

基于 GEE 和 Sentinel-2 影像的杭州城市湿地 精细化分类研究*

潘宸^{1,2)} 侯浩^{1,2)†} 唐伟³⁾ 蒋卫国⁴⁾ 王品^{1,2)} 胡潭高^{1,2)}

(1) 杭州师范大学遥感与地球科学研究院, 浙江杭州;

2) 浙江省城市湿地与区域变化研究重点实验室, 浙江杭州;

3) 浙江省测绘科学技术研究院, 浙江杭州; 4) 北京师范大学地理科学学部, 北京)

摘要 基于 Google Earth Engine(GEE)云平台, 使用 Sentinel-2 卫星影像数据, 采用随机森林算法面向杭州城市湿地类型进行遥感分类研究. 研究表明: 在进行湿地类型精细化分类时, 相比于使用单一特征信息分类, 多特征信息的组合可以显著提高分类精度, 最优特征组合包括传统光谱特征、红边光谱特征、变换特征、纹理特征和地形特征, 其总体精度为 81.2%, Kappa 系数为 0.75; 不同特征信息对于不同湿地类型的提取具有重要作用, 传统光谱特征对于滩地和养殖池的识别非常有利, 而红边光谱特征和变换特征分别对草本沼泽和湖泊的识别更为有效. 此外, 红边光谱特征和纹理特征的组合有利于运河/水渠的识别. 通过特征对比实验可为湿地精细化分类的案例研究提供参考, 湿地分类结果可为杭州城市湿地信息的遥感识别提供数据支撑.

关键词 Google Earth Engine (GEE)云平台; Sentinel-2 影像; 随机森林算法; 湿地精细化分类

中图分类号 K901.4/TP79

DOI: 10.12202/j.0476-0301.2023136

0 引言

湿地是自然界中生物多样性最丰富的生态系统之一, 在储蓄水资源、调节河川径流、补给地下水资源、改善水源水质和维持区域水循环等方面发挥着重大作用^[1]. 2022年11月5日武汉与日内瓦联合召开《湿地公约》第十四届缔约方大会, 会议以“珍爱湿地, 人与自然和谐共生”为主题, 强调了保护、修复和持续管理湿地资源, 以及建设可持续发展湿地城市的重要性^[2]. 杭州作为一个湿地资源丰富的城市, 拥有众多的水系和河流. 然而, 随着城市化的快速推进, 杭州的湿地面积不断萎缩或消失、生态系统质量下降、功能退化等问题也逐渐浮现^[3]. 因此, 对杭州市土地覆盖和湿地类型进行分类, 能够为湿地的可持续发展以及保护治理政策的制定提供重要参考.

为了实现湿地资源的针对性保护, 需要采用更优化、精细的分类方案获取详细的湿地数据集. 然而, 目前许多土地覆盖数据集仅包含全球或国家层面的大规模数据, 这些数据集未对湿地进行详细分类. 例如: Brown等^[4]基于深度学习开发的全球10m分辨率土地覆盖分类产品 Dynamic World, 仅包含水体和淹

没植被2种湿地类型. 湿地类型的复杂性与多样性导致分类困难; 邓雅文等^[5]在对国际湿地城市常德市进行湿地精细分类的研究中指出, 目前缺乏面向国际湿地城市的湿地精细分类研究, 同时现有研究中的湿地分类类别相对较为粗糙, 未能满足湿地资源管理和保护的需要. 因此, 设计高精度、详细的分类方案, 开展更精细的湿地分类研究至关重要.

近年来, 随着遥感技术的不断发展, 针对不同地区和时序的湿地监测与分类已经成为可能. 特别是 GEE(Google Earth Engine)云平台^[6]的引入, 为大尺度长时序遥感分析提供了先进的机器学习算法和并行处理、快速图像处理能力等优势^[7-8]. Liu等^[9]基于 GEE 云平台提出了一种利用 Sentinel-2 时序影像进行不同类型热带湿地制图的综合方法, 成功获得了包含 10 种湿地类型的越南湿地覆盖图; Ghorbanian 等^[10]应用随机森林算法, 生成了具有 13 个地表覆盖类的伊朗 10 m 空间分辨率土地覆盖图. 然而, 这些研究在提取指数特征时仅限于传统光谱指数特征, 忽略了其他波段(如短波红外、中红外等)提供的信息, 导致分类精度偏低. 因此, 改进与完善特征提取过程对于获得准确的分类结果至关重要.

* 国家自然科学基金资助项目(42201106)

† 通信作者: 侯浩, 男, 博士, 副教授. 研究方向: 城市景观生态. E-mail: houhao@hznu.edu.cn

收稿日期: 2023-07-05

本文旨在基于 GEE 云平台和 Sentinel-2A 时序遥感数据,采用随机森林算法实现杭州市土地覆盖分类,并进一步得到杭州市湿地精细分类图.为实现这一目的,本文研究包括:1)构建高质量的密集时序 Sentinel-2 影像堆栈;2)设计多种湿地信息分类方案,利用随机森林算法进行分类;3)验证各方案的分类精度和可行性,评估不同方案的分类效果.研究结果可为杭州市湿地监测与保护提供良好支持,为湿地城市可持续发展做出贡献.

1 研究区与数据

1.1 研究区概况 杭州市位于中国东南沿海地区,地

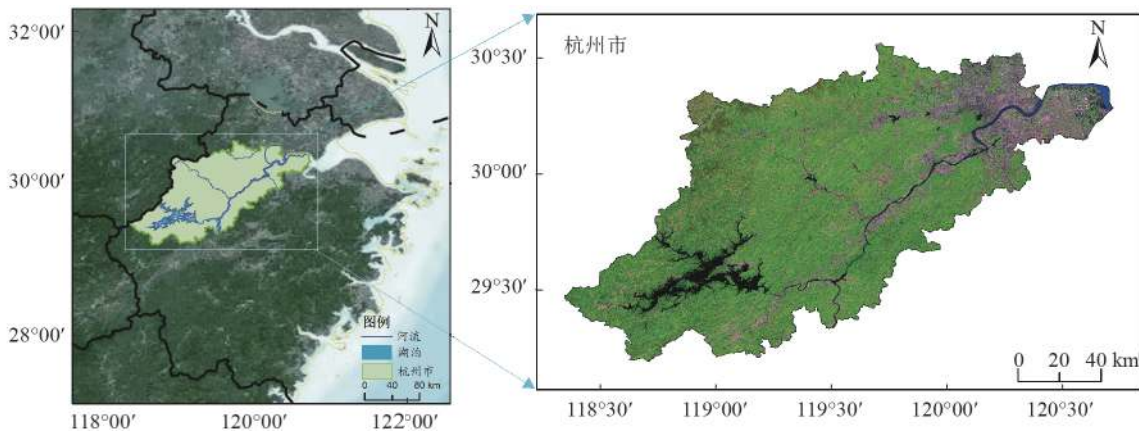


图 1 基于 Sentinel-2 影像的研究区示意

1.2 数据来源和预处理 Sentinel-2A 数据集是由欧洲航天局开发的一款高分辨率卫星影像数据集,可以通过 GEE 平台进行访问和处理.该数据集包含 13 个波段,如可见光、近红外、短波红外等,可以用于地表覆盖类型分类、植被指数计算、水资源监测等应用^[13].本文使用 2020 年 Sentinel-2A 数据作为数据源,对其进行预处理,包括去云、去阴影和大气校正等步骤.

