

# 中国新增 1598 个传统村落空间分布数据集

余亮<sup>1\*</sup>, 丁雨倩<sup>1</sup>, 唐铭婕<sup>1</sup>, 曹倩颖<sup>1</sup>, 付蒙<sup>1</sup>, 杨旭辉<sup>1</sup>, 沈军鑫<sup>2</sup>

1. 苏州大学建筑学院, 苏州 215123; 2. 同济大学浙江学院, 嘉兴 314000

**摘要:** 国家住房和城乡建设部、文化部和财政部继 2012 年至 2014 的三次公布 2555 个中国传统村落名录后, 于 2016 年 12 月公布了第四批 1598 个中国传统村落名录。目前总计 4,153 个村落纳入到中国传统村落名录。中国新增 1598 个传统村落空间分布数据集参用了“中国 2555 个中国传统村落空间分布数据集”研发方法, 空间位置数据的坐标提取参考了百度地图和 Google Earth 影像图, 依据村落名称的文字提示、村落影像图形的几何中心或者易辨认的村落位置。对于非常规出现且难辨认的村落名称, 采取查询相关资料、或者与该村进行直接联系等一些确认措施, 来确定村落的名称和位置。对于地图和影像上无该村落信息的情况, 则以村落所辖的上一级行政单位的坐标为该村落坐标。最后, 获取中国 1,598 个传统村落空间分布数据集, 数据集存储为.shp 和.kmz 格式, 由 7 个数据文件组成, 数据量为 4.50 MB(压缩为 2 个文件, 258 KB)。

**关键词:** 中国; 传统村落; 名录; 空间分布

**DOI:** 10.3974/geodp.2019.02.05

## 1 前言

传统村落是宝贵的历史文化遗产, 是人类早期的居住与空间聚合形式和载体, 遍布世界各地且地域特色明显, 作为生活必要的空间形态, 是人类尊重自然, 与自然和谐共生理念在乡村格局的应用体现, 由于人类的活动需要, 传统村落(以下简称村落)中穿插着大量的传统民居, 民居和村落不能分离, 有着水滴与水塘的密切关系, 图 1 显示的是笔者(第一作者)调查过的浙江温州泰顺县筱村镇徐岙村和北京市房山区南窖乡南窖村的民居建筑, 两者相距遥远, 山区环境, 建筑的建造有共性, 均使用了大量的当地石材资源, 审美别致, 展现了和谐有序的村落空间格局。以此可知, 分析比较村落空间时, 不能缺少特有民居建筑的形态思考, 通过比较不同村落的民居建筑, 可进一步梳理村落宏观层面的空间分布特性。刘大均等<sup>[1]</sup>和康璟瑶等<sup>[2]</sup>分别对住房和城乡建设部先期公布的 3 批共计 2,555 个村落作过村落的空间分布研究。中国新增 1598 个传统村落空间分布数据集, 是分析解读村落地理位置及空间分布特点的基础数据, 是继前次中国 2555 个村落空间分布数据集<sup>[3-4]</sup>的再次整理汇总, 能为其他研究提供从省市到村落层级较准确完整的行政隶属关系和称法, 村落既是空间领域, 又是与自然相适应的乡村劳作组织和生活场景记录, 通过系统地空间数据表

**收稿日期:** 2019-04-24; **修订日期:** 2019-06-05; **出版日期:** 2019-06-25

**基金项目:** 国家自然科学基金(41371173)

**\*通讯作者:** 余亮 F-8099-2018, 苏州大学建筑学院, yuliang\_163cn@163.com

**数据引用方式:** [1]余亮, 丁雨倩, 唐铭婕等. 中国新增 1598 个传统村落空间分布数据集[J]. 全球变化数据学报, 2019, 3(2): 155-170. DOI: 10.3974/geodp.2019.02.05.

[2] 余亮, 丁雨倩, 唐铭婕等. 中国新增 1598 个传统村落空间分布数据集[DB/OL]. 全球变化科学研究数据出版系统, 2019. DOI: 10.3974/geodb.2019.01.19.V1.

