



文献 DOI:
10.11922/csdata.170.2015.0002

文献分类: 地球科学

收稿日期: 2015 年 7 月 16 日
发表日期: 2016 年 3 月 31 日

中国 27 省乡镇(街道)级人口密度数据集(2000 年)

柏中强^{1,2}, 王卷乐^{1,3*}

摘要: 人口密度是表现人口分布最主要的形式和衡量人口分布地区差异的主要指标。乡镇(街道)处于我国行政体系的基层位置,也是中国人口普查数据公开发布的最小行政单元,乡镇级人口密度数据能够客观、精细地表征中国人口分布的空间格局和态势,对于资源、环境及人口问题的研究具有十分重要的意义。本文规范化处理了中国辽宁、吉林、内蒙古(部分地区)、北京、天津、上海、河北、河南、陕西、宁夏、山西、山东、安徽、江苏、湖南、湖北、江西、浙江、福建、广东、海南、云南、贵州、四川、重庆、青海、西藏这 27 个省(直辖市、自治区)乡镇(街道)级行政界线数据及第五次人口普查乡镇(街道)级人口统计数据,保证乡镇边界和县(区)界一致、各乡镇单元空间范围和人口普查信息一一对应。在此基础上,准确地匹配了各乡镇的空间范围及人口普查信息,计算了公里格网尺度的各乡镇平均人口密度,形成了本数据集。

关键词: 人口分布; 密度; 乡镇; 中国

数据库(集)基本信息简介

数据库(集)名称	中国 27 省乡镇(街道)级人口密度数据集(2000 年)		
数据库(集)英文名称	A dataset of population density at township level for 27 provinces of China (2000)		
通讯作者	王卷乐 (wangjl@reis.ac.cn)		
数据作者	柏中强、王卷乐		
地理区域	辽宁省、吉林省、内蒙古自治区(部分地区)、北京市、天津市、上海市、河北省、河南省、陕西省、宁夏回族自治区、山西省、山东省、安徽省、江苏省、湖南省、湖北省、江西省、浙江省、福建省、广东省、海南省、云南省、贵州省、四川省、重庆市、青海省、西藏自治区		
数据时间范围	2000 年	空间分辨率	1 km
数据格式	TIF	数据量	58.5 MB
数据服务系统	http://www.scidb.cn/dataSet/handle/2		
基金项目	中科院信息化专项项目(XXH12504101), 国家科技基础性工作专项重点项目(2011FY110400, 2013FY114600)		
数据库(集)组成	数据集包含 4 个文件, 其中: population_density.tif (59912 KB) 是一幅展示中国 27 省镇级平均人口密度的栅格图像, population_density.tfw (1 KB) 存储了空间参考信息; population_density.ovr (2226 KB) 是金字塔文件, 方便在 ArcGIS 中快速加载和展示; population_density.aux (2 KB) 是一个存储其他辅助信息的 XML 文件		

引言

中国拥有世界上最多的人口总数, 人口分布不均衡, 资源环境承载压力大, 使得人口问题研究在中国极具现实意义^[1]。空间分布是人口的重要属性, 人口密

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. 江苏省地理信息资源开发与利用协同创新中心, 江苏, 南京 210023

* 通讯作者 (Email: wangjl@reis.ac.cn)

度则是表现人口空间分布最主要的形式和衡量人口分布地区差异的主要指标^[2]。准确、精细的人口密度数据能够为区域规划与开发、灾害风险救援、健康科学和资源环境承载力评估等研究和应用提供支撑。

当前，县级平均人口密度是人口问题研究中常用的指标。乡镇（街道）处于中国行政体系的基层位置，也是全国人口普查数据公开发布的最小单元。较县级平均人口密度而言，乡镇（街道）级平均人口密度能够更加客观细致地刻画中国人口分布的格局和态势，也可为流域、县域尺度的人口问题研究提供数据支撑，具有重要的科学价值和广泛的应用领域。目前，中国尚无可共享的较大范围的乡镇（街道）人口密度空间数据集，主要原因包括乡镇（街道）界线数据的限制使用、乡镇（街道）人口统计数据的缺乏及乡镇行政区划调整等^[3]。国家科技基础条件平台——地球系统科学数据共享平台收集了中国 2000 年的乡镇（街道）级行政界线数据，国家统计局发布了第五次人口普查乡镇（街道）级人口统计数据，本文基于这两个基础数据库，经过标准化处理之后形成中国 27 省乡镇级人口空间基础数据，计算获得各乡镇（街道）级平均人口密度，形成了本数据集。

