

DOI:10.3969/j.issn.1003-5060.2025.03.011

# 基于大数据和熵权-随机森林的城市地下空间需求评价

葛睿雅, 李晓晖, 袁峰, 窦帆帆, 熊芸莹, 薛晨

(合肥工业大学 资源与环境工程学院, 安徽 合肥 230009)

**摘要:**科学评估地下空间开发需求潜力是缓解城市化问题和合理拓展有限区域的重要基础工作。目前地下空间评价中的社会经济数据多来自于传统官方文件,其全面完整性和时空精度并不理想;此外主观赋权方法的使用,一定程度上存在主观性强和受数据干扰等不足。文章以多源大数据支持的指标体系为基础,构建熵权-随机森林耦合的地下空间需求评价模型。该模型基于熵权法确定负样本,将总样本和指标因子导入随机森林算法中,挖掘社会经济指标与现有地下设施间的复杂非线性关系。研究表明,经过网格搜索调优后的模型 AUC(area under curve)精度达到 0.979,其中 77.45%的现有设施落入评价的高需求区内,证明所采用模型有较强的准确性和可靠性,其精细化评价结果可为今后地下建设选址提供更符合实际的借鉴。

**关键词:**熵权-随机森林模型;多源地理大数据;社会经济指标因子;地下空间需求评价

**中图分类号:**P208.2 **文献标志码:**A **文章编号:**1003-5060(2025)03-0360-09

## Urban underground space demand evaluation based on big data and entropy-random forest

GE Ruiya, LI Xiaohui, YUAN Feng, DOU Fanfan, XIONG Yunying, XUE Chen

(School of Resources and Environmental Engineering, Hefei University of Technology, Hefei 230009, China)

**Abstract:** Scientific assessment of underground space development demand potential is an important basic work to alleviate urbanization problems and reasonably expand limited areas. The current socio-economic data in underground space evaluation are mostly from traditional official documents, and their comprehensive completeness and spatio-temporal accuracy are not ideal. In addition, the use of subjective and objective assignment methods is to a certain extent subjective and subject to data interference. This paper constructs an entropy-random forest coupled underground space demand evaluation model based on the index system supported by multi-source big data. The model is based on the entropy weight method to determine the negative sample, and the total sample and index factors are imported into the random forest algorithm to explore the complex nonlinear relationship between socio-economic indices and existing underground facilities. The case study shows that the area under curve(AUC) accuracy of the model after grid search tuning reaches 0.979, in which 77.45% of the existing facilities fall into the high demand area of the evaluation, which proves that the adopted model has strong accuracy and reliability, and its refined evaluation results can be used as a practical reference for future underground construction site selection.

**Key words:** entropy-random forest model; multi-source geographic big data; socio-economic index factors; underground space demand evaluation

**收稿日期:**2022-10-11; **修回日期:**2022-11-14

**基金项目:**安徽省自然科学基金资助项目(1808085QD116);安徽省公益性地质调查工作资助项目(2023-g-1-18)和中央高校基本科研业务费专项资金资助项目(PA2019GDZC0093)

**作者简介:**葛睿雅(1997—),女,安徽合肥人,合肥工业大学硕士生;

李晓晖(1986—),男,河北定州人,博士,合肥工业大学副教授,硕士生导师,通信作者, E-mail:lxhlixiaohui@163.com;

袁峰(1971—),男,广西桂林人,博士,合肥工业大学教授,博士生导师。

## 0 引言

随着工业化和城镇化进程的加快,大量人群快速涌入城市,造成饱和的地表容量和密集的高层建筑,带来空气污染、出行不便及内部拥挤等问题,降低了人们的生活质量<sup>[1-2]</sup>。为了从根本上解决高度城市化带来的日益恶劣的土地紧缺和供需矛盾,21世纪以来人们逐渐转向对地下空间资源的探索<sup>[3]</sup>。我国城市地下空间利用起步较晚且多是盲目碎片式开发,地下工程又具有不可逆、技术性强、容量有限等特点<sup>[4]</sup>,因此,需要明确开发的地下建筑位置格局、形态规模和类型,完善地下空间开发需求评价模式,从而满足资源合理配置和长远发展。

国内外学者从不同视角出发,提出地下空间评价的内容主要包括开发容量<sup>[5]</sup>、适宜性<sup>[6]</sup>、质量及效益<sup>[7]</sup>、需求评价<sup>[8]</sup>等。其中地下空间资源需求评价通过分析城市社会经济因素,得到地下空间需求潜在价值在城市区位间的差异。当前地下空间需求评价已成为城市规划效应和可持续发展的重要课题,已有研究主要从指标含义和选取、权重计算和确定、模型尺度和构建、结果展示和表达等角度进行了探索和应用实践<sup>[9]</sup>。其中,合理的确权方式和评价方法是城市地下空间资源需求评价中的重要环节,将极大影响整个评估结果准确度。目前广泛应用层次分析法等主观赋权<sup>[10]</sup>和熵权法等客观赋权法<sup>[11]</sup>确定相关因素重要性,并结合多目标线性加权函数法<sup>[12-13]</sup>、模糊综合评判法<sup>[14-15]</sup>、灰色评估法<sup>[16]</sup>等评价方法得到结果。

然而现有研究仍然存在一些局限:①传统数据多来源于官方文件,获取难度大、时间相对滞后,大部分集中于街道或地块等中宏观粗糙尺度;②常用的主观赋权法易受人为经验的干扰,而客观赋权法因仅从客观数据出发,难免会受到数据波动干扰,有时与实际不符。

