

# 基于超分辨率制图的中国大陆分省多尺度 土地利用数据集 (2005)

葛咏<sup>1,2\*</sup>, 贾远信<sup>1,2</sup>, 陈跃红<sup>3</sup>, 徐新良<sup>1</sup>, 江东<sup>1</sup>, 张蝶<sup>1,2</sup>

1. 中国科学院地理科学与资源研究所 资源与环境信息系统国家重点实验室, 北京 100101;
2. 中国科学院大学, 北京 100049; 3. 河海大学球科学与工程学院, 南京 210098

**摘要:** 遥感影像的空间分辨率、时间分辨率和光谱分辨率的相互制约导致了遥感影像中普遍存在着混合像元。由于混合像元的存在, 在利用遥感影像获取土地利用时, 结果精度会降低。为解决这一问题, 超分辨率制图的方法被提出, 它是指利用类别比例影像和地物空间结构规律, 确定混合像元内地物类别的空间位置, 得到更高空间分辨率的土地利用图。利用 2005 年全国 1,000 m 分辨率土地利用类别比例影像, 选择基于矢量边界的超分辨率制图方法 (Vectorial Boundary-based Super-resolution Mapping, VBSRM), 分别采用降尺度因子 2、3、4、5、6、7、8、9 和 10, 得到空间分辨率分别为 500 m、333 m、250 m、200 m、167 m、143 m、125 m、111 m 和 100 m 的中国大陆 31 个省级行政单位的多尺度土地利用数据集。以中国科学院资源环境科学数据中心提供的中国大陆各个省份空间分辨率为 30 m 土地利用数据为参考, 得到了各个省级行政单位多尺度土地利用数据集的总体精度及各个类别的制图精度和用户精度。从总体精度来看, 新疆和西藏的总体精度明显优于其他省份, 最高精度能够达到 90%; 山西、陕西、重庆、贵州和福建的总体精度较差, 在 70% 左右。从类别精度来看, 耕地、林地和草地的制图及用户精度较高, 建筑用地精度较低, 未利用土地及水域精度存在明显的省际差异。

**关键词:** 多尺度土地利用图集; 中国大陆; 超分辨率制图; 2005 年

**DOI:** 10.3974/geodp.2018.03.11

## 1 前言

人类不断地发掘和利用地表资源, 以满足我们对高质量生活的追求, 这造成了地球表面形态特征的持续变化。随着生产力的不断提升, 尤其是近两个世纪以来的工业革命和信息革命, 我们改造地表的能力和意愿也愈发强烈, 地表的形态发生了剧烈变化。土地覆被/利用数据是描述地表特征的最直接形式之一, 快速获取一个地区的土地覆被/利用图, 是研究区域乃至全球地表形态变化的基础。遥感科学与技术作为一种能瞬时获取大范围地表图像并提供处理技术的科学技术手段, 完美契合快速获取土地覆被/利用的需求。

传统获取土地覆被/利用产品的方法是通过分类技术 (硬分类) 将遥感影像的像元赋予一种类别。然而在遥感影像获取过程中, 由于传感器自身的局限性以及地表特征的复杂性,

收稿日期: 2018-08-02; 修订日期: 2018-09-16; 出版日期: 2018-09-25

基金项目: 国家自然科学基金 (41471296); 中华人民共和国科学技术部 (2012BAH33B01)

\*通讯作者: 葛咏 L-4781-2016, 中国科学院地理科学与资源研究所, gey@lreis.ac.cn

数据引用方式: [1] 葛咏, 贾远信, 陈跃红等. 基于超分辨率制图的中国大陆分省多尺度土地利用数据集(2005) [J]. 全球变化数据学报, 2018, 2(3): 323-330. DOI: 10.3974/geodp.2018.03.11.

[2] 葛咏, 贾远信, 陈跃红等. 基于超分辨率制图的中国大陆分省多尺度土地利用数据集(2005) [DB/OL]. 全球变化科学研究数据出版系统, 2018. DOI: 10.3974/geodb.2018.05.17.V1.

