

天津静海区村庄宅基地再利用分类 指标研究数据集

张 钰¹, 蔡为民^{2*}

1. 天津工业大学经济与管理学院, 天津 300387;

2. 天津工业大学环境科学与工程学院, 天津 300387

摘 要: 本数据集是在静海区政府对 56 个宅基地试点村的摸底调查工作所获统计数据的基础上, 构建“村-地”系统耦合特征分析框架, 建立村庄识别指标体系和宅基地利用特征识别指标体系所形成的原创数据集, 数据存储为.xlsx 和.shp 格式, 由 2 个数据文件组成, 包括宅改试点村位置、宅改试点村的村庄及宅基地利用特征指标等内容, 数据量为 614 KB (压缩后为 1 个文件 480 KB)。

关键词: 宅基地再利用; “村-地”系统特征; 分类指标; 天津市静海区

DOI: <https://doi.org/10.3974/geodp.2022.04.06>

CSTR: <https://cstr.escience.org.cn/CSTR:20146.14.2022.04.06>

数据可用性声明:

本文关联实体数据集已在《全球变化数据仓储电子杂志 (中英文)》出版, 可获取:

<https://doi.org/10.3974/geodb.2022.06.07.V1> 或 <https://cstr.escience.org.cn/CSTR:20146.11.2022.06.07.V1>.

1 前言

静海区位于天津市西南部, 下辖 2 个街道、18 个乡镇、383 个行政村, 区域面积 1,475.68 km²。2020 年, 静海区被确定为宅基地制度改革试点区之一, 自 2020 年 12 月 9 日至 2021 年 2 月 4 日分两批次开展了宅基地摸底调查试点工作, 摸底调查涉及了 18 个乡镇、56 个村、2 万户村民。农民权益是我国各级政府及社会各界人士关注的重点, 长期以来, 宅基地作为农民的居住场所和重要权益, 在保障农民生活和维护社会稳定等方面发挥了重要作用。因此, 集合 56 个试点村的村庄及宅基地数据对于宅基地及宅基地制度改革等研究具有重要意义。

目前, 关于宅基地的研究内容主要集中在宅基地“三权分置”^[1,2]、宅基地制度改革^[3,4]及宅基地退出^[5,6]等方面, 对于宅基地再利用的研究较少, 作者依据本数据集研究了宅基地

收稿日期: 2022-07-13; 修订日期: 2022-10-26; 出版日期: 2022-12-24

基金项目: 国家自然科学基金 (41801193)

*通讯作者: 蔡为民 GLN-5852-2022, 天津工业大学环境科学与工程学院, caiweimin@tiangong.edu.cn

数据引用方式: [1] 张钰, 蔡为民. 天津市静海区村庄宅基地再利用分类指标研究数据集[J]. 全球变化数据学报, 2022, 6(4): 557–565. <https://doi.org/10.3974/geodp.2022.04.06>. <https://cstr.escience.org.cn/CSTR:20146.14.2022.04.06>.

[2] 张钰, 蔡为民. 天津市静海区村庄宅基地再利用分类指标研究数据集[J/DB/OL]. 全球变化数据仓储电子杂志, 2022. <https://doi.org/10.3974/geodb.2022.06.07.V1>. <https://cstr.escience.org.cn/CSTR:20146.11.2022.06.07.V1>.

再利用类型选择的问题，研究方法上大多采用案例研究^[7,8]和理论分析^[9]等方法，较少运用空间量化分析，究其原因，主要是因为微观视角下宅基地空间数据较为匮乏且难以获取，阻碍了微尺度研究。

鉴于此，本数据集收集和整理了天津静海区 56 个宅改试点村的村庄及宅基地数据，并通过运用 ArcGIS10.8 软件将数据可视化，实现对宅基地的空间分析。该数据集包含试点村庄经济、生态、文化、管理、人口、土地等多方面信息，对研究中国农村具有重要意义。

2 数据集元数据简介

《天津市静海区村庄宅基地再利用分类指标研究数据集》^[10]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 《天津市静海区村庄宅基地再利用分类指标研究数据集》元数据简表