目前,多源大数据的出现为城市研究提供了大量可靠的数据支持,此外机器学习算法因其具有高效、低成本和自动精确的优越性,也逐渐被相关学者用来研究复杂的对象评价类问题,包括滑坡易发性评价<sup>[17]</sup>、城市空气质量评价<sup>[18]</sup>、山洪灾害风险评价<sup>[19]</sup>、资源与生态安全评价<sup>[20]</sup>、城市市政用地评价<sup>[21]</sup>等。因此,本文引入开放多源地理大数据,以杭州市上城区为实例,运用熵权-随机森林法建立现有地下设

施与社会经济指标因素之间的训练模型,从网格尺度进行精细化地下空间需求评价,圈画出新的高需求潜力区位,旨在为城市合理配置地下资源以及规划治理提供更加精准可靠的数据来源和方法依据。

## 1 研究区概况

截至2021年底,杭州市地下空间开发总量已突破 $1 \times 10^8 \text{ m}^2$ 。杭州地下空间综合实力位居全国第四,地下空间治理体系和建设指标位居全国第一<sup>[22]</sup>。上城区位于浙江省杭州市的中南部,作为杭州市内面积最小的城区,同时又是中心老城区,其核心优质区域建成比例高。受其内含众多的历史街区和文化保护区、西湖和钱塘江水源保护区等限制,上城区可进行进一步地面开发的空间有限。因此,转向地下开发成为其未来扩展空间的重要方式。

目前上城区地下空间开发尚未饱和,因此亟需研究已有地下建设的分布特点,根据区块间社会经济水平差异,选择今后可建设的地下空间,以期与现有地下设施间共同形成更高品质和聚集度的地下空间,提升城区的综合承载能力和立体城市格局。

## 2 数据来源与指标体系

当前有关地下空间的研究多采用政府公布的文件数据,其年份跨度一般较大,且以市县、地块等中宏观尺度为单元,多源大数据的出现极大地提供了地下空间评价方面的数据支持。本文选用的多种来源大数据主要包括开放数据、遥感数据和基础地理数据3类,其中开放数据是指任何人在较小限制下自由挖掘最新数据,在开源端口免费获取互联网公开信息平台与网站提供的信息。

### 2.1 现有地下设施

近年来,随着互联网和数据挖掘技术迅猛发展,基于地理位置服务产生了大量准确真实的实体兴趣点(point of interest, POI)数据,在多个科学研究领域尤其是城市精细地理分析方面都有广泛应用<sup>[23]</sup>,文献<sup>[24-25]</sup>分别使用网络采集的地下POI数据作为数据源,取得了良好的应用效果。本文以2021年3月为时间节点,利用高德开发平台(<https://lbs.amap.com/>)抓取上城区范围内现有地下设施POI点,结合人工识别筛选出674个有效数据,统计结果见表1所列。

表 1 上城区地下空间 POI 点统计

| 地下用地类型 | 功能              | POI 数量/个 |
|--------|-----------------|----------|
| 商业     | 生活服务、餐饮、购物、体育休闲 | 211      |
| 办公     | 物流仓储、公司企业       | 6        |
| 交通设施   | 交通服务、公共停车场      | 452      |
| 公用设施   | 防空地下室           | 5        |

## 2.2 指标数据来源

指标体系的构建是评价城市地下空间潜力需求的关键。在城市内部不同区域,由于改善和解决城市问题的动力差异,地下空间开发的需求存在显著的空间分异性,这与其交通状况、土地资源利用和经济规模密切相关<sup>[26]</sup>。因此本文遵循科学性、广泛性、代表性和系统性 4 个原则,综合文献<sup>[27]</sup>,按照研究区实际情况筛选出可操作性强且数据易收集的指标,研究共从 3 个方面共选取 6 个指标因子。

1) 混乱拥堵的交通状况是制约城市运转效率的关键因素,将地面交通运输转移至地下成为解决交通问题、疏散人流的有效方式。区域交通顺畅程度与是否需要开发地下交通有很强的关联性,因此研究选取路网密度<sup>[28]</sup>和交通拥堵<sup>[29]</sup>指

标来体现交通联系便利性。

2) 当地面无法承载更多的建设时,向地下扩展成为缓解地面建设用地紧张的重要手段,土地资源利用状况对地下需求有很大的影响<sup>[30]</sup>。其中,不同用地类型对地下空间开发的推动作用和经济社会效益有所差异<sup>[31]</sup>,建筑密度越高的地区往往也是地下空间开发需求较大的地区<sup>[32]</sup>,因此选择用地类型和建筑密度作为指标因子。

3) 地下空间作为城市化发展伴随产物,经济越繁华的区域地下空间资源开发价值越高<sup>[33]</sup>。集中的人口和经济活动的密集场所,其基础设施和空间负担较大,迫切需要开发地下空间优化布局,同时保持可持续发展<sup>[34]</sup>。为了表达人口密度和经济活力,研究选取人口数据集和夜间灯光数据来衡量,目前已有研究通过夜间灯光表征经济水平的空间异质性,取得良好的应用效果<sup>[35]</sup>。

由于难以保证所有数据获取的时空尺度相同,尽量选取时间相近且空间分辨率合理的数据,部分数据演变较为缓慢,足以支撑研究得出相对科学合理的成果<sup>[36]</sup>。具体的数据描述及对应指标见表 2 所列。