1 数据采集和处理方法

中国 27 省乡镇（街道）级人口密度数据集利用的基础数据集包括国家科技基础条件平台——地球系统科学数据共享平台提供的 2000 年全国 27 省乡镇（街道）行政界线数据集（<http://www.geodata.cn>），及国家统计局发布的第五次全国人口普查乡镇（街道）级人口统计数据（<http://www.stats.gov.cn/tjsj/ndsj/renkoupucha/2000jiedao/jiedao.htm>）。前者通过离线共享获得，后者为在线下载资源。处理流程包括乡镇（街道）界线数据处理、乡镇（街道）人口统计数据处理和平均人口密度计算 3 个部分。乡镇（街道）界线数据的处理包括原始数据分析、空间参考统一和几何纠正，乡镇（街道）人口统计数据的处理主要是虚拟乡镇标定。在这两项工作基础上，综合处理乡镇（街道）界线数据和乡镇（街道）人口统计数据，

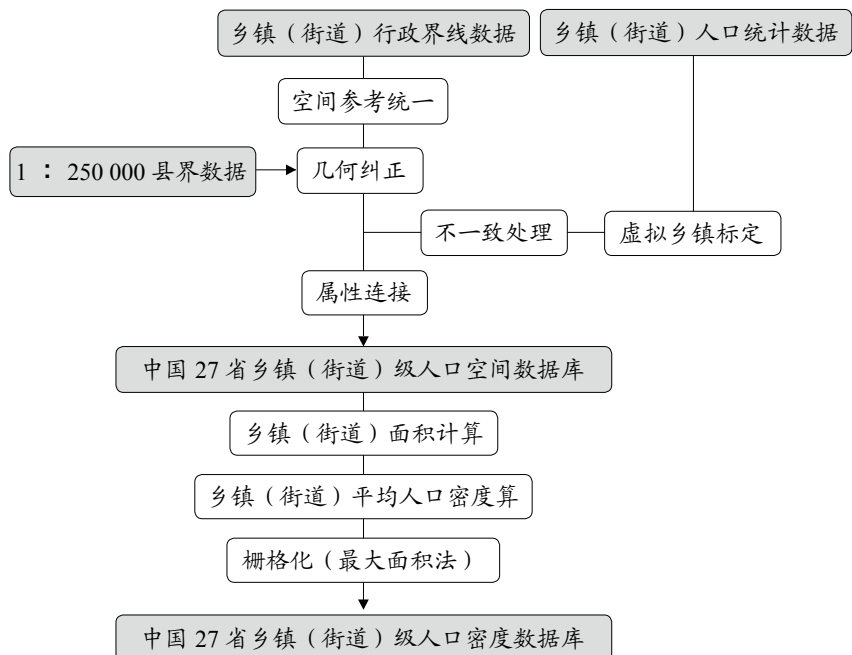


图 1 数据集生产技术路线图

消除二者不一致的情况，经过编码、属性连接等形成中国 27 省乡镇（街道）级人口空间数据库，计算各乡镇（街道）面积，进而计算各乡镇（街道）平均人口密度，栅格化形成本数据集。主要的技术流程如图 1 所示。

1.1 乡镇（街道）界线数据处理

1. 初始资料获取与分析。首先向国家科技基础条件平台——地球系统科学数据共享平台申请乡镇界线数据，空间范围覆盖辽宁、吉林、内蒙古（部分地区）、北京、天津、上海、河北、河南、陕西、宁夏、山西、山东、安徽、江苏、湖南、湖北、江西、浙江、福建、广东、海南、云南、贵州、四川、重庆、青海、西藏共 27 个省（直辖市、自治区）。该数据集以 1 : 25 万、2000 年的数据为参考原则，数据质量良好，但其数据来源、时间和比例尺并不完全相同。