条 目	描 述
数据集名称	天津市静海区村庄宅基地再利用分类指标研究数据集
数据集短名	VillagesJinghai
作者信息	张钰 GLN-4098-2022, 天津工业大学经济与管理学院, 15822580019@163.com 蔡为民 GLN-5852-2022, 天津工业大学环境科学与工程学院, caiweimin@tiangong.edu.cn
地理区域	天津市静海区
数据年代	2021 年
数据格式	.xlsx、.shp
数据量	614 KB（压缩后为 1 个文件 480 KB）
数据集组成	数据集由两个部分组成： 天津静海区 56 个宅改试点村地理位置分布数据（.shp） 天津静海区村庄宅基地再利用分类指标数据（.xlsx）
基金项目	国家自然科学基金（41801193）
数据计算环境	ArcGIS、Excel
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）、通过《全球变化数据仓储电子杂志（中英文）》发表的实体数据集和通过《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 ^[11]
数据和论文检索系统	DOI, CSTR, Crossref, DCI, CSCD, CNKI, SciEngine, WDS/ISC, GEOSS

3 数据研发方法

3.1 数据来源

宅基地利用特征识别指标体系中所有数据均来自天津市静海区人民政府公开的宅基地摸底调查数据，村庄特征识别指标体系数据来源见表 2。

表 2 村庄特征识别指标体系数据及来源统计表

维度	指标	数据来源
生活 A	服务完善度 B1	POI 数据（ https://lbs.amap.com/ ）
	购物服务设施密度 B2	
	生活服务设施密度 B3	
	城镇化率 B4	天津市静海区人民政府（ http://www.tjjh.gov.cn ）公开的宅基地摸底调查数据
生态 C	河流流经长度均值 D1	静海区政府水务局官网（ http://www.tjjh.gov.cn/jhqzf/zwgk_28985/zfxxgk_qjjgl/swjl/fdzdgknr16/swjwj1/ ）
	水文调节指数 D2	静海区政府水务局官网（ http://www.tjjh.gov.cn/jhqzf/zwgk_28985/zfxxgk_qjjgl/swjl/fdzdgknr16/swjwj1/ ）
	交通设施密度 D3	POI 数据（ https://lbs.amap.com/ ）
	区位优势度 D4	OpenStreetMap（ http://www.openhistoricalmap.org ）数据
乡风 E	科教文化服务密度 F1	POI 数据（ https://lbs.amap.com/ ）
	负面资讯率 F2	今日头条（ http://www.toutiao.com ）数据爬取
	村庄特色资源丰富度 F3	静海区政府文旅局官网（ http://www.tjjh.gov.cn/jhqzf/zwgk_28985/zfxxgk_qjjgl/wlj1/fdzdgknr18/wljwj1/ ）；2020 年静海统计年鉴
	村庄网站曝光度 F4	今日头条（ http://www.toutiao.com ）数据爬取
产业 G	企业密度 H1	POI 数据（ https://lbs.amap.com/ ）
	企业注册资本均值 H2	天眼查（ http://www.tianyancha.com ）公开数据
	人口密度 H3	天津市静海区人民政府（ http://www.tjjh.gov.cn ）公开的宅基地摸底调查数据
	金融保险服务设施密度 H4	POI 数据（ https://lbs.amap.com/ ）
管理 I	非集体组织成员比例 J1	天津市静海区人民政府（ http://www.tjjh.gov.cn ）公开的宅基地摸底调查数据
	宅基地唯一率 J2	
	宅基地批准率 J3	
	政府机构及社会团体密度 J4	POI 数据（ https://lbs.amap.com/ ）

3.2 数据处理

在数据收集、整理和记录过程中，出现少部分数据值的缺失，为保证样本量的完整性，对缺失数据值进行填充。样本中“宅基地使用年限均值”变量缺失数据占比 1.8%、“容积率”变量缺失数据占比 5.4%，数据缺失比例较小且指标均为连续数值型数据，故采用均值插值法进行缺失值填充^[12]。

由于静海区政府公开的试点村宅基地摸底调查原始数据是分两批次进行的，各村所收

集的数据指标存在细微差异,本数据集通过运用 SPSS.26 软件,对原始数据进行整理和计算后得到所需数据,静海区宅基地制度改革专班调查积累了丰富的实地数据,为微观尺度的分析提供了数据支撑。

POI 数据来源于对 2021 年高德地图 API 的抓取结果,运用 ArcGIS10.8 软件对 POI 数据进行裁剪,并与静海区矢量图层叠加,运用空间连接工具,统计 56 个试点村不同类型 POI 的数量。本文所用 POI 分类数据均与高德地图官方公布的 POI 分类编码保持一致。