表 2 数据来源与描述

| 特征指标         | 数据类型   | 数据来源   | 数据获取 |
|--------------|--------|--|------|
| 路网密度 $X_1$   | 道路数据   | Open Street Map 网站( <a href="http://www.openstreetmap.org">http://www.openstreetmap.org</a> )  | 下载   |
| 交通拥堵 $X_2$   | 道路拥堵指数 | 杭州市交通拥堵指数实时监测平台网站( <a href="http://www.hzjtydzs.com/">http://www.hzjtydzs.com/</a> )   | 实时截取 |
| 用地类型 $X_3$   | 土地利用数据 | 高德地图 AOI(Area of Interest)数据( <a href="https://lbs.amap.com/">https://lbs.amap.com/</a> )  | 网络爬虫 |
| 建筑密度 $X_4$   | 建筑轮廓数据 | Open Street Map 网站( <a href="http://www.openstreetmap.org">http://www.openstreetmap.org</a> )  | 下载   |
| 人口数据 $X_5$   | 人口数据   | WorldPop 全球高分辨率人口项目数据集 <sup>[37]</sup> ( <a href="http://www.worldpop.org">www.worldpop.org</a> )  | 下载   |
| 夜间灯光强度 $X_6$ | 夜光遥感数据 | NPP VIIRS Annual VNL V2 夜间灯光影像 <sup>[38]</sup> ( <a href="https://eogdata.mines.edu/products/vnl/">https://eogdata.mines.edu/products/vnl/</a> ) | 下载   |

## 2.3 评价指标建立

原始数据整理与处理是指标数据集建立的前提,将多源数据进行清洗、裁剪等预处理,经

过地理配准、转换要素、空间分析、重采样等步骤得到指标因素,具体的评价指标构建见表 3 所列。

表 3 评价指标构建

| 特征指标  | 数据预处理                         | 指标计算                |
|-------|-------------------------------|---------------------|
| $X_1$ | 简化道路数据,消除拓扑错误                 | 每个网格内道路总长度除以网格面积    |
| $X_2$ | 图像校准                          | 矢量化拥堵数据             |
| $X_3$ | 按照文献 <sup>[39]</sup> 划分 11 大类 | 按用地类型代码赋值,将定性转为定量指标 |
| $X_4$ | 对照百度地图补充缺失                    | 每个网格内建筑面积除以网格面积     |
| $X_5$ | 裁剪                            | 直接使用数据集             |
| $X_6$ | 裁剪、去除异常值                      | 直接使用数据集             |