2. 空间参考统一。将各省数据集的空间参考统一为 Albers Equal Area 投影，地理参考为 WGS84 坐标系。

3. 几何纠正。以 1 : 25 万全国县界数据（2000 年）为标准数据，分省对原始乡镇界线数据进行几何纠正。根据省区面积各省约选取 300 ~ 600 个配准点，保证纠正完成输出的乡镇界线和县界空间偏差不超过 500 m。

1.2 乡镇（街道）人口统计数据处理

1. 虚拟乡镇标定。对于普查数据中的经济开发区、虚拟普查片区等无明确行政边界范围的虚拟乡镇单元，通过查阅文献及空间位置判定两种手段，将其人口归并到具体乡镇（街道）单元中。

2. 乡镇空间单元与统计数据不一致情况的处理。分省逐县区核对统计乡镇空间单元与统计数据对应情况，对于二者不一致的情况，分以下三种情况处理：

1) 乡镇界线数据表征时间晚于 2000 年且该县在 2000 年之后发生过乡镇撤并和拆分，查阅资料，依据调整情况合并乡镇人口统计数据以匹配乡镇界线数据；

2) 乡镇界线数据表征时间早于 2000 年且该县在该表征时间至 2000 年之间发生过乡镇撤并和拆分，查阅资料，依据调整情况合并乡镇界线数据以匹配人口统计数据；

3) 某些县（区）乡镇（街道）由于缺少详细的边界数据、划分过粗，查阅资料，合并统计数据以匹配乡镇界线数据。这种情况多发生在面积较小且人口密集或边界模糊的街道。

1.3 平均人口密度计算

1. 分别对乡镇空间界线数据和人口统计数据按照省名称—县名称—乡镇（街道）名称顺序排序，并按顺序赋予各乡镇空间单元及统计属性唯一乡镇代码；

2. 利用乡镇代码为关键字，将各乡镇的人口统计属性赋予其空间多边形单元；

3. 计算各乡镇单元的面积；

4. 依据公式（1），计算各乡镇平均人口密度：

$$PD=Np/A \quad (1)$$

式中：PD 为人口密度（人/平方千米）；Np 为乡镇（街道）人口数量（人）；A 为乡镇（街道）面积（平方千米）。

1.4 栅格化处理

在 ArcGIS 中，基于上一步计算的乡镇平均人口密度，计算 100 m×100 m 网格尺寸上的人口分布，将人口密度图层栅格化为百米格网人口密度图层；基于百米格网人口密度图层，利用栅格聚合功能，生成 1 km×1 km 的乡镇（街道）级平均人口密度数据。

