

多源多特征集成的南美洲典型地区湿地制图

黄玉玲¹, 杨刚¹, 孙伟伟¹, 朱琳¹, 黄可¹, 孟祥超²

1. 宁波大学 地理与空间信息技术系, 宁波 315211;

2. 宁波大学 信息科学与工程学院, 宁波 315211

摘要: 南美洲湿地面积广且类型多样, 但湿地制图相关研究匮乏, 通过遥感手段可为南美洲全域湿地制图提供科学技术支撑。本研究依托GEE (Google Earth Engine) 平台面向南美洲湿地提出一种多源多特征集成的湿地制图方法。研究选取南美洲典型湿地地区为研究区, 首先利用已有土地覆盖数据集提出一种有效的湿地样本采集流程以保证样本质量, 其次结合哨兵1号、哨兵2号和SRTM数据构建多源特征集合, 并基于随机森林的递归特征消除算法 (RF_RFE) 进行特征优选, 构建不同特征组合方案对比多源特征对湿地分类结果的影响, 最后采用随机森林算法对研究区湿地进行分类提取。研究结果表明, 设计样本采集方案可有效提高样本质量, 多源特征集合能够提升湿地分类精度, 特征优选能够减少特征冗余并提升分类精度。研究区分类总体精度为85.62%, Kappa系数为0.8333, 其中湿地类别的精度最低为69.85%, 最高为95.18%。

关键词: 遥感, 哨兵1号 (Sentinel-1), 哨兵2号 (Sentinel-2), Google Earth Engine, 湿地分类, 特征选择, 南美洲中图分类号: P2

引用格式: 黄玉玲, 杨刚, 孙伟伟, 朱琳, 黄可, 孟祥超. 2023. 多源多特征集成的南美洲典型地区湿地制图. 遥感学报, 27(6): 1300-1319

Huang Y L, Yang G, Sun W W, Zhu L, Huang K and Meng X C. 2023. Wetlands mapping in typical regions of South America with multi-source and multi-feature integration. National Remote Sensing Bulletin, 27(6): 1300-1319[DOI:10.11834/jrs.20232265]

1 引言

湿地作为地球上对人类和自然最有价值的生态系统之一, 在维持生态系统多样性、涵养水源、保持土壤、促进全球碳和甲烷循环等方面发挥了重要作用 (Adam 等, 2010; Mahdavi 等, 2018)。但由于人类活动等因素的影响 (Junk, 1993) 湿地在全球范围内呈现了不同程度的退化趋势, 因此对不同湿地类型进行大规模监测至关重要 (Hu 等, 2017)。

与实地调查相比, 遥感技术具有覆盖范围广、信息量丰富等优势, 被广泛应用于各种土地覆盖、湿地制图与监测等研究中 (Amani 等, 2021; Mao 等, 2020; 张磊 等, 2019)。目前, 用于湿地识别

和分类的遥感数据主要包括光学数据和雷达数据。光学数据中多光谱数据虽然具备长时序、大尺度等优势, 但其光谱分辨率较低且易受云层覆盖影响; 高光谱数据的光谱分辨率较高可用于湿地精细分类, 但目前还无法提供足够的数据进行长时间、大尺度观测。SAR数据对地物的介电特性敏感, 可提供光学影像不具备的独特信息, 且不受云层覆盖影响, 但噪声现象明显且极化方式较少。如何结合多源遥感数据进行大尺度湿地精确制图仍是一项具有挑战性的任务。

研究表明, 结合光学和雷达数据能够提高湿地的分类精度 (Li和Chen, 2005; Wang等, 1998)。如Amani等 (2017) 协同使用RapidEye、Landsat光学影像和Radarsat-2 SAR影像进行加拿大纽芬兰5

收稿日期: 2022-05-23; 预印本: 2022-11-10

基金项目: 国家自然科学基金 (编号: 42122009, 41971296); 宁波市科技创新 2025 重大专项 (编号: 2021Z107); 宁波市公益项目 (编号: 2021S089); 中国博士后科学基金 (编号: 2020M670440); 浙江省省属高校基本科研业务费专项资金资助 (编号: SJLZ2022002); 浙江省自然科学基金 (编号: LR19D010001)

第一作者简介: 黄玉玲, 研究方向为湿地遥感信息提取。E-mail: 18339161167@163.com

通信作者简介: 杨刚, 研究方向为遥感影像数据质量改善与信息提取理论和方法、遥感滨海健康监测技术与应用。E-mail: yanggang@nbu.edu.cn

个研究区的湿地制图，分类的总体精度在81%—91%。但限于数据获取困难且处理过程繁琐，融合光学和雷达数据进行分类主要集中于小区域。欧洲航天局发布的哨兵1号和哨兵2号数据为协同使用光学和雷达影像进行大尺度湿地分类提供了更好的机会。哨兵1号卫星携带C波段合成孔径雷达系统，以10 m分辨率获取地表的后向散射系数，哨兵2号卫星具有红边波段和短波红外波段，结合两种数据的优势进行土地覆盖制图、湿地制图成为当前研究热点（Borges等，2020；Whyte等，2018；常文涛等，2020）。Cai等（2020）以洞庭湖湿地为研究区，结合哨兵1号和哨兵2号数据，利用面向对象的随机森林分类方法进行湿地分类，最终总体精度和Kappa系数分别为92.46%和0.92。Slagter等（2020）以南非St. Lucia湿地为研究区，结合哨兵1号、哨兵2号数据和随机森林方法构建了湿地分层分类框架，结果表明与使用单一数据相比，结合两种数据能够显著提升分类精度。其他研究表明光学、雷达以及地形数据的结合可进一步土地覆盖分类和湿地分类精度（Borges等，2020；Ruiz等，2021；Whyte等，2018）。云平台的出现为实现大尺度、多时相的湿地制图提供了新的机遇。GEE（Google Earth Engine）具备PB级的公用遥感影像数据和高速并行计算能力，用户可以通过GEE提供的编程接口（API）实现海量数据和算法的访问，极大减少了数据获取和处理的工作量（Tamiminia等，2020）。已有诸多学者利用GEE平台做出大尺度湿地制图尝试（Amani等，2019a，2019b；Mahdianpari等，2021）。多源多特征组合虽然有利于湿地类型的识别，但过多特征会造成数据冗余和维度灾难（Schratz等，2021）。递归特征消除随机森林算法（RF_RFE）是一种比较常见的特征优选方法，如周小成等（2021）发现RF_RFE算法可选出有利于林分类型分类的特征，提高了分类精度。

监督分类是遥感地表覆盖分类常用的一种方法，其中随机森林算法在湿地分类方面具有一定的优势（Amani等，2019a）。监督分类结果的好坏与样本质量直接相关，为保证分类结果的准确性，应选择具有典型性、代表性的纯净像元作为样本（Yan和Zhou，2021；郑利林等，2019）。采样方法主要包括现场实地采样和基于高分影像进行人工目视解译。实地采样具有最高的准确性，但由

于本文研究区的特殊性，无法进行实地采样，因此采用人工目视解译方法采集样本。传统的人工目视解译方法是基于Google Earth平台进行目视判别，但这种方法获取样本费时费力。同时，对于湿地而言，Google Earth能提供的影像数据类型、时相有限，无法保证湿地样本的准确性，因此，亟需提出一套有效的湿地样本选取流程以保证样本的高质性。当前已发布的多种尺度的全球土地覆盖数据集为湿地样本点的选取提供了丰富的辅助数据。Collect Earth Online平台是由美国国家航空航天局主导开发的专业、开源在线影像解译平台，主要用于土地覆盖数据的收集、管理，该平台具备数字地球（Digital Globe）、Bing Maps及GEE平台上提供的多种遥感影像源，同时还可进行光谱指数（如NDVI）时间序列分析，可为大尺度湿地样本的采集提供平台支撑（Saah等，2019）。

本研究选取南美洲典型湿地地区为研究区研发大尺度湿地分类提取方法。首先，基于现有土地覆盖数据集和湿地数据集制定大尺度湿地样本采样流程，然后结合光学数据、SAR数据及DEM数据构建多特征集合，设计多特征组合方案，结合随机森林分类器和RF_RFE特征优选方法确定适合南美地区湿地制图的分类方案，为完成南美洲全域湿地制图提供技术支撑。

2 研究区及数据源

2.1 研究区概况

南美洲大部分地区属于热带雨林气候和热带草原气候，西部有呈带状分布的热带沙漠气候和地中海气候，东南部则为亚热带季风和季风性湿润气候，年降水量为几十毫米到5000 mm不等（Junk，1993）。南美洲湿地面积占其陆地总面积的20%以上（Junk，1993；Junk等，2013；Lehner和Döll，2004）。主要包括与亚马逊河、奥里诺科等大河相关的大量洪泛湿地，由地下水、当地降雨和融雪形成的内陆湿地，以及数千公里的沿海湿地（如红树林等）（Junk等，2014；Kandus等，2018）。

综合考虑南美洲主要湿地类型和其生态区位置，本研究选取了4个5°×5°格网大小的区域作为研究区（图1）。研究区1主要气候类型为亚热带

季风和季风性湿润气候，研究区2为热带草原气候，研究区3为高原山地气候和热带草原气候，研究区4为热带雨林气候。图1中研究区矢量边界来源于ArcGIS Online世界矢量边界，一级生态区矢量为USGS EROS数据中心和美国EPA NHEERL-WED所绘，具体信息见网页 http://www.ecologicalregions.info/data/sa/sa_eco_l3.htm [2022-05-23]。

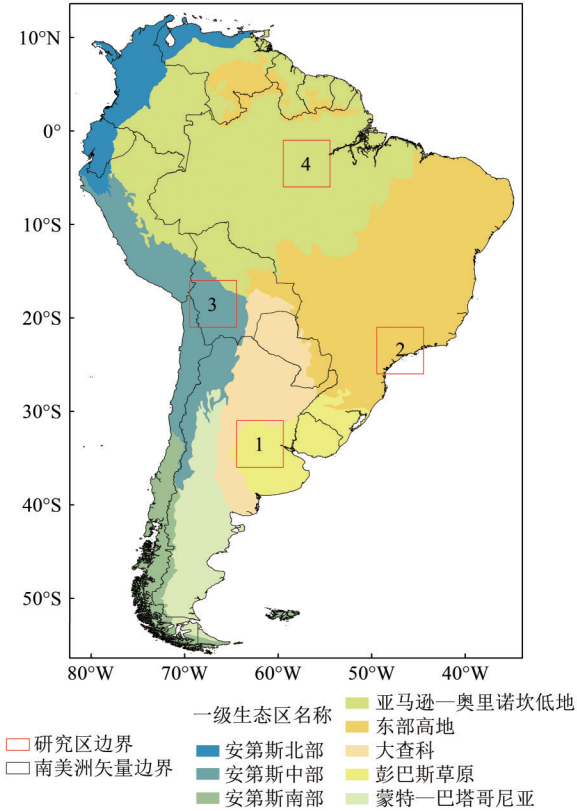
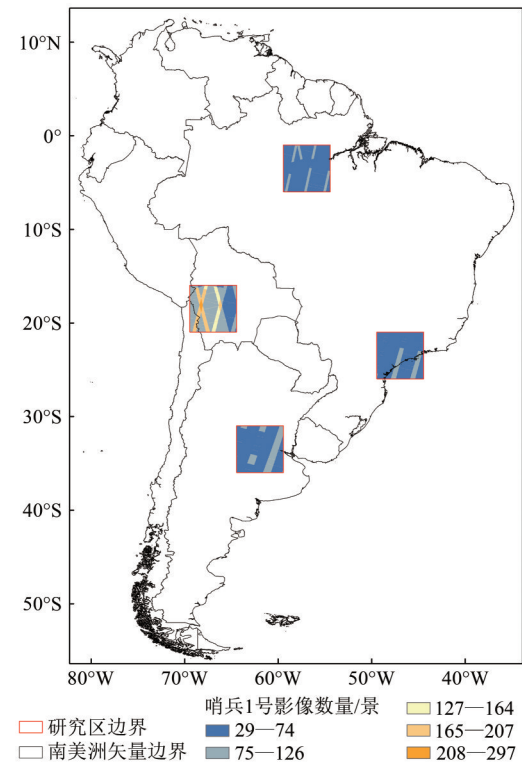