栅格数据作为具有属性数值的空间均匀分割的像元数据格式,便于后续网格计算和精细化评

价。本文在参考广州<sup>[24]</sup>、杭州<sup>[40]</sup>等地研究的基础上,结合研究区实际范围、原始数据分辨率和数

据分布情况等,最终选择软件工具划分 200 m×200 m 的分辨率的网格,并进一步将不同来源的

专题数据整合栅格形式数据集<sup>[41]</sup>,相关数据集结果如图 1 所示。

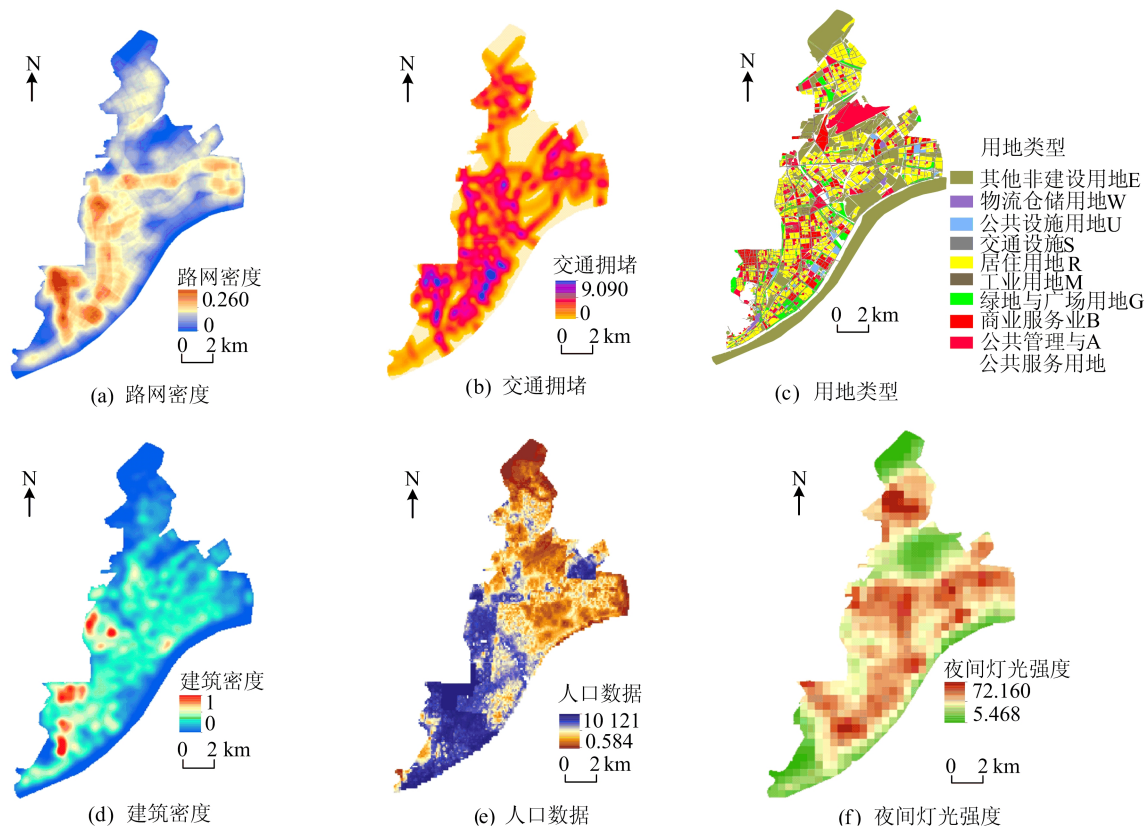


图 1 评价指标数据集

### 3 研究方法

#### 3.1 总体技术路线

本研究的技术路线如图 2 所示。

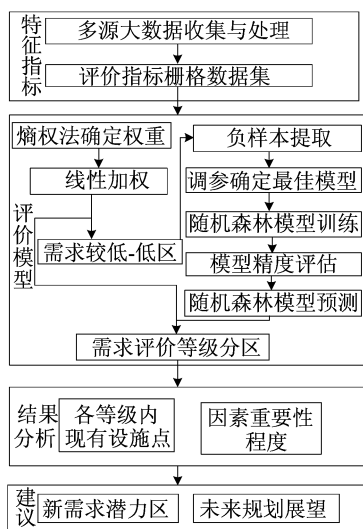


图 2 总体技术路线

具体研究方法为:① 选取多源地理大数据,考虑社会经济因素的多方面构建评价指标体系,

转化为栅格指标数据集,提取属性值至点要素;② 采用熵权法确定指标权重,结合线性加权计算地下空间需求值;③ 以现有地下设施点为正样本,并从熵权-线性加权结果的需求较低和低区域选取等量负样本,构建随机森林分类模型,通过网格搜索、精度检验等训练最优模型;④ 输入研究区待测的网格数据,进行地下空间需求概率评估,利用自然断点法划分等级,从各等级内现有设施数和因素重要程度方面对结果进行分析;⑤ 依据各网格需求值高低排序确定新的需求潜力区位置,结合地上、地下布局特征对城市地下空间的未来开发提出建议。

#### 3.2 熵权法

熵权法属于客观赋权法,能深刻反映出指标的离散程度,可以合理规避人工经验造成的主观性。其基本原理是区分指标变异程度的高低,计算相关熵权,修正得到权重<sup>[42]</sup>。

设有  $m$  个评价方案,  $n$  个评价指标,熵权法的具体步骤<sup>[43]</sup>为:

1) 计算第  $j$  个指标下第  $i$  个方案指标值的比重  $f_{ij}$  为:

$$f_{ij} = \frac{x_{ij}}{\sum_{i=1}^m x_{ij}} \quad (1)$$

2) 计算第  $j$  项指标的熵值  $e_j$  为:

$$e_j = -k \sum_{i=1}^m (f_{ij} \ln f_{ij}) \quad (2)$$

其中:  $k = \frac{1}{\ln m}$ ;  $e_j \geq 0$ 。

3) 各指标的权重  $w_j$  为:

$$w_j = \frac{1 - e_j}{n - \sum_{j=1}^n e_j} \quad (3)$$

其中,  $\sum_{j=1}^n w_j = 1$ 。

### 3.3 线性加权组合法

本文在使用熵权法计算出各指标的权重后,以网格为单元对各指标的属性数据集进行线性加权,计算公式为:

$$A = \sum_{i=1}^n a_i b_i \quad (4)$$

其中:  $A$  为地下空间需求评价价值;  $a_i$  为第  $i$  个指标因子的权重;  $b_i$  为第  $i$  个指标的属性值。  $A$  值越高,该单元格区域对应地下空间的需求程度越高。

### 3.4 随机森林算法

随机森林算法是基于多个决策树分类器的非参数集成算法<sup>[44]</sup>。从训练样本中有放回地重复随机抽取  $n$  个训练子集,以此构建  $n$  棵决策分类树,用不纯度来度量节点的最佳分割,最终通过投票统计按照少数服从多数原则得到分类结果<sup>[45]</sup>。以往研究表明,本文选用的随机森林算法在解决地理应用中的高维度、非线性和复杂多变量的空间分布预测上具有较好的表现,在稳定性、泛化性及避免过拟合等方面相较传统方法有较大优势<sup>[46]</sup>。其中,不纯度是通过基尼(Gini)系数计算的,此外还可以比较不同特征的平均 Gini 系数改变量占所有特征改变量总和的百分比来判断重要性权重,得到每个指标在随机森林中每棵分类树上的贡献大小,能够有效减少人为因素对因子权重赋予的干扰<sup>[35]</sup>。计算公式为:

$$G(t) = 1 - \sum_{i=0}^{c-1} p(i|t)^2 \quad (5)$$

其中:  $G$  为 Gini 系数;  $c$  为类别数;  $t$  为给定的节点;  $i$  为样本的任意分类;  $p(i|t)$  为样本分类  $i$  在节点  $t$  上所占的比例。

在机器学习中,算法本身会对模型精度造成影响,参数选取也同样重要,默认的参数不能保证机器学习算法达到最佳性能,因此本文采用网格

搜索法优化参数。该算法是一种对模型中参数集遍历优化的方法,主要通过建立待选参数集,然后组合不同属性的待选参数集,通过试错逐步缩小参数的可选择范围<sup>[45]</sup>。

## 4 评价结果

### 4.1 熵权-线性加权模型

利用熵权法计算得到上城区各指标权重大小,从高到低排序见表 4 所列。并将每个格网的指标属性值线性加权,得到最终上城区地下空间需求评价价值。使用自然断点法将熵权-线性加权模型的结果值从高到低进行划分,如图 3 所示,可以看出,共划分为高需求区、较高需求区、中需求区、较低需求区、低需求区 5 个等级。