2 研究方法

湿地精细化分类流程如图 2 所示,主要包括 5 个部分,即数据预处理、样本数据构建、分类特征信息提取与筛选、随机森林分类和精度评价与对比分析.数据预处理包括去云去阴影、大气校正、波段重采样和数据筛选等步骤.样本数据构建按照地物类型划分采样点,构建训练样本数据和验证样本数据.分类特征信息提取与筛选步骤主要提取 Sentinel-2A 数据的特征信息.随机森林分类利用训练样本数据和提取的特征信息进行分类.最后,对分类结果进行精度评价,并对各分类方案进行对比分析.

理位置为 118°21'~120°30'E, 29°11'~30°33'N(图 1).杭州市湿地地形地貌复杂,水系纵横,河口湾泾渭分明,湿地景观类型丰富多样,包括了河口滩涂、河口三角洲湿地、河口平原湿地、内陆水体湿地、人工湿地等多种类型^[11].杭州市属于亚热带季风气候,年平均气温约 17.8℃,平均相对湿度 70.3%,年降水量 1 454 mm,降水主要集中在 6—7 月份,年内季节分配不均匀^[12].由于城市化和人类活动等原因,杭州市湿地面积不断缩小,湿地生态环境受到严重破坏,因此对于杭州城市湿地的分类研究具有重要的现实意义.

2.1 土地覆盖和湿地分类方案 参考《中国土地覆盖遥感监测》与湿地公约中的分类体系^[14],结合对杭州湿地类型长时序变化的监测研究^[15],本文遵循湿地分类基本原则^[16],在研究区影像资料的基础上,建立了杭州市土地覆盖和湿地分类方案.将杭州市土地覆盖类型分为湿地和非湿地,其中湿地又分为自然湿地和人工湿地,具体分类细则如表 1 所示.

2.2 分类特征信息提取与筛选 为提高土地覆盖类型的分类精度,本研究提取了 4 种类型特征信息,包括光谱特征、变换特征、纹理特征、地形特征^[17],如表 2 所示.首先选择原始波段的地表反射率作为基本光谱特征,同时选取传统光谱指数(NDVI、MNDWI、BSI、MSAVI、LSWI、EVI、NDBI)作为光谱特征的补充.其次,提取部分红边光谱指数,包括 MCARI、PSSRI、NDI45、NDre1 和 NDre2,以进一步增强光谱特征的描述能力^[18].光谱特征详细描述如表 3 所示.本研究采用了主成分分析(PCA)和缨帽变换(TCT) 2 种特征变换方法^[19]提取影像信息.使用 PCA 法提取影像的前 3 个主成分,以最好地解释影像的变异性.通过 TCT 提取亮度(brightness)、绿度(greenness)

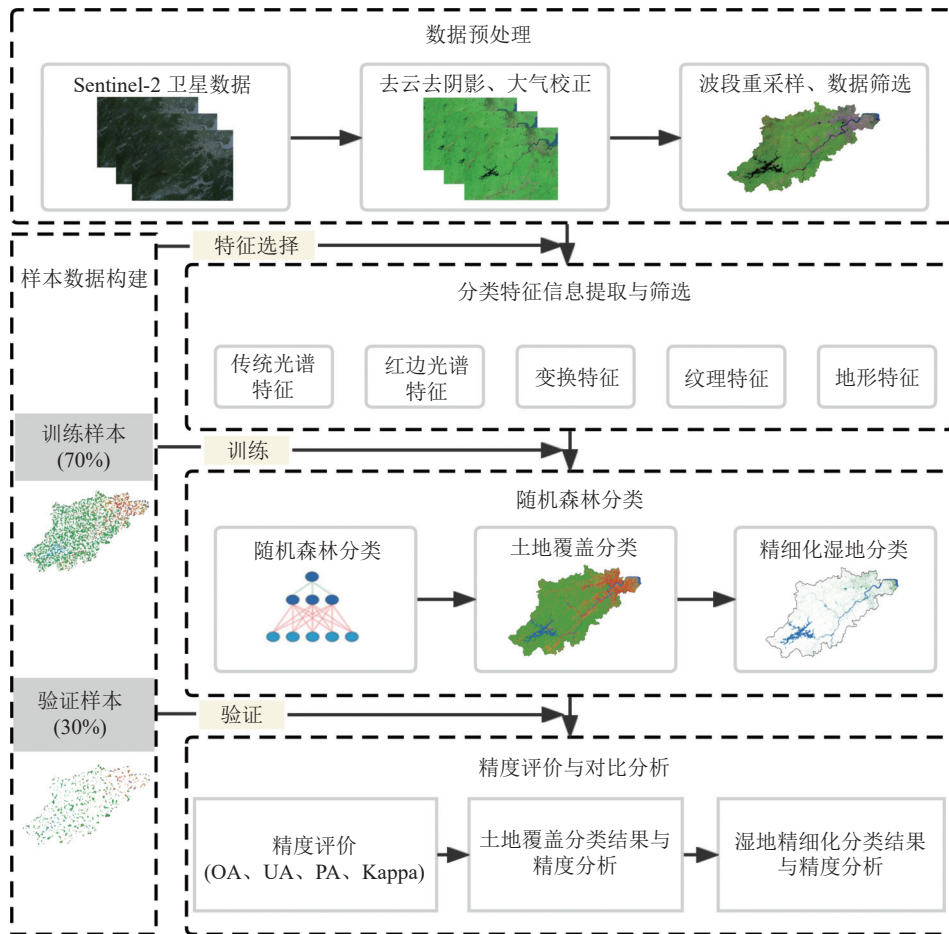


图 2 研究流程

和湿度(wetness)分量作为特征变量,代表影像中的特征信息.同时,使用灰度共生矩阵(GLCM)提取纹理特征.为了避免纹理特征数量较多而导致信息冗余,对原始地表反射率图像灰度级进行了压缩,选取使用角二阶距、对比度、相关度、方差、逆差矩、和平均、和方差、熵、和熵共 9 种纹理特征.另外将 NASADEM 30 m 数字高程数据集的坡度和海拔特征作为 2 个独立的波段添加至影像中,以提取地形特征.