“今日头条”作为 Web 3.0 时代的互联网资讯代表,运用大数据和算法精准分发资讯,借助智能手机的普及,为小城镇和农村人口获取信息提供了便利,满足了低学历、低年龄、低收入人群的信息需求^[13]。并且百度指数显示,“今日头条”的百度搜索量远高于网易、腾讯、搜狐、新浪四大新闻平台。因此,本文选取“今日头条”网站作为村庄报道资讯数据的来源。首先,以“天津市静海区××镇××村”为搜索关键词,在“今日头条”网站中资讯栏目下搜索 56 个试点村的相关报道,抓取所有资讯的标题、内容简介及来源网站等数据;其次,考虑到不同来源网站的报道数量越多,村庄影响力越大,故对于报道内容相同但报道来源不同的资讯未作删减,一并计入总报道量;第三,村庄影响力会带动所在乡镇的影响力,即村庄报道越多,所在乡镇的报道量越多,因此,对搜索内容中关于所在乡镇及镇政府的报道内容进行了删减,并删除了非研究区域内的同名村庄或事物等无效报道内容。

天眼查作为一家全国企业信息查询网,主要提供专业的企业信息查询、企业关系挖掘服务,其网站数据来源于全国企业信用信息公示系统、中国裁判文书网、中国执行信息公开网、国家知识产权局、商标局、版权局等国家权威网站^[14]。本文通过对 POI 数据中的公司企业数据进行整理分类,运用天眼查抓取各公司的注册资本数据以筛查公司经营状态,反映企业规模和体量,剔除经营地址、经营状态(吊销、注销等)发生变化的企业、天眼查上未找到的企业,个体工商户(规模较小,对产业维度的影响不突出)以及未披露注册资本的企业,最终获得 231 条有效数据。

3.3 技术路线

为研究宅基地再利用中的类型选择问题,分别收集村庄特征和宅基地利用特征 2 方面数据,借助 Microsoft office Excel 2013 和 ArcGIS 10.8 软件,对数据进行分析、整理和分类(图 1)。首先,以村庄特征指标体系计算结果为依据,运用村庄特征识别模型对试点村进行分类;其次,根据宅基地摸底调查数据计算结果,运用宅基地利用特征识别模型以村庄为单位对宅基地进行分类;最终,综合考虑“村-地”分类及空间分布情况,得出宅基地再利用的类型。

4 数据结果与验证

4.1 数据集组成

数据集共包括两个部分:天津静海区 56 个宅改试点村地理位置分布数据(.shp)和天津静海区村庄及宅基地利用特征指标数据(.xlsx)。其中,村庄及宅基地利用特征指标数据包括宅改试点村及宅基地利用特征指标得分、村庄分类指标标准化及权重得分、村庄特征识别指标得分汇总、村庄分类识别模型计算得分、村庄分类、宅基地权重得分、宅基地利