遥感影像不可避免地同时包含纯像元和混合像元。由于混合像元中不止包含一类地物,通过传统分类法将混合像元赋予一种类别的处理过程,会降低土地覆被/利用产品精度。为解决混合像元中多种地物共存的判别问题,许多学者通过混合像元分解法(或称软分类方法),求出混合像元内各个类别的面积比例,并获取整幅遥感影像的类别比例图。混合像元分解的方法解决了像元内部类别不确定性的局限,但是混合像元内部类别位置的不确定性依然存在。因此,超分辨率制图的概念被提出,它是指利用类别比例数据和地物空间结构特征,确定混合像元内部类别的位置信息。

超分辨率制图作为一种获取土地覆被/利用的方法,它不仅能够同时兼顾软硬分类的优点,而且作为一种尺度转换方法,它能够获取更高空间分辨率的土地利用产品。对于高分辨率影像缺失的地区,它可以使我们从同时期同区域的免费中低分辨率遥感影像中解译得到与高分辨率一致的土地覆被/利用产品<sup>[1]</sup>。因此,本数据集利用超分辨率制图法,以粗分辨率的中国类别比例数据为输入数据,获取更精细的多尺度土地覆被/利用产品,为相关科研工作者提供可供选择的土地利用数据产品。

## 2 数据集元数据简介

基于超分辨率制图的中国大陆分省多尺度土地利用数据集(2005)<sup>[2]</sup>的名称、作者、地理区域、数据年代、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。

## 3 数据采集及研发方法

本数据集的输入数据是空间分辨率为1,000 m的2005年中国土地利用类别比例数据,由30 m的土地利用图通过网格叠加和区域统计获取生成。为验证本数据集的精度,我们通过面积最优的尺度上推法,对30 m的土地利用数据进行升尺度转换,得到与超分辨率制图产品相对应分辨率的土地利用图,作为参考数据。30 m的土地利用数据是由中国科学院资源环境科学数据中心(<http://www.resdc.cn>)提供。

表1 中国大陆分省多尺度土地利用数据集(2005)元数据简表

条目	描述
数据集名称	基于超分辨率制图的中国大陆分省多尺度土地利用数据集(2005)
数据集短名	MLUD_China_Prov2005
作者信息	葛咏 L-4781-2016, 中国科学院地理科学与资源研究所 资源与环境信息系统国家重点实验室, <a href="mailto:gey@lreis.ac.cn">gey@lreis.ac.cn</a> 贾远信 L-4527-2016, 中国科学院地理科学与资源研究所 资源与环境信息系统国家重点实验室, 中国科学院大学, <a href="mailto:jiayx@lreis.ac.cn">jiayx@lreis.ac.cn</a> 陈跃红, 河海大学 地球科学与工程学院, <a href="mailto:chenyh@lreis.ac.cn">chenyh@lreis.ac.cn</a> 徐新良 L-5222-2016, 中国科学院地理科学与资源研究所 资源与环境信息系统国家重点实验室, <a href="mailto:xuxl@lreis.ac.cn">xuxl@lreis.ac.cn</a> 江东, 中国科学院地理科学与资源研究所, <a href="mailto:jiangd@igsnr.ac.cn">jiangd@igsnr.ac.cn</a> 张蝶, 中国科学院地理科学与资源研究所 资源与环境信息系统国家重点实验室, 中国科学院大学, <a href="mailto:zhangd@lreis.ac.cn">zhangd@lreis.ac.cn</a>
地理区域	中国大陆22个省(不包括台湾省、香港、澳门)、4个直辖市和5个自治区