表 4 熵权法计算权重结果

| 指标    | 信息熵值  | 权重    |
|-------|-------|-------|
| $X_5$ | 0.899 | 0.429 |
| $X_2$ | 0.954 | 0.195 |
| $X_4$ | 0.968 | 0.136 |
| $X_3$ | 0.973 | 0.113 |
| $X_1$ | 0.980 | 0.084 |
| $X_6$ | 0.990 | 0.042 |

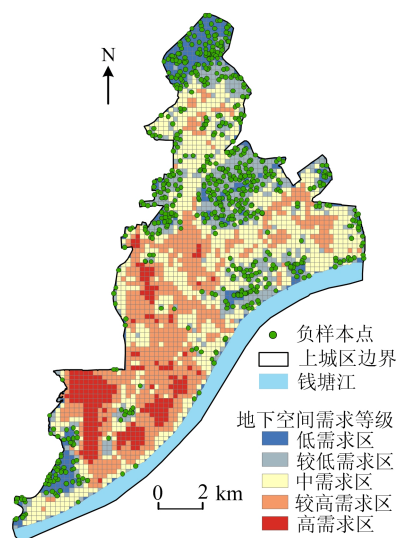


图 3 熵权-线性加权模型需求等级分区及负样本选取

### 4.2 熵权-随机森林模型

#### 4.2.1 正负样本的选取

将研究区已有地下设施点作为正样本,即认为现有地下设施点所在位置为地下空间需求高的地区。为克服负样本选取的主观性,提高负样本的纯度和误差,部分学者尝试利用聚类<sup>[47]</sup>、熵值法<sup>[48]</sup>等方法事先进行初步的分区,再进行负样本的选取<sup>[49]</sup>。本文在前人选取方式的基础上,为了

进一步解决主观因素带来的问题,在较低-低需求区内,利用随机均匀采样方法选取了等量的负样本点。

#### 4.2.2 训练集和测试集

为了提高模型泛化能力,从总样本数据集中随机选取 70% 的数据作为训练样本,剩余 30% 作为测试样本。

#### 4.2.3 网格搜索和参数调优

为了使模型性能达到最优值,保证算法的收敛性和良好的预测结果,选择对模型的结果影响较大的 2 个重要参数<sup>[50]</sup>,即决策树的数量( $n\_estimators$ )和分类树的最大深度( $max\_depth$ )进行网格搜索遍历,其中树的数量  $n\_estimators$  的取值范围设定为  $[0, 300]$ ,树的最大深度  $max\_depth$  的取值范围设定为  $[0, 20]$ ,均以 1 为搜索步长。

调参过程曲线如图 4 所示。由图 4 可知:参数组合为  $['n\_estimators': 113, 'max\_depths': 10]$  时,模型精度最高,其余参数固定统一设置为默认。

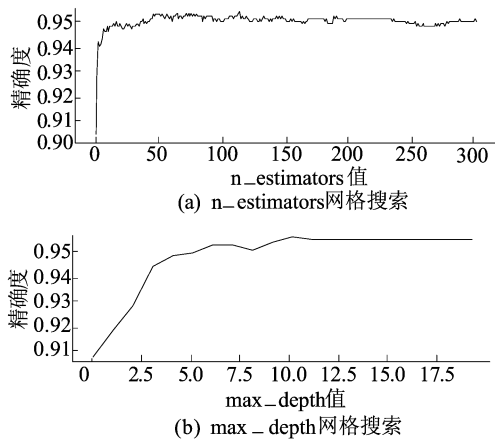


图 4 重要参数调优过程

#### 4.2.4 模型训练和精度检验

受试者工作特征(receiver operating characteristic, ROC)曲线使用敏感性作为 Y 轴,即被正确预测的概率;以 1-特异性作为 X 轴,表示被错误预测的概率。引入 ROC 曲线下方与 X 轴围成的面积 AUC(area under curve)值作为衡量模型预测精度的标准,值越大预测能力越好<sup>[51]</sup>。本文 AUC 值为 0.979,大于 0.900,说明该模型精度很好,适用于反映地下空间建设与社会经济因素之间的关系。

#### 4.2.5 模型预测和需求评价

将划分的研究区待测网格点的指标属性数据

输入训练好的模型中,得到地下建设需求可能的  $[0, 1]$  概率形式作为每个网格的输出值,利用自然断点法得到研究区地下空间需求评价等级,结果如图 5 所示。

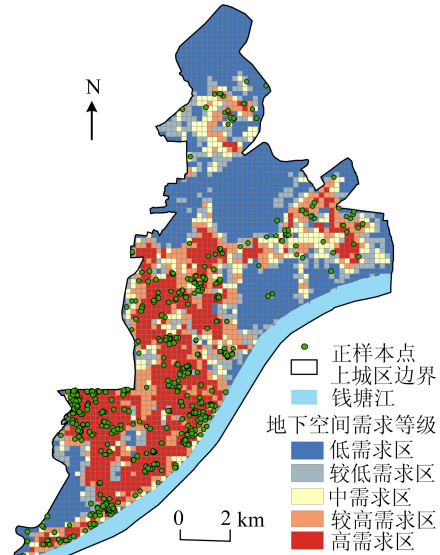


图 5 熵权-随机森林模型需求评价等级结果

#### 4.2.6 指标贡献率

在城市地下空间需求评价熵权-随机森林模型的建立中,随机森林算法对自变量与因变量之间的关系处理相当于一个黑盒,通过 Gini 系数计算出各指标因素对模型效果的贡献率,其从大到小依次为:路网密度  $X_1$  (0.284)、交通拥堵  $X_2$  (0.245)、人口数据  $X_5$  (0.208)、用地类型  $X_3$  (0.111)、夜间灯光强度  $X_6$  (0.089)、建筑密度  $X_4$  (0.062)。