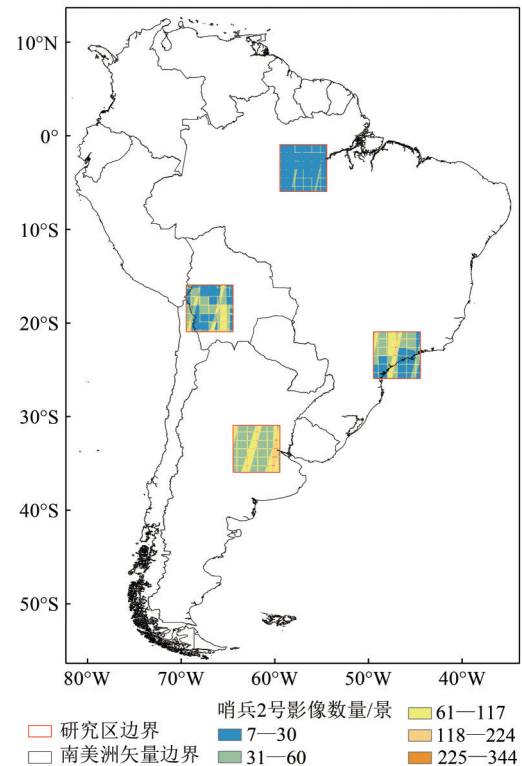
图1 研究区位置
Fig. 1 Location of the study area

2.2 卫星遥感数据

本研究基于GEE平台获取了研究区哨兵1号、哨兵2号、SRTM数据。哨兵1号数据为GEE平台提供的宽幅（IW）地距模式（GRD）的双极化数据，经过热噪声消除、辐射定标和地形校正等预处理。哨兵2号数据为GEE平台Level-2A产品，筛选得到2020年云量小于20%的影像之后利用质量评估（QA）波段进行了云掩膜处理，研究区哨兵1号和哨兵2号影像覆盖情况如图2所示。SRTM数据为由NASA JPL提供的SRTM1 3.0版本（SRTM Plus）数据，具体信息见表1。



(a) 哨兵1号影像统计图
(a) Sentinel-1 image statistics map



(b) 哨兵2号影像统计图
(b) Sentinel-2 image statistics map

图2 研究区影像数量统计

Fig. 2 Statistics of the number of images in the study area

表 1 卫星遥感数据
Table 1 Satellite remote sensing data

数据名称	传感器类型	影像获取时间	分辨率/m	重访周期/d
哨兵 1 号	C 波段雷达	2020-1-1 至 2021-1-1	10	约 6
哨兵 2 号	光学	2020-1-1 至 2021-1-1	10—60	约 5
SRTM	C/X 波段雷达	2005	30	—

2.3 辅助数据集

本研究利用 3 个土地覆盖数据集和 3 个湿地数据集作为辅助数据集。GlobeLand30 全球地表覆盖数据是由中国研制的 30 m 空间分辨率的土地覆盖数据，包括 10 个一级类型，其中 GlobeLand30 V2020 数据总体精度为 85.72%，Kappa 系数为 0.82。Esri Landcover 是 Esri 公司基于 10 m 的哨兵数据使用深度学习方法制作的全球土地覆盖数据，

该数据一共分为 10 类，据验证该数据的总体精度为 85%。ESA World Cover 土地覆盖数据集是欧州航天局主导结合哨兵 1 号数据和哨兵 2 号数据制作，包括 11 个地物类别，数据的总体精度为 74.4%。全球湖泊湿地数据集（GLWD）是世界野生生物基金会（WWF）和德国卡塞尔大学环境系统研究中心结合各种已有信息所绘，数据包括较大的水库和湖泊、较小的水体以及湿地 3 个一级类别。全球潮间带数据集（Global-Intertidal）是由马里兰大学基于 Landsat 数据生成的全球潮滩分布数据，经验证数据集的总体精度为 82.3%。全球红树林数据集（GMW）是日本航空航天局基于 ALOS PALSAR、Landsat 数据和随机森林分类器生成的红树林全球范围分布图，精度为 93.6%—94.5%。数据具体来源见表 2。

表 2 辅助数据集
Table 2 Auxiliary dataset

数据集名称	选用时间	分辨率/m	数据集来源	参考文献
GlobeLand30	2020 年	30	http://globeland30.org/defaults.html?type=data&src=/Scripts/map/defaults/browse.html&head=browse&type=data	(Chen 和 Chen, 2018)
Esri LandCover	2020 年	10	https://www.arcgis.com/apps/instant/media/index.html?appid=fc92d38533d440078f17678ebc20e8e2	—
ESA WorldCover	2020 年	10	https://esa-worldcover.org/en	—
全球湖泊湿地数据集	2004 年	约 900	https://www.worldwildlife.org/publications/global-lakes-and-wetlands-database-lakes-and-wetlands-grid-level-3	(Lehner 和 Döll, 2004)
全球潮间带数据集	2016 年	30	https://www.intertidal.app/	(Murray 等, 2019)
全球红树林数据集	2016 年	30	https://data.unep-wcmc.org/datasets/45	(Bunting 等, 2018)

2.4 湿地分类体系

参考《湿地公约》及其他相关文献资料（张海英 等，2017），结合研究区湿地分布情况，制定以下分类体系（表 3），包括木本沼泽、草本沼泽、

水体、洪泛湿地、盐田、滩涂、森林或灌丛、草地、耕地、不透水面或裸地、雪地 11 类地物，其中前 6 项为湿地类，后 5 项为非湿地类。

表 3 湿地分类体系及描述
Table 3 Wetland classification system and description

一级类别	二级类别	描述
湿地	木本沼泽	以木本植物为主的永久或季节性咸淡水沼泽
	草本沼泽	以草本植物为主的永久或季节性咸淡水沼泽
	水体	包括河流、湖泊、河口水域、水库以及城市娱乐景观、污水处理场等人工水体，主要指示为永久水体
	洪泛平原	洪泛平原是指在丰水季节有洪水泛滥的河滩、河心洲、河谷，季节性泛滥的草地以及保持了常年或季节性被水浸润的内陆三角洲(植被覆盖度低于 10%)，指河流和湖泊周围漫滩、河流冲积扇等
	滩涂	包括潮间带泥滩，岩石和沙滩部分，植被覆盖度低于 10%
非湿地	盐田	被盐蒸发岩覆盖的浅层地表，通常形成于末端湖盆或干旱环境中的洼地
	森林/灌丛	包括森林和灌丛，乔木覆盖且树冠盖度超过 30% 的土地和灌木覆盖且灌丛覆盖度高于 30% 的土地
	草地	天然草本植被覆盖，且盖度大于 10% 的土地，包括草原、草甸、稀树草原、荒漠草原，以及城市人工草地等
	耕地	用于种植农作物的土地，包括水田、灌溉旱地、雨养旱地、菜地、牧草种植地、大棚用地、以种植农作物为主间有果树及其他经济乔木的土地，以及茶园、咖啡园等灌木类经济作物种植地
	不透水面/裸地	包括人工建设形成的非植被覆盖的地表(各类建筑、道路、施工地等)以及无植被覆盖、自然存在的裸地(基岩、裸沙、裸土等)
	雪	指常年积雪

3 方法

3.1 技术路线

本研究首先依托GEE平台对哨兵1号、哨兵2号和SRTM数据进行预处理,并构建特征变量集合;其次结合辅助数据集和Google Earth Engine、Google Earth、Collect Earth Online平台构建多源数据样本采集流程获取样本数据;最后,结合不同类型特征组合和基于随机森林递归特征消除法(RF_RFE)获得的优选特征集合设置7组实验,使用随机森林分类器进行分类,以获得最佳分类方案。总体的技术路线如图3所示。

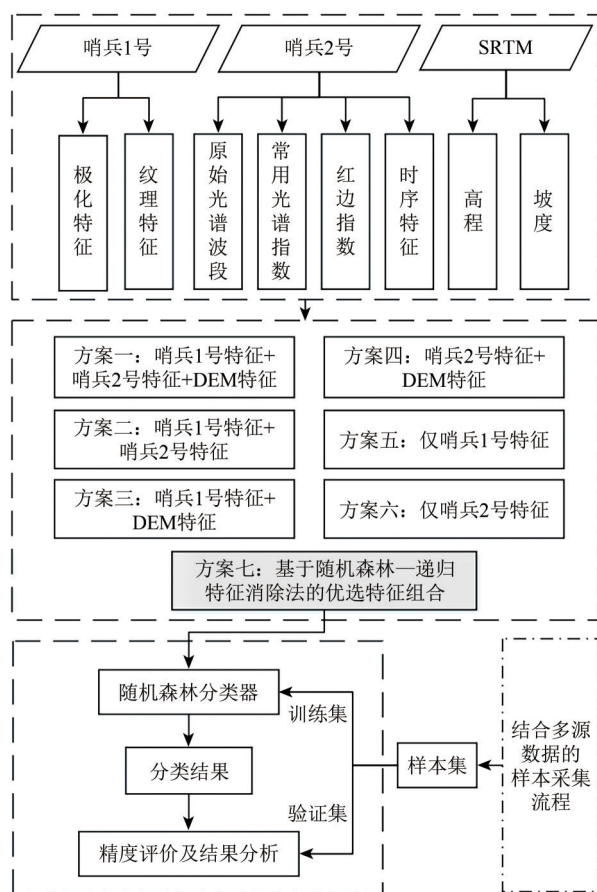


图3 技术路线

Fig. 3 Technical scheme

3.2 样本采集流程

本研究结合已有辅助数据和人工目视判别来提升湿地样本准确性。本研究首先利用ArcGIS工具在研究区范围内产生随机点,其次结合GlobeLand 30 2020年30 m的土地覆盖数据提取湿地、水体、非湿地样本点,并基于GEE平台和

Collect earth Online平台对湿地样本进行目视解译。然后利用ESRI Land Cover 2020数据、ESA Land Cover 2020和GLWD数据对非湿地和湿地样本点进行样本筛选以保证样本点的准确性和一致性,接着基于谷歌地球影像进一步对湿地样本点进行目视判别。最后为保证湿地样本数量充足,本研究结合已有专题数据和平台进行湿地样本的补充以生成最终样本集(图4)。各研究区所用各类别训练和验证样本统计如下表4所示。

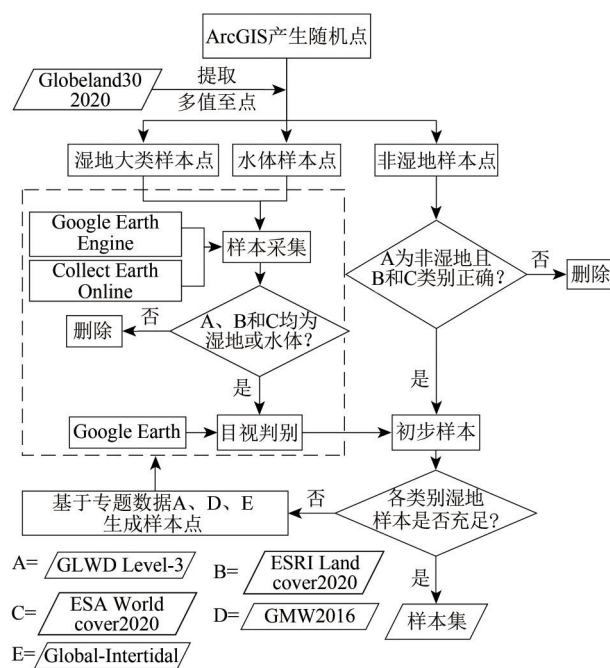


图4 结合多源数据的样本采集流程

Fig. 4 Sample collection process that combines multi-source data

3.3 湿地分类流程

3.3.1 特征提取

本研究联合多源数据派生了6类特征,共计82个特征变量(表5)。为突出哨兵2号数据多光谱波段和红边波段的优势,构建光谱特征变量、植被指数、水体指数、红边指数及NDVI时序特征;针对哨兵1号雷达数据,构建雷达极化特征和由其极化特征生成的灰度共生矩阵GLCM(Gray-Level Co-occurrence Matrix)纹理特征([https://developers.google.com/earth-engine/apidocs/ee-image-glcmtxture?hl=en\[2022-05-23\]](https://developers.google.com/earth-engine/apidocs/ee-image-glcmtxture?hl=en[2022-05-23]))。针对SRTM数据,提取其高程和坡度特征。本研究对哨兵1号和哨兵2号数据采用了年度均值合成和中值合成,年度中值合成