续表

条目	描述
数据年代	2005年
空间分辨率	500 m、333 m、250 m、200 m、167 m、143 m、125 m、111 m、100 m
数据文件个数	310 数据格式 .tif、.xlsx 数据量 335 MB
数据集组成	<p>本数据集按照省级行政区共分为31个部分，命名规则为“MLUD_China(2005)_省级行政单位拼音(首字母大写)”，如MLUD_China(2005)_Anhui表示安徽省在2005年多尺度土地利用数据集。每部分包括9种分辨率的土地利用数据，命名规则为“省级行政单位拼音_分辨率_年份.tif”，如anhui_100m_2005.tif表示安徽省100米土地利用数据。数据格式是GeoTiff。各个省份的数据量跟省份的面积相关，上海市的数据量最小，为0.4 MB；内蒙古自治区的数据量最大，为61.4 MB</p> <p>精度定量评价结果在“精度评价”文件夹中，每个省份有一个.xlsx格式的精度评价表，精度评价内容包括每个数据集的总体精度和各个类别的用户及制图精度</p> <p>数据集采用中国土地资源分类系的一级类别分类体系<sup>[3]</sup>，共包括6类，分别是耕地、林地、草地、水域、建设用地和未利用土地</p>
基金项目	国家自然科学基金（41471296）；中华人民共和国科学技术部（2012BAH33B01）
出版与共享服务平台	全球变化科学研究数据出版系统 <a href="http://www.geodoi.ac.cn">http://www.geodoi.ac.cn</a>
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	<p>全球变化科学研究出版系统出版的“数据”包括元数据、实体数据和数据论文。数据共享政策如下：（1）数据（元数据、实体数据、数据论文）以最便利的方式免费向全社会开放，用户免费浏览、免费下载；（2）用户使用数据需要按照引用格式标注数据来源；（3）用户以任何形式散发数据需要与中国科学院地理科学与资源研究所签署书面协议；（4）用户应用部分数据记录产生新数据需要遵循10%原则<sup>[4]</sup></p>

本数据集的研发方法采用Ge等<sup>[5]</sup>提出的基于矢量边界的超分辨率制图法（Vectorial boundary-based super-resolution mapping method, VB SRM）。该方法利用混合像元及邻域像元的类别比例信息，计算像元内部类别的矢量边界，然后确定混合像元内部像元类别及位置。主要包括两个步骤，第一，提取矢量边界：在获取输入数据之后，根据公式（1）和类别比例数据，首先计算矢量边界线段长度及位置，随后连接线段形成初始多边形，然后对其进行拓扑调整，最终获取各类别地物的矢量多边形；第二，确定像元类别属性：设定放大尺度因子，用射线法判断混合像元内亚像元中心点与矢量多边形的位置关系，继而判断每个亚像元的类别属性，得到亚像元（代表更高分辨率）级别的地物类别图。

$$V_{itc} = V_{itc}^{(0)} + \begin{cases} (1 - P_{ijc}) \frac{P_{(i+1)c}}{P_{(i+1)c} + P_{(i-1)c}}, P_{(i+1)c} \neq P_{(i-1)c} \\ (1 - P_{ijc}), P_{(i+1)c} = P_{(i-1)c} \end{cases} \quad (1)$$

式中， $V_{itc}^{(0)}$  和  $V_{itc}$  是线段  $t (t=1, 2, \dots, 10)$  对应于邻域  $i (i=1, 2, \dots, N \times M)$  中类别  $c (c=1, 2, \dots, C)$  的开始和结束位置。 $P_{ijc}$ 、 $P_{(i+1)c}$  和  $P_{(i-1)c}$  分别是像元  $i$  及其左右像元  $i+1$  和  $i-1$  中类别  $c$  的面积比例信息。本数据集共选取了9种降尺度因子，分别是2, 3, 4, ..., 10，得到了对应空间分辨率为500 m, 333 m, 250 m, ..., 100 m的中国大陆分省多尺度土地利用数据集。

## 4 数据结果与验证

### 4.1 数据集组成

2005年中国多尺度土地利用数据集由31个中国大陆省级行政单位组成，每部分包括9种空间分辨率的土地利用数据，9种分辨率分别为500 m、333 m、250 m、200 m、167 m、143 m、125 m、111 m和100 m。土地利用图采用.tif的存储格式，图像中填充值表示土地类型，其中1表示耕地，2表示林地，3表示草地，4表示水域，5表示城乡工矿居民地，6表示未利用土地。由于本数据集中文件较多，本文仅以各个省级行政区250 m的制图结果为例进行展示（图1，图2）。