2.3 随机森林分类 随机森林算法是一种基于决策树的集成学习方法,通过从训练数据中随机抽取特征来训练多棵决策树,根据结果进行分类和回归分析.本研究利用不同的特征信息来训练随机森林模型,设计了 16 种实验方案,如表 4 所示.首先,在函数中指定输入训练数据、特征名称和标签名称,设置决策树的数量为 100,每棵树的袋装样本比例为 0.7,此参数设置通过预实验确定,可生成稳定的分类精度;其次,进行第 1 次随机森林分类,获得土地覆盖类型的粗分类结果,在此基础上,进行第 2 次随机森林分类,进一步细分湿地类型;最后,利用 explain() 函数输出每个特征的重要性分数,以分析不同特征信息在随机森林

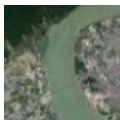


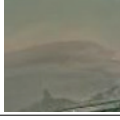
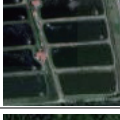
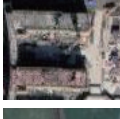

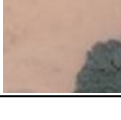
分类器分类决策过程中的重要性,调整特征组合,优化分类器的性能.

2.4 精度评价与对比分析 精度评价基于混淆矩阵计算总体精度(OA)、生产者精度(PA)、用户精度(UA)和 Kappa 系数,用于评估不同实验方案下湿地类型分类的表现,比较不同特征组合方案对精细化湿地类型识别的效果.本研究在 ArcGIS 中生成 500 个随机点,在 Google Earth Pro 中选取 2020 年高分辨率影像进行目视标记,用于分类结果的精度验证.

3 结果与分析

3.1 土地覆盖分类结果与精度分析 基于不同特征组合方案的土地覆盖分类结果如图 3 所示,包括了湿地、林地、建设用地、耕地、草地和裸地 6 种基本土地覆盖类型.表 5 展示了不同特征组合方案下土地覆盖类型分类的总体精度对比.使用单一地形特征时(实验方案 1),总体分类精度与 Kappa 系数均最低.当结合 2 种特征信息时(实验方案 2~5),分类精度显著提高.其中地形特征与传统光谱特征组合(实验方案 2)的总体精度提高 9.6%,Kappa 系数提高 0.13,在 2 个特征信息组合的方案中达到最高水平;地形特征

表 1 杭州市土地覆盖分类体系

一级类型	二级类型	类型描述	GEE影像图例
自然湿地	河流	有流水的天然线性水体	
	湖泊	有积水且面积>8 hm ² 的天然多边形水体	
	草本沼泽	草本植物占优势的自然湿地	
	滩地	在正常水位与河湖洪水位之间或低于季节性洪水位的滩面	
人造湿地	水库/坑塘	有积水且有明显坝体的人工多边形水体	
	运河/水渠	有流水且有明显坝体或直线边界的人工线性水体	
	养殖池	用于水产养殖的多边形水体	
非湿地	林地	包括森林和灌木的天然木本植被覆盖区域	
	建设用地	包括城镇、工矿、交通设施等各类住宅区且人工建筑活动形成的区域	
	耕地	以种植农作物为主的土地区域	
	草地	天然草本植物覆盖区域	
	裸地	植被覆盖率低于10%的自然覆盖区域	

与纹理特征组合(实验方案 5)的总体精度提高最少,当结合 3 种特征信息时(实验方案 6~11),相比于 2 种特征信息组合情况,分类精度提高不明显.其中,地形特征、红边光谱特征和变换特征组合(实验方案 9)的总体精度最高,达到 85.6%.在使用更多特征信息组合的情况下(实验方案 12~16),所有 5 种特征信息

组合(实验方案 16)的总体精度在所有方案中最高,为 93.0%, Kappa 系数也达到了 0.88.

3.2 湿地精细化分类结果与精度分析 不同特征组合下湿地类型的第 2 次随机森林分类结果如表 6 所示.在使用单一地形特征时(实验方案 1),总体分类精度与 Kappa 系数均最低,分别为 64.4%与 0.53.当

表 2 特征信息表

特征类型	特征信息	特征描述
光谱特征	影像自带波段	$B_2、B_3、B_4、B_5、B_6、B_7、B_8、B_{11}、B_{12}、B_{8A}$
	传统光谱特征	NDVI、MNDWI、BSI、MSAVI、LSWI、EVI、NDBI
	红边光谱特征	MCARI、PSSRI、NDI45、NDre1、NDre2
变换特征	主成分分析	提取影像主成分变换的前3个主成分
	缨帽变换	亮度、绿色、湿度分量
纹理特征	灰度共生矩阵	角二阶距、对比度、相关度、方差、逆差矩、和平均、和方差、熵、和熵
地形特征	坡度、海拔	NASADEM数据的地形特征

表 3 光谱特征描述

指数简称	指数全称	计算公式
NDVI	normalized difference vegetation index	$(B_8 - B_4) / (B_8 + B_4)$
MNDWI	modified normalized difference water index	$(B_3 - B_{11}) / (B_3 + B_{11})$
BSI	bare soil index	$((B_4 + B_{11}) - (B_8 + B_2)) / ((B_4 + B_{11}) + (B_8 + B_2))$
MSAVI	modified soil adjusted vegetation index	$[2 \times B_8 + 1 - \sqrt{(2 \times B_8 + 1)^2 - 8 \times (B_8 - B_4)}] / 2$
LSWI	land surface water index	$(B_8 - B_{11}) / (B_8 + B_{11})$
EVI	enhanced vegetation index	$2.5 \times (B_8 - B_4) / (B_8 + 6.0 \times B_4 - 7.5 \times B_2 + 1)$
NDBI	normalized difference built-up index	$(B_{11} - B_8) / (B_{11} + B_8)$
MCARI	modified chlorophyll absorption ratio index	$[(B_5 - B_4) - 0.2 \times (B_5 - B_3)] \times (B_5 - B_4)$
PSSRI	pigment specific simple ratio index	B_7 / B_4
NDI45	normalized difference index	$(B_5 - B_4) / (B_5 + B_4)$
NDre1	normalized difference red-edge1	$(B_6 - B_5) / (B_6 + B_5)$
NDre2	normalized difference red-edge2	$(B_7 - B_5) / (B_7 + B_5)$

表 4 特征组合方案

实验方案	特征组合	实验方案	特征组合
1	TF	9	TF、RESF、TRF
2	TF、TSF	10	TF、RESF、TEF
3	TF、RESF	11	TF、TRF、TEF
4	TF、TRF	12	TF、TSF、TRF、TEF
5	TF、TEF	13	TF、RESF、TRF、TEF
6	TF、TSF、RESF	14	TF、TSF、RESF、TEF
7	TF、TSF、TRF	15	TF、TSF、RESF、TRF
8	TF、TSF、TEF	16	TF、TSF、RESF、TRF、TEF

注: TF、TSF、RESF、TRF、TEF分别表示地形特征、传统光谱特征、红边光谱特征、变换特征、纹理特征。

结合 2 种特征信息时(实验方案 2~5), 分类精度显著提高. 其中地形特征与变换特征组合(实验方案 4)的总体精度在 2 个特征信息组合的方案中最高, 相较于实验方案 1 提高了 6.8%, Kappa 系数提高了 0.1, 而地形特征与纹理特征组合(实验方案 5)的总体精度提高最少. 当结合 3 种特征信息时(实验方案 6~11), 分类精度相较于 2 种特征信息组合情况提高缓慢. 其中地形特征、红边光谱特征和变换特征组合(实验方案

9) 的总体精度最高, 为 73.7%. 然而, 方案 8 和方案 11 在特征信息增加的情况下出现了分类精度不提高反而降低的情况. 在更多特征信息组合的情况下(实验方案 12~16), 所有 5 种特征信息组合(实验方案 16)最优, 总体精度为 81.2%, Kappa 系数为 0.75.