和均值合成是计算每个匹配波段像素的中值和均值，本研究通过这种方式增加数据可用性并提高各类别地物可分性。

表 4 各研究区样本数量统计
Table 4 Statistics of samples in each study area

研究区	木本沼泽	草本沼泽	水体	洪泛湿地	盐田	滩涂	森林/灌丛	草地	耕地	不透水面/裸地	雪	合计
1	213/95	114/43	295/113	55/19	0/0	0/0	33/27	69/33	360/160	51/13	0/0	1693
2	89/48	82/39	124/59	13/5	0/0	20/7	230/90	72/38	151/59	104/46	0/0	1276
3	30/11	13/3	62/23	126/30	112/66	0/0	108/56	241/106	30/5	69/30	12/3	1137
4	187/86	105/46	224/101	12/2	0/0	0/0	260/101	25/6	0/0	0/0	0/0	1155

注：表格中样本数量表示方式为训练/验证。

表 5 特征信息
Table 5 Feature information

数据源	特征名称	详细描述	说明或计算公式	参考文献	合成类型
哨兵 2 号	光谱波段	B2, B3, B4, B5, B6, B7, B8, B8A, B11, B12	光谱波段	—	均值(mean)、 中值 (median)
	常用光谱 指数特征	NDVI	$(B8-B4)/(B8+B4)$	(Deering, 1978)	
		NDBI	$(B11-B8)/(B11+B8)$	(查勇 等, 2003)	
		NDWI	$(B3-B8)/(B3+B8)$	(McFeeters, 1996)	
		MNDWI	$(B3-B11)/(B3+B11)$	(徐涵秋, 2005)	
	红边指数 特征	NDVIre1	$(B8A-B5)/(B8A+B5)$	(Dong 等, 2019; Fernández-Manso 等, 2016)	
		NDVIre2	$(B8A-B6)/(B8A+B6)$		
		NDVIre3	$(B8A-B7)/(B8A+B7)$		
		NDre1	$(B6-B5)/(B6+B5)$		
		NDre2	$(B7-B5)/(B7+B5)$		
	NDVI时序 特征	NDVI 10—25 分位数均值(NDVI_1025)、NDVI 25—50 分位数均值(NDVI_2550)、NDVI 50—75 分位数均值(NDVI_5075)、NDVI 75—90 分位数(NDVI_7590)		分位数特征	(Wu 等, 2021)
哨兵 1 号	雷达特征	极化特征	VV、VH 极化后向 散射系数	—	均值(mean)、 中值(median)
		纹理特征	VV、VH 极化特征均值 合成灰度共生矩阵	(Haralick 等, 1973)	—
SRTM	地形特征	高程、坡度	—	—	—

3.3.2 特征优选

特征选择方法可以减少特征空间中可能的噪声并提高预测性能 (Schratz 等, 2021)。递归特征消除 (RFE) 是一种集成方法, 结合递归特征消除和随机森林方法在特征选择上可以达到较优的性能 (Demarchi 等, 2020)。其基本过程是在随机森林特征重要性的基础上对特征进行排序, 每次从特征集合中去掉重要性小的特征, 逐步迭代到特征集为空 (杨珺雯 等, 2015)。在此过程中, 采用分层交叉验证将样本数据等分为 10 份, 即每个特征集分类时依次使用其中 9 份作为训练数据、1 份作为验证数据, 10 次分类精度的均值作为该特征集的分类精度, 最后对比不同特征集下的分类精度确定最佳特征子集 (周小成 等, 2021)。本研究

通过基于随机森林的递归特征消除法进行特征优选。

3.3.3 随机森林分类

本研究采用随机森林方法进行湿地分类提取。随机森林 (Random Forest) 是以决策树为基本分类器的一种集成学习方法 (Breiman, 2001)。随机森林建立的步骤主要分为 3 步, 首先从原始样本中通过有放回的方式抽取 N 个样本建立多个子样本集, 然后基于子样本集分别建立 N 棵决策树, 在决策树生长过程中, 每棵树的每个节点处随机抽取 m 个特征 (总特征数为 M , $m \leq M$), 根据 Gini 系数最小原则选择一个最具有分类能力的特征在决策树内部进行节点分裂, 最后将生成的决策树组成随机森林分类器, 采用投票的方式进行分类

(杨珺雯 等, 2015; 夏盈 等, 2021)。在随机森林建立的步骤中主要需要设置两个参数: 决策树的数量 N 和每棵树的每个节点的特征数 m (张磊 等, 2019)。本研究通过实验确定决策树的数目为 245, 分类节点的特征数目默认为参与分类总特征数目的平方根。

3.3.4 精度评价

常用的精度验证方法利用验证样本点计算混淆矩阵 (Confusion Matrix)。本研究将 70% 的样本数据作为训练数据, 30% 的样本数据为验证数据, 利用混淆矩阵计算总体精度 OA (Overall Accuracy) 和 Kappa 系数、用户精度 UA (User's Accuracy) 和生产者精度 PA (Producer's Accuracy), 分别对实验方案的分类结果进行精度评价。

4 实验与分析

4.1 实验对比方案

本研究设计了 7 种方案对比多源特征变量对湿地信息提取精度的影响, 其中前 6 种方案作为对比方案, 方案七为特征优选方案, 各方案中的哨兵 1 号特征为极化特征和纹理特征, 哨兵 2 号特征为原始光谱波段特征、常用光谱指数特征、红边指数特征和 NDVI 时间序列特征, 地形特征包括高程特征和坡度特征, 特征变量及方案设计见表 6。

表 6 特征组合方案

Table 6 Feature combination scheme

实验方案	特征集合
方案一	哨兵 1 号特征+哨兵 2 号特征+DEM 特征
方案二	哨兵 1 号特征+哨兵 2 号特征
方案三	哨兵 1 号特征+DEM 特征
方案四	哨兵 2 号特征+DEM 特征
方案五	仅哨兵 1 号特征
方案六	仅哨兵 2 号特征
方案七	优选特征组合

4.2 特征优选结果

从各研究区特征数量与交叉验证分数之间的关系图 (图 5) 可看出, 随着参与分类特征个数的增加, 贡献率高的变量先输入随机森林模型, 此时交叉验证分数急速上升, 到中期上升速度降低, 后期提升速度趋于平缓。研究区 1 优选特征个数为 37 个, 研究区 2 优选特征个数为 11 个, 研究区 3 优选特征个数为 40 个, 研究区 4 优选特征个数为 30 个。

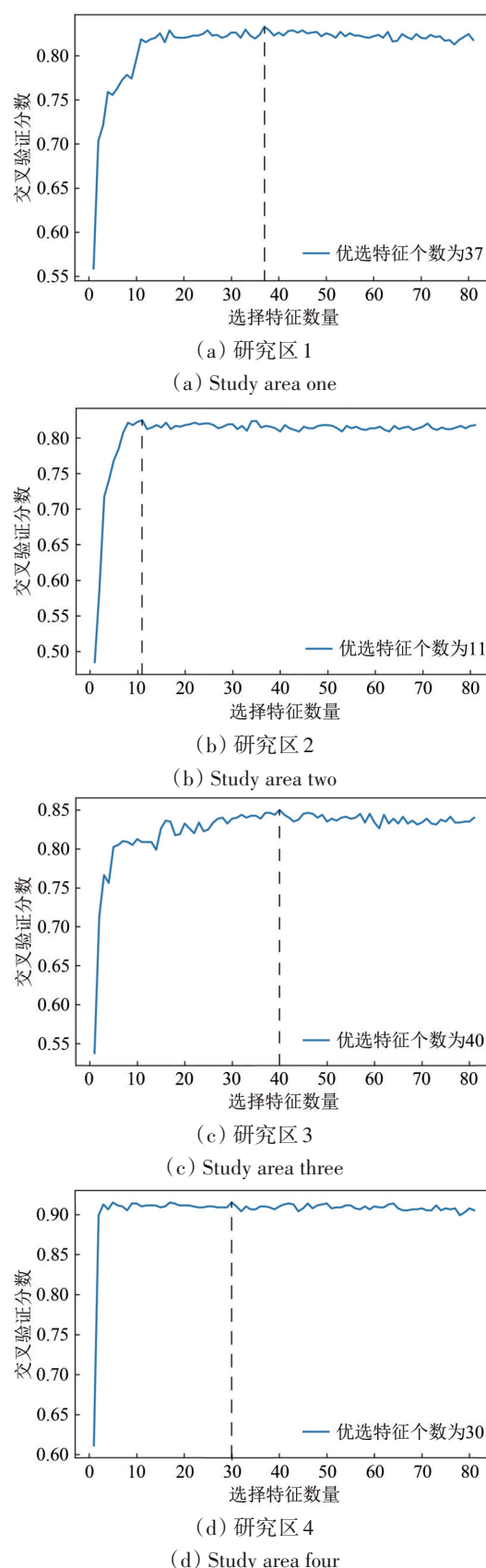


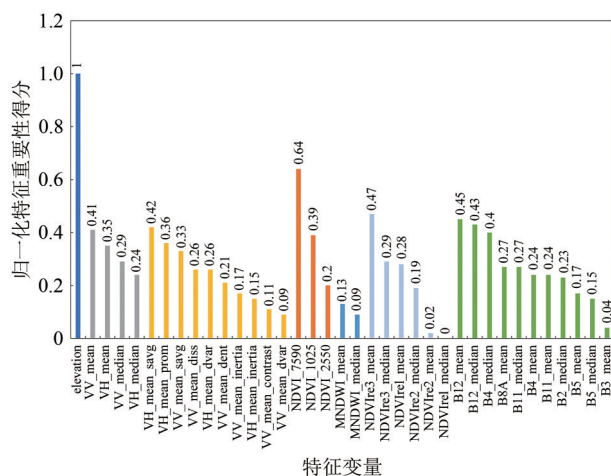
图 5 特征选择结果

Fig. 5 Feature selection results

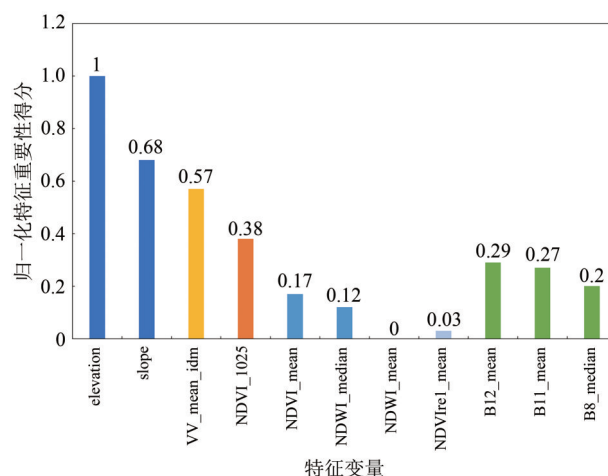
本研究根据归一化 Gini 指数重要性来评估各类特征变量重要性排序 (Nembrini 等, 2018)。

从各研究区特征变量重要性排序图（图6）可看出，（1）各研究区中高程特征均为重要性得分最高的特征。（2）各研究区的优选特征均包括3种数据源获取的特征，说明加入多源特征有助于分类。（3）就各个研究区优选特征分布来看，研究区1优势特征主要为哨兵2号原始光谱波段特征和哨兵1号纹理特征，其次为哨兵2号红边指数特征。研究区2优势特征为地形特征，其次为哨兵2号原始光谱波段特征和哨兵1号纹理特征。研究区3优势特征为哨兵2号原始光谱波段特征和哨兵2号红边指数特征，其次为哨兵1号纹理特征和地形特征。研究区4优势特征为地形特征和哨兵2号原始光谱波段，其次为哨兵2号常用光谱指数特征和哨兵2号红边指数特征。（4）就研究区整体来看，各类