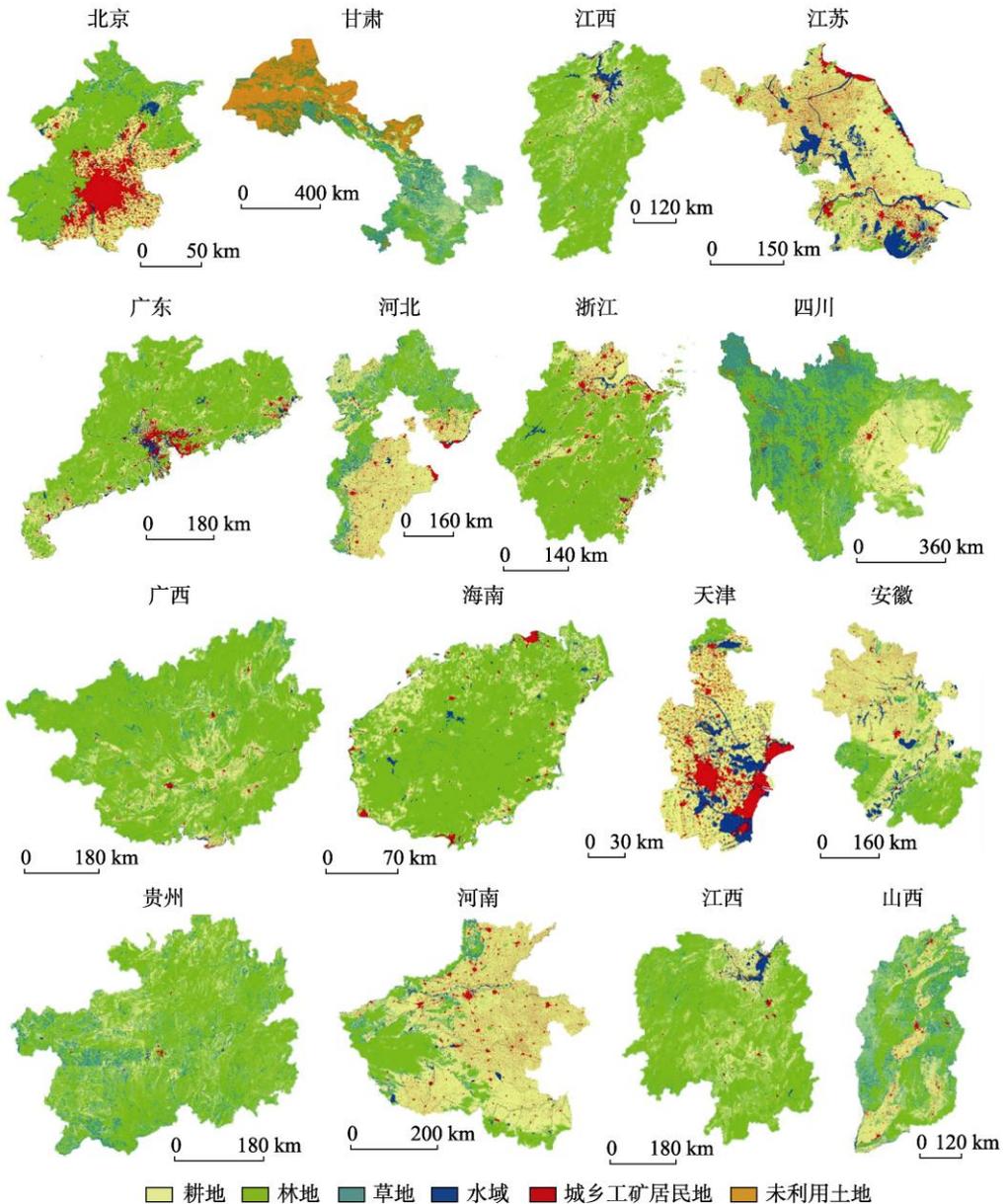


图1 2005年中国大陆分省土地利用图（空间分辨率250 m）（A）