## 5 讨 论

本文以上城区为实例,使用熵权-随机森林模型进行地下空间需求评价,从各等级分区及指标重要性差异方面对评价结果进行分析,依据其结果划出高开发需求区,从而为该地区今后规划发展提供借鉴。

### 5.1 需求等级分区

本文讨论熵权-随机森林模型各等级分区分布情况,并与传统熵权-线性加权模型进行对比,结果发现:

1) 图 4、图 5 直观可视化显示,2 种模型结果的评价等级单元格空间分布较为一致,较高和高需求区主要集中于上城区中南部,呈现组团式聚集分布。

2) 对各等级内现有地下设施点定量统计, 结果见表 5 所列, 熵权-随机森林模型中有 77.45% 的现有设施落入评价的高需求区内, 远多于熵权-线性加权模型的 31.01%, 说明其评价等级分区更为符合真实设施空间特征, 其算法准确性和可行性更优。

表 5 各等级内现有地下设施点数量

| 需求等级 | 熵权-线性加权模型 |       | 熵权-随机森林模型 |       |
|------|-----------|-------|-----------|-------|
|      | 数量/个      | 占比/%  | 数量/个      | 占比/%  |
| 高    | 209       | 31.01 | 522       | 77.45 |
| 较高   | 318       | 47.18 | 92        | 13.65 |
| 中    | 117       | 17.36 | 29        | 4.30  |
| 较低   | 19        | 2.82  | 13        | 1.93  |
| 低    | 11        | 1.63  | 18        | 2.67  |

## 5.2 指标重要性

在 4.1 和 4.2 节中分别使用熵权法和 Gini 系数计算出了指标因素的权重大小, 虽然权重大小有所差异, 但仍有相似之处, 综合可看出: ① 交

通拥堵  $X_2$  因素对于地下空间需求的贡献率较大, 该指标主要与人们日常出行密切相关, 城市早期向周边郊区大肆扩张, 形成扁平式城市结构, 而外来人口多居住在远离市中心的位置, 上下班极易造成早晚高峰拥堵情况, 人们对于出行便捷畅通的需求使得地下轨道交通建设显得尤为重要。② 夜间灯光强度  $X_6$  因素影响程度较小, 该指标主要与第二、三产业密切相关, 能客观反映社会经济活力, 而上城区作为杭州市重点核心区, 经济发展程度较为成熟、平均, 可能计算权重时从中学习不到足够的空间异质性特征, 从而其特征重要性相对较低。

## 5.3 未来规划建议

以熵权-随机森林模型得到的上城区实例研究结果为导向, 删去现有设施点所覆盖的网格, 在未开发区域圈出结果值在前 0~25% 的新需求潜力区大致位置①~位置⑥, 对照在线电子地图, 观察周边标志性实际地物, 对上城区未来规划提出如下建议, 如图 6 所示。

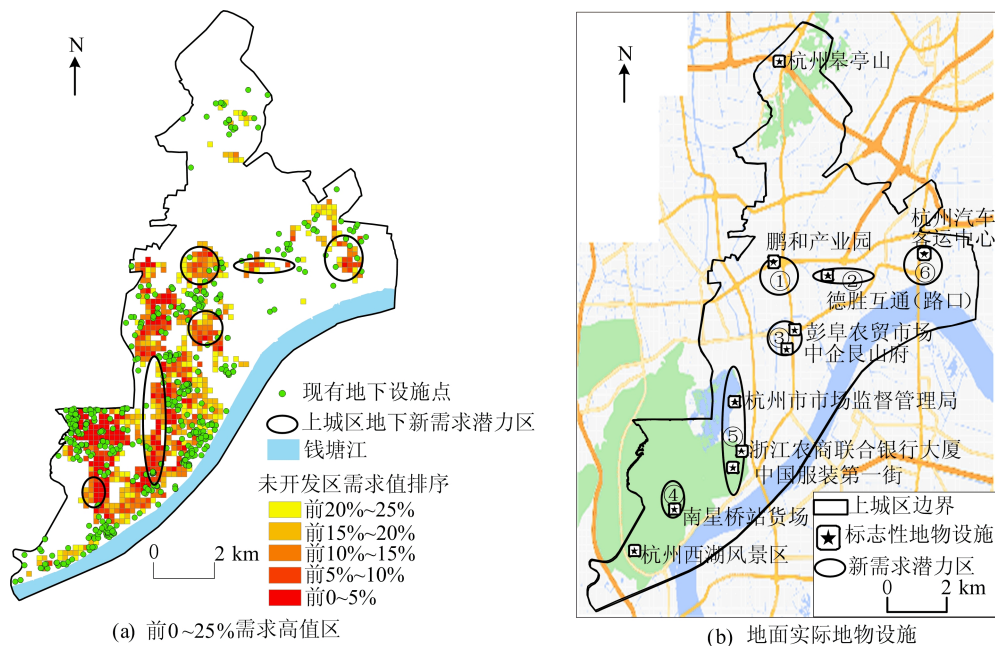


图 6 未开发区域地下需求评价结果

1) 结合需求评价结果和城市当地的区域发展确定地下空间扩展方向。根据需求度排序差异可以看出, 新需求高红色区块多位于现有地下设施周边, 说明该区域地下设施尚未饱和, 是今后进一步开发的优先建设区。对于待开发的地下空间需要与现有独立设施相结合, 通过有效整合地铁 1 号~8 号线形成的地下网络框架连接东西南北