2 数据样本描述

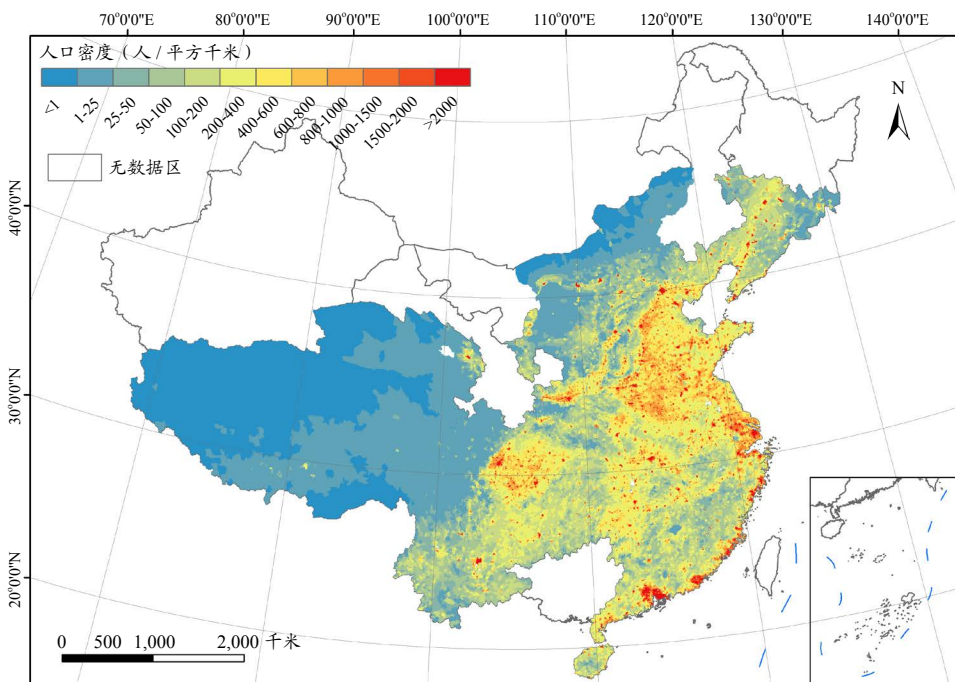


图 2 中国 27 省 2000 年乡镇（街道）级平均人口密度图

经过处理，中国 27 省 2000 年乡镇（街道）级平均人口密度图如图 2 所示；数据样本：北京市 2000 年乡镇（街道）级平均人口密度数据（图 3）。

3 数据质量控制和评估

主要通过以下手段对数据集进行质量控制：

1. 原始中国 27 省乡镇界线数据在生产时具有严格的质量控制规范（参见其数据文档）；第五次人口普查乡镇（街道）级人口统计数据由中国统计局发布，保证了其可靠性和规范性。
2. 分省对乡镇界线进行几何纠正所参考的 1 : 25 万县界数据集来源于中国测绘与地理信息局，纠正过程中，选择了充足的特征点（各省依据面积在 300 ~ 600 个），保证纠正后的乡镇边界与县界的误差不超过 500 m。
3. 严格匹配了乡镇单元空间界线与人口统计属性，对于不一致情况进行了规范化处理，处理依据中国民政部或各级政府组织所发布的乡镇行政界线变迁文件。

数据质量评估：

经统计，本文研究区的原始人口普查数为 1 103 880 920，栅格化后的总人口数为 1 103 493 334.3，栅格化过程损失人口为 387 586，损失量约为 0.035%。可以认为，该数据集质量可靠，精度较高。