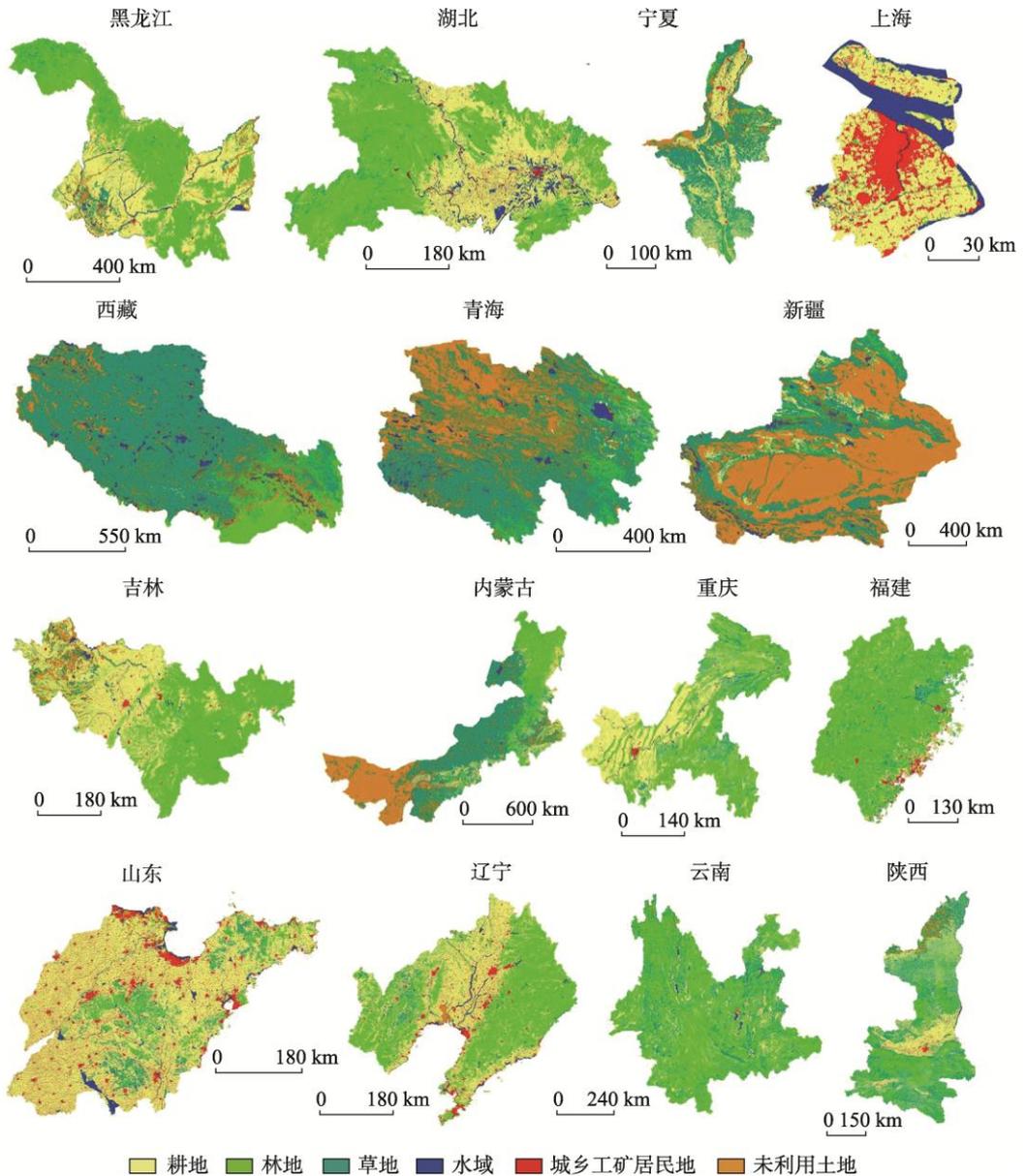


图 2 2005 年中国大陆分省土地利用图（空间分辨率 250 m）(B)