用特征指标贡献度得分、宅基地指标分类、宅基地贡献度较大指标分值汇总、不同类型宅基地分类汇总、宅基地再利用类型分类汇总等内容。

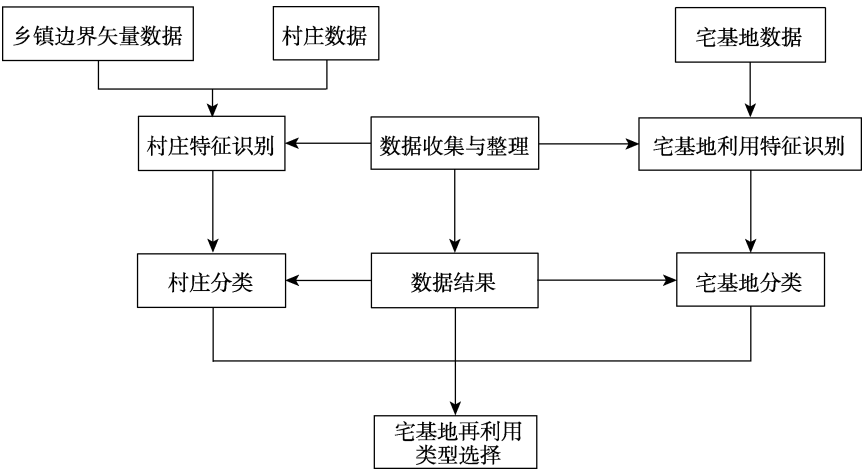


图 1 数据集研发技术路线图

4.2 数据结果

4.2.1 村庄特征识别

静海区 56 个村庄可分为产业主导型、资源主导型和均衡发展型 3 类，资源主导型和均衡发展型分别占比 42.86% 和 41.07%，产业主导型占比 16.07%，表明目前以产业发展带动乡村发展的比例还较小。其中，产业主导型包括四党口后村、满意庄村、二街村、杨小庄村、袁村、下圈村、顺民屯村、王庄子村、中明庄村 9 个村庄；资源主导型包括丰普村、惠丰西村、王家村、丁家房子村、西高庄村、八堡村、宫家屯村、前邓村、王千户村、中翟庄村、增福堂村、吕官屯村、高官屯村、十一堡村、刘家房子村、东港村、孟庄子村、北二堡村、刘祥庄村、南坝台村、东柳木村、后小屯村、林庄子村、朴楼村等 24 个村庄；均衡发展型包括北万营村、王匡村、堂上村、圈里村、许庄子村、东寨村、南万营村、西滩头村、陆家村、后邓村、蔡庄子村、曾家河村、姜家场村、刘下道村、良辛庄村、罗庄子村、韩庄子村、东房子村、南二堡村、小黄洼村、后明庄村、潘村、十里堡村 23 个村庄。

4.2.2 宅基地利用特征识别

静海区 56 个村庄可分为单功能主导型宅基地和多功能主导型宅基地 2 类，单功能主导型宅基地占比 50%，包括 28 个村庄，分别是林庄子村、良辛庄村、刘下道村、满意庄、吕官屯村、高官屯村、袁村、潘村、二街村、王庄子村、下圈村、八堡村、罗庄子村、曾家河村、北万营村、西高庄村、刘家房子村、丁家房子村、东房子村、刘祥庄村、惠丰西村、四党口后村、增福堂村、朴楼村、杨小庄村、中翟庄村、顺民屯村、十里堡村，单功能主导型宅基地多分布于运河沿岸和静海区南部，主要原因是运河沿岸为保护区，严格管控非