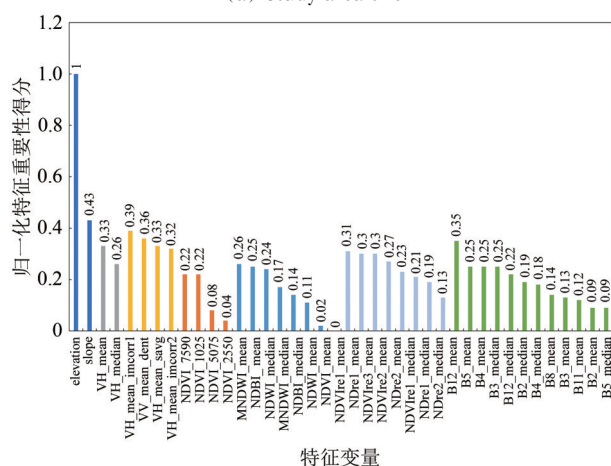
特征重要性排序为哨兵2号原始光谱波段特征>地形特征>哨兵1号纹理特征>哨兵2号红边指数特征>哨兵2号NDVI时序特征>哨兵2号常用光谱指数特征>哨兵1号极化特征。（5）研究区1归一化特征重要性得分排序前5的特征为elevation, NDVI_7590, NDVIre3_mean, B12_mean, B12_median。研究区2归一化特征重要性得分排序前5的特征为elevation, slope, VV_mean_idm, NDVI_1025, B12_mean。研究区3归一化特征重要性得分排序前5的特征为elevation, slope, VH_mean_imcorr1, VV_mean_dent, B12_mean。研究区4归一化特征重要性得分排序前5的特征为elevation, B5_median, slope, VV_mean_savg, B6_median。



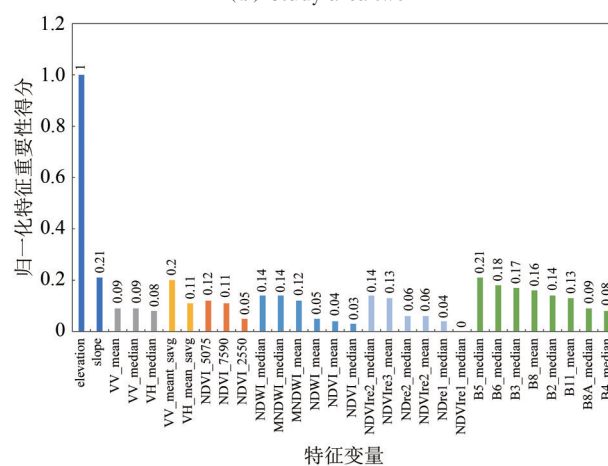
(a) 研究区1
(a) Study area one



(b) 研究区2
(b) Study area two



(c) 研究区3
(c) Study area three



(d) 研究区4
(d) Study area four

■ 哨兵2号原始光谱波段特征 ■ 哨兵2号红边指数特征 ■ 哨兵2号常用光谱指数特征 ■ 哨兵1号纹理特征
■ 哨兵2号NDVI时序特征 ■ 哨兵1号极化特征 ■ 地形特征

图6 特征变量重要性排序图

Fig. 6 Feature variable importance ranking map

4.3 分类结果及精度评价

4.3.1 分类结果比较

研究区1主要湿地类型为木本沼泽、洪泛湿地，从研究区1整体分类效果来看（图7），几种方案

整体分类效果相似。但方案三（图7（c））和方案五（图7（e））的湿地识别效果较差，存在木本沼泽和森林误分情况；方案四（图7（d））和方案六（图7（f））的不透水面/裸地这一类别错分较多；方案一、方案二和方案七整体来看分类效果相似。

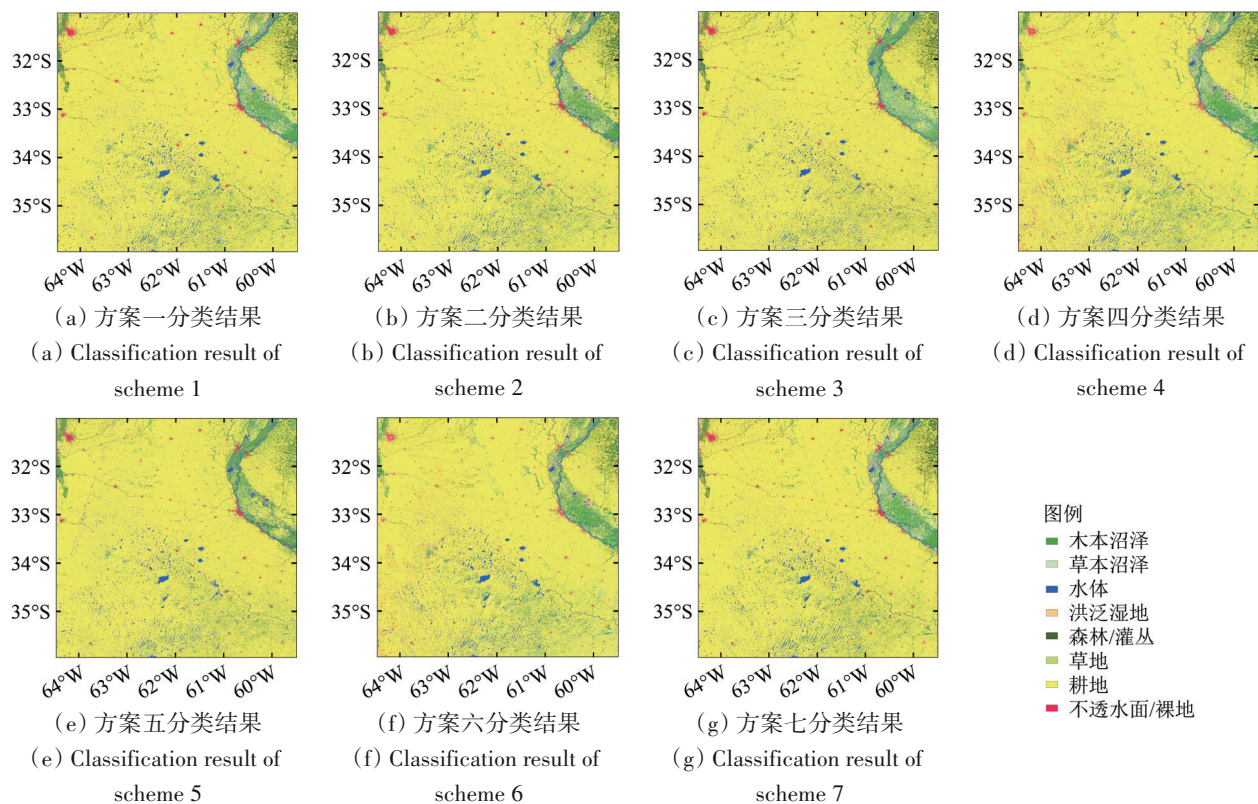


图7 研究区1的7种方案分类结果

Fig. 7 Classification results of seven schemes in study area 1

图8研究区1示例区不同方案分类结果，方案三（图8（d））和方案五（图8（f））分类效果最差，存在洪泛湿地漏分，同时草本沼泽和木本沼泽的误分也较严重；方案四（图8（c））和方案六（图9（g））中洪泛湿地与不透水面/裸地之间误分较为严重；方案二（图8（e））中耕地这一类别误分较多，方案一（图8（b））和方案七（图8（h））整体表现较为相似，但方案七水体识别目视效果更佳。

研究区2主要湿地类型滩涂及滨海的木本、草本沼泽。研究区2整体分类结果（图9）显示，方案三（图9（c））和方案五（图9（e））将滨海水体错分为滩涂；方案四（图9（d））和方案六（图9（f））整体分类效果相近，但都将耕地这一类别错分为草地。方案一、方案二和方案七整体分类效果相似。

从研究区2示例区细节图可以看出，方案三（图10（d））和方案五（图10（f））分类效果差；方案四（图10（e））和方案六（图10（g））分类效果类似，不透水面/裸地和耕地这两类别漏分严重，同时将洪泛湿地误分为不透水面/裸地；方案一（图10（b））、方案二（图10（c））和方案七（图10（h））分类表现相似，基本都能正确分出木本沼泽、草本沼泽、滩涂这几类主要湿地。

研究区3主要湿地类型为盐田。从该研究区整体分类效果来看（图11），方案三（图11（c））和方案五（图11（e））分类效果最差，盐田、不透水面/裸地、草地这三类错分、漏分较多；方案二（图11（b））和方案六（图11（f））主要将人工地表/裸地误分为洪泛湿地；方案一（图11（a））和方案七（图11（g））整体分类效果相似。

研究区3示例区的分类结果（图12）显示，

方案三（图 12（d））和方案五（图 12（f））分类效果较差，水体、不透水面/裸地和盐田误分严重，方案四（图 12（e））和方案六（图 12（g））中洪泛

湿地和不透水面/裸地误分严重；方案一（图 12（b））和方案二（图 12（c））的整体分类效果较为相似，但方案一存在洪泛湿地与水体误分现象。

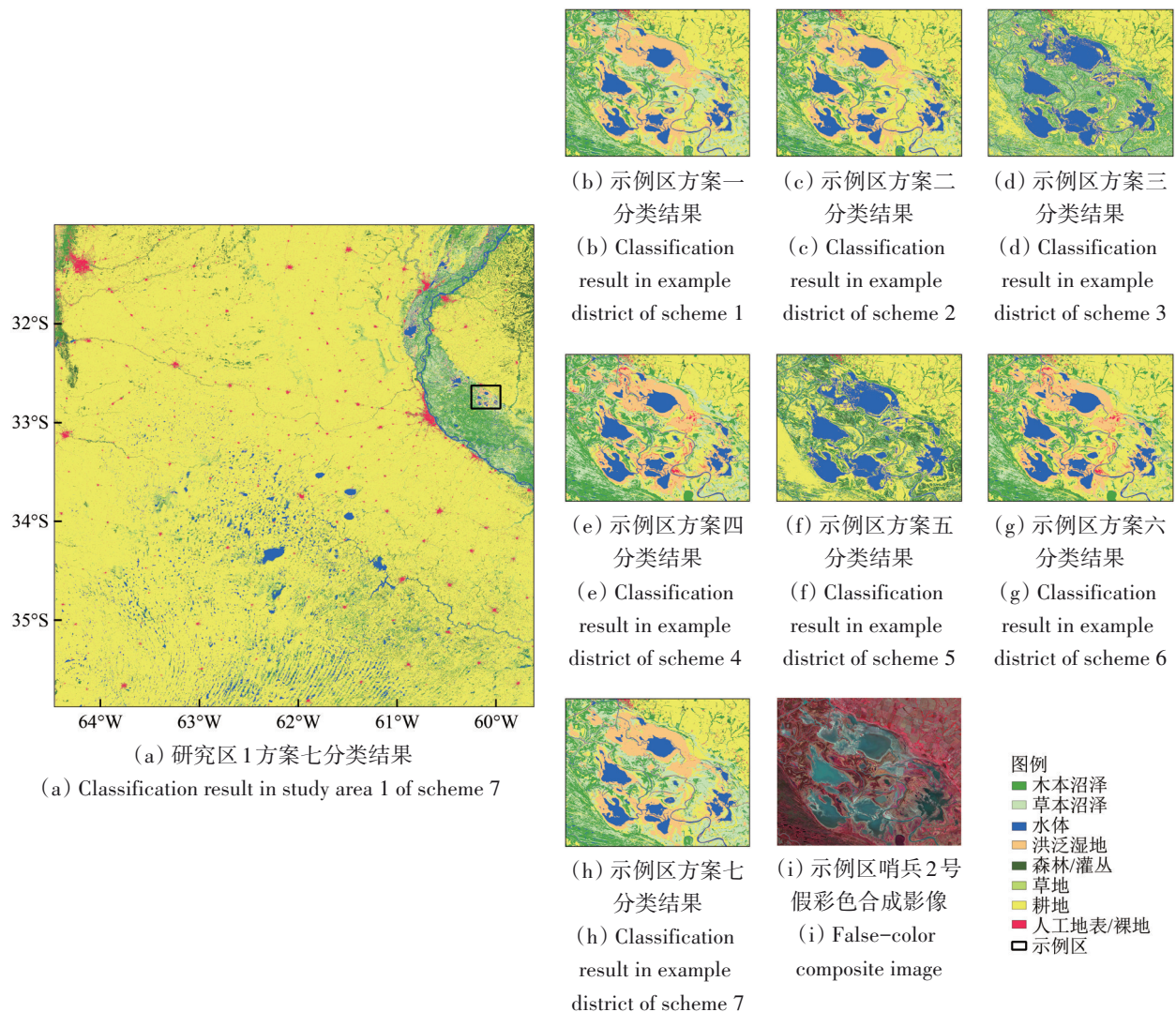
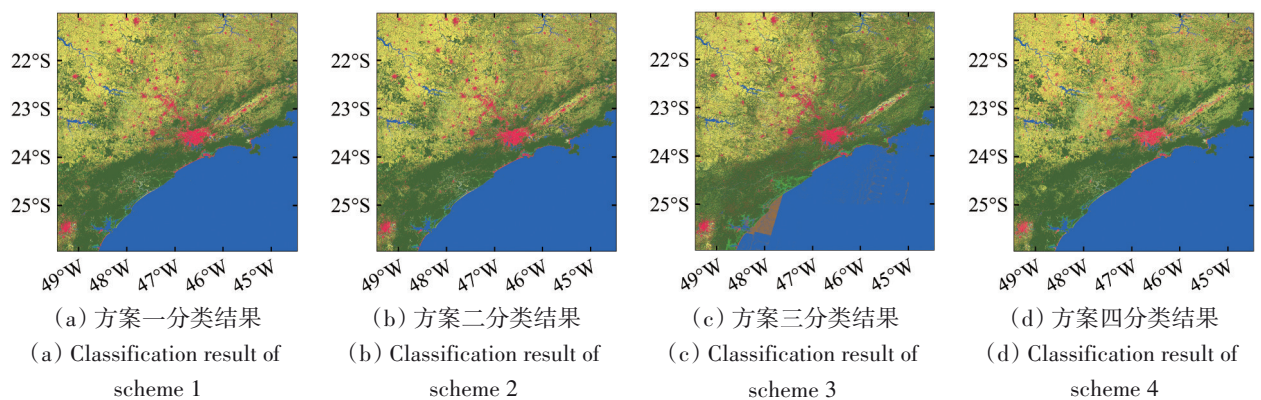