达与比较应用，可以更好地展现优秀的乡村文化沉积，为乡村振兴的内涵提升发挥应有的数据作用。



图 1 地域特色鲜明的传统村落（左：浙江温州泰顺县筱村镇徐岙村；右：北京市房山区南窖乡南窖村）

2 数据集元数据简介

中国新增 1598 个传统村落空间分布数据集<sup>[5]</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。

表 1 中国新增 1598 个传统村落空间分布数据集元数据简表

条 目	描 述
数据集名称	中国新增 1598 个传统村落空间分布数据集
数据集短名	VillagesChina1598
作者信息	余亮 F-8099-2018, 苏州大学建筑学院, yuliang_163cn@163.com 丁雨倩 O-6445-2018, 苏州大学建筑学院, 442801422@qq.com 曹倩颖 O-6478-2018, 苏州大学建筑学院, 604459330@qq.com 唐铭婕 O-6467-2018, 苏州大学建筑学院, 361988267@qq.com 付蒙 O-6455-2018, 苏州大学建筑学院, 821064405@qq.com 杨旭辉, 苏州大学建筑学院, 654712015 @qq.com 沈军鑫, 同济大学浙江学院, 409209740 @qq.com
地理区域	中国, 覆盖 31 个省级行政区（不包括港澳台）。
数据年代	2012–2014 年
数据格式	.shp、.kmz
数据量	4.50 MB
数据集组成	2 个文件（VillagesChina1598.kmz + VillagesChina1598.rar）
基金项目	国家自然科学基金（41371173）
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	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 <sup>[6]</sup>

3 村落空间点数据处理

村落的空间点数据非信手拈来之物，公布的村落名录中不显示地理坐标，有时也难辨别村落名关联的行政隶属关系。从分析空间中的村落形体可知，村落由人集聚而成，既是空间形态，又是社会形式，在村落空间的认知范围内，除植物、水塘河流等自然物体外，最大和最突出的物体应属民居建筑物。民居作为人类起居生活工作必需的场所，有遮蔽隔绝外界且易辨认的空间（主要是屋顶）突起，因而容易成为村落空间点数据的获取线索。在村落坐标获取上，利用名录自带的村落名称信息，通过辨别梳理信息特征，以及空间定位辨认确定坐标位置。坐标的获取原则，方法上与前次基本相同，尽管有些地方的对应处理不同<sup>[4]</sup>，图 2 是村落点数据确定的思考与程序步骤。

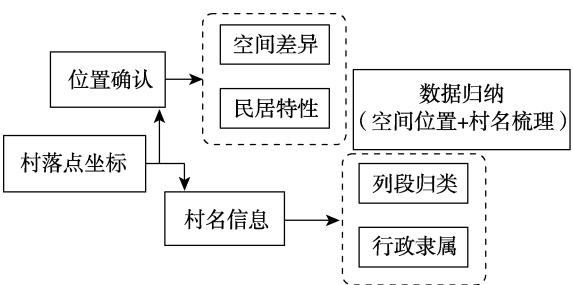


图 2 村落点数据确定的思考与程序步骤

3.1 数据集范围和村落原始数据

数据集范围：覆盖中国 31 个省级行政区（不包括港澳台），村落原始数据：“第四批列入中国传统村落名录的村落名单”的 1598 个村落（2016 年公布，继 2012 年起国家住房和城乡建设部、文化部、财政部等公布三批全国传统村落名录基础上的第四批次公布）。名录的村落总数已达到 4153 个，从各批次中国传统村落名录的村落数中可见，第一批到第四批有明显的数量递增趋势（图 3）。

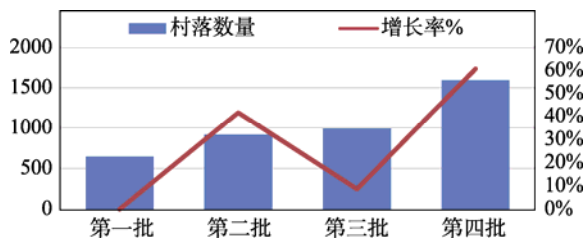


图 3 第一批至第四批村落数量

3.2 村落点数据获取和村落名

村落空间点的数据获取包括：一是空间认知，关注空间中突起易识别的个体，确认建筑或民居的空间存在，并区分与自然物体的差异，获取村落在广域空间的点位置；二是村名辨别，确定村落与上级部门的行政隶属关系，以归入不同的数据列段。方法上，先利用百度地图获取村落空间的点数据坐标（经纬度），考虑到百度地图和 Google Earth 的位置差异，特别是数据集需在 Google Earth 的全球共享，在百度地图数据获取的基础上，对 Google Earth 的线上影像进行村落位置的手动目测微调，图 4 是浙江台州市临海市河头镇殿前村的微调手艺，调整前 Google Earth



图 4 依据 Google Earth 影像的村落位置微调（浙江台州市临海市河头镇殿前村）

的线上位置被标示在田间，调整后挪动到了建筑上。有关 Google Earth 的线上影像，帅菲等<sup>[6]</sup>指出其在地学领域较广泛的应用。而对非常规出现且难辨认的村落名称，采取查询相关资料、或者与该村进行直接联系等一些确认措施，确定村落的名称和位置，除此的一些民居及建筑上的空间定位，基本参照先前方法<sup>[4]</sup>。全部数据检查判断没有问题后形成.kmz 格式后，转成.shp 格式。