在湿地精细化分类的过程中, 增加特征可以提高分类精度, 但随着特征数量的增加, 精度提升逐渐减小. 在面对更大范围的区域应用时, 需要在提高效率保障精度的同时适度减少特征的计算参与. 图 4 展示了湿地精细分类的总精度和 Kappa 系数, 同时引入了蓝、黄、红 3 条横线, 展示不同实验方案, 满足不同尺度下湿地分类的高中低精度需求, 可为未来其他区域的分类方法应用提供参考.

图 5 展示了在不同特征组合方案下, 2020 年杭州市湿地类型的详细分类结果, 涵盖了河流、草本沼泽、滩地、湖泊、水库/坑塘、运河/水渠、养殖池 7 种湿地类型. 图 6 展示了在不同特征组合下的分类精度. 对于河流的分类精度, 在使用所有特征组合进行分类时(实验方案 16), 生产者精度和用户精度分别为 85.71% 和 77.64%, 达到最高值. 对于草本沼泽, 其生产者精度在地形特征、红边光谱特征和纹理特征

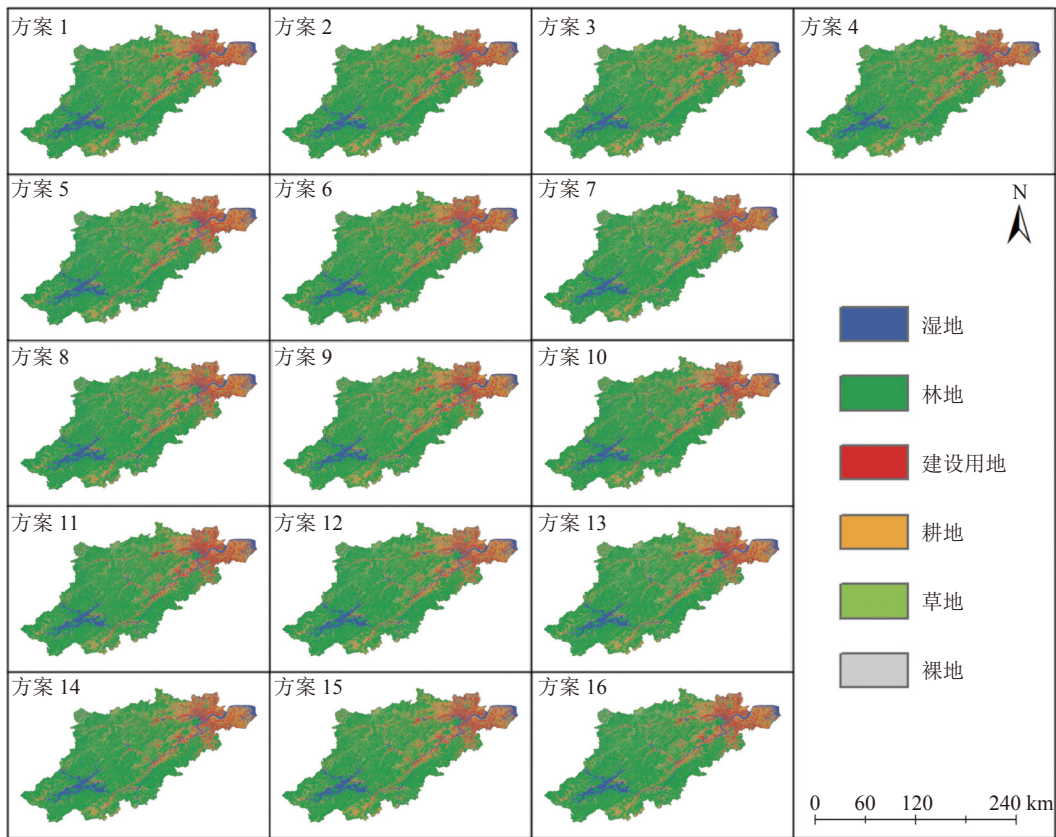


图 3 不同实验方案下 2020 年杭州市土地覆盖类型分类

表 5 不同分类方案的土地覆盖类型分类总精度对比

实验方案	特征组合	OA/%	Kappa系数	实验方案	特征组合	OA/%	Kappa系数
1	TF	71.8	0.56	9	TF、RESF、TRF	85.6	0.75
2	TF、TSF	81.2	0.69	10	TF、RESF、TEF	82.2	0.70
3	TF、RESF	79.8	0.67	11	TF、TRF、TEF	83.0	0.71
4	TF、TRF	80.6	0.68	12	TF、TSF、TRF、TEF	88.8	0.80
5	TF、TEF	75.2	0.62	13	TF、RESF、TRF、TEF	85.0	0.75
6	TF、TSF、RESF	83.8	0.72	14	TF、TSF、RESF、TEF	87.4	0.78
7	TF、TSF、TRF	84.8	0.73	15	TF、TSF、RESF、TRF	85.0	0.74
8	TF、TSF、TEF	84.0	0.72	16	TF、TSF、RESF、TRF、TEF	93.0	0.88

注: TF、TSF、RESF、TRF、TEF分别表示地形特征、传统光谱特征、红边光谱特征、变换特征、纹理特征, OA为总体精度, 下同。

组合时(实验方案 10)最高, 为 97.62%。相比于只使用地形特征和纹理特征组合时(实验方案 5), 生产者精度显著提高了 8.73%。对于湖泊, 其生产者精度在地形特征、红边光谱特征和变换特征组合时(实验方案 9)最高, 为 75.44%, 相比于只使用地形特征和红边光谱特征组合时(实验方案 3), 生产者精度显著提高了 18.03%。水库/坑塘的生产者精度与用户精度在所有实验方案中都较低, 主要是由于其样本量较少。对于养殖池, 其生产者精度在地形特征、传统光谱特征、变换特征和纹理特征组合时(实验方案 12)达到最高值, 为 71.43%, 相比于其最低值时, 即只使用地形特

征、变换特征和纹理特征组合时(实验方案 11), 生产者精度显著提高了 24.76%。

4 讨论

在使用随机森林算法对土地覆盖类型进行分类时, 不同特征组合对分类结果的重要性可能不同。在随机森林模型中, 每棵决策树的建立都是基于一部分特征子集进行的, 因此不同的特征组合会导致不同的决策树生成, 从而影响最终的分类结果^[20]。目前较多的研究提出多特征信息在地物分类方面的重要性^[21-22], 然而并未阐明如何合理选取优选特征组合, 以充分发