图 8 研究区 1 示例区分类结果

Fig. 8 Example area classification results in study area one



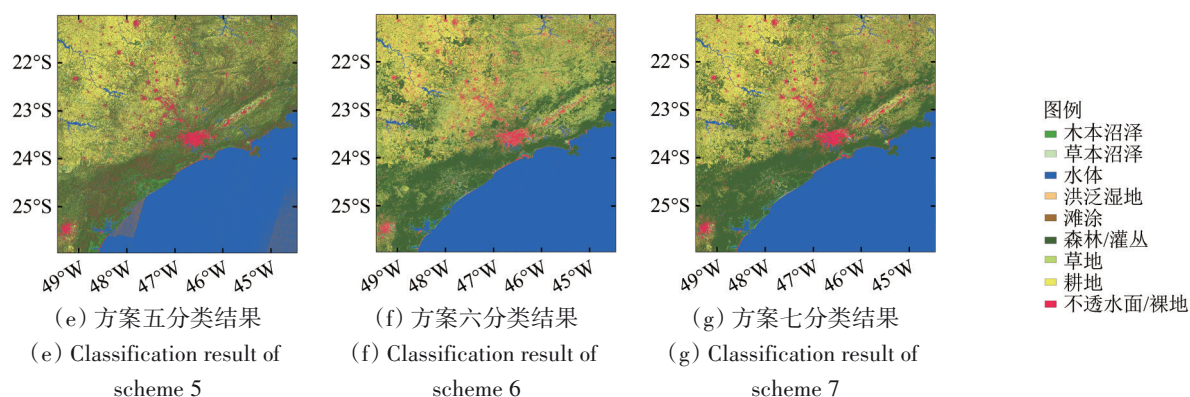


图9 研究区2七种方案分类结果

Fig. 9 Classification results of seven schemes in study area two

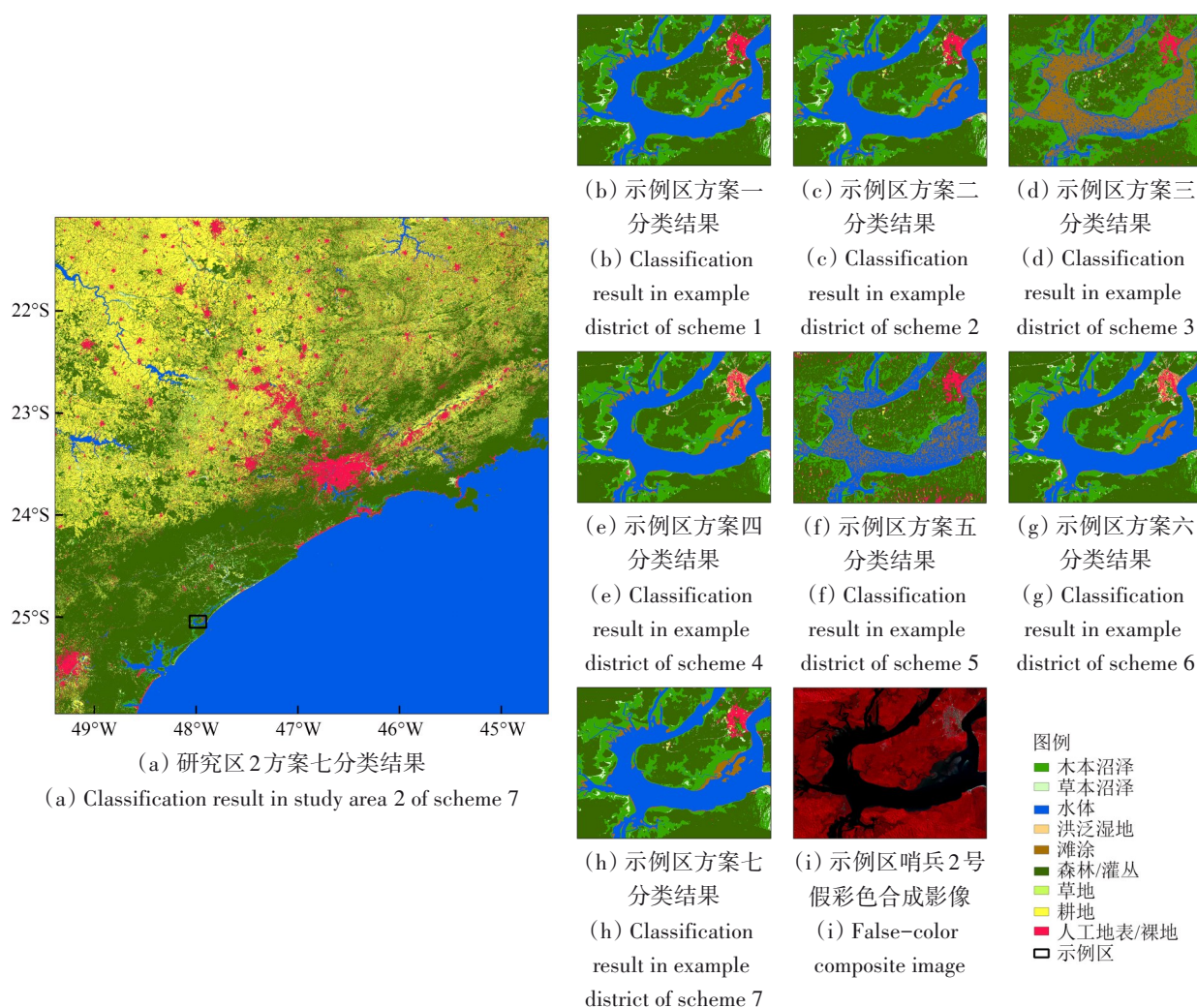


图10 研究区2示例区分类结果

Fig. 10 Example area classification results in study area two

研究区4位于亚马逊河流域, 主要湿地类型为木本沼泽和草本沼泽。受天气状态影响, 该研究区光学影像成像质量较差, 方案二(图13(b))和方案六(图13(f))分类效果都受到光学影像质量影响, 方案四(图13(d))中利用哨兵2号

特征和地形特征分类后效果改善较大, 方案三(图13(c))为利用哨兵1号特征和地形特征进行分类所得到的效果, 相比与方案五(图13(e))噪声现象也有明显改善, 方案一和方案七分类效果较好。

