, 结果显示河南省土地覆被解译二级类型总体精度为 91.19%, Kappa 系数为 0.886。

关键词: 河南省; 土地覆被; Landsat 8; 目视解译; 2015 年

数据库 (集) 基本信息简介

数据集名称	2015 年河南省土地覆被栅格数据
数据作者	李宁、朱筠、秦奋
数据通信作者	秦奋 (qinfen@henu.edu.cn)
数据时间范围	2015 年
地理区域	河南省
空间分辨率	100 m、50 m
数据量	24.84 MB
数据格式	Geotiff
数据服务系统网址	http://www.sciencedb.cn/dataSet/handle/716
基金项目	国家科技基础条件平台-地球系统科学数据共享平台项目“2015 年河南地区 1:10 万土地覆被遥感解译制图”; 国家科技基础条件平台-地球系统科学数据共享平台项目“2015 年河南省区域 1:10 万土地覆被质量验证”。
数据库 (集) 组成	本数据集包含一个数据库, 名称为“shujushiti.gdb”, 数据库中有 3 条数据, 分别是: 2015 年河南省土地覆被 100 米栅格数据、野外采样点、地面验证点。(1) 数据库中“hns_tdfb100m”为 2015 年河南省土地覆被 100 米栅格数据, 栅格 value 是地类代码;(2) 数据库中“hns_tdfb50m”为 2015 年河南省土地覆被 50 米栅格数据, 栅格 value 是地类代码;(3) “yangdian”为野外采样点, DLDM 字段为类型代码;(4) “yanzhengdian”为地面验证点数据, DLMC1_samp 字段为一级地类名称, DLMC2_samp 字段为二级地类名称。

* 论文通信作者

秦奋: qinfen@henu.edu.cn

引言

土地利用/覆盖变化是全球环境变化研究的前沿和热点问题^{[1][2][3]}。土地覆被空间分布及其变化反映了人类社会经济活动过程,被视为全球陆地生态系统变化最重要的影响因素之一^[4]。它是生态系统水循环^{[5][6]}、气候变化^[7]、土壤侵蚀^[8]、生物多样性^[9]、人地关系以及区域可持续^[10]等研究的基础数据,其质量好坏直接影响研究成果。随着土地利用/覆被变化研究的深入,致使政府、基层组织和居民对生态环境变化的响应成为研究热点。2010 年,覆盖全球的 30 米分辨率的陆地卫星图被免费应用,全球尺度 100 米分辨率的陆地覆盖地图得以实现^[11]。由国家地理信息中心联合北京师范大学、清华大学、中科院遥感所等 18 家单位,研制出世界首套两期(2000、2010 年)30 米全球地表覆盖数据产品 GlobeLand30,分类级别为一级,分类精度总体为 83.5%^[12]。宫鹏团队利用 Landsat 8 完成了 2010、2015、2017 年全球 30 米土地覆被数据 FROM-GLC30,2017 年数据分类级别为一级,总体精度为 70.2%^[13]。2019 年宫鹏团队利用 Google Earth Engine 基于 Sentinel-2 影像完成了首套全球 10 米 2017 年土地覆被数据 FROM-GLC10,分类级别为一级,总体精度为 72.76%^[14]。全球尺度上较早的 5 套土地覆被数据集(GLCC、UMD、GLC2000、MODIS LC、GlobCover)一致性评价精度最高的 MODIS LC 的精度为 65.6%^[15]。此外还有由国际地圈-生物圈计划的全球 1 千米土地覆被数据 ISLSCP II IGBP DISCover and SiB Land Cover, 1992–1993^[16]。在国家尺度上,中科院基于中国 HJ-A/B 环境一号卫星影像生产出 2010 年全国土地覆被数据(ChinaCover),一级类总体精度为 91%,二级类总体精度为 82%。美国地质调查局 USGS 生产的 30 米国家土地覆盖数据库(NLCD),跨越 1992、2001、2007、2011 年 4 个时期^[17]。2010 年河南省区域土地覆被数据 GlobeLand30 一级类总体精度为 83.33%^[18]。此外还有一些商业土地覆被数据产品,如地理国情监测云平台有全国长时间序列的土地覆被数据,该数据是基于 Landsat TM/ETM/OLI 遥感影像生产,数据分级到二级类,但是数据收费,获取代价较大。

随着大数据、超算技术的发展,机器学习、深度学习等遥感数据处理方法的涌现,不同时间和空间尺度的土地覆被数据越来越多。这些数据的分类和精度评价多数到一级地类,少数到二级地类的数据限于某个年份,且精度不是很高。精确的土地覆被数据在提高碳循环、水文循环、气候变化监测、土壤退化等多种科学研究的准确性方面起着至关重要的作用^{[19][20][21]}。本条数据通过建立大量的野外采样点数据、建立解译样区等方法,充分了解研究区地类的光谱特征、纹理特征、形状特征、空间分布、NDVI 特征等,建立研究区的解译标志库,最终完成研究区二级分类体系下的土地覆被解译。本数据虽然在空间分辨率上略低,但是在分类精细,且具有较高的精度,可为河南省的生态文明建设、粮食生产核心区建设、区域模拟和可持续发展等提供科学数据支撑,同时为河南省后期的土地覆被解译提供参考依据。

1 数据采集和处理方法

1.1 区域范围

河南省位于 31°23'N–36°22'N, 110°21'E–116°39'E 之间,地处淮河、黄河、长江、海河流域,东连山东、安徽,西邻陕西,北与河北、山西相接,南临湖北,总面积 16.7 万平方千米。河南省地势西高东低,北、西、南三面由太行山、伏牛山、桐柏山、大别山沿省界呈半环形分布;中、东部为黄

淮海冲积平原；西南部为南阳盆地。平原和盆地、山地、丘陵分别占总面积的 55.7%、26.6%、17.7%。灵宝市境内的老鸦岔为全省最高峰，海拔 2413.8 米；海拔最低处在固始县淮河出省处，仅 23.2 米。河南省大部分地处暖温带，南部跨亚热带，属北亚热带向暖温带过渡的大陆性季风气候，同时还具有自东向西由平原向丘陵山地气候过渡的特征，具有四季分明、雨热同期、复杂多样和气象灾害频繁的特点。

1.2 数据源

以 2015 的 Landsat 8 OLI 光谱数据为主，数据来源于从美国地质调查局（United States Geological Survey, USGS）网站（<http://glovis.usgs.gov/>），根据河南省所覆盖的行列号下载区内春、夏、秋、冬 4 个时相纹理特征明显，云量 < 10% 的影像数据，共 86 景；对于较难识别的地物将图斑勾勒出来后，叠加到 Google Earth 影像辅助识别；国家地球系统科学数据共享平台-黄河下游科学数据中心的“黄河下游典型地区无人机遥感影像数据集（2015 年）”，辅助黄河下游地区的解译工作；河南省西部山区、南部水田分布广泛的地区的高分一号卫星各 1 景，辅助判别图斑属性。Landsat 8 OLI 影像数据空间分布如图 1 所示。

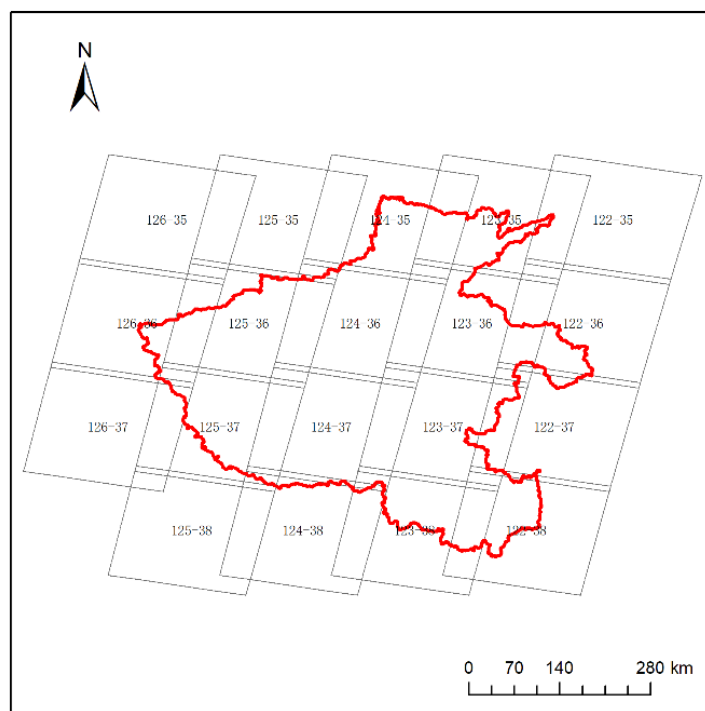


图 1 河南省影像数据空间分布图

1.3 数据处理

本条数据集由目视判读、人工手绘的方式完成解译。首先根据研究区地理环境、地势地貌等特征，建立解译样区，并对样区内土地覆被类型进行实地考察，建立野外考察矢量数据库。根据野外样点不同类型的土地覆被特征对样区进行解译，邀请遥感解译方向专家审阅样区解译结果，指导修正易混淆类型、不易判读的地类。样区解译完成之后，野外实地考察剩余区域并建立完整的研究区的样点数据库，根据样区以及野外考察样点数据库建立解译标志库。然后对整个区域目视解译勾绘，同时建立属性数据库，形成初步的矢量解译成果。初步解译完成后，检查结果的拓扑错误，最后进