表 6 不同分类方案的湿地类型分类总精度对比

实验方案	特征组合	OA/%	Kappa系数	实验方案	特征组合	OA/%	Kappa系数
1	TF	64.4	0.53	9	TF、RESF、TRF	73.7	0.66
2	TF、TSF	68.9	0.60	10	TF、RESF、TEF	72.5	0.64
3	TF、RESF	70.0	0.62	11	TF、TRF、TEF	69.7	0.60
4	TF、TRF	71.2	0.63	12	TF、TSF、TRF、TEF	72.5	0.65
5	TF、TEF	66.3	0.56	13	TF、RESF、TRF、TEF	77.4	0.71
6	TF、TSF、RESF	72.2	0.64	14	TF、TSF、RESF、TEF	73.4	0.66
7	TF、TSF、TRF	70.7	0.62	15	TF、TSF、RESF、TRF	75.6	0.69
8	TF、TSF、TEF	68.5	0.60	16	TF、TSF、RESF、TRF、TEF	81.2	0.75

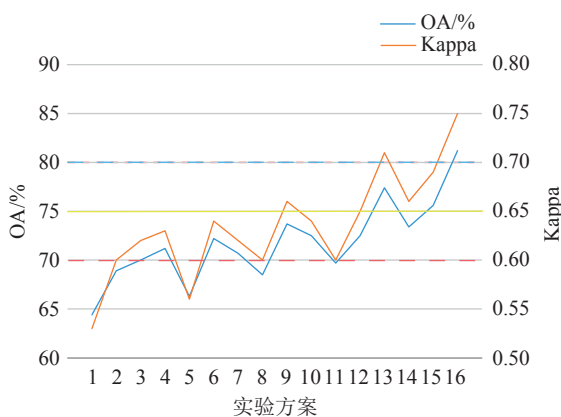


图 4 不同分类方案的湿地分类精度需求示意

挥分类不同地物时的优势。实验发现在湿地资源丰富的杭州市, 相对于只使用单一特征的实验方案, 使用多个特征组合的实验方案能够显著提高湿地类型的分类精度。然而, 一味地增加特征信息并不能持续提高分类精度。例如, 与使用 3 种特征信息组合的实验方案 9 相比, 实验方案 13 和 15 增加了 1 种特征信息, 但总体精度反而分别下降了 0.60% 和 0.58%, 这是因为这 2 种实验方案所组合的特征存在多重共线性问题, 即高度相关的特征组合可能会导致信息冗余, 使训练模型过于关注这些相关特征, 从而影响其他特征的分类能力。尽管随机森林模型通常具备处理高维特征的能力, 但多重共线性问题仍然存在, 导

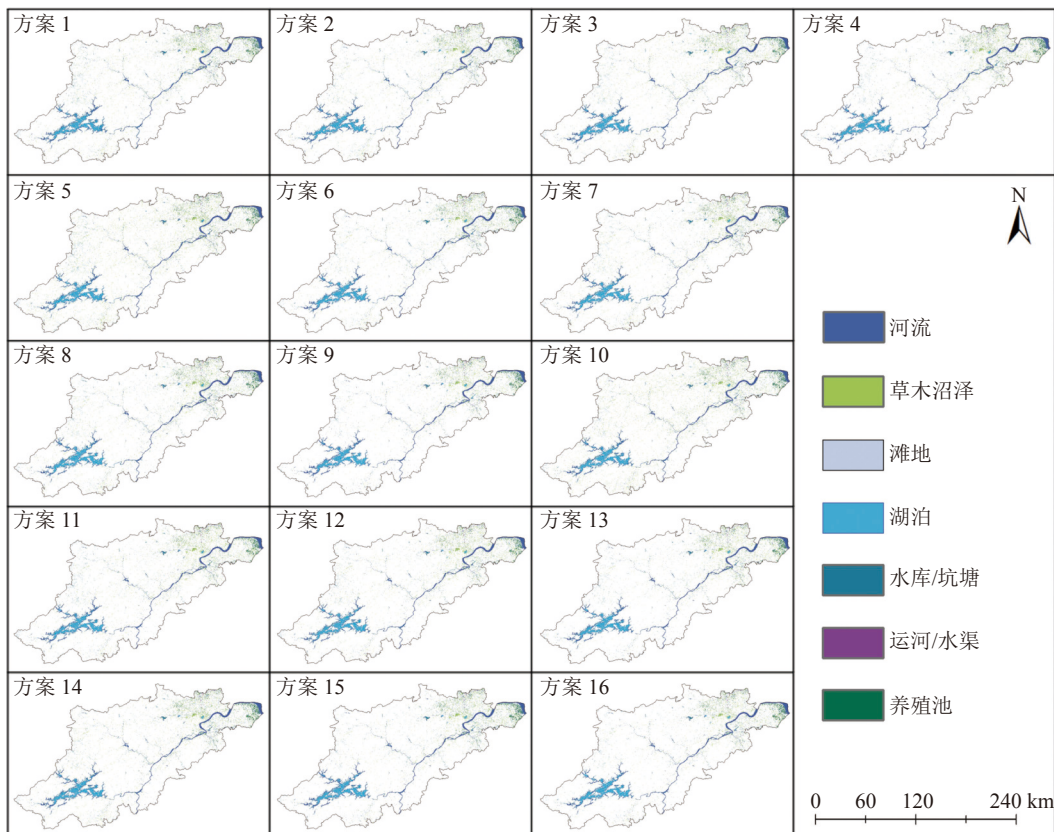


图 5 不同实验方案下 2020 年杭州市湿地类型详细分类

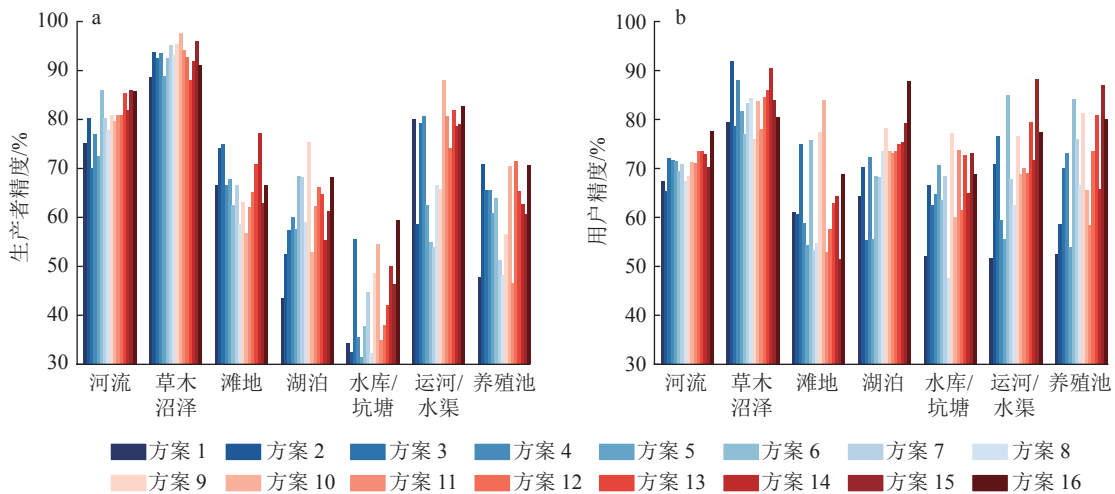


图 6 不同湿地类型的生产者精度与用户精度

致模型难以准确评估各个特征的重要性。何云等^[23]在随机森林分类中的特征优选研究中指出,多个特征之间的相关性会导致信息的冗余,进而降低分类精度并减慢分类速度,这一观点与本研究相呼应。因此,本研究进行了特征组合比对,以寻求能在降低模型复杂度的同时保证分类精度最优特征组合。