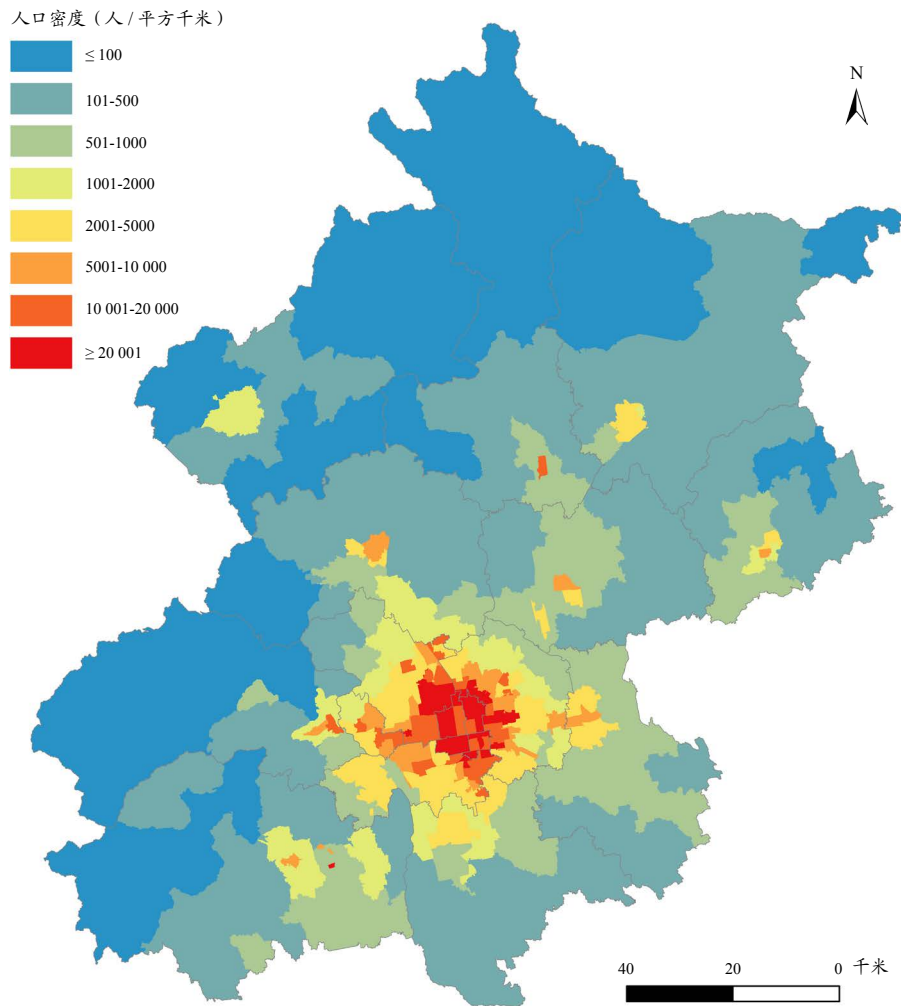


图 3 北京市 2000 年乡镇（街道）级平均人口密度图

4 数据使用方法和建议

中国 2000 年 27 省乡镇级人口空间数据集可用 ArcGIS、ENVI 等地理信息系统软件打开。数据集空间分辨率较高，且未经统计建模，适用性广。该数据集可直接用于表征和分析乡镇尺度上中国人口空间分布格局及地区分异规律，同时也是开展中国人口、资源、环境、灾害等科学研究的重要基础数据集，能够为区域尺度（如流域、县域）的人类活动模拟、灾害风险评估、区域规划与发展、传染病防治等研究提供基础数据。

致 谢

感谢国家科技基础条件平台——地球系统科学数据共享平台提供中国 27 省 2000 年乡镇界线数据，感谢中国统计局提供中国第五次人口普查乡镇（街道）级人口统计数据。

作者分工职责

柏中强（1988—），男，湖北十堰人，博士生，主要从事精细尺度人口分布模拟及夜间灯光遥感应用研究。主要承担工作：方法设计及数据集生产。

王卷乐（1976—），男，河南洛阳人，博士，副研究员，主要从事资源环境数据格网与遥感应用研究。主要承担工作：数据收集及技术指导。

参考文献

- [1] 葛美玲, 封志明. 基于 GIS 的中国 2000 年人口之分布格局研究——兼与胡焕庸 1935 年之研究对比 [J]. 人口研究, 2008, 32(1): 51 ~ 57.
- [2] 祝卓. 人口地理学 [M]. 北京: 中国人民大学出版社, 1991.
- [3] 柏中强, 王卷乐, 杨雅萍, 等. 基于乡镇尺度的中国 25 省区人口分布特征及影响因素 [J]. 地理学报, 2015, 70(8): 1229 ~ 1242.

引用数据

- [1] 柏中强, 王卷乐. 中国 27 省乡镇（街道）级人口密度数据集 [DB/OL]. Science Data Bank. DOI: 10.11922/sciencedb.170.2.

A dataset of population density at township level for 27 provinces of China (2000)

Bai Zhongqiang, Wang Juanle

ABSTRACT Population density is the key indicator to characterize the pattern of population distribution and to measure regional differences in this respect. In China, township level is the basic administrative unit and the finest level for which census data are publicly available. Average population density data at township level can accurately describe the spatial pattern and trend of population distribution. Moreover, it can supplement adequate information for studies on resources, environment and population problems. In this paper, we systematically processed the township boundary data and the township-level demographic data in China's 27 provinces. Then, a database of GIS-linked circa 2000 census data at township level covering 27 provinces was constructed. Based on the database, the average population density at township level was calculated and the dataset of population distribution at township level for China's 27 provinces (2000) was generated.

KEYWORDS population distribution; density; township; China

引文格式: 柏中强, 王卷乐. 中国 27 省乡镇（街道）级人口密度数据集（2000 年） [J/OL]. 中国科学数据, 2016, 1(1). <http://www.nsdata.cn/paperView?id=2>. DOI: 10.11922/csdata.170.2015.0002.