行精度评价并生成栅格数据。耕地、林地、草地水体的勾绘以夏季影像为主，由于建设用地、道路在夏季影像上可能会被旁边的树覆盖，尤其是西部山区，所以此类地物的勾绘以冬季影像为主。而落叶林、常绿林则同时依靠夏冬两个季节的影像来区分，采用的波段组合主要是 5、4、3 假彩色和 4、3、2 真彩色。夏季影像质量不好或者有云的地方，采用质量较好的春季、秋季的影像替代。在 4 季影像均有云的部分区域，同时采用 2014 和 2016 年的影像判读，如果两年地类、区域均未发生变化，则直接勾绘，个别地类或者斑块大小发生了变化，则参考 Google Earth 影像或者实地走访调查的方式判定。解译详细步骤如下：

(1)样区选取：根据河南省的地形地貌选取了山区、平原以及山区平原过渡带建立了 4 个样区，如图 2 所示，共 3.4 万平方千米，涉及濮阳、三门峡、南阳、信阳 4 个城市，面积分别为 0.7、0.4、1.5、0.8 万平方千米。

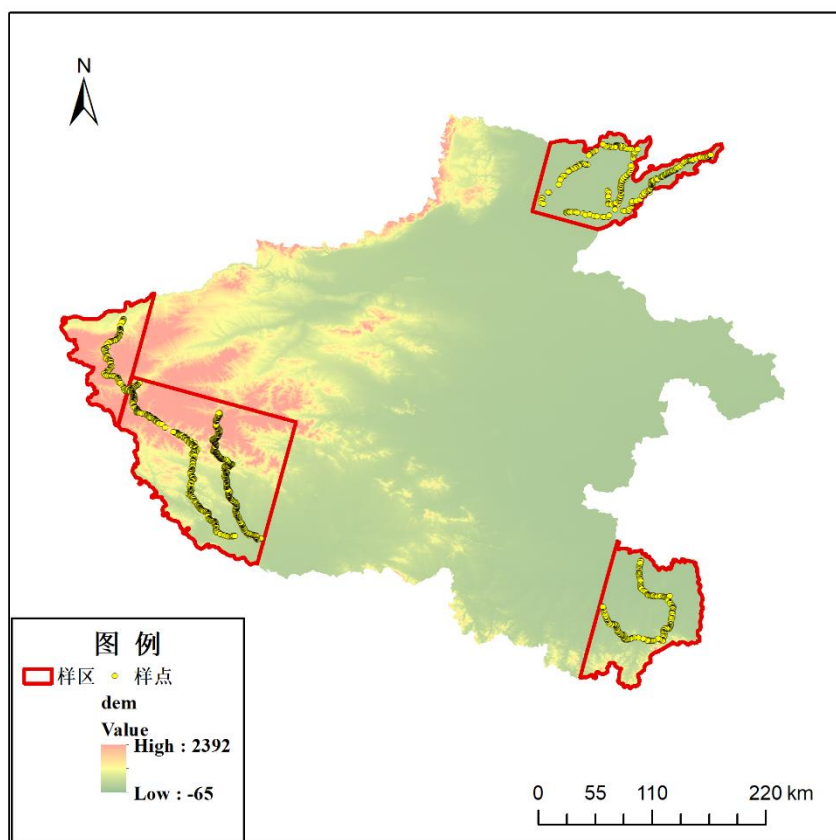


图 2 样区位置图

(2)样区野外考察：采用车载 GPS、手持 GPS、ArcMap、Landsat 影像图、大疆无人机精灵 4 考察样区所包含的所有土地覆被类型，并进行拍照和电子标记。人和车无法到达的地区借助于无人机拍摄。最后整理成完整的样区样点数据库，共 1518 个样点，详细信息如表 1 所示。

表 1 样区各类型样点数量（单位：个）

地类名称	地类代码	数量
常绿针叶林	11	49
落叶阔叶林	14	354

地类名称	地类代码	数量
针阔混交林	15	48
灌丛	16	82
草丛	22	56
草甸	23	4
灌丛草地	24	15
水田	31	57
旱地	32	266
园地	33	84
城镇建设用地	41	70
农村居民地	42	149
交通用地	43	13
工矿用地	44	151
河流	61	39
湖泊	62	2
水库/坑塘	63	27
河湖滩地	64	26
裸岩	72	12
裸地	73	13
苔地	74	1

(3) 解译人员培训：目视解译是地学研究和遥感应用的一项基本技能。本次解译人员全部为地图学与地理信息系统专业硕、博士。在解译之前，请遥感学教授培训所有工作人员地表景观-成像过程-遥感影像的过程及其逆过程遥感影像-目视解译-地表景观；培训省内涉及到的二级分类下的地类在某种或不同波段组合下的色调、颜色、大小、阴影、纹理、图型、位置、相关布局以及所在地貌、土壤等特征；培训解译图斑的勾绘、节点的疏密，保证不同工作人员生产数据的一致性。

(4) 样区解译：根据野外采样点以及遥感影像的光谱信息、纹理、形态、空间分布、NDVI 等信息，分别用影像的 4、3、2 波段和 6、5、4 波段组合，结合不同时相的影像判别每个地类斑块儿属性，含糊不清的地类斑块通过 ArcMap 的 Layer to KML 工具将地类斑块变为 KML 图层叠加至 Google Earth 影像上辅助判别。样区解译过程中，反复展示和研讨各个分区已形成的解译成果，及时发现、修正和克服图斑勾绘、属性判别以及解译人员解译过程中记录的重点难点问题，最终得到研究区的样区解译成果。

(5) 样区修正：邀请遥感解译专家审阅、考核样区解译结果，着重指导并修正易判别错误的地类。重点训练存在解译问题的相关解译人员。

(6) 研究区野外考察：在样区解译的基础上野外考察采样河南省剩余区域。共获取 4883 个样点，野外样点空间分布图如图 3，详细信息如表 2。

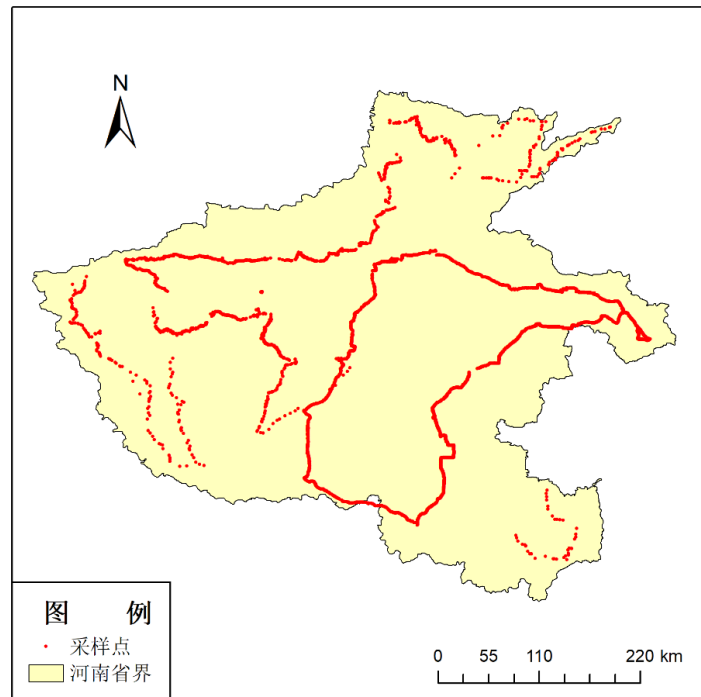


图3 野外考察样点空间分布图

表2 研究区各类型样点数量（单位：个）

地类名称	地类代码	数量
常绿针叶林	11	76
常绿阔叶林	12	4
落叶阔叶林	14	604
针阔混交林	15	54
灌丛	16	83
草丛	22	90
草甸	23	8
灌丛草地	24	16
水田	31	151
旱地	32	1345
园地	33	264
城镇建设用地	41	521
农村居民地	42	809
交通用地	43	77
工矿用地	44	487
河流	61	171
湖泊	62	3
水库/坑塘	63	51