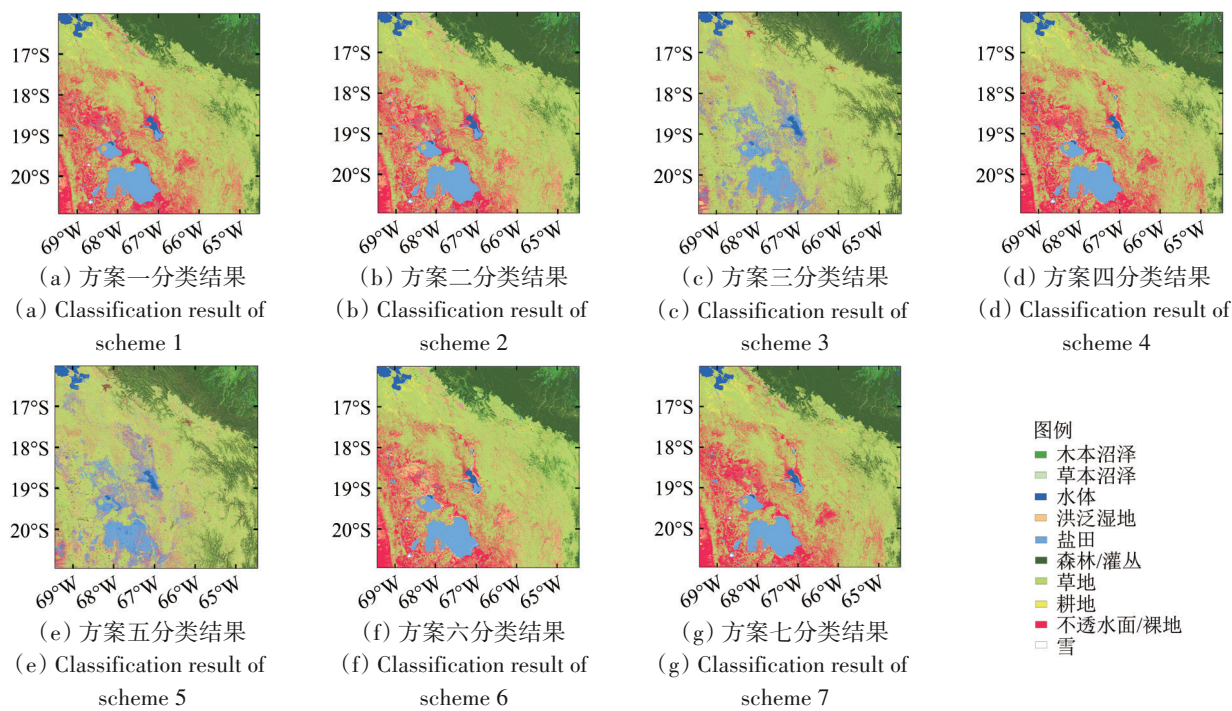


图 11 研究区 3 七种方案分类结果

Fig. 11 Classification results of seven schemes in study area three

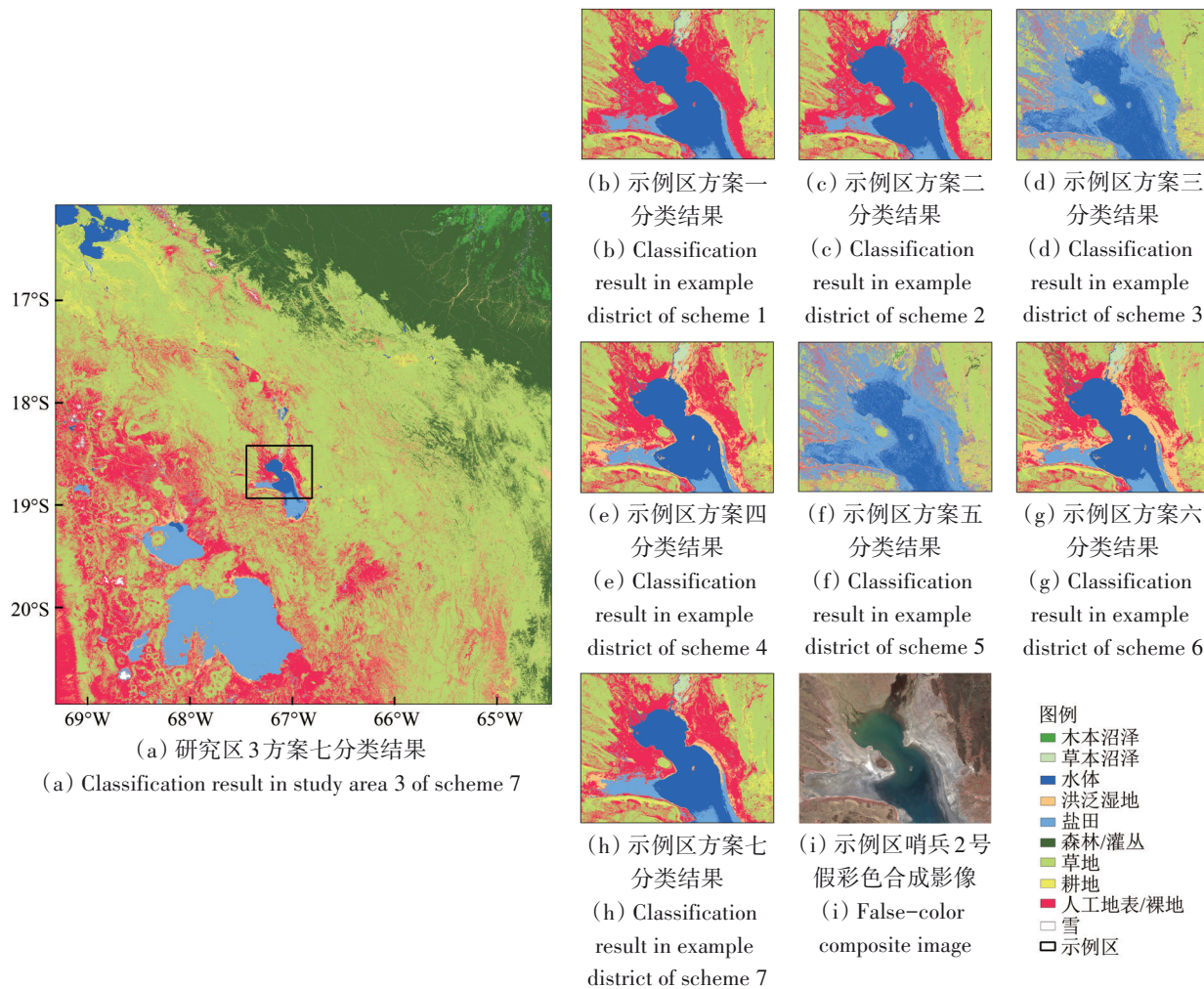


图 12 研究区 3 示例区分类结果

Fig. 12 Example area classification results in study area three

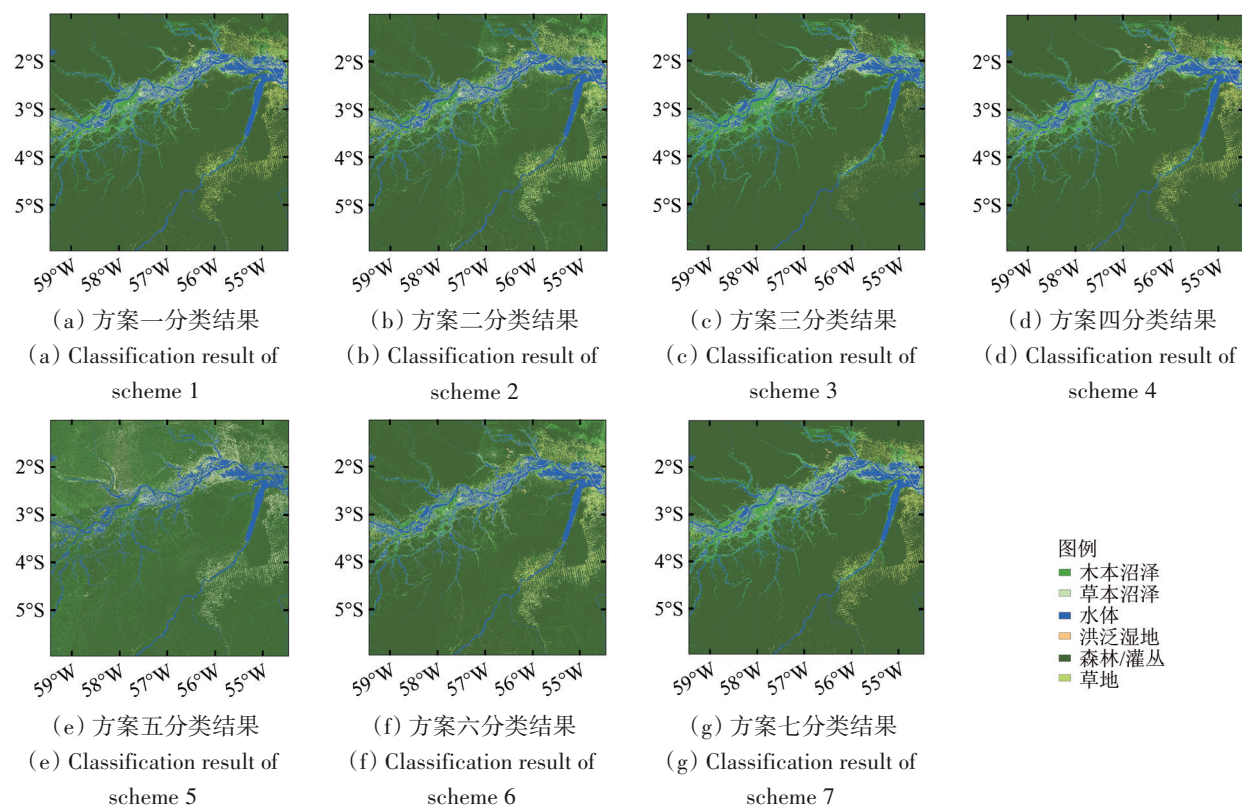


图 13 研究区 4 七种方案分类结果

Fig. 13 Classification results of seven schemes in study area four

从研究区 4 示例区分类结果可以看出, 方案五 (图 14 (e)) 分类图存在明显噪声现象, 方案三 (图 14 (c)) 在方案五 (图 14 (e)) 的基础上分类效果有所改善, 但将洪泛湿地误分为水体。方案六 (图 14 (f)) 木本沼泽与森林误分较多, 方案四 (图 14 (e)) 中草本沼泽与水体误分较多。与方案一 (图 14 (b)) 和方案七 (图 14 (h)) 相比, 方案二 (图 14 (c)) 的森林/灌丛类别漏分较多, 方案七和方案一的整体表现较为相似。

4.3.2 分类精度比较

从研究区 1 的各方案精度对比表格 (表 7) 可以看出, 方案二和方案六对比, 哨兵 1 号雷达特征加入后, 湿地类别中只有木本沼泽的 UA 提升了 20.47%, 其他湿地类别精度提升不大甚至有所降低, 非湿地类别中耕地和不透水面/裸地两类的精度有大幅提升, 而加入 DEM 特征后的方案一的总体表现要明显优于方案二, 特别是草本沼泽这一类别的精度提升较大, 方案七的整体表现要优于方

案一, 木本沼泽、水体、洪泛湿地这 3 种湿地类型的精度都表现较好。但由于非湿地样本均为辅助数据集迁移所得, 样本质量相对较差, 非湿地类别精度普遍较低。研究区 1 中森林/灌丛、草地这两类非湿地类别的制图精度低。研究区 1 中方案七精度表现最好, 总体精度为 82.11%, Kappa 系数为 0.7714。

研究区 2 的各方案精度对比表格见表 8, 方案二和方案六精度对比可发现, 木本沼泽的 UA 由 88.24% 提升至 93.33%, 草本沼泽的 UA 由 64.58% 提升至 72.09%, 非湿地类别中的草地、耕地、不透水面/裸地这 3 类的 UA 较大, 说明在研究区 2 雷达特征的加入提升了木本沼泽和草本沼泽的可分性, 且对非湿地类别的精度提升也较为明显。方案七的湿地类别精度整体表现优于其他方案, 与方案一对比可发现, 草本沼泽的 UA 和 PA 提升了 3.26% 和 7.69%, 滩涂这一类别的 UA 虽然有所下降, 但 PA 由 57.14% 提升至 85.71%。研究区 2 中方案七精度表现最好, 总体精度为 83.89%, Kappa 系数为 0.8101。

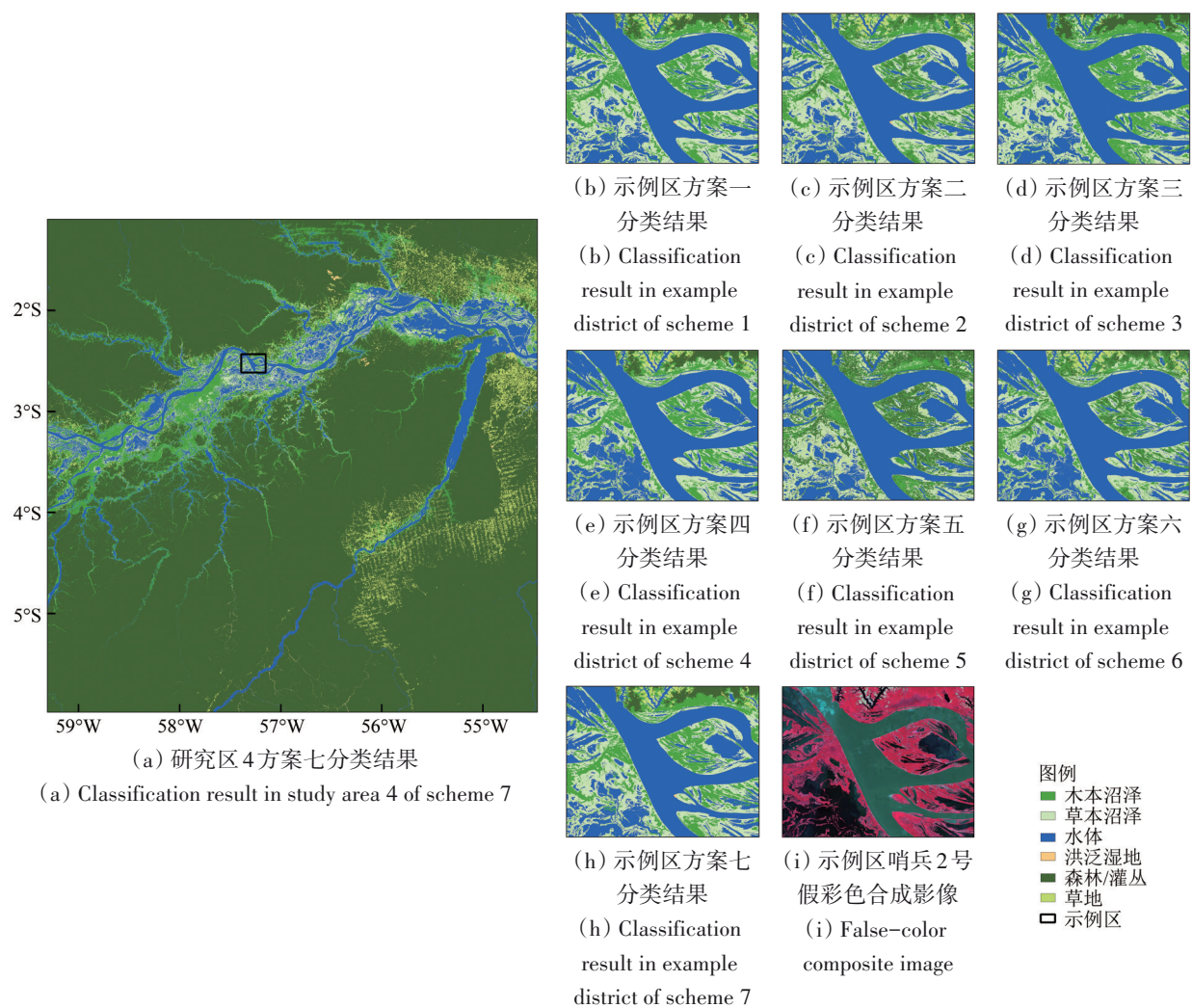


图 14 研究区4示例区分类结果

Fig. 14 Example area classification results in study area four

表 7 研究区1各方案精度统计

Table 7 Accuracy statistics of each scheme in study area one

类别	方案一		方案二		方案三		方案四		方案五		方案六		方案七	
	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%
木本沼泽	70.83	89.47	86.32	68.91	72.45	74.74	67.50	85.26	69.31	73.68	65.85	85.26	71.19	88.42
草本沼泽	68.75	51.16	39.53	53.13	50.00	65.12	53.57	34.88	47.92	53.49	38.46	23.26	63.89	53.49
水体	96.30	92.04	92.92	95.45	92.11	92.92	93.86	94.69	91.15	91.15	93.75	92.92	96.33	92.92
洪泛湿地	70.00	73.68	68.42	72.22	100.00	42.11	85.71	63.16	100.00	31.58	72.22	68.42	83.33	78.95
森林/灌丛	72.73	29.63	22.22	54.55	75.00	22.22	71.43	18.52	50.00	14.81	60.00	11.11	70.00	25.93
草地	50.00	33.33	33.33	52.38	66.67	12.12	46.15	36.36	40.00	6.06	42.31	33.33	60.00	36.36
耕地	86.93	95.63	96.25	86.52	77.50	96.88	87.64	97.50	73.81	96.88	85.88	95.00	86.52	96.25
不透水面/裸地	92.86	100.00	100.00	92.86	92.31	92.31	75.00	92.31	91.67	84.62	75.00	92.31	92.86	100.00
总体精度/%	81.51		79.72		77.34		79.52		74.35		76.94		82.11	
Kappa系数	0.7641		0.7408		0.7069		0.7377		0.6661		0.7048		0.7714	