#### 4.2 数据结果统计

对于每个省份 9 种尺度的土地利用数据，按照土地覆被的类别，分别获取了 31 个省份的各种土地覆被的平均面积（表 2）。从土地覆被类别的角度来看，在所有省份中，黑龙江省具有最大的耕地面积，面积为  $167,402.2 \text{ km}^2$ ，四川、内蒙古、山东、河南和河北的耕地面积也超过了  $10 \times 10^4 \text{ km}^2$ ；云南的林地面积为  $226,098 \text{ km}^2$ ，黑龙江的林地面积为  $200,211 \text{ km}^2$ ，是所有省份中林地面积最多的两个省份；西藏自治区的草地面积为  $844,232 \text{ km}^2$ ，位居各个省份之首，新疆维吾尔自治区、内蒙古自治区和青海省的草地面积都超过了  $35 \times 10^4 \text{ km}^2$ ；西藏自治区水域面积最大，为  $52,738 \text{ km}^2$ ；新疆维吾尔自治区的未利用土地面积最大，

表2 各个省份土地利用产品的面积 (单位: km<sup>2</sup>)

省份	耕地	林地	草地	水域	建设用地	未利用土地
安徽	86,279.1	33,194.8	7,504.0	7,006.2	6,308.4	3.7
北京	4,502.2	7,651.8	1,114.9	390.5	2,741.3	0.3
重庆	39,111.8	33,165.7	8,770.3	808.3	573.8	8.1
福建	20,291.3	80,659.5	16,158.6	1,434.1	3,032.6	65.9
甘肃	68,978.5	37,995.8	141,165.0	3,172.2	2,156.7	172,371.8
广东	40,795.4	115,026.5	3,538.2	7,307.1	9,328.8	84.5
广西	52,858.8	161,677.3	15,802.8	3,013.3	3,060.1	18.3
贵州	47,320.9	98,171.7	29,664.4	340.3	479.9	24.8
海南	6,851.9	24,993.1	507.1	815.3	696.1	62.5
河北	102,157.4	40,630.1	28,768.2	3,371.6	11,296.5	1,400.5
黑龙江	167,402.2	200,211.0	33,258.5	11,471.6	6,538.2	34,083.6
河南	113,384.0	29,424.2	7,905.4	3,446.8	11,446.6	12.6
湖北	70,072.0	95,792.8	6,263.3	9,868.7	3,646.8	387.9
湖南	59,652.2	136,964.1	6,322.5	6,283.1	2,131.7	663.0
江苏	73,495.8	3,093.3	788.0	12,791.8	11,795.8	13.7
江西	44,832.6	107,508.1	6,063.5	6,298.6	1,791.9	595.1
吉林	79,142.0	84,946.8	6,771.4	3,809.4	4,456.0	11,887.9
辽宁	65,086.0	62,561.0	4,378.9	4,463.3	8,122.7	1,450.5
内蒙古	118,514.0	182,216.0	475,450.1	12,613.5	8,981.2	348,019.8
宁夏	18,432.0	2,561.0	24,081.1	789.9	793.5	5,111.7
青海	8,651.0	27,520.3	379,391.6	26,897.3	915.1	253,515.8
山东	113,647.0	9,730.2	8,639.7	6,037.4	16,214.5	1,958.4
上海	4,385.8	89.5	14.5	1,632.1	1,798.7	0.0
山西	62,163.0	44,412.0	44,892.4	1,419.4	3,696.1	84.0
陕西	73,578.2	47,711.4	76,190.1	1,557.9	2,153.0	4,695.4
四川	123,857.1	169,004.6	170,945.7	3,212.5	2,550.3	16,446.8
天津	7,108.9	378.5	98.1	1,712.0	2,277.4	36.0
新疆	68,886.6	35,283.9	480,205.4	45,596.3	4,005.0	997,680.9
西藏	5,464.6	125,230.1	844,232.1	52,738.9	153.8	174,684.3
云南	66,457.6	226,098.1	84,287.2	3,001.8	1,404.6	2,034.7
浙江	27,198.6	67,073.9	1,560.3	2,811.3	4,815.7	47.4