地类名称	地类代码	数量
河湖滩地	64	28
裸岩	72	12
裸地	73	17
苔地	74	12

（7）建立解译标志库：根据野外考察样点和样区解译结果，对研究区内所涉及到的地类归纳总结，建立解译标志库，描述每种地类的空间分布、所在生态区、影像形态、色调、纹理以及多时期影像示意信息、所对应的 Google Map 点位、NDVI 信息等。

（8）研究区划分：由于人工手绘解译方法工作量大，解译前依据解译人员对解译区熟悉、地形地貌尽量统一、单景遥感图像完整性、界限两侧地物容易识别（如河流、水库等）的原则将样区进行了划分。具体划分成果如图 4 所示。

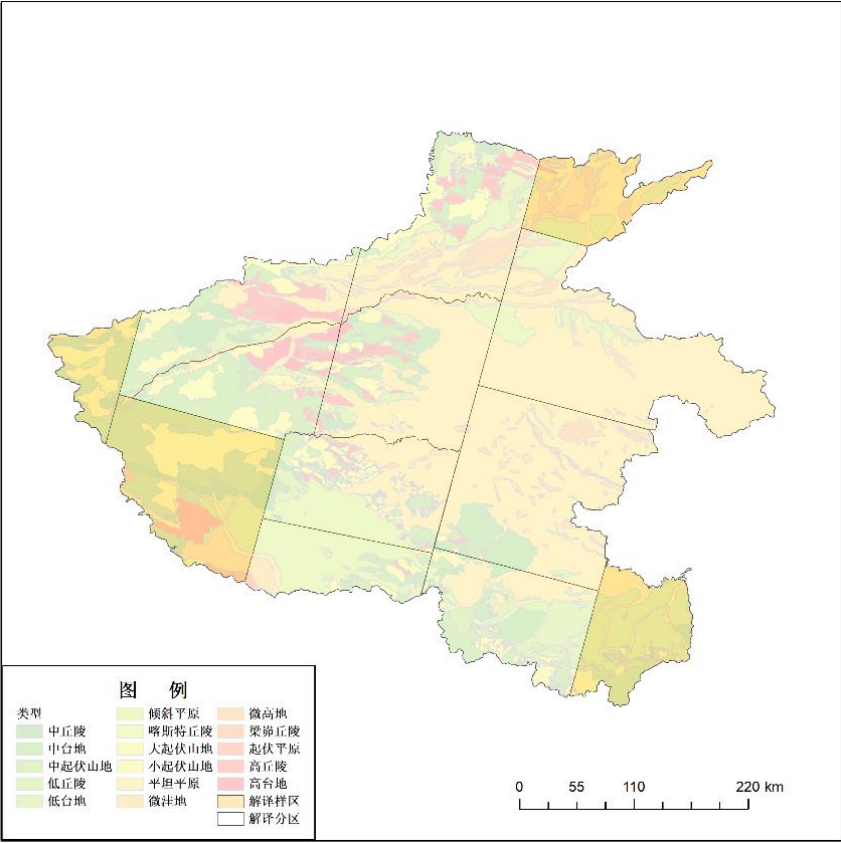


图 4 研究区任务分割图

（9）判读解译：根据已有知识，参考所建立的解译标志库完成研究区影像解译勾绘工作。解译中间出现的问题，多次进行研讨和训练。使用高分一号影像、Google Earth 影像、无人机影像辅助解译。对于不能确定类型的复杂地区可将勾画的图斑叠加到对应位置高分影像（分辨率 2 米）上辅助判别，如图 5 所示；对应地区没有高分影像的或者借助高分影像仍无法判别的，将勾画的图斑转成 KML 文件添加到 Google Earth 影像（分辨率为 0.49 米），并将 Google Earth 影像时间调至 2015 年，或者下载 2015 年相应位置的 Google Earth 影像添加到 ArcMap 中与勾画出的图斑叠合，如图 6 所示；Google Earth 影像仍无法判别的借助于无人机影像或者实地调查的方式辅助判别，如图 7 所示。