3.3 村落名称整理

村落名称是记载村落线索的文字信息，原始公布的村落名，如北京市房山区南窖乡南窖村，一串文字不仅表示了村落名称，同时还显示出直属的行政隶属关系，从村落到省级层次，一般有：省—>市—>县—>镇—>村的隶属架构，这些隶属关系需要根据各自的行政级别分别在不同的列段表示，如北京市房山区南窖乡南窖村，因是直辖市，可分为市、区、乡和村的 4 级，而在一般省，如河北省邯郸市涉县固新镇原曲村，则需省、市、县、镇和村的 5 级，为使行政隶属的列段数统一，本数据集以 5 列段为公约数，因而有的列段会重复，如直辖市北京，必须省和市级的名称相同，不能在北京市下还有市级存在，类似的还有，天津市蓟县蓟州区下营镇黄崖关村等，此现象在直辖市中较常见。

其次，涉及到具体的级别文字表述，列段的隶属关系有不易辨别的特点，有如下情况：1) 名称与功能的结合，名录中有“郊区”标记，所谓“郊区”，既是由多个行政区域组成的地域范围和为城市提供农副产品的功能定位，又是称法，这里仅有“铜陵市郊区大通镇和悦村”的 1 个标记，前 3 批村落共有 5 个，“郊区”出现在“市”级后，作用相当于县，归类在“县”段中；2) 行政单位名称的字位数，以“县”为例，一般为两字，如河北省张家口市怀安县的“怀安”，也有单字，同为河北省张家口市蔚县的“蔚”，使用单字，由中文转英文时，容易影响理解，以行政单位名为单字，且为“县”级时保留行政单位称法，列段直接使用蔚县，译为 Yuxian；3) 村落不仅是基层的村级组织，也是村落名录的标记尾部，但不一定以“村”为标记，由于我国多民族和多气候类型等特点，使得村级的称法多而丰富，南方或少数民族有用“寨”、“组”、“屯”的，如桂林市龙胜各族自治县江底乡建新村委矮岭红瑶组的“组”，北方或少数民族的“查”（均未细究用词的合理性），如巴彦淖尔市乌拉特后旗新忽热苏木莫仁嘎查的“查”等，经归纳可知 1598 个村落中有村、社区、居委会、村委会、台、屯、寨、组和查的多种表述（表 2）。其中，“村”级中的“社区”、“居委会”表述，一般认为已非乡村，而是城市范畴，住民大多从事非农业劳动，有明显的服务城市性质，这些亦可通过空间距离得到进一步验证，带“社区”和“居委会”标记的“村落”，往往离市区较近，固有的村落形态易发生变化，如笔者调查过的，四川省成都

表 2 不同行政级别组织的内含表示

级次	简称	内含及表示
1	省	自治区、直辖市、特别行政区
2	市	地区、州、盟及直辖市辖区和县
3	县	市辖区、县级市、旗
4	镇	乡、镇、街道
5	村	村、社区、居委会、村委会，台、屯、寨、查



市龙泉驿区洛带镇老街社区，因离成都市近，商业活动多而发达，感觉已非传统的村落概念（图 5（左））。同样还有，福建省泉州市晋江市新塘街道梧林社区，这片明朝开始的古村落，林立着闽南大厝、西洋楼、番仔楼等建筑，珍藏着百年古井、盘根老树等，看到这些，就难以想象还属“农耕文化”的村落（图 5（右））。第 4 批 1598 个村落中带“社区”标记的 22 个，带“居委会”的 1 个。



图 5 城市周边的“村落”社区

（来源：左：四川省成都市龙泉驿区洛带镇老街社区，<https://image.baidu.com/search/>；右：福建省泉州市晋江市新塘街道梧林社区，[http://qz.fjsen.com/2017-08/07/content\\_19900355\\_all.htm](http://qz.fjsen.com/2017-08/07/content_19900355_all.htm)）

## 4 数据集构成及村落空间分布

### 4.1 数据集构成

整理完成后的数据集由 2 个子集构成，分别是中国新增的 1598 个传统村落空间分布点数据的 ArcGIS 的.shp 版以及谷歌的.kmz 版（名称分别为：VillagesChina1598.kmz + VillagesChina1598.rar），共包括 7 个数据文件，数据大小为 4.50 MB（压缩为 2 个文件，258 KB）。