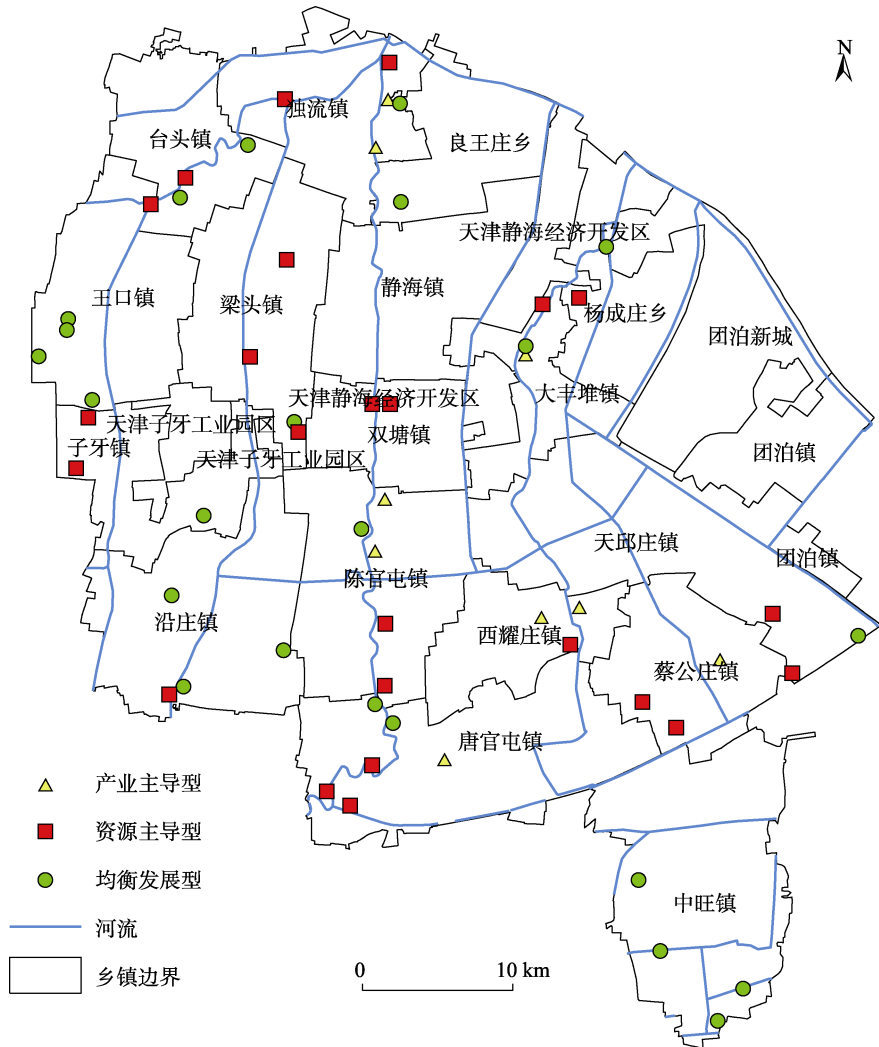


图2 天津市静海区村庄类型分布图

农建设活动，致使沿岸村庄多为单功能主导型宅基地；多功能主导型宅基地占比 50%，包括 28 个村庄，分别是王千户村、南二堡村、北二堡村、南万营村、堂上村、王家村、中明庄村、后明庄村、东港村、西滩头村、小黄洼村、王匡村、后小屯村、十一堡村、南坝台村、姜家场村、韩庄子村、蔡庄子村、许庄子村、后邓村、孟庄子村、东柳木村、丰普村、陆家村、官家屯村、东寨村、圈里村、前邓村，多功能主导型宅基地主要分布在静海区西部及西北部，主要原因在于子牙经济技术开发区和林海循环经济示范区分别坐落于静海区西部和西北部，覆盖大部分乡镇，具有一定的辐射带动作用。

4.2.3 基于“村-地”系统特征的宅基地再利用类型识别

宅基地再利用类型可分为多元主体类、要素流动类、特色资源类 3 类，多元主体类占比 28.57%，包括中明庄村、王匡村、堂上村、圈里村、许庄子村、东寨村、南万营村、西滩头村、后邓村、陆家村、蔡庄子村、姜家场村、韩庄子村、南二堡村、小黄洼村、后明