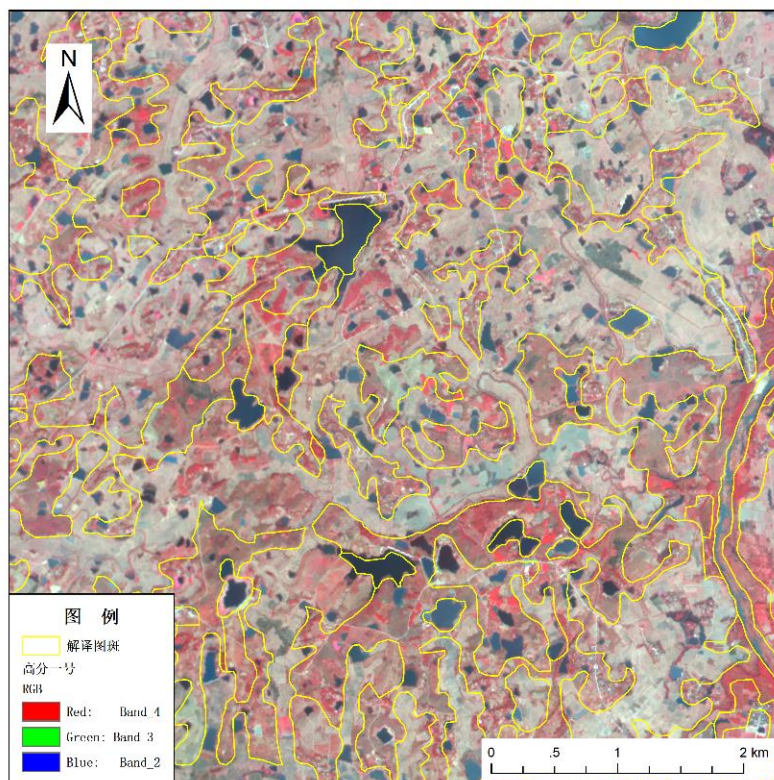


图 5 高分影像辅助解译示例

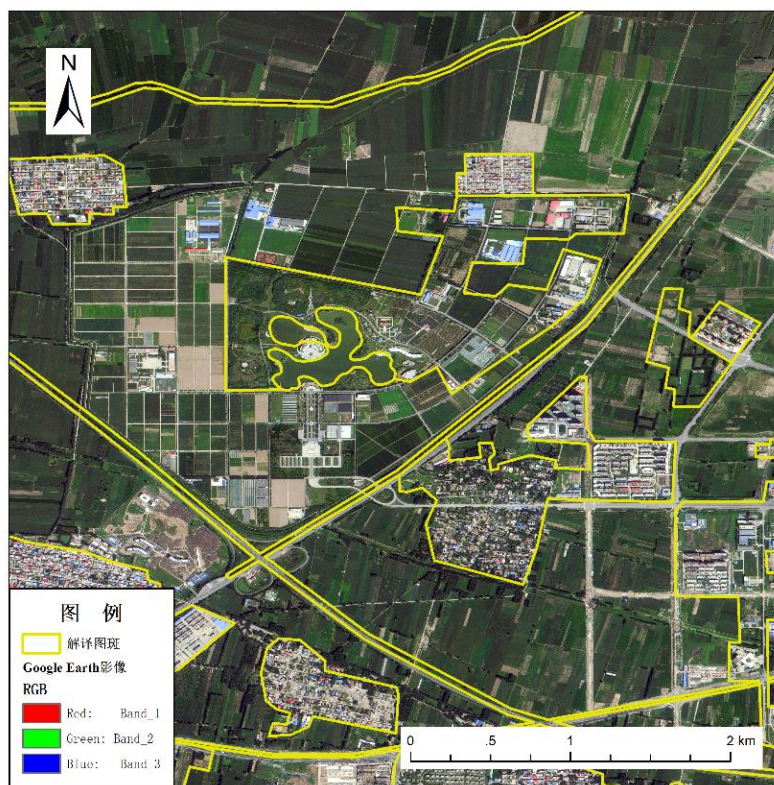


图 6 谷歌影像辅助解译示例

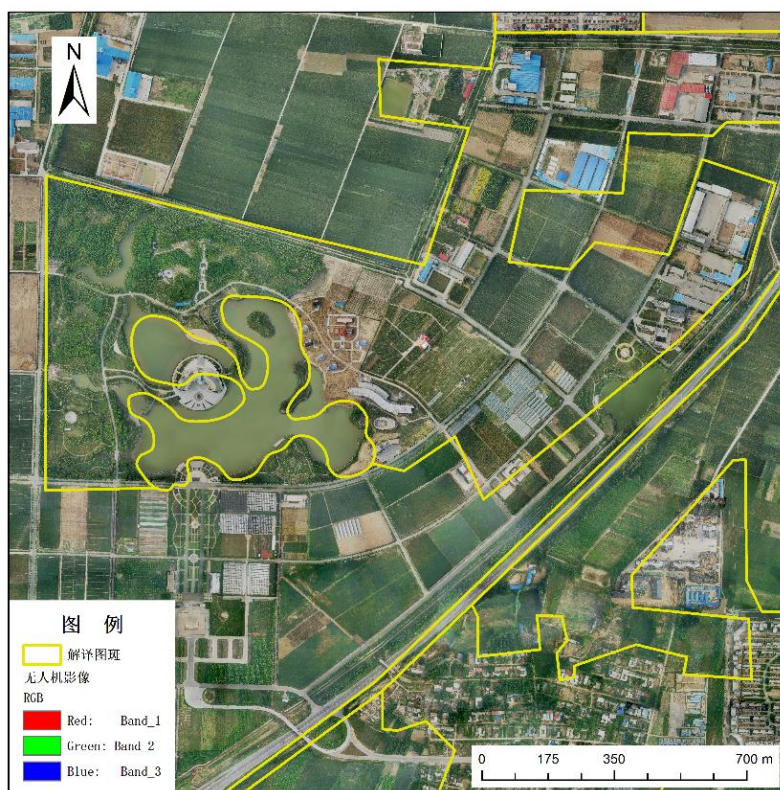


图 7 无人机影像辅助解译示例

(10) 接边: 每块区域边界向外做 1000 米的缓冲区, 以保证边界两侧同时被解译。边界接边工作由相邻两组解译人员共同完成。

(11) 拓扑错误纠正: 利用 ArcMap Toolbox 中的 Topology 工具检查拓扑错误, 拓扑规则有 “Must Not Overlap” “Must Not Have Gaps”, 检查并修正属性数据完整性。

(12) 精度评价: 基于混淆矩阵和 Kappa 系数的方法, 对研究区的解译结果进行精度评价。采用全球随机正六边形格网的方法随机选取验证样点和野外验证点在研究区共选取了 1895 个验证点以验证结果。

(13) 栅格数据生成: 此种方法得到的原始数据是 Shapefile 矢量数据, 由于判读影像的分辨率为 30 米, 所以结果可以转换成 ≥ 30 米任何大小的栅格, 本文分别选择了 50 米和 100 米的栅格作为产品发布。利用 ArcMap 中 Tool box 的 To Raster 模块下的 Polygon to Raster 功能, 按照地理代码/地类名称的属性, 分别设置 Cellsize 大小为 100 和 50, 将解译结果转换成分辨率为 100 米和 50 米的栅格数据。

2 数据样本描述

2.1 分类描述

此次分类按照国家科学基础条件平台-地球系统科学数据共享平台的《2015 年全国 1:10 万土地覆被遥感解译技术规范》^[22]中的分类体系, 分为林地、草地、耕地、建设用地、湿地、水体以及其他地类 7 种一级土地覆被类型, 28 种二级土地覆被类型 (表 3)。

表 3 土地覆被分类体系

一级类型	二级类型	编码	含义（定义、主要特征、主要分布）
林地	常绿针叶林	11	定义：是以针叶树为建群种所组成的各类森林的总称。 主要特征：属于裸子植物的乔木林，具有典型的针状叶；一年没有落叶或少量落叶时期的物候特征。郁闭度>30%，高度>3米，包括常绿针叶天然林和人工林。 主要分布：分布广泛，从寒温带至中、低纬度亚高山地区均有分布。
	常绿阔叶林	12	定义：是亚热带湿润地区由常绿阔叶树种组成的地带性森林类型。 主要特征：一般属于双子叶、被子植被的乔木林，叶型扁平、较宽；一年没有落叶或少量落叶时期的物候特征。属于亚热带季风气候的植被，是具有热带至温带之间的过渡性质的森林类型。郁闭度>30%，高度>3米，包括常绿阔叶天然林和人工林。 主要分布：大致分布在南、北纬度22°-34°之间，其中以中国长江流域南部的常绿阔叶林最为典型，面积也最大。
	落叶针叶林	13	定义：由落叶松柏类为主的针叶树所构成的森林。 主要特征：属于裸子植物的乔木，具有典型的针状叶；一年中有明显落叶时期的物候特征。林冠的色彩非常单调一致，结构简单，乔木层通常一层。郁闭度>30%，高度>3米，包括有一定落叶周期、针状叶片天然林和人工林。 主要分布：我国落叶针叶林分布在大兴安岭北部山地。主要树种有兴安落叶松。
	落叶阔叶林	14	定义：是由冬季落叶的阔叶乔木组成的森林群落。 主要特征：属于双子叶、被子植被的乔木林，叶型扁平、较宽；一年中有明显落叶时期的物候特征。郁闭度>30%，高度>3米，包括有一定落叶周期、具有较宽叶片天然林和人工林，主要树种有砾木、山毛榉、槭、栎、椴、桦树等。 主要分布：这种森林几乎完全分布在北半球受海洋性气候影响的温暖地区，在中国地区主要分布在东北地区南部和中国北方各省。
	针阔混交林	15	定义：是寒温带针叶林和夏绿阔叶林间的过渡类型。 主要特征：通常由砾属、槭属、椴属等阔叶树种与云杉、冷杉、松属的树种混合组成。郁闭度>30%，高度>3米，针阔混交林中的每种类型的覆盖度不超过75%。 主要分布：在欧亚大陆的中高纬度形成一条不连续的混交林带。在中国主要分布于东北和西南各省山地，其他地区受人类活动影响现在也有分布。
	灌木	16	定义：是一种以灌木占优势的植被类型。 主要特征：多以灌丛建群种，多为簇生的矮树或灌木丛林。郁闭度>30%，高度<3米，通常丛生、无明显主干的木本植物，但有时也有明显主干。 主要分布：灌木在我国分布很广，如温带的落叶阔叶灌丛有榛属、蔷薇属等。

一级类型	二级类型	编码	含义（定义、主要特征、主要分布）
草地	草原	21	定义：指温带半干旱气候下的有旱生草本植物组成的植被。 主要特征：植被特点鲜明、类型单一，属地带性植被。宏观上在地球表面处于湿润的森林区域和干旱的荒漠区域之间，通常包括稀疏草原和温带草原。 主要分布：在我国分布北起松辽平原和呼伦贝尔草原，呈连续的带状，向西南延伸至青藏高原地区。此外，在荒漠和森林地区也有分布。
	草丛	22	定义：指中生和旱生多年草本植物为主要建群种的植物群落。 主要特征：在大多数情况下，群落中散生着稀疏的矮小灌丛，属地带性植被。 主要分布：大多数情况下是由森林或灌丛经破坏后形成的次生植被。原生草丛主要分布在滨海、河漫滩、湖边或湖中淤积地。在我国多分布于东部、南方地区。
	草甸	23	定义：指生长在低温和温凉气候、中度湿润条件下的多年生中生草本植物为优势的植被群落。 主要特征：属中生植物，也包括旱中生植物，属非地带性植被。草群生长繁茂、密集，覆盖度大，层次结构简单但层片结构复杂，生产力较高。 主要分布：在我国主要分布于秦岭—淮河一线以北的温带森林区、半干旱草原区和干旱荒漠区以及青藏高原地区，此外在亚热带的山地上部和湖滨湿地也有少量分布。
	灌丛草地	24	草地中灌丛郁闭度<30%，灌丛高度<2 米
耕地	水田	31	定义：指有水源保证和灌溉设施，在一般年景能正常灌溉，用以种植水稻、莲藕等水生农作物的耕地。在多类作物轮作中，只要有一季节为水稻或水生作物，则视为水田。 主要特征：水田按水源情况分为灌溉水田和望天田两类。 主要分布：在中国地区，主要分布于东部季风区内。
	旱地	32	定义：一般降水量大于 250—400 毫米以上地区主要种植旱生作物、菜地等可以获得一定产量的耕地，包括没有灌溉、引洪淤灌的耕地。 主要特征：以种植旱生作物、菜地等为主。 主要分布：全国各地都有分布，其中东北、华北和西南这 3 个大区旱地面积之和占全国的 62% 左右。
	园地	33	定义：种植以采集果、叶、干、茎、汁为主的集约经营的多年生木本和草本植物。 主要特征：主要包括果园、茶园、桑园、橡胶、苗圃、观赏园林等，一般覆盖度在 50% 以上或每亩株数大于合理株数 70% 以上。 主要分布：全国范围内均有种植，但因园地种类而异，种植在一定的海拔之内，主要分布在温带、亚热带、热带地区。
建设用地	城镇建设用	41	定义：指镇建制以上的城镇及附近公共配套设施等建设用地。 主要特征：包括普通住宅、公寓、别墅等用地。 主要分布：全国各地区均有分布。