注：加粗数值说明精度表现较好。

表8 研究区2各方案精度统计
Table 8 Accuracy statistics of each scheme in study area two

类别	方案一		方案二		方案三		方案四		方案五		方案六		方案七	
	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%
木本沼泽	96.77	62.50	93.33	58.33	81.58	64.58	96.97	66.67	63.33	39.58	88.24	62.50	94.59	72.92
草本沼泽	70.21	84.62	72.09	79.49	58.97	58.97	66.67	87.18	52.17	30.77	64.58	79.49	73.47	92.31
水体	93.55	98.31	92.06	98.31	85.07	96.61	98.31	98.31	78.08	96.61	95.00	96.61	98.33	100.00
洪泛湿地	66.67	80.00	42.86	60.00	100.00	20.00	66.67	40.00	0.00	0.00	50.00	20.00	75.00	60.00
滩涂	100.00	57.14	100.00	42.86	75.00	42.86	66.67	85.71	0.00	0.00	75.00	85.71	75.00	85.71
森林/灌丛	78.10	91.11	75.93	91.11	60.98	83.33	81.19	91.11	51.06	80.00	75.96	87.78	80.39	91.11
草地	87.50	55.26	84.00	55.26	71.43	13.16	69.70	60.53	23.08	7.89	67.74	55.26	84.62	57.89
耕地	76.92	84.75	73.13	83.05	73.13	83.05	64.79	77.97	68.66	77.97	61.97	74.58	75.00	86.44
不透水面/裸地	82.98	84.78	80.00	78.26	77.78	76.09	83.87	56.52	74.42	69.57	81.82	58.70	91.89	73.91
总体精度/%	82.10		79.54		71.36		79.03		61.64		75.70		83.89	
Kappa系数	0.7886		0.7581		0.6584		0.7530		0.5386		0.7132		0.8101	

注：加粗数值说明精度表现较好。

研究区3地物类型复杂，该研究区各方案精度对比表格（表9）可看出，方案三和方案五湿地类别精度表现最差，其他各方案湿地类别精度相差较小，方案一本本沼泽和洪泛湿地精度表现最好，方案七盐田精度表现最好。方案四的总体精度表现最好，比方案七总体精度高0.3%，Kappa系数高0.004，分析发现方案七与方案四的第一个区别在于方案七中具有SAR极化特征和由极化特征生成的纹理特征，且对洪泛湿地和盐田两类湿地的分类精度要高于方案四，说明SAR本身的极化特

征和纹理特征对湿地识别具有贡献，能够为光学特征提供有效补充；第二个区别在于方案七的光学特征变量总数为32个，方案四光学特征变量总数为42个，方案七的非湿地类别精度表现和总体精度表现低于方案四，说明丰富的光谱特征有利于非湿地类别分类，同时说明特征优选方案也可能存在特征冗余，导致特征优选方案没有获得最佳精度表现。研究区3方案四总体精度表现最好，总体精度为86.49%，Kappa系数为0.8322。

表9 研究区3各方案精度统计
Table 9 Accuracy statistics of each scheme in study area three

类别	方案一		方案二		方案三		方案四		方案五		方案六		方案七	
	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%
木本沼泽	66.67	72.73	66.67	72.73	47.06	72.73	66.67	72.73	33.33	45.45	61.54	72.73	53.85	63.64
草本沼泽	60.00	100.00	60.00	100.00	0.00	0.00	60.00	100.00	0.00	0.00	60.00	100.00	60.00	100.00
水体	82.61	82.61	82.61	82.61	75.00	65.22	82.61	82.61	63.64	60.87	82.61	82.61	82.61	82.61
洪泛湿地	83.33	83.33	73.53	83.33	60.98	83.33	75.76	83.33	39.47	50.00	64.71	73.33	80.65	83.33
盐田	95.45	98.44	92.65	98.44	82.61	86.36	94.03	98.44	80.00	78.79	91.30	98.44	94.12	100.00
森林/灌丛	92.86	69.64	92.86	69.64	81.40	62.50	95.24	71.43	85.00	60.71	95.12	69.64	92.86	69.64
草地	83.47	95.28	82.79	95.28	72.95	83.96	85.59	95.28	62.31	76.42	84.87	95.28	84.87	95.28
耕地	66.67	40.00	50.00	20.00	66.67	40.00	80.00	80.00	0.00	0.00	100.00	60.00	100.00	60.00
不透水面/裸地	81.48	73.33	90.48	63.33	47.06	26.67	87.50	70.00	28.57	20.00	83.33	66.67	88.00	73.33
雪	100.00	80.00	100.00	80.00	100.00	60.00	100.00	80.00	100.00	40.00	100.00	40.00	100.00	80.00
总体精度/%	85.89		84.68		72.24		86.49		62.39		84.08		86.19	
Kappa系数	0.8243		0.8090		0.6535		0.8322		0.5281		0.8019		0.8282	

注：加粗数值说明精度表现较好。

研究区4地物类别简单，从其精度表格（表10）中可以看出，方案七木本沼泽这一类别的UA较高，同时非湿地类别的精度表现也比其他方案表现要好，但草本沼泽、洪泛湿地这两类湿地精度表现不如方案一，总体而言方案七精度最优。研究区4方案七精度表现最好，总体精度为92.16%，Kappa系数为0.8954。

表 10 研究区4各方案分类精度统计
Table 10 Accuracy statistics of each scheme in study area four

类别	方案一		方案二		方案三		方案四		方案五		方案六		方案七	
	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%
木本沼泽	87.65	82.56	75.68	61.54	87.64	85.71	88.61	76.92	50.00	31.87	75.68	61.54	90.59	84.62
草本沼泽	77.78	76.09	60.38	69.57	75.56	73.91	65.31	69.57	57.41	67.39	64.58	67.39	71.74	71.74
水体	97.09	99.01	94.96	100.00	94.17	100.00	94.96	100.00	94.17	100.00	94.92	99.12	94.96	100.00
洪泛湿地	100.00	100.00	100.00	75.00	0.00	0.00	100.00	75.00	0.00	0.00	100.00	75.00	100.00	75.00
耕地	95.05	95.05	84.40	89.32	91.89	99.03	95.37	100.00	58.39	77.67	83.04	90.29	98.10	100.00
不透水面/裸地	60.00	100.00	91.67	84.62	100.00	38.46	91.67	84.62	100.00	7.69	80.00	92.31	100.00	92.31
总体精度/%	90.64		82.97		89.73		89.73		68.65		82.97		92.16	
Kappa系数	0.8746		0.7734		0.8617		0.8631		0.5770		0.7734		0.8954	

注：加粗数值说明精度表现较好。

通过对比各个研究区方案一和方案二的精度表现可以看出，去掉地形特征后总体精度和Kappa系数整体呈下降趋势。分析各研究区方案一、方案三和方案四精度表现可以发现，去掉哨兵1号或者哨兵2号特征之后总体精度和Kappa系数都有所降低，但去掉哨兵2号特征后总体精度和Kappa系数下降幅度更大。总体来看以多源特征为基础的方案一总体精度和Kappa系数高于基于单源特征的方案二到方案六，但低于进行特征优选后的方案

七，由此可见，对比单源特征，多源特征的加入有利于提高分类的总体精度，但存在特征冗余，在多源特征的基础上进行特征优选后总体精度得到提高。

综合考虑各研究区整体分类表现，本研究选取方案七为最终分类方案。对4个研究区的方案七进行总体精度评价后得到，总体精度为85.62%，Kappa系数为0.8333，具体各类别精度见表11。

表 11 4个研究区方案七总体精度统计
Table 11 Overall accuracy statistics of program seven in four study areas

类别	用户精度/%	生产者精度/%	类别	用户精度/%	生产者精度/%
木本沼泽	80.24	82.86	草地	83.05	77.37
草本沼泽	69.85	72.52	耕地	83.53	92.86
水体	95.18	96.1	不透水面/裸地	90.79	77.53
洪泛湿地	82.14	79.31	雪	100	80
盐田	94.12	100	总体精度/%	85.62	
滩涂	75	85.71	Kappa系数	0.8333	
森林/灌丛	89.19	83.7			

5 结 论

本研究结合已有数据集提出了适用于大区域样本采集的方法，并以南美洲4个典型湿地研究区为例，构建了适用于该地区湿地分类的多源数据特征集合，结果表明：（1）样本采集方案可有效

提高样本质量，可为大尺度分类样本采集提供参考。（2）哨兵1号和哨兵2号数据结合能够提升土地覆盖制图精度，地形特征有助于大幅度提升分类总体精度和各类别地物的精度。对湿地类别而言，多源数据特征的加入能够提升湿地类别的可分性。（3）特征优选结果表明，各种类型特征按

照重要程度排序为哨兵2号原始光谱波段特征>地形特征>哨兵1号纹理特征>哨兵2号红边指数特征>哨兵2号NDVI时序特征>哨兵2号常用光谱指数特征>哨兵1号极化特征, SAR极化特征和派生获得的纹理特征可作为光学特征的有效补充, 但优势特征占比较少。其中地形特征按重要性排序为高程特征>坡度特征; 哨兵2号特征中优势特征为短波红外波段(B12)、红边波段(B6、B5), 其次为红边指数特征(NDVIre3)和NDVI时序特征(NDVI_1025, NDVI_7590); 哨兵1号特征中较有优势的特征为GLCM纹理特征(VV_mean_idm, VV_mean_dent, VV_mean_savg, VH_mean_imcorr1)。(4) 基于随机森林的递归特征消除方法可提升总体精度和Kappa系数, 减少特征冗余。

尽管本研究特征优选方案在3个研究区得到了最高的总体精度和Kappa系数, 但某些研究区的非湿地精度并不理想, 这可能是分类时所采用的非湿地样本为辅助数据集迁移所得导致。本研究采用的为C波段SAR数据, 结果表明该数据的重要性相对较低, 可能是因为该数据的探测能力有限, 有论文指出X/L波段的SAR数据对湿地类别的识别效果更好(Rapinel等, 2020)。本研究重点探究了不同特征对湿地信息提取的影响, 也尝试利用NDVI时序特征来获取不同湿地类别的物候信息, 但缺乏深入思考, 后续将进一步挖掘时间序列数据对湿地信息提取的贡献。下一步工作中, 将在现有研究的基础上, 结合已有方法解决当前存在的问题, 以实现南美洲全域的湿地制图。

参考文献(References)