本研究中的全国土地覆盖分类体系在大多数地方具有通用性和可比性,但在涉及精细的小区域如杭州市时,可能存在一定程度的偏差。这一偏差主要由杭州市独特的地理、气候和生态环境,以及人类活动对土地利用的影响所导致。例如,杭州市拥有独特的湿地景观,包括西湖和千岛湖等自然湖泊,以及众多的人工湿地,如西溪湿地公园。这些独特情境可能无法完全符合全国通用的土地覆盖分类标准,因此,为了更好地适应杭州市的地理和生态环境,本研究在进行土地覆盖类型分类时进行了优化与改进。首先,将通用的“水体”分类更改为更具代表性的“湿地”类型,因为杭州市的许多水体以及部分周围的草本植被类型在本质上可以归类为湿地。其次,在后续的分类工作中,进一步将湿地类型精细化划分为符合杭州市湿地特点的 7 个类别,分别为河流、湖泊、草本沼泽、滩地、水库/坑塘、运河/水渠和养殖池。这一改进能更好地反映杭州市土地覆盖的实际情况,从而更好地支持杭州市土地管理和湿地相关保护决策的制定。

本研究利用 Sentinel-2A 影像的 5 种特征信息在不同程度上提高了地物区分能力,并且在不同湿地类型的识别方面也具有不同的优势。图 7 对不同湿地类型的最优与最差识别结果进行局部放大和对比。在湖泊类型识别中,实验方案 3(图 7-b)将部分湖泊误识为河流和养殖池,最优识别方案 9(图 7-a)通过应用 PCA 和 TCT 额外提取了变换特征,提高了分类精

度。在滩地和养殖池湿地类型的识别中,最优识别方案(图 7-c 和图 7-i)相较于最差识别方案(图 7-d 和图 7-j),提取了传统光谱特征,从而改善了识别效果。在识别养殖池时,传统光谱能捕捉到水体的特征反射。通过计算相应的光谱指数,如 MNDWI(modified normalized difference water index)可以提取水体特征^[24],从而辅助养殖池的识别。此外,养殖池周围可能存在植被或其他背景环境,传统光谱特征有助于分离植物与背景之间的光谱差异。在草本沼泽的识别中,最优识别方案 10(图 7-e)相较于最差识别方案 5(图 7-f)进一步提取了红边光谱特征,有效地改善了识别效果。草本沼泽通常由大量的湿生草本植被组成,而草本植被在红边波段通常表现出较高的反射率。红边光谱指数则利用红边波段的反射差异反映草本植被的生理特征^[25],因此在草本沼泽识别中具有一定优势。此外,草本沼泽地区通常包含水体,而红边波段的光谱反射对水体也具有一定敏感性,可以提取水体特征并与草本植被进行区分,这有助于进一步提高草本沼泽的识别精度。在运河/水渠的识别中,最优识别方案 10(图 7-g)通过组合红边光谱特征和纹理特征的方式,成功地减少了最差识别方案 7(图 7-h)中运河与河流严重混淆的问题。通过引入纹理特征,计算像素间的灰度差异或局部纹理模式^[26],可以揭示运河/水渠表面的空间变化和纹理细节,进而更好地识别运河/水渠。

由于湿地类型的复杂性和多样性,湿地提取过程中存在许多不确定性^[27-28]。在实验中,观察到 2 种湿地类型之间的混淆问题,如图 8 所示。其中,图 8-a 和图 8-c 为原始影像局部放大图,可以发现一些湖泊被误识别为河流(图 8-b),一些河流被误识别为水库(图 8-d)。该混淆问题的存在是由于这些湿地类型可能具有相