一级类型	二级类型	编码	含义（定义、主要特征、主要分布）
	农村居民地	42	定义：居民以农业为主要活动形式的聚落，包括农村居民点、定居放牧点等。 主要特征：具有农舍、牲畜棚圈、仓库场院、农村道路、水渠、宅旁绿地，以及特定环境 and 专业化生产条件下的附属设施。 主要分布：分布于农村地区。随着现代城市化的发展，在城镇郊区还出现了城市化村这种类似城市的农村聚落。
	交通用地	43	定义：指居民点以外的各种道路及其附属设施和民用机场港口用地。 主要特征：包括高速公路、国道、省道、县乡道、铁道、机场、港口等。 主要分布：全国范围均有分布。
	工矿用地	44	定义：指工业、采矿、仓储业等工业生产及附属设施的用地。 主要特征：包括工业生产及附属设施、采矿、采石、采沙场、仓储、盐田、砖瓦窑等。 主要分布：全国范围均有分布。
水体	沼泽	51	定义：植被覆盖度高的湿生草地以及地势平坦低洼、排水不畅、长期潮湿、多积水且表层生长湿生草本植被的土地。 主要特征：地表生长着湿性植物和沼泽植物，有泥炭累积或虽无泥炭累积但有潜育层存在。按分布区域可分为森林湿地、湖泊湿地、灌丛湿地、草本湿地等。 主要分布：中国的沼泽面积 11 万平方千米，三江平原、大小兴安岭和长白山沼泽最多，青藏高原次之，其余则零散分布。
	近海湿地	52	定义：各种近海及海岸的海涂、珊瑚礁、红树林沼泽等。 主要特征：生物多样性丰富、生产力高，主要有红树林沼泽、海草湿地、潮间盐沼、潮间淤泥质海滩。 主要分布：多分布在河口三角洲、沙丘间洼地、堤外洼地、泻湖及潮间、潮下带，在我国沿海 11 个省（市、自治区）和港澳台沿海地区均有分布，并以杭州湾为界分为南、北两部分。
	河流	61	定义：通常是指陆地河流经常或间歇地沿着狭长凹地流动的水流地带。 主要特征：狭长，特征明显。一般是在高处作源头，沿地势向下流，一直流入湖泊或海洋，包括天然河流、溪流和人工运河等流动水体。 主要分布：全国各地均有分布。
	湖泊	62	定义：指地表相对封闭可蓄水的天然洼池及其所容纳的水体。 主要特征：天然形成、特征明显、较容易辨识。 主要分布：中国湖泊众多，但在地区分布上很不均匀。总的来说，东部季风区，特别是长江中下游地区，分布着中国最大的淡水湖群；西部以青藏高原湖泊较为集中，多为内陆咸水湖。
	水库 / 坑塘	63	定义：水库指在山沟或河流的狭口处建造拦河坝形成的人工水面；坑塘指人工开挖或天然形成的蓄水量小于 10 万立方米的常水位以下的土地及其承载水体。

一级类型	二级类型	编码	含义（定义、主要特征、主要分布）
			主要特征：人工建造形成，水面边界清楚，特征明显。 主要分布：水库一般分布在山沟或河流的狭口处；坑塘主要分布在中国东部的中南区和华东区（长江以南平原湖区周围和丘陵稻作区，以及广东的珠江三角洲地区），这两大区坑塘水面面积占全国坑塘水面总面积的 73.6%，广大的西北区坑塘水面仅占全国的 2.6%。
	河湖滩地	64	定义：指河湖边淤积成的平地，是水陆互相过渡的地带，主要有沙洲、滩头、滩涂、河滩、湖滩等。 主要特征：土层深厚、土质肥沃、地势平坦。 主要分布：河流沿岸或湖泊周边，包括边滩、心滩等。
	冰川积雪	65	定义：常年积雪或冰川覆盖的土地。 主要特征：常年积雪或冰川覆盖（雪被）。 主要分布：主要分布在中、高纬度的高山区。
其他	沙漠、戈壁	71	定义：地面完全被松散沙粒或碎砾石所覆盖、植被覆盖度小于 4% 的土地。 主要特征：沙漠地域大多是沙滩、沙丘或碎砾石也经常出现。泥土稀薄、植物很少。有些沙漠是盐滩，完全没有草木，沙漠一般是风成地貌。 主要分布：中国西北部、东北西部，有大片沙丘覆盖的沙质荒漠，由砾石、碎石组成的戈壁、砾漠，以及称之为岩漠或石质荒漠的岩石裸露的山地。它们主要位于北纬 35°–50°、东经 75°–125° 之间，分布在新疆、青海、甘肃、内蒙古、陕西、吉林和黑龙江 7 个省区。
	裸岩	72	定义：地表以岩石或石砾为主、植被覆盖度在 5% 以下的裸露石山等无植被地段。 主要特征：裸露石山等无植被地段。 主要分布：裸岩主要分布在西北区，其中以新疆最多，占全国的 46.9%。
	裸地	73	定义：指地表为土质、植被覆盖度低、风霜雨雪等可直接作用的裸露地面。 主要特征：地表为土质、植被覆盖度在 5% 以下的裸土地、盐碱地等无植被地段。 主要分布：没有植被覆盖的地方。
	苔地	74	定义：指以耐寒的北极和北极–高山成分的藓类、地衣、小灌木及多年生草本植物为主组成的植物群落，包括苔原、苔地等。 主要特征：区域内冬季寒冷漫长，夏季凉爽短促，降水量 200–300 毫米，蒸发量极小，风力强劲，土壤下面常有永冻层存在。 主要分布：苔原多处于极圈内的极地东风带内以及温带、寒温带的高山树木线以上。苔原绝大部分都围绕北冰洋分布。在我国长白山和阿尔泰山上的山地苔原是亚洲最靠南的苔原，也是我国仅有的两片苔原。