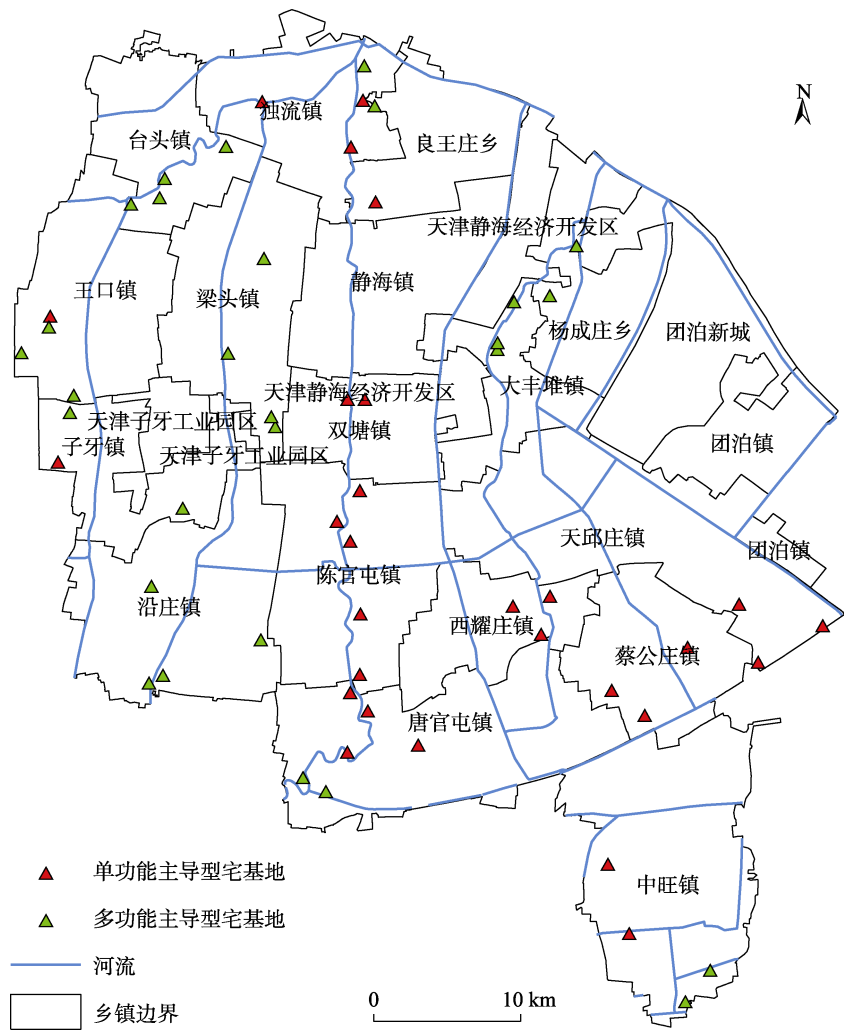


图 3 天津市静海区宅基地类型分布图

庄村等 16 个村庄，多元主体类村庄主要分布在静海区西部，表明静海区西部市场经济发展较好，多元主体活跃度较高，利用多元主体盘活资源是该类村庄的重要发展方向；要素流动类占比 35.71%，包括王千户村、北二堡村、王家村、东港村、后小屯村、十一堡村、南坝台村、孟庄子村、东柳木村、丰普村、宫家屯村、前邓村、四党口后村、满意庄村、二街村、杨小庄村、袁村、下圈村、顺民屯村、王庄子村等 20 个村庄，要素流动类村庄主要沿静海区河流分布，表明多分支河流是互补村庄实现要素流动的关键考量因素，充分利用自身优势要素是该类村庄的重要发展方向；特色资源类占比 35.71%，包括北万营村、曾家河村、刘下道村、良辛庄村、罗庄子村、东房子村、潘村、十里堡村、林庄子村、吕官屯村、高官屯村、八堡村、西高庄村、刘家房子村、丁家房子村、刘祥庄村、惠丰西村、增福堂村、朴楼村、中翟庄村等 20 个村庄，特色资源类村庄主要分布在南运河下游和静海区东南部，表明该区域资源禀赋优越，充分利用自身资源是该类村庄实现乡村振兴的重要发展方向。