区域, 连接现有的地下破碎单体, 建设地上地下一体化交通系统、地下商业街与文化娱乐、停车等多功能综合体, 为地下与城市整体同步发展制定长期指导方针。

2) 根据城市地面功能特点, 对不同类型地下设施开发布局差异化设计。以鹏和产业园为主的 ① 区和以南星桥站货场为主的 ④ 区, 应适时开

发地下容纳货物存放及承载物流运输的空间;以德胜互通(路口)为交通枢纽的②区和以杭州汽车客运中心为周边的⑥区,应无缝衔接轨道交通,汇集和疏散人流、物流,增强人群出行便捷性;以中企艮山府等住宅和彭阜农贸市场为集聚的③区,开发地下空间应及时满足居民停车和日常生活购物需要;以杭州市市场监督管理局、浙江农商联合银行大厦、中国服装第一街为主的行政商务核心⑤区,可以构建连续的步行网络系统,辅之以休憩和商业功能,使得地块间联系更紧密,实现内部配套资源平衡和办公需要。除此之外,以杭州市西湖、皋亭山风景区为代表的自然文化保护区,如果能发展地下设施将文化保护与旅游商贸结合,那么将极大提升游客的出游体验。

## 6 结 论

本文立足于多源地理大数据,提出一种切实可行的熵权-随机森林模型,实现对上城区网格尺度下高精度分类和定量评估,得到如下结论。

1) 多源大数据有力地弥补了传统数据时空精度低、时效性差的缺陷,也在一定程度上完善了因数据问题造成的指标构建不全面。

2) 构建的熵权-随机森林模型有较高的准确性和可行性,相比于熵权-线性加权传统方法,其评价结果更加贴近实际情况,通过增加真实地下设施参照,有效反映了社会经济指标与地下空间需求之间的复杂非线性关系和特征规律,可以为今后积极开展专项规划、扩充城区发展空间提供参考。

后续研究工作可在本文研究结果的基础上,从地质环境稳定的维护角度出发,考虑不良地质体和工程实施安全等约束,从而更好降低建设成本,避免地质环境恶化和地质灾害的发生。

## [参 考 文 献]

- [1] ZHAO J W, PENG F L, WANG T Q, et al. Advances in master planning of urban underground space (UUS) in China[J]. *Tunnelling and Underground Space Technology Incorporating Trenchless Technology Research*, 2016, 55(5):290-307.
- [2] DURMISEVIC S. The future of the underground space[J]. *Cities*, 1999, 16(4):233-245.
- [3] XU Y J, CHEN X S. Quantitative analysis of spatial vitality and spatial characteristics of urban underground space (UUS) in metro area[J]. *Tunnelling and Underground Space Technology*, 2021, 111(5):103875.
- [4] 黄莉,王直民,鲍海君,等. 城市地下空间的资源属性与开发特性分析[J]. *上海国土资源*, 2018, 39(2):37-40.
- [5] HE L, SONG Y, DAI S Z, et al. Quantitative research on the capacity of urban underground space: the case of Shanghai, China[J]. *Tunnelling and Underground Space Technology*, 2012, 32(6):168-179.
- [6] STERLING R, HAN A, BOBYLEV N, et al. Sustainability issues for underground space in urban areas[J]. *Urban Design & Planning*, 2012, 165(4):241-254.
- [7] WU L X, JIANG Y, CHE D F, et al. Fuzzy synthesis evaluation and 3D visualization for resource quality of urban underground space[J]. *Journal of China University of Mining & Technology*, 2007, 36(1):97-102.
- [8] 王睿. 城市地下空间开发的需求与前景研究[D]. 南京: 南京大学, 2018.
- [9] DOYLE M R. From hydro/geology to the streetscape: evaluating urban underground resource potential[J]. *Tunnelling and Underground Space Technology incorporating Trenchless Technology Research*, 2016, 55(5):83-95.
- [10] 王振宇,朱太宜,王星华. 长沙城市地下空间开发利用的适宜性评价体系研究[J]. *铁道科学与工程学报*, 2019, 16(5):1274-1281.
- [11] HOU W, YANG L, DENG D, et al. Assessing quality of urban underground spaces by coupling 3D geological models: the case study of Foshan city, South China[J]. *Computers & Geosciences*, 2016, 89(4):1-11.
- [12] 刘健,魏永耀,高立,等. 苏州城市规划区地下空间开发适宜性评价[J]. *地质学刊*, 2014, 38(1):94-97.
- [13] ZHOU C A, REN H, CHEN C, et al. GIS-based urban underground space resources evaluation toward three-dimensional land planning: a case study in Nantong, China[J]. *Tunnelling and Underground Space Technology*, 2019, 84(2):1-10.
- [14] 王红杰,卞孝东,邓晓伟,等. 基于模糊数学理论的煤矿地下空间开发利用适宜性评价:以白源煤矿为例[J]. *西北地质*, 2021, 54(4):156-170.
- [15] ZHOU C, REN H, ZHOU J. Fuzzy comprehensive evaluation on multiple constraint factors of city underground space development and utilization[C]//International Conference on Management Science & Engineering. [S. l.]: IEEE, 2013:357-360.
- [16] 彭苗枝,包海玲,刘道彬,等. 地下空间工程地质适宜性灰色评估[J]. *采矿技术*, 2010, 10(3):114-116.
- [17] SONG Y X, NIU R Q, XU S L, et al. Landslide susceptibility mapping based on weighted gradient boosting decision tree in Wanzhou Section of the Three Gorges Reservoir Area (China)[J]. *International Journal of Geo-Information*, 2019, 8(1):1-19.
- [18] 杨瑞君,赵楠,凡耀峰,等. 基于随机森林模型的城市空气质量评价[J]. *计算机工程与设计*, 2017, 38(11):3151-3156.
- [19] 王倩丽,马细霞,刘欣欣,等. 基于随机森林的山洪灾害风险评价方法及应用[J]. *人民黄河*, 2022, 44(4):63-66, 73.