2.2 数据样本

河南省的土地覆被解译完全依靠人工完成，充分考虑了光谱、时间、空间、邻接地类等各种特

征进行判别，且建立了完整的研究区解译标志库，野外调研了大量的样点，解译过程中经历了多次检查、问题反馈、数据修正。经验证，解译结果准确可靠（图 8）。

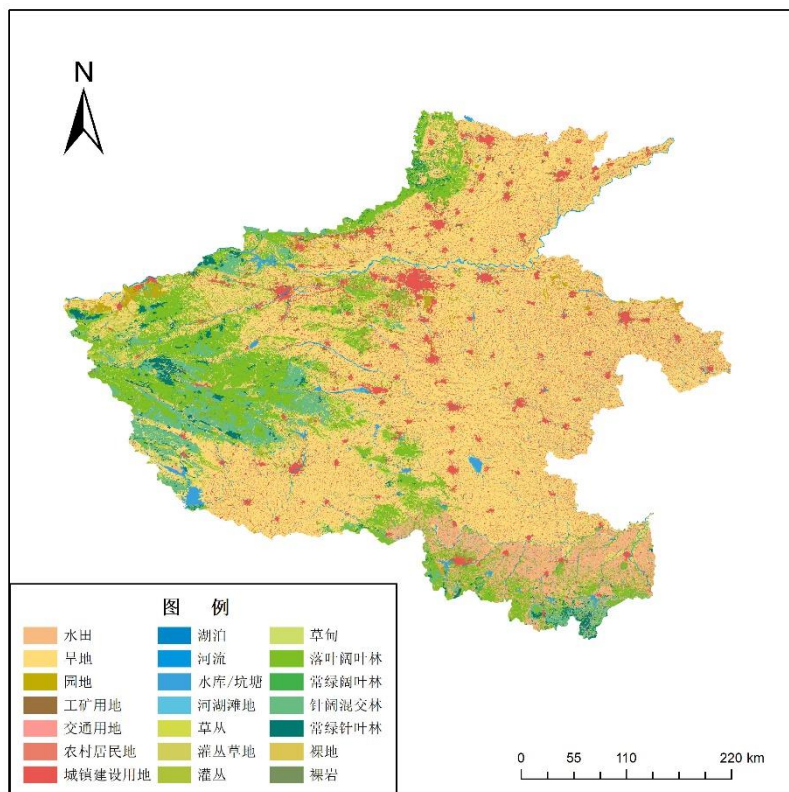


图 8 2015 年河南省土地覆被分类图

3 数据质量控制和评估

在数据处理的各个环节控制数据产品质量，以确保数据精度。主要采取的措施有：样区的选取覆盖全省所有类型；明确每种地类的直接解译标志和间接解译标志；确定图斑大小并保证不同工作人员勾绘的图斑疏密程度相同；确保地物判别无差异；勾绘过程中对不明确类型对照 Google Earth 影像，参照 Google Earth 影像仍无法判别的地物汇总之后进行野外实地调研或者邀请专家以会议研讨的形式确定；根据解译人员对研究区的熟悉程度、地貌特征、单景遥感图像完整性、容易判别地物作为边界等原则划分研究区；样区解译过程中对已形成的解译成果反复研讨，及时发现和解决问题，充分训练解译人员的工作能力；对每块区域边界多解译 1000 米缓冲区，保证接边工作进行顺利；过程以及解译完成之后反复交叉检查等。

采用野外调查验证点与全球随机正六边形网格的方法对解译结果进行验证。利用 DGGRID Version 6.2b 软件生成第 12 级的全球随机正六边形网格，在 ArcMap 中用 Clip 工具裁剪出河南的正六边形格网。根据 Kevin Sahr 的正六边形生成软件说明书可知每个格网的面积约 95.97785 平方千米。河南省的格网（包含不完整的格网）一共有 1895 个，其中正六边形格网内有野外验证点的以其为验证点，野外验证点共 217 个，格网内没有野外验证点的以格网中心点为验证点。河南省的验证点随机分布在全省。邀请 3 位遥感专业教授、副教授（非解译工作人员）对照 Google Earth 解译非野外验证点，对有出入的点的地类进行实地调研，确保每个点位的地类准确无误。然后由工作人员对所生产的数据进行一致性检验计算。1895 个验证点在研究区各个空间位置如图 9。对研究区的土地覆

被二级地类精度验证结果如表 4 所示。

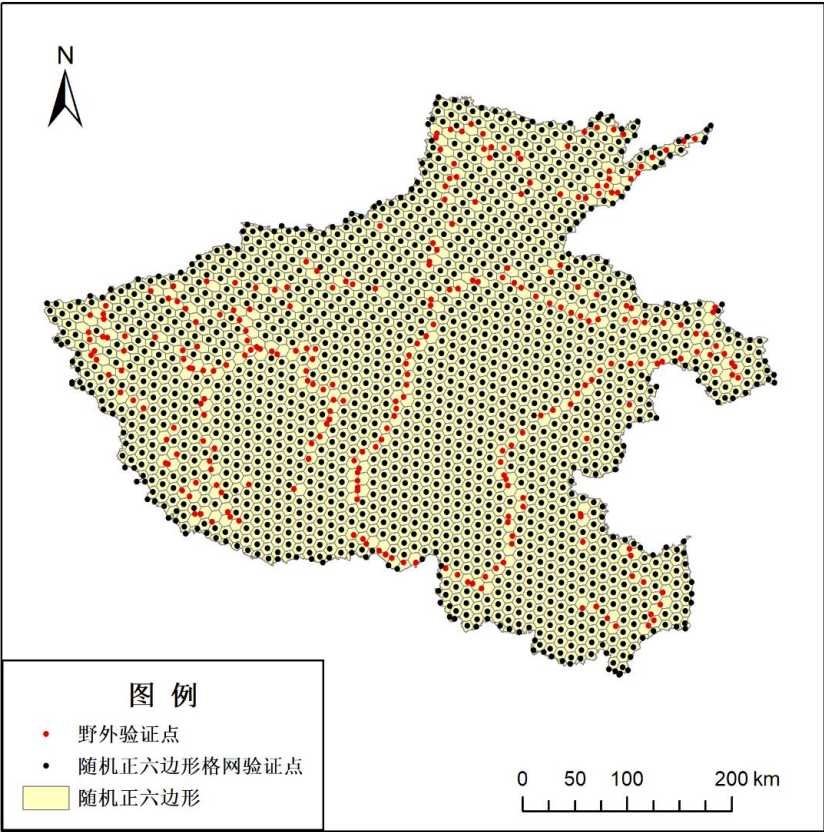


图 9 精度验证样点空间分布图

本次验证样点共 1895 个，其中常绿针叶林 56 个、常绿阔叶林 4 个、落叶阔叶林 293 个、针阔混交林 69 个、灌丛 19 个、草丛 23 个、草甸 3 个、灌丛草地 38 个、水田 117 个、旱地 806 个、园地 63 个、城镇建设用地 84 个、农村居民点 195 个、交通用地 16 个、工矿用地 34 个、河流 36 个、水库/坑塘 24 个、河湖滩地 8 个、裸岩 3 个、裸地 4 个。验证结果如表 4 所示。

表 4 2015 年河南省土地覆被解译精度验证表（单位：个）

地类名称	常绿针叶林	常绿阔叶林	落叶阔叶林	针阔混交林	灌丛	草丛	草甸	灌丛草地	水田	旱地	园地	城镇建设用地	农村居民点	交通用地	工矿用地	河流	水库 / 坑塘	河湖滩地	裸岩	裸地	汇总	精度 (%)
常绿针叶林	45	0	4	5	0	0	0	0	1	0	0	1	0	0	0	0	0	0	0	0	56	80.36

地类名称	常绿针叶林	常绿阔叶林	落叶阔叶林	针阔混交林	灌丛	草丛	草甸	灌丛草地	水田	旱地	园地	城镇建设用地	农村居民地	交通用地	工矿用地	河流	水库/坑塘	河湖滩地	裸岩	裸地	汇总	精度(%)
常绿阔叶林	0	3	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	4	75.00
落叶阔叶林	4	0	247	20	2	2	0	0	1	11	2	0	1	0	0	0	0	0	0	0	293	84.30
针阔混交林	0	6	6	55	0	0	0	1	1	0	0	0	0	0	0	0	0	0	0	0	69	79.71
灌丛	0	0	1	2	15	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	19	78.95
草丛	0	0	1	0	0	20	0	2	0	0	0	0	0	0	0	0	0	0	0	0	23	86.96
草甸	0	0	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0	0	0	0	3	100
灌丛草地	0	0	5	0	0	0	0	32	0	1	0	0	0	0	0	0	0	0	0	0	38	84.21
水田	0	0	4	2	0	0	0	0	97	13	0	0	0	0	0	1	0	0	0	0	117	82.91
旱地	0	0	7	0	0	0	0	0	2	783	1	1	4	1	0	5	1	0	0	1	806	97.15
园地	0	0	0	0	0	0	0	0	0	7	56	0	0	0	0	0	0	0	0	0	63	88.89
城镇建设用地	0	0	0	0	0	0	0	0	0	0	0	82	1	1	0	0	0	0	0	0	84	97.62
农村居民地	0	0	2	0	0	0	0	0	2	3	0	0	176	0	0	0	0	0	0	0	195	90.26
交通用地	0	0	0	0	0	0	0	0	0	0	0	3	0	13	0	0	0	0	0	0	16	81.25
工矿用地	0	0	0	0	0	0	0	0	0	0	0	2	0	0	31	0	0	0	0	1	34	91.18
河流	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	36	0	0	0	0	36	100
水库/坑塘	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	23	0	0	0	24	95.83
河湖滩地	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	1	0	5	0	1	8	62.50

地类名称	常绿针叶林	常绿阔叶林	落叶阔叶林	针阔混交林	灌丛	草丛	草甸	灌丛草地	水田	旱地	园地	城镇建设用地	农村居民地	交通用地	工矿用地	河流	水库/坑塘	河湖滩地	裸岩	裸地	汇总	精度(%)
裸岩	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	3	100
裸地	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	4	75