为 997,680 km<sup>2</sup>, 内蒙古自治区、青海省、西藏自治区、甘肃省的未利用土地面积均超过了 17 × 10<sup>4</sup> km<sup>2</sup>; 山东省、江苏省、河南省、河北省的建设用地面积超过了 10 × 10<sup>4</sup> km<sup>2</sup>, 上海市、天津市、北京市、江苏省和山东省具有较高的建设用地比例。

### 4.3 数据验证

本数据集采用总体精度和类别的用户及制图精度, 以30 m全国土地利用数据为参考数据, 获取了省级行政区的9种分辨率土地覆被的总体精度和类别精度, 并对9种分辨率的总体精度和类别精度求平均值, 得到了31个省级行政单位的平均精度评价结果(表3)。由表3可以看出, 就总体精度而言, 新疆和西藏的总体精度明显优于其他省份, 最高精度能够达到90%, 山西、陕西、重庆和贵州的总体精度较差, 在70%左右; 就类别精度而言, 耕地、

表3 各个省份土地利用产品的平均精度（%）

	西藏	新疆	内蒙古	浙江	江苏	吉林	四川	青海
总体精度	92.0	90.0	76.4	79.7	80.7	79.5	75.8	82.1
耕地	47.0	81.2	64.8	68.3	87.6	80.7	79.4	67.1
林地	79.1	60.0	78.1	89.1	68.3	87.7	74.0	59.7
草地	95.2	88.1	78.6	38.0	58.8	42.3	77.3	84.8
水域	84.1	73.7	59.8	50.5	82.7	55.2	47.0	73.3
建设用地	46.1	52.1	33.0	50.9	48.6	33.0	49.2	56.1
未利用土地	80.2	94.9	71.9	22.4	29.3	67.9	61.5	82.3
	上海	湖北	广西	广东	甘肃	河南	安徽	江西
总体精度	83.6	78.3	77.8	76.2	77.8	79.6	79.3	77.1
耕地	86.3	76.7	65.0	63.1	65.4	86.7	85.6	65.4
林地	49.9	85.1	86.8	86.5	68.9	84.6	82.0	85.6
草地	59.9	52.7	52.7	37.6	72.1	59.0	63.9	47.7
水域	90.8	65.0	40.5	56.5	46.5	52.8	73.9	67.9
建设用地	73.3	33.6	37.5	57.4	33.6	38.8	31.0	36.3
未利用土地	无数据	53.9	31.7	30.9	92.3	29.8	15.6	70.9
	天津	黑龙江	湖南	河北	云南	山东	宁夏	辽宁
总体精度	75.7	76.5	75.1	74.1	74.8	72.9	73.0	71.1
耕地	84.4	83.1	64.0	83.0	61.5	84.2	73.1	75.1
林地	73.7	86.4	83.2	76.4	83.5	59.5	57.0	78.1
草地	44.7	31.9	47.6	60.4	64.1	45.1	76.8	31.9
水域	66.2	54.0	60.5	44.3	58.1	52.5	49.4	48.1
建设用地	60.3	35.7	38.9	45.6	36.7	45.0	39.9	41.3
未利用土地	36.4	43.9	68.9	49.0	72.4	25.5	73.5	67.1
	海南	贵州	福建	重庆	北京	山西	陕西	
总体精度	71.6	68.4	69.6	69.6	76.8	67.7	63.3	
耕地	52.9	54.5	53.1	74.1	71.5	70.3	63.8	
林地	82.4	77.5	81.0	71.3	88.1	75.3	69.2	
草地	33.1	63.5	44.5	50.5	53.7	60.0	60.2	
水域	46.8	38.1	45.4	53.6	51.4	36.7	47.8	
建设用地	46.6	47.2	49.7	54.6	68.8	43.8	36.8	
未利用土地	33.9	41.3	33.5	24.1	3.2	33.3	67.0	