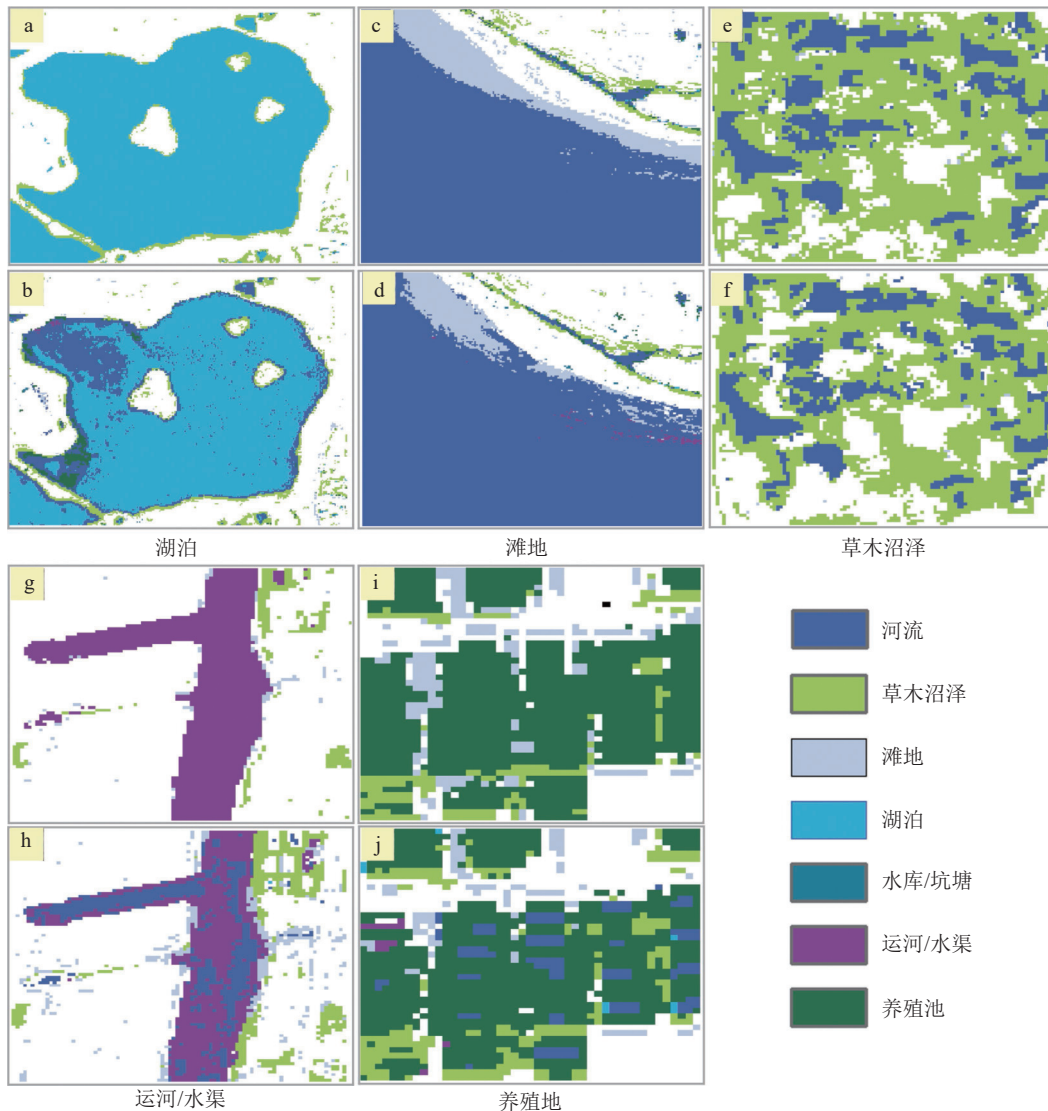


图7 不同湿地类型最优与最差识别结果局部放大

似的形状、大小和颜色等特征,例如湖泊、河流和水库的水体颜色在遥感影像中呈现相似的蓝色调或绿色调,仅通过颜色信息来区分它们存在困难^[29]。此外,水体的纹理特征相似,它们可能都具有平静的水面,难以通过纹理差异来明确分类。而且,水位、水体流动和季节性变化等因素会导致外观与特征发生变化,这些因素导致湿地的外观和特征在不同季节和时段之间发生变化^[30],算法需要识别这些变化来准确分类湿地类型,增加了分类模型的构建难度。实验发现,增加变换特征与纹理特征,可以显著提高湖泊和水库的分类正确率。通过组合多个特征信息进行分类,能够在提高分类精度的同时,有效地避免同物异谱、同谱异物和椒盐噪声等现象。

5 结论

本研究选取 Sentinel-2A 影像为数据源,提取了地

形特征、传统光谱特征、红边光谱特征、变换特征和纹理特征,通过不同特征的组合,运用随机森林算法训练模型,对 2020 年杭州市地区进行土地覆盖分类和湿地类型分类,分析不同特征组合对湿地类型分类的影响,得到以下结论:

1) 仅依赖单一特征信息的方法会导致杭州市湿地类型的严重混淆。相反,通过结合多个特征,如地形特征、传统光谱特征、红边光谱特征、变换特征和纹理特征,可以显著提高杭州市湿地复杂下垫面类型的分类精度,从而降低遥感影像特征选取对湿地分类的不确定性。综合这 5 种特征可达到最高的分类精度,总体精度为 81.2%, Kappa 系数为 0.75。

2) 特征信息在湿地类型提取中起着重要作用。实验结果显示,传统光谱特征在滩地和养殖池的识别中发挥了重要作用;红边光谱特征有利于区分草本沼泽;变换特征有利于湖泊的识别。此外,红边光谱特