注：总体分类精度=91.19%，Kappa 系数=0.886

在表 4 中，精度是指所解译的某一土地覆被类型的正确数量与该类全部验证样点个数的比值；总体分类精度是指验证样点中，正确的分类数量与全部验证样点数量的比值；Kappa 系数也是精度验证的一种方式，描述的是土地覆被解译的一致性，其值越接近 1，精度越高。基于表 4，建立混淆矩阵，计算数据的总体精度和 Kappa 系数，计算公式如下：

$$k = \frac{p_o - p_e}{1 - p_e} \quad (1)$$

其中 p_o 为实际一致率，即总体精度，是分类正确的总数与样点总数的比值； p_e 为理论一致性，计算公式如下：

$$p_e = \frac{1}{n^2} \sum_{k=1}^N (\sum_{i=1}^N a_{ki} * \sum_{i=1}^N a_{ik}) \quad (2)$$

其中 n 为验证点数， i 、 k 分别混淆矩阵行列数， N 为分类总数。

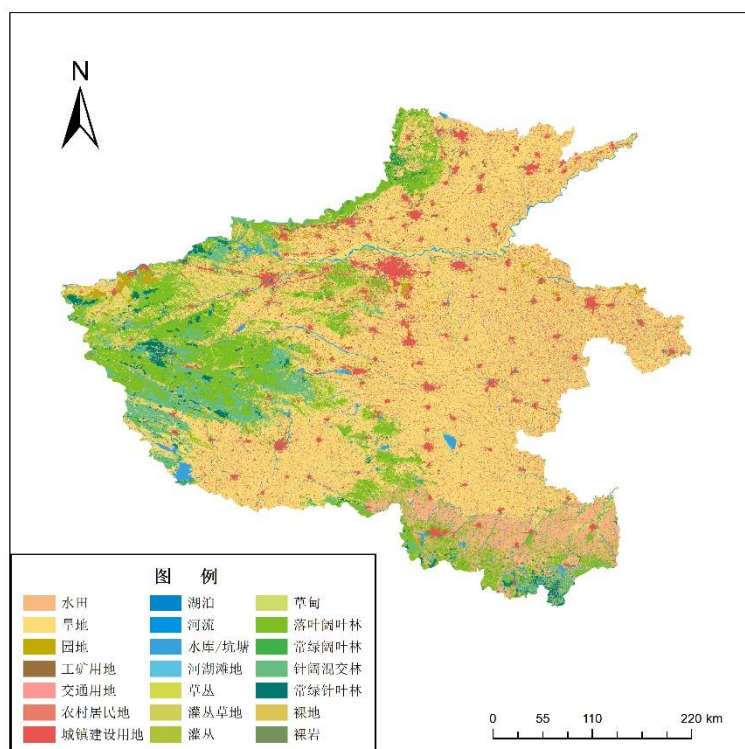
河南省土地覆被解译数据二级分类总体分类精度为 91.19%，Kappa 系数为 0.886。

为了验证数据的可信度，本文分别用中国科学院资源环境数据中心的分辨率为 1 千米的植被类型图和宫鹏团队 10 米空间分辨率的 Sentinel-2 数据生产的 2017 年土地覆被数据（全球整体精度为 72.76%）与本数据进行对比。

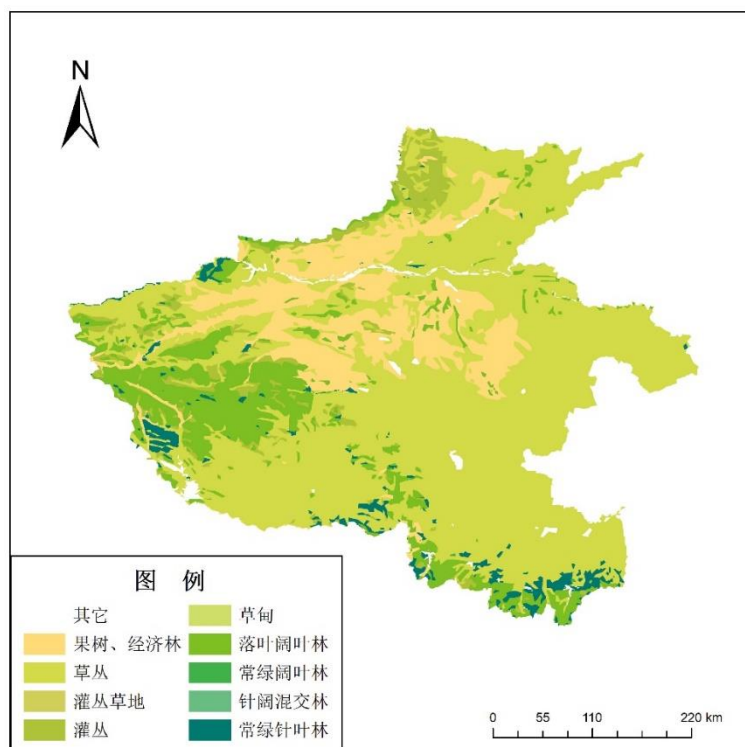
由于植被类型图种类繁多，先将其按照本数据的分类体系进行梳理，然后进行对比。由于两幅图的空间、时间尺度差异较大，且梳理后的分类体系也不完全相同，所以本文只做简单的空间分布对比。图 10 中 a、b 图可知，常绿针叶林、落叶阔叶林、针阔混交林以及常绿阔叶林（河南省主要是竹林）在空间分布上整体一致，分布在河南省的西部以及南部。由于两图的性质不同，本数据中的旱地、水田在植被类型图上显示为草丛，而建设用地在植被类型图上并未体现。逐一对比各类相同类型发现本数据比植被类型图多一些破碎斑块，这主要是由空间尺度差异引起的，但是其整体空间分布具有很好的一致性。

2017 年 10 米的土地覆被数据的分类体系包含耕地、林地、草地、灌丛、湿地、水体、苔地、不透水面、裸地、冰川 10 类一级地类。在河南省没有冰川，只有 9 类，如图 10c 所示。将本文中的数据合并至对比数据的一级分类体系，其中灌丛草地类转为灌丛，然后将数据转换成 10 米的栅格，对两条数据进行一致性评价。尽管目标数据本身的精度、数据源时间、空间分辨率的差异以及分类体系（灌丛草地划分为灌丛导致的误差）等因素会导致分类结果存在一定的误差，但是两个数据的一致性为 76.88%（一致栅格数为 1 274 914 106，总栅格数为 1 658 338 437），从而验证了本数据的可

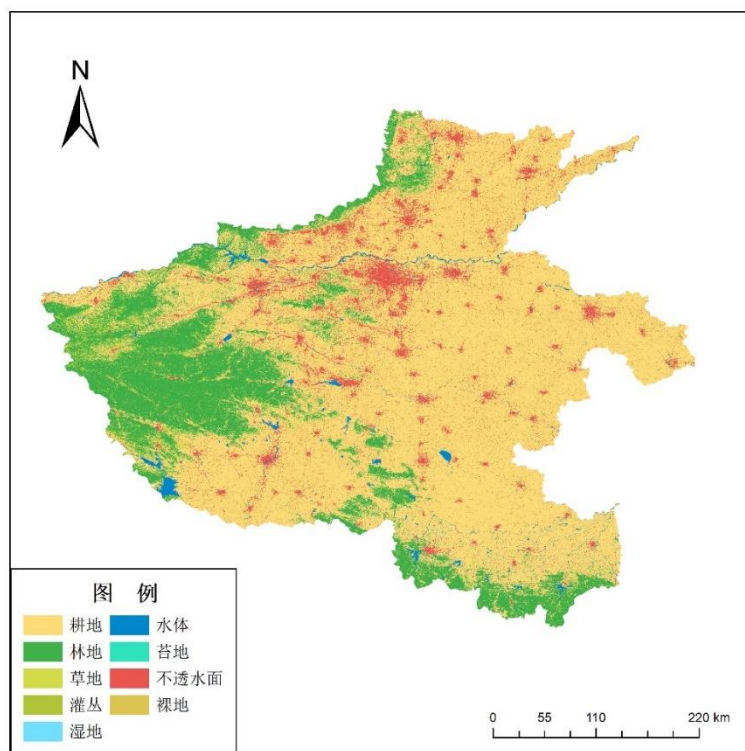
靠性。



(a)



(b)



(c)