林地和草地的制图及用户精度较高，建筑用地精度较低，未利用土地及水域精度存在明显的省际差异。Ge等研究发现，随着地表异质性和空间分辨率的增加，土地利用产品的精度会降低<sup>[6]</sup>，本文的精度评价结果与之类似。西藏和新疆两地区存在大范围的同质地物类型，（新疆主要以未利用土地为主，西藏以草地为主），因此这两个地区的总体精度最高；山西、陕西、贵州、重庆和福建五个省份，由于位于地表易破碎的地区，地表沟壑纵横，异质程度高，总体精度相对较低。耕地、林地和草地多以大斑块地物出现，地表类型相对均一，类别精度较高。建筑用地尤其是农村用地，多以散点形状分布，并与其他地物混合出现，地表异质性较高，类别精度相对较低。不同省份类水域及未利用土地的分布形态差异较大。因此，两个类别在不同省份之间存在较大精度差异，如未利用土地在新疆地区呈现

状地物分布, 异质性较低, 类别精度能高达96%; 而在浙江地区多以线状地物或点状地物分布, 破碎程度高, 类别精度较低, 最低只有20%左右。

## 5 讨论和结论

土地覆被/利用作为诸多地学模型与应用的研究基础, 快速获取该产品是一项重要的需求。目前获取土地覆被产品的常规法是解译遥感影像, 但遥感影像中混合像元的存在影响了土地覆被产品的精度。在此背景下, 超分辨率制图方法被提出, 它是指利用类别比例影像和地物空间结构, 判断混合像元内地物位置, 得到更高空间分辨率的土地利用图。本数据集采用超分辨率制图方法, 对1,000 m空间分辨率的全国土地利用类别比例数据进行降尺度, 得到了空间分辨率分别为500 m、333 m、250 m、200 m、167 m、143 m、125 m、111 m和100 m的全国土地利用数据集, 并对数据集按省份进行了划分。我们以30 m空间分辨率的土地利用产品为参考, 按照省份对各个子数据集进行了精度评价。通过精度评价的结果可以看出, 全国多尺度土地利用数据集与参考数据保持了较好的一致性。本数据集中包含精度评价结果, 便于用户使用者按精度与地区进行查询, 以满足其特定需求。

**致谢:** 本数据集中用到的2005年中国30 m土地利用数据来自中国科学院资源环境科学数据中心, 在此表示感谢。

## 参考文献

- [1] 江昱, 葛咏, 陈跃红等. 亚像元制图适应性分析与评价——以天津市津南区和北京市海淀区土地覆被制图为例[J]. 地球信息科学, 2015, 17(10): 1215-1223.
- [2] 葛咏, 贾远信, 陈跃红等. 基于超分辨率制图的中国大陆分省多尺度土地利用数据集(2005) [DB/OL]. 全球变化科学研究数据出版系统, 2018. DOI: 10.3974/geodb.2018.05.17.V1.
- [3] 江昱. 亚像元制图放大因子适应性分析与评价[D]. 北京: 中国科学院大学, 2015.
- [4] 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).
- [5] Ge, Y., Chen, Y., Li, S., *et al.* Vectorial boundary-based sub-pixel mapping method for remote-sensing imagery [J]. *International Journal of Remote Sensing*, 2014, 35(5): 1756-1768.
- [6] Ge, Y., Jiang, Y., Chen, Y., *et al.* Designing an experiment to investigate subpixel mapping as an alternative method to obtain land use/land cover maps [J]. *Remote Sensing*, 2016, 8: 360. DOI: 10.3390/rs8050360.