- Adam E, Mutanga O and Rugege D. 2010. Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review. *Wetlands Ecology and Management*, 18(3): 281-296 [DOI: 10.1007/s11273-009-9169-z]
- Amani M, Brisco B, Afshar M, Mirmazloumi S M, Mahdavi S, Mirzadeh S M J, Huang W M and Granger J. 2019a. A generalized supervised classification scheme to produce provincial wetland inventory maps: an application of Google Earth Engine for big geo data processing. *Big Earth Data*, 3(4): 378-394 [DOI: 10.1080/20964471.2019.1690404]
- Amani M, Brisco B, Mahdavi S, Ghorbanian A, Moghimi A, DeLancey E R, Merchant M, Jahncke R, Fedorchuk L, Mui A, Fisette T, Kakooei M, Ahmadi S A, Leblon B and Larocque A. 2021. Evaluation of the landsat-based Canadian wetland inventory map using multiple sources: challenges of large-scale wetland classification using remote sensing. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14: 32-52 [DOI: 10.1109/jstars.2020.3036802]
- Amani M, Mahdavi S, Afshar M, Brisco B, Huang W M, Mirzadeh S M J, White L, Banks S, Montgomery J and Hopkinson C. 2019b. Canadian wetland inventory using google earth engine: the first map and preliminary results. *Remote Sensing*, 11(7): 842 [DOI: 10.3390/rs11070842]
- Amani M, Salehi B, Mahdavi S, Granger J and Brisco B. 2017. Wetland classification in Newfoundland and Labrador using multi-source SAR and optical data integration. *GIScience and Remote Sensing*, 54(6): 779-796 [DOI: 10.1080/15481603.2017.1331510]
- Borges J, Higginbottom T P, Symeonakis E and Jones M. 2020. Sentinel-1 and Sentinel-2 Data for savannah land cover mapping: optimising the combination of sensors and seasons. *Remote Sensing*, 12(23): 3862 [DOI: 10.3390/rs12233862]
- Breiman L. 2001. Random forests. *Machine Learning*, 45(1): 5-32 [DOI: 10.1023/A:1010933404324]
- Bunting P, Rosenqvist A, Lucas R M, Rebelo L M, Hilarides L, Thomas N, Hardy A, Itoh T, Shimada M and Finlayson C M. 2018. The global mangrove watch—a new 2010 global baseline of mangrove extent. *Remote Sensing*, 10(10): 1669 [DOI: 10.3390/rs10101669]
- Cai Y T, Li X Y, Zhang M and Lin H. 2020. Mapping wetland using the object-based stacked generalization method based on multi-temporal optical and SAR data. *International Journal of Applied Earth Observation and Geoinformation*, 92: 102164 [DOI: 10.1016/j.jag.2020.102164]
- Chang W T, Wang H, Ning X G and Zhang H C. 2020. Extraction of Zhalong wetlands information based on images of sentinel-2 red edge bands and Sentinel-1 radar bands. *Wetland Science*, 18(1): 10-19 (常文涛, 王浩, 宁晓刚, 张翰超. 2020. 融合 Sentinel-2 红边波段和 Sentinel-1 雷达波段影像的扎龙湿地信息提取. *湿地科学*, 18(1): 10-19) [DOI: 10.13248/j.cnki.wetlandsci.2020.01.002]
- Chen J and Chen J. 2018. GlobeLand30: operational global land cover mapping and big-data analysis. *Science China Earth Sciences*, 61(10): 1533-1534 [DOI: 10.1007/s11430-018-9255-3]
- Deering D W. 1978. *Rangeland Reflectance Characteristics Measured by Aircraft and Spacecraft Sensors*. College Station, TX: Texas A and M University
- Demarchi L, Kania A, Ciężkowski W, Piórkowski H, Oświecimska-Piaszko Z and Chormański J. 2020. Recursive feature elimination and random forest classification of natura 2000 grasslands in Lowland River Valleys of Poland based on airborne hyperspectral and LiDAR data fusion. *Remote Sensing*, 12(11): 1842 [DOI: 10.3390/rs12111842]
- Dong T F, Liu J G, Shang J L, Qian B D, Ma B L, Kovacs J M, Walters D, Jiao X F, Geng X Y and Shi Y C. 2019. Assessment of red-

- edge vegetation indices for crop leaf area index estimation. *Remote Sensing of Environment*, 222: 133-143 [DOI: 10.1016/j.rse.2018.12.032]
- Fernández-Manso A, Fernández-Manso O and Quintano C. 2016. SENTINEL-2A red-edge spectral indices suitability for discriminating burn severity. *International Journal of Applied Earth Observation and Geoinformation*, 50: 170-175 [DOI: 10.1016/j.jag.2016.03.005]
- Haralick R M, Shanmugam K and Dinstein I. 1973. Textural features for image classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6): 610-621 [DOI: 10.1109/TSMC.1973.4309314]
- Hu S J, Niu Z G, Chen Y F, Li L F and Zhang H Y. 2017. Global wetlands: potential distribution, wetland loss, and status. *Science of the Total Environment*, 586: 319-327 [DOI: 10.1016/j.scitotenv.2017.02.001]
- Junk W J. 1993. Wetlands of tropical South America//Whigham D F, Dykxjová D and Hejný S, eds. *Wetlands of the World: Inventory, Ecology and Management Volume I*. Dordrecht: Springer: 679-739 [DOI: 10.1007/978-94-015-8212-4_14]
- Junk W J, An S Q, Finlayson C M, Gopal B, Květ J, Mitchell S A, Mitsch W J and Robarts R D. 2013. Current state of knowledge regarding the world's wetlands and their future under global climate change: a synthesis. *Aquatic Sciences*, 75(1): 151-167 [DOI: 10.1007/s00027-012-0278-z]
- Junk W J, Piedade M T F, Lourival R, Wittmann F, Kandus P, Lacerda L D, Bozelli R L, Esteves F A, Nunes Da cunha C, Maltchik L, Schöngart J, Schaeffer-Novelli Y and Agostinho A A. 2014. Brazilian wetlands: their definition, delineation, and classification for research, sustainable management, and protection. *Aquatic Conservation: Marine and Freshwater Ecosystems*, 24(1): 5-22 [DOI: 10.1002/aqc.2386]
- Kandus P, Minotti P G, Morandeira N S, Grimson R, Trilla G G, González E B, San Martín L and Gayol M P. 2018. Remote sensing of wetlands in South America: status and challenges. *International Journal of Remote Sensing*, 39(4): 993-1016 [DOI: 10.1080/01431161.2017.1395971]
- Lehner B and Döll P. 2004. Development and validation of a global database of lakes, reservoirs and wetlands. *Journal of Hydrology*, 296(1/4): 1-22 [DOI: 10.1016/j.jhydrol.2004.03.028]
- Li J H and Chen W J. 2005. A rule-based method for mapping Canada's wetlands using optical, radar and DEM data. *International Journal of Remote Sensing*, 26(22): 5051-5069 [DOI: 10.1080/01431160500166516]
- Mahdavi S, Salehi B, Granger J, Amani M, Brisco B and Huang W M. 2018. Remote sensing for wetland classification: a comprehensive review. *GIScience and Remote Sensing*, 55(5): 623-658 [DOI: 10.1080/15481603.2017.1419602]
- Mahdianpari M, Brisco B, Granger J, Mohammadimanesh F, Salehi B, Homayouni S and Bourgeau-Chavez L. 2021. The third generation of pan-canadian wetland map at 10 m resolution using multi-source earth observation data on cloud computing platform. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14: 8789-8803 [DOI: 10.1109/jstars.2021.3105645]
- Mao D H, Wang Z M, Du B J, Li L, Tian Y L, Jia M M, Zeng Y, Song K S, Jiang M and Wang Y Q. 2020. National wetland mapping in China: a new product resulting from object-based and hierarchical classification of Landsat 8 OLI images. *ISPRS Journal of Photogrammetry and Remote Sensing*, 164: 11-25 [DOI: 10.1016/j.isprsjprs.2020.03.020]
- McFeeters S K. 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17(7): 1425-1432 [DOI: 10.1080/01431169608948714]
- Murray N J, Phinn S R, DeWitt M, Ferrari R, Johnston R, Lyons M B, Clinton N, Thau D and Fuller R A. 2019. The global distribution and trajectory of tidal flats. *Nature*, 565(7738): 222-225 [DOI: 10.1038/s41586-018-0805-8]
- Nembrini S, König I R and Wright M N. 2018. The revival of the Gini importance?. *Bioinformatics*, 34(21): 3711-3718 [DOI: 10.1093/bioinformatics/bty373]
- Rapinel S, Betbeder J, Denize J, Fabre E, Pottier É and Hubert-Moy L. 2020. SAR analysis of wetland ecosystems: effects of band frequency, polarization mode and acquisition dates. *ISPRS Journal of Photogrammetry and Remote Sensing*, 170: 103-113 [DOI: 10.1016/j.isprsjprs.2020.10.007]
- Ruiz L F C, Guasselli L A, Simioni J P D, Belloli T F and Fernandes P C B. 2021. Object-based classification of vegetation species in a sub-tropical wetland using Sentinel-1 and Sentinel-2A images. *Science of Remote Sensing*, 3: 100017 [DOI: 10.1016/j.srs.2021.100017]
- Saah D, Johnson G, Ashmall B, Tondapu G, Tenneson K, Patterson M, Poortinga A, Markert K, Quyen N H, San Aung K, Schlichting L, Matin M, Uddin K, Aryal R R, Dilger J, Lee Ellenburg W, Flores-Anderson A I, Wiell D, Lindquist E, Goldstein J, Clinton N and Chishtie F. 2019. Collect Earth: an online tool for systematic reference data collection in land cover and use applications. *Environmental Modelling and Software*, 118: 166-171 [DOI: 10.1016/j.envsoft.2019.05.004]
- Schratz P, Muenchow J, Iturrutxa E, Cortés J, Bischl B and Brenning A. 2021. Monitoring forest health using hyperspectral imagery: does feature selection improve the performance of machine-learning techniques?. *Remote Sensing*, 13(23): 4832 [DOI: 10.3390/rs13234832]
- Slagter B, Tsendbazar N E, Vollrath A and Reiche J. 2020. Mapping wetland characteristics using temporally dense Sentinel-1 and Sentinel-2 data: a case study in the St. Lucia wetlands, South Africa. *International Journal of Applied Earth Observation and Geoinformation*, 86: 102009 [DOI: 10.1016/j.jag.2019.102009]
- Tamiminia H, Salehi B, Mahdianpari M, Quackenbush L, Adeli S and Brisco B. 2020. Google Earth Engine for geo-big data applica-

- tions: a meta-analysis and systematic review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 164: 152-170 [DOI: 10.1016/j.isprsjprs.2020.04.001]
- Wang J, Shang J, Brisco B and Brown R J. 1998. Evaluation of multi-date ERS-1 and multispectral landsat imagery for wetland detection in Southern Ontario. *Canadian Journal of Remote Sensing*, 24(1): 60-68 [DOI: 10.1080/07038992.1998.10874692]
- Whyte A, Ferentinos K P and Petropoulos G P. 2018. A new synergistic approach for monitoring wetlands using Sentinels -1 and 2 data with object-based machine learning algorithms. *Environmental Modelling and Software*, 104: 40-54 [DOI: 10.1016/j.envsoft.2018.01.023]
- Wu N, Shi R H, Zhuo W, Zhang C, Zhou B C, Xia Z L, Tao Z, Gao W and Tian B. 2021. A classification of tidal flat wetland vegetation combining phenological features with Google earth engine. *Remote Sensing*, 13(3): 443 [DOI: 10.3390/rs13030443]
- Xia Y, Li E H, Wang X L, Zhang Y Y, Yang J and Zhou R. 2021. Application of random forest algorithm based on feature optimization in wetland information extraction: a case study of Honghu Wetland Nature Reserve in Hubei Province. *Journal of Central China Normal University (Natural Sciences)*, 55(4): 639-648, 660 (夏盈, 厉恩华, 王学雷, 张莹莹, 杨娇, 周瑞. 2021. 基于特征优选的随机森林算法在湿地信息提取中的应用——以湖北洪湖湿地自然保护区为例. *华中师范大学学报(自然科学版)*, 55(4): 639-648, 660) [DOI: 10.19603/j.cnki.1000-1190.2021.04.019]
- Xu H Q. 2005. A study on information extraction of water body with the Modified Normalized Difference Water