图 10 本数据 (a)、植被类型图 (b)、10 米分辨率的土地覆被图 (c) 对比图

4 数据使用方法和建议

本数据集可用于河南省生态系统服务评估、生态环境指数计算、生态系统固碳模拟、土地覆被格局及其水文资源效应研究。本数据集生产耗费大量的人力、物力、财力，具有足够多的野外实测样点，经过多次专家审阅和论证，反复修改纠正，具有较高的精度，可用于检验和评价 2015 年相同区域的其他数据分类精度。本数据集可作为后期土地覆被数据更新的本底数据。本数据集中的野外样点和野外验证数据可作为其他数据集的样本库。数据空间分辨率分别为 50 米和 100 米，适用于中级及以上尺度的研究。

数据作者分工职责

李宁（1987—），女，河南省郑州市人，在读博士研究生，助理实验师，研究方向为土地覆被变化及其生态环境效应。主要承担工作：数据准备、数据采样、数据生产、数据验证。

朱筠（1983—），女，河南省开封市人，在读博士研究生，研究方向为生态环境遥感。主要承担工作：数据准备、数据采样、数据生产、数据验证。

秦奋（1966—），男，河南省开封市人，博士，教授，主要研究方向为地理信息集成分析与共享应用。主要承担工作：组织实施、数据质量控制。

参考文献

- [1] TURNER B L I, MOSS R H, SKOLE D L. Relating land use and global land-cover change: a proposal for an IGBP-HDP core project. A report from the IGBP/HDP Working Group on Land-Use/Land-Cover Change[J]. Global Change Report, 1993.
- [2] 李秀彬. 全球环境变化研究的核心领域: 土地利用/土地覆被变化的国际研究动向[J]. 地理学报, 1996, 51(6): 553-558.
- [3] 史培军, 潘耀忠, 陈晋. 深圳市土地利用/覆盖变化与生态环境安全分析[J]. 自然资源学报, 1999, 14(4): 293-299.
- [4] 陈军, 陈晋, 宫鹏, 等. 全球地表覆盖高分辨率遥感制图[J]. 地理信息世界, 2011(2): 12-14.
- [5] CRAMER W, KICKLIGHTER D W, BONDEAU A, et al. Comparing global models of terrestrial net primary productivity (NPP): Overview and key results[J]. Global Change Biology, 1999, 5(Suppl.1): 1-15.
- [6] FOLEY J A, DEFRIES R, ASNER G P, et al. Global consequences of land use[J]. Science, 2005, 309(5734): 570-574.
- [7] BOUNOUA L, DEFRIES R, COLLATZ G J, et al. Effects of land cover conversion on surface climate[J]. Climatic Change, 2002, 52(1-2): 29-64.
- [8] DOUGLAS I. Hydrological investigations of forest disturbance and land cover impacts in South-East Asia: A review[J]. Philosophical Transactions of the Royal Society B-Biological Sciences, 1999, 354(1391): 1725-1738.
- [9] CHAPIN F S, ZAVALITA E S, EVINER V T, et al. Consequences of changing biodiversity[J]. Nature, 2000, 405(6783): 234-242.
- [10] LAMBIN E F, GEIST H J, LEPERS E. Dynamics of land-use and land-cover change in tropical regions[J]. Annual Review of Environment and Resources. 2003, 28: 205-241.
- [11] GONG P, WANG J, YU L, et al. Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data[J]. International Journal of Remote Sensing, 2013, 34(7): 2607-2654.
- [12] 陈军, 廖安平, 陈晋, 等. 全球 30 m 地表覆盖遥感数据产品-Globe Land30[J]. 地理信息世界, 2017, 24(1): 1-8.
- [13] LI C, GONG P, WANG J, et al. The first all-season sample set for mapping global land cover with Landsat-8 data[J]. Science Bulletin, 2017, 62(7):508-515.
- [14] GONG P, HAN L, ZHANG M, et al. Stable classification with limited sample: transferring a 30-m resolution sample set collected in 2015 to mapping 10-m resolution global land cover in 2017[J]. Science Bulletin, 2019, 64: 370-373.
- [15] 白燕, 冯敏. 全球尺度多源土地覆被数据融合与评价研究[J]. 地理学报, 2018, 73(11): 2223-2235.
- [16] LOVELAND T R, BROWN J, OHLEN D, et al. ISLSCP II IGBP DISCover and SiB Land Cover, 1992–1993[DB/OL]. ORNL DAAC, 2009. DOI:10.3334/ORNLDAAAC/930.
- [17] HOMER CG, DEWITZ J A, YANG L, et al. Completion of the 2011 National Land Cover Database for the conterminous United States-Representing a decade of land cover change information[J]. Photogrammetric Engineering and Remote Sensing, 2015, 81(5): 345-354.
- [18] 黄亚博, 廖顺宝. 首套全球 30 m 分辨率土地覆被产品区域尺度精度评价——以河南省为例[J]. 地理研究, 2016, 35(8): 1433-1446.
- [19] FRITZ S, SEE L, YOU L, et al. The need for improved maps of global cropland[J]. Eos Transactions, 2013, 94(3): 31-32.
- [20] 王慧敏, 曾永年. 青海高原东部土地利用的低碳优化模拟: 以海东市为例[J]. 地理研究, 2015, 34(7): 1270-1284.
- [21] 廖顺宝, 秦耀辰. 草地理论载畜量调查数据空间化方法及应用[J]. 地理研究, 2014, 33(1): 179-

190.

[22] 国家科学基础条件平台-地球系统科学数据共享平台. 2015 年全国 1:10 万土地覆被遥感解译技术规范[S]. 2015.

论文引用格式

李宁, 朱筠, 秦奋. 2015 年河南省土地覆被栅格数据[J/OL]. 中国科学数据, 2019, 4(4). (2019-11-29). DOI: 10.11922/csdata.2018.0097.zh.

数据引用格式

李宁, 朱筠, 秦奋. 2015 年河南省土地覆被栅格数据[DB/OL]. Science Data Bank, 2018. (2019-06-13). DOI: 10.11922/sciencedb.716.

Grid data on land cover of Henan Province (2015)

Li Ning^{1,2}, Zhu Jun^{1,2}, Qin Fen^{1,2*}

1. College of Environment and Planning, Henan University, Kaifeng 475004, China

2. Laboratory of Geospatial Technology for the Middle and Lower Yellow River Regions, Henan University, Kaifeng 475004, China

*Email: qinfen@henu.edu.cn

Abstract: This study draws sources from 86-scene Landsat 8 OLI satellite images with cloud volume less than 10%, high-temporal resolution GF-1 WFV wide-viewing angle multispectral camera, UAV (unmanned aerial vehicle) images, and Google Earth images. According to the the topographical and land cover features of Henan Province, four sample regions with a full range of land cover types were selected for conducting a comprehensive field survey, through which we obtained 1518 samples. Then ArcMap 10.3 was used to visually interpret the sample area (to the secondary type) based on the characteristics of sample points and land cover features. Experts were asked to evaluate and correct the interpretation results. On the basis of the interpretation of sample area, another field survey was conducted in the remaining land of Henan Province, and a total of 4,883 field samples were obtained at last. Finally, 1,895 ground verification points were selected through the global random regular hexagon grid (12 levels) method to evaluate the interpretation accuracy. The results showed that the overall interpretation accuracy of the land cover in Henan Province was 91.19%, and the Kappa coefficient was 0.886.

Keywords: Henan Province; land cover; Landsat 8; visual interpretation; 2015

Dataset Profile

Title	Grid data on land cover of Henan Province (2015)
Data authors	Li Ning, Zhu Jun, Qin Fen
Data corresponding author	Qin Fen (qinfen@henu.edu.cn)
Time range	2015

Geographical scope	Henan Province
Spatial resolution	100 m, 50 m
Data volume	24.84 MB
Data format	Geotiff
Data service system	< http://www.sciencedb.cn/dataSet/handle/716 >
Sources of funding	National Earth System Science Data Sharing Infrastructure, National Science & Technology Infrastructure of China: “Remote sensing interpretation and mapping of 1:100,000 land covers in Henan in 2015”; National Earth System Science Data Sharing Infrastructure, National Science & Technology Infrastructure of China: “Quality verification of the 1:100,000 land cover interpretation in Henan in 2015”.
Dataset composition	This dataset includes land cover raster data of Henan Province, field sampling points, and ground verification points. The dataset contains the following files: (1) “hns_tdfb100m” layer is the 100m resolution land cover raster data of Henan Province in 2015, and the raster value is the classification code. (2) “hns_tdfb50m” layer is the 50m resolution land cover raster data of Henan Province in 2015, and the raster value is the classification code. (3) “yanzhengdian” layer is the ground verification points. (4) “yangdian” layer is the field sampling points.