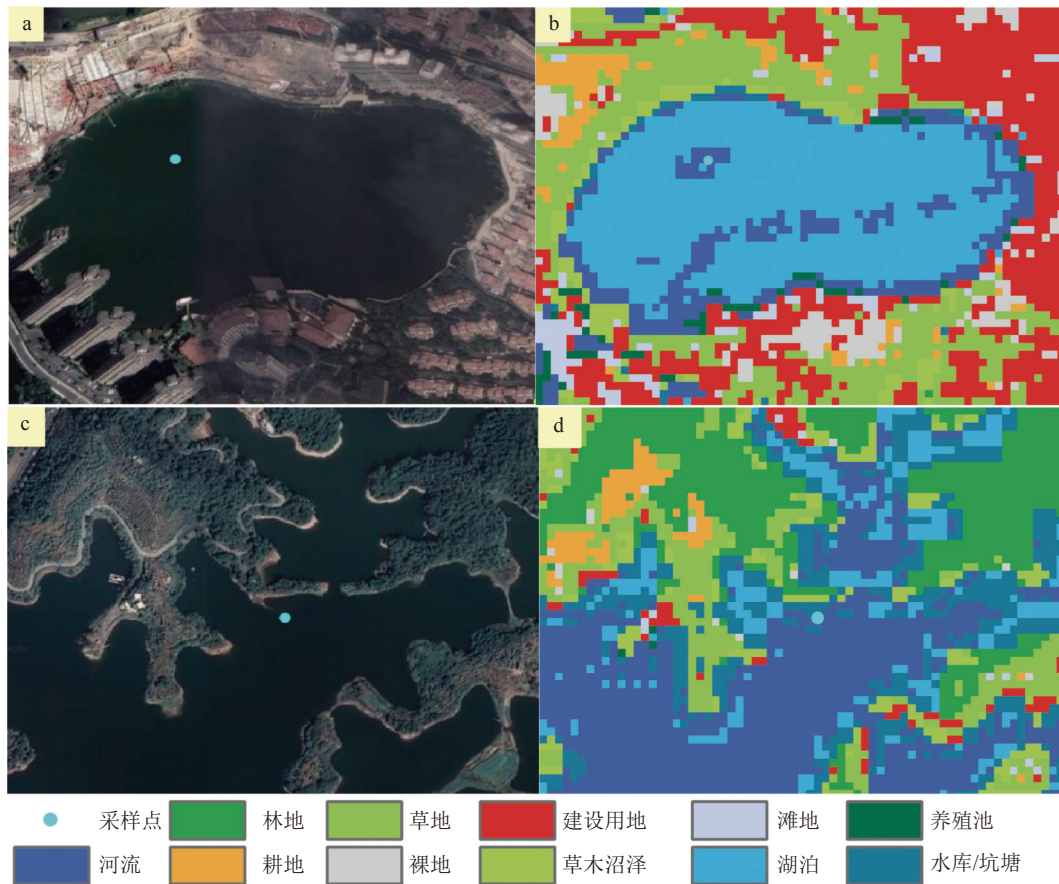


图 8 湿地类别间的混淆案例

征和纹理特征的组合有助于运河/水渠的识别。

6 参考文献

- [1] 裴理鑫, 叶思源, 何磊, 等. 中国湿地资源与开发保护现状及其管理建议 [J]. 中国地质, 2023, 50(2): 459
- [2] 符超, 金韬. 共同努力, 谱写全球湿地保护新篇章:《湿地公约》第十四届缔约方大会综述 [J]. 绿色中国, 2022(21): 10
- [3] 李智, 张慧, 刘光, 等. 近 50 年来杭州城西湿地时空变化分析 [J]. 科技通报, 2018, 34(8): 255
- [4] BROWN C F, BRUMBY S P, GUZDER-WILLIAMS B, et al. Dynamic world, near real-time global 10 m land use land cover mapping[J]. *Scientific Data*, 2022, 9: 251
- [5] 邓雅文, 蒋卫国, 王晓雅, 等. 基于随机森林算法和知识规则的国际湿地城市精细湿地分类: 以常德市为例 [J]. 遥感学报, 2023, 27(6): 1426
- [6] GORELICK N, HANCHER M, DIXON M, et al. Google earth engine: planetary-scale geospatial analysis for everyone[J]. *Remote Sensing of Environment*, 2017, 202: 18
- [7] 郝斌飞, 韩旭军, 马明国, 等. Google Earth Engine 在地球科学与环境科学中的应用研究进展 [J]. 遥感技术与应用, 2018, 33(4): 600
- [8] 王小娜, 田金炎, 李小娟, 等. Google Earth Engine 云平台对遥感发展的改变 [J]. 遥感学报, 2022, 26(2): 299
- [9] LIU Y, ZHANG H Q, ZHANG M, et al. Vietnam wetland cover map: using hydro-periods Sentinel-2 images and Google Earth Engine to explore the mapping method of tropical wetland[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2022, 115: 103122
- [10] GHORBANIAN A, KAKOOEIM, AMANI M, et al. Improved land cover map of Iran using Sentinel imagery within Google Earth Engine and a novel automatic workflow for land cover classification using migrated training samples[J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2020, 167: 276
- [11] 田晓晖, 孙庆来, 李国志, 等. 杭州市湿地资源保护对策 [J]. 湿地科学与管理, 2020, 16(3): 33
- [12] 胡云丽, 陈斌. 杭州市四季气候变化特征分析 [J]. 浙江气象, 2020, 41(2): 1
- [13] 田颖, 陈卓奇, 惠凤鸣, 等. 欧空局哨兵卫星 Sentinel-2A/B 数据特征及应用前景分析 [J]. 北京师范大学学报(自然科学版), 2019, 55(1): 57
- [14] LÓPEZ-CALATAYUD N, MÁRQUEZ-ROMANCE A M, GUEVARA-PÉREZ E. Wetland management modeling in the frame of Ramsar convention[J]. *Journal of Applied Water Engineering and Research*, 2022, 10(4): 296
- [15] 郑建华. 基于 Landsat TM/OLI 遥感影像的杭州土地利用变化时空特征研究 [D]. 杭州: 浙江农林大学, 2018.

- [16] 王翀, 林慧龙. 中国内陆天然湿地的类型特征及分布规律—I类的划分[J]. *草业学报*, 2012, 21(1): 262
- [17] HILL M J. Vegetation index suites as indicators of vegetation state in grassland and savanna: an analysis with simulated SENTINEL 2 data for a North American transect[J]. *Remote Sensing of Environment*, 2013, 137: 94
- [18] 张磊, 宫兆宁, 王启为, 等. Sentinel-2 影像多特征优选的黄河三角洲湿地信息提取[J]. *遥感学报*, 2019, 23(2): 313
- [19] MACINTYRE P, NIEKERK A, MUCINA L. Efficacy of multi-season Sentinel-2 imagery for compositional vegetation classification[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2020, 85: 101980
- [20] 马慧娟, 高小红, 谷晓天. 随机森林方法支持的复杂地形区土地利用/土地覆被分类研究[J]. *地球信息科学学报*, 2019, 21(3): 359
- [21] 杨仙保, 张王菲, 孙斌, 等. 基于 GEE 和 Sentinel-2 时序数据的呼伦贝尔沙地及其周边植被类型识别研究[J]. *遥感技术与应用*, 2022, 37(4): 982
- [22] WANG X Y, JIANG W G, PENG K F, et al. A framework for fine classification of urban wetlands based on random forest and knowledge rules: taking the wetland cities of Haikou and Yinchuan as examples[J]. *GIScience & Remote Sensing*, 2022, 59(1): 2144
- [23] 何云, 黄翀, 李贺, 等. 基于 Sentinel-2A 影像特征优选的随机森林土地覆盖分类[J]. *资源科学*, 2019, 41(5): 992
- [24] LAONAMSAI J, JULPHUNTHONG P, SAPRATHET T, et al. Utilizing NDWI, MNDWI, SAVI, WRI, and AWEI for estimating erosion and deposition in Ping River in Thailand[J]. *Hydrology*, 2023, 10(3): 70
- [25] 姚茂林, 江洪, 张丽玉. 基于 Sentinel-2 影像红边光谱指数与特征优选的竹林提取研究[J]. *海南大学学报(自然科学版)*, 2022, 40(4): 373
- [26] 闫国东, 左雪漫, 陈瑾, 等. 基于多特征优选的 Sentinel-2 遥感影像林分类型分类[J]. *森林工程*, 2023, 39(3): 12
- [27] 程丽娜, 钟才荣, 李晓燕, 等. Sentinel-2 密集时间序列数据和 Google Earth Engine 的潮间带湿地快速自动分类[J]. *遥感学报*, 2022, 26(2): 348
- [28] 梁爽, 宫兆宁, 赵文吉, 等. 基于多季相 Sentinel-2 影像的白洋淀湿地信息提取[J]. *遥感技术与应用*, 2021, 36(4): 777
- [29] 徐娜, 项超超. 基于 Sentinel-2 影像的河流湖泊遥感提取[J]. *测绘与空间地理信息*, 2023, 46(5): 120
- [30] SIYAMTHANDA G, TIMOTHY D, DOMINIC M. Leveraging Google Earth Engine platform to characterize and map small seasonal wetlands in the semi-arid environments of South Africa[J]. *Science of the Total Environment*, 2022, 803: 150139

Refined classification of the wetlands in Hangzhou by Google Earth Engine (GEE) and Sentinel-2 imagery

PAN Chen^{1, 2)} HOU Hao^{1, 2)} TANG Wei³⁾ JIANG Weiguo⁴⁾ WANG Pin^{1, 2)} HU Tangao^{1, 2)}

(1) Institute of Remote Sensing and Earth Sciences, Hangzhou Normal University, Hangzhou, Zhejiang, China;

2) Zhejiang Provincial Key Laboratory of Urban Wetlands and Regional Change, Hangzhou, Zhejiang, China;

3) Zhejiang Institute of Surveying and Mapping Science and Technology, Hangzhou, Zhejiang, China;

4) Faculty of Geographical Science, Beijing Normal University, Beijing, China)

Abstract Hangzhou possesses diverse and abundant wetland resources. With rapid urbanization, however, the total area of wetland has been shrinking. Issues such as degradation of ecosystem quality and functional decline have become increasingly severe. Consequently, it is crucial to conduct refined land cover classification of wetland types in Hangzhou, for effective wetland protection and management. In the present work, Google Earth Engine (GEE) cloud platform and Sentinel-2 satellite imagery data were used to categorize wetlands in Hangzhou with Random Forest Algorithm. A combination of multiple feature information substantially enhanced classification accuracy, compared to single-feature classification. Feature combination encompassed traditional spectral features, red-edge spectral features, transformed features, texture features, and topographic features, yielding an overall accuracy of 81.2% and a Kappa coefficient of 0.75. Different feature information contributed significantly to the extraction of various wetland types. Traditional spectral features were particularly advantageous for identifying tidal flats and aquaculture ponds. Red-edge spectral features and transformed features were more effective for recognizing herbaceous marshes and lakes

respectively. Further, combining red-edge spectral features and texture features proved beneficial for identifying canals and water channels. The feature comparison experiment in this study is useful for future studies on refined wetland classification. The wetland classification results provide data support for remote sensing identification of wetland information in Hangzhou.

Keywords Google Earth Engine (GEE) cloud platform; Sentinel-2 imagery; Random Forest Algorithm; wetland refined classification

【责任编辑:刘先勤】