- [20] 康有,陈元芳,顾圣华,等.基于随机森林的区域水资源可持续利用评价[J].水电能源科学,2014,32(3):34-38.
- [21] 范峻恺,徐建刚,胡宏.基于BP神经网络模型的海绵城市建设适宜性评价:以福建省长汀县为例[J].生态经济,2019,35(11):222-229.
- [22] 中国工程院战略咨询中心,中国岩石力学与工程学会.地下空间分会中国城市规划学会中国城市地下空间发展蓝皮书2021[EB/OL].(2021-12-26)[2022-08-16].<http://www.cs-rme.com/Academic/Content/index/cateid/274.do>.
- [23] 龙瀛,田乐,史肖杰.(新)城市科学:利用新数据、新方法和新技术研究"新"城市[J].景观设计学,2019,7(2):8-21.
- [24] 纪蔓梓.基于兴趣点(POI)的广州市地下空间需求预测及开发策略研究[D].广州:暨南大学,2020.
- [25] 起晓星,李建春.基于互联网地图POI数据的城市地下空间需求度评价:以济南市为例[J].中国土地科学,2018,32(5):36-44.
- [26] 吴文博.苏州城市地下空间资源评估研究[D].南京:南京大学,2012.
- [27] CUI J Q, BROERE W, LIN D. Underground space utilisation for urban renewal[J]. Tunnelling and Underground Space Technology, 2021, 108(3): 103726.
- [28] 梁朋朋.城市轨道交通站点地区地上空间与地下空间的功能耦合研究[D].成都:西南交通大学,2018.
- [29] 雷升祥,申艳军,肖清华,等.城市地下空间开发利用现状及未来发展理念[J].地下空间与工程学报,2019,15(4):965-979.
- [30] 张瑞达.基于社会经济因素的城市地下设施需求量预测研究[D].北京:北京建筑大学,2020.
- [31] 宋敏聪.基于社会经济因素的城市地下空间开发潜力评估[D].天津:天津商业大学,2012.
- [32] 赵怡婷.城市中心地区地下空间开发潜力量化评价研究[J].中华建设,2019(9):39-41.
- [33] WANG X, ZHEN F, HUANG X J, et al. Factors influencing the development potential of urban underground space: structural equation model approach[J]. Tunnelling and Underground Space Technology incorporating Trenchless Technology Research, 2013, 38(6): 235-243.
- [34] 王波.城市地下空间开发利用问题的探索与实践[D].北京:中国地质大学(北京),2013.
- [35] 林奕晨,周鹏,潘悦,等.荆州市洪涝灾害影响因子探究及风险评估:基于随机森林和XGBoost算法[J].中国农村水利水电,2022(6):125-132.
- [36] 赵智勇.融合多源地理大数据的杭州市功能区识别和空间优化研究[D].杭州:浙江大学,2018.
- [37] WorldPop. China population 2020[EB/OL].(2020-09-16)[2021-11-01].<http://www.worldpop.org/geodata/summary?id=49919>.
- [38] ELVIDGE C D, ZHIZHIN M, GHOSH T, et al. Annual time series of global VIIRS nighttime lights derived from monthly averages: 2012 to 2019 [J]. Remote Sensing, 2021, 13(5): 922.
- [39] 中华人民共和国建设部.城市规划基本术语标准:GB/T 50280—98[S].北京:中国建筑工业出版社,2010:1-28.
- [40] 蒋杰.杭州市钱塘新区地下空间开发地质适宜性评价[D].南京:中国地质科学院,2021.
- [41] 程朋根,黄毅,郭福生,等.基于多源数据的城市洪涝灾害风险评估[J].灾害学,2022,37(3):69-76.
- [42] 王富喜,毛爱华,李赫龙,等.基于熵值法的山东省城镇化质量测度及空间差异分析[J].地理科学,2013,33(11):1323-1329.
- [43] 高伟,殷小曼.基于熵值-BP神经网络的机场旅客吞吐量预测[J].计算机仿真,2021,38(10):64-67.
- [44] BREIMAN L. Random forests [J]. Machine Learning, 2001, 45(1): 5-32.
- [45] 冯磊.基于随机森林的城市生活便利性评价算法研究[D].重庆:重庆大学,2018.
- [46] GE G, SHI Z J, ZHU Y J, et al. Land use/cover classification in an arid desert-oasis mosaic landscape of China using remotesensed imagery: performance assessment of four machine learning algorithms[J]. Global Ecology and Conservation, 2020, 22(2): 971-983.
- [47] 黄发明,殷坤龙,蒋水华,等.基于聚类分析和支持向量机的滑坡易发性评价[J].岩石力学与工程学报,2018,37(1):156-167.
- [48] 白志刚,刘启蒙,刘瑜.基于熵指数与随机森林模型的滑坡易发性评价[J].人民长江,2022,53(10):95-102.
- [49] 谈树成,赵娟娟,杨林,等.基于GIS和信息量快速聚类模型的滑坡易发性研究:以云南省福贡县为例[J].云南大学学报(自然科学版),2018,40(6):1148-1158.
- [50] 朱蓬丹.基于随机森林模型的湘乡市生态安全评价及格局优化调控[D].长沙:湖南师范大学,2021.
- [51] 吴小娟.基于遥感及机器学习的油气资源潜力空间评价方法研究[D].武汉:武汉大学,2016.

(责任编辑 吴亮)