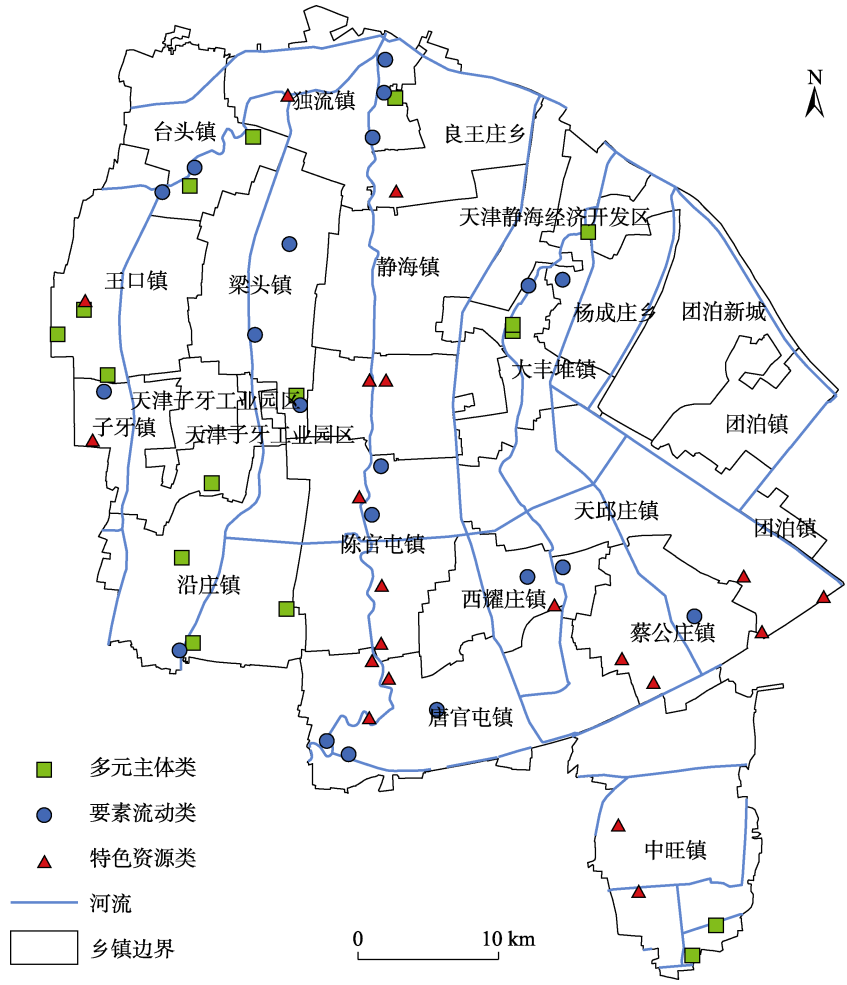


图4 天津市静海区宅基地再利用类型分布图

5 讨论和总结

本研究所构建的数据集主要是通过“村-地”特征视角整理和集合了天津市静海区村庄宅基地再利用分类指标的数据，为破除宅基地再利用相关问题和障碍提供理论参考。数据集从解决实际问题出发，不仅基于乡村尺度提供了分类治理思路，而且从宅基地尺度和宅基地再利用角度提出了具有可行性和可操作性的分类治理方案，与同类型研究相比，该数据集采用多源数据，运用跨尺度研究方法，探索了一条可行的解决现实问题的路径，具有一定的创新性和可推广性，可为中国东部地区农村居民点、宅基地、乡村振兴等研究提供数据支撑。不足之处是该数据集所选指标有限，且与人口具有较大的相关性，因此在数据使用过程中应注意地域性和时效性。

作者分工：张钰采集和处理了数据并撰写了数据论文；蔡为民审核了数据论文等。

利益冲突声明：本研究不存在研究者以及与公开研究成果有关的利益冲突。

参考文献

- [1] 崔久富, 郭贯成, 韩述等. 农村集体土地“三权分置”的逻辑审视及实现形式[J]. 中国土地科学, 2022, 36(4): 8–15.
- [2] 陈胜祥. 农村宅基地“三权”分置: 问题导向、分置逻辑与实现路径[J]. 南京农业大学学报(社会科学版), 2022, 22(2): 147–158.
- [3] 陈卫华, 吕萍. 宅基地制度改革创新动力: 困局与突围——基于对两轮试点调研的分析[J]. 农村经济, 2022(5): 30–39.
- [4] 邱丽. “地权集体化”: 传统农区宅基地制度改革的逻辑[J]. 西北农林科技大学学报(社会科学版), 2022, 22(3): 82–89.
- [5] 吴郁玲, 于亿亿, 洪建国. 产权让渡、价值实现与宅基地退出收益分享——基于金寨、余江的实地调查[J]. 中国农村经济, 2022(4): 42–63.
- [6] 闫啸, 李录堂, 李晗. 宅基地退出降低了农户的贫困脆弱性吗?——来自安徽金寨的证据[J]. 中国土地科学, 2022, 36(4): 38–48.
- [7] 魏程琳, 钟晓华. 空间再组织: 城乡接合部闲置农房产权整合与社会有效治理——上海农房再利用案例研究[J]. 中国农村经济, 2022(4): 23–41.
- [8] 杨璐璐, 王航航. 宅基地整治盘活与乡村产业发展的路径选择——基于两个直辖市七个典型试点村的研究[J]. 西北大学学报(哲学社会科学版), 2022, 52(3): 63–79.
- [9] 汪洋. “三块地”改革背景下集体建设用地使用权的再体系化[J]. 云南社会科学, 2022(3): 137–149.
- [10] 张钰, 蔡为民. 天津市静海区村庄宅基地再利用分类指标研究数据集[J/DB/OL]. 全球变化数据仓储电子杂志, 2022. <https://doi.org/10.3974/geodb.2022.06.07.V1>. <https://cstr.escience.org.cn/CSTR:20146.11.2022.06.07.V1>.
- [11] 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]. <https://doi.org/10.3974/dp.policy.2014.05> (2017 年更新).
- [12] 李金玲. 基于纳税信息的小微企业信用动态评价[D]. 南京: 南京邮电大学, 2020.
- [13] 欧阳日辉. 从新闻门户到社交媒体: 门户网站的商业模式变迁与发展路径[J]. 新闻与写作, 2019(2): 11–17.
- [14] 胡健. 企业组织视角下区域关联网特征研究[D]. 武汉: 武汉大学, 2018.