### 4.2 村落空间分布

在 Google Earth 中呈现的第四批 1598 个村落点数据及空间分布见图 6，由图可知村落在地理空间的分布并不均匀——村落主要集中在东南部，西北和东北均少，而一些区域则无村落。村落数量超 100 个的省市有 7 个，其中浙江最多，228 个，约占总数的 15%，湖南有 166 个和山西有 150 个，最少的是西藏、宁夏和新疆等省市，均少于 10 个。由此可见村落的分

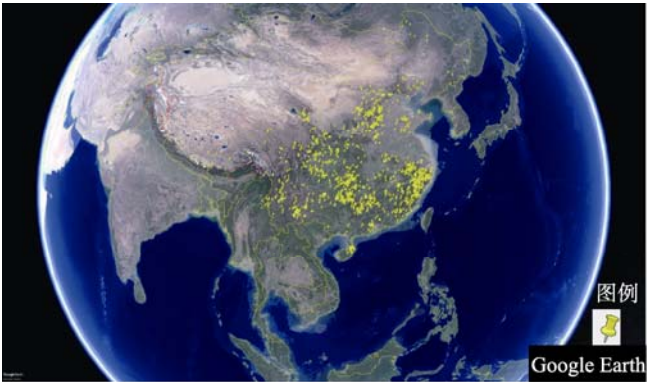


图 6 第四批 1598 个村落的空间分布（Google Earth）

布状态基本与人口的数量分布吻合,分布密度高的区域说明自然资源较丰富,且气候,地理条件等都较适宜人类居住和生活,易使人集中,从而衍生出种类多样且适应当地自然人文环境的,富有各地特色的村落与民居。人口密度低的则反之。

## 5 讨论和总结

在数据集的村落点坐标获取上,利用了百度地图和 Google earth 影像<sup>[7]</sup>,通过坐标提取使村落得到明确的空间位置,利用该数据集能够了解该批次中国传统村落的空间分布关系。其次,第四批 1598 个村落中占比最大的是浙江省,入选的村落较多,足以说明当地百姓对守住自己“祖业”有根深蒂固的意识基础,加之地方政府保护工作的重视和一贯性,是“绿水青山就是金山银山”生态发展理念在浙江的成功贯彻和实践的最好证明,随着我国社会经济水平的不断提升,政府通过对村落及民居保护的关注和重视,增加了国民对传统文化的认同感和自信心,一定程度上遏制了老宅老村拆旧造新的行为,使民居和村落消失的速度得以放缓。湖南省和山西省紧跟其后,保护工作也很有效。相反,上海市已连续 3 批次没有村落被收录,说明城市化的影响较强,而减少了村落的存量,故被收录的村落数量少了。类似现象同样在天津也有所显现,可见如何在城市化进程及传统民居和村落保护的数量中寻求平衡,是一个永恒的话题,也检验着政府、民众的保护施策能力和智慧。由于村落的生成与发展受太多的自然、社会地理条件制约,情况复杂,因而上述通过村落数据讨论的内容仅是众多村落现象表现的一部分,村落细部和干系包罗万象,而太多的奥妙隐藏于数据之中,要想真正地解码释意,后续还需更多更广的数据以细究思考。

**作者分工:** 余亮对数据集的获取开发做了总体设计,并撰写了数据论文等;丁雨倩,和唐铭婕整理了关键的数据;曹倩颖,付蒙,杨旭晖和沈军鑫等采集和处理了本数据集。

## 参考文献

- [1] 刘大均,胡静,陈君子等.中国传统村落的空间分布格局研究[J].中国人口·资源与环境,2014,24(4): 157-162.
- [2] 康璟瑶,章锦河,胡欢等.中国传统村落空间分布特征分析[J].地理科学进展,2016,35(7): 839-850.
- [3] 余亮,刘佳,丁雨倩等.中国 2555 个传统村落空间分布数据集[DB/OL].全球变化科学研究数据出版系统,2018. DOI: 10.3974/geodb.2018.04.06.V1.
- [4] 余亮,刘佳,丁雨倩等.中国 2555 个传统村落空间分布数据集[J].全球变化数据学报,2018,2(2): 144-150. DOI: 10.3974/geodp.2018.02.03
- [5] 余亮,丁雨倩,唐铭婕等.中国新增 1598 个传统村落空间分布数据集[DB/OL].全球变化科学研究数据出版系统,2019. DOI: 10.3974/geodb.2018.04.06.V1.
- [6] 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).
- [7] 帅菲,肖根如,揭志强. Google Earth 在地学研究中的应用[J]. 东华理工大学学报(自然科学版), 2011, 34(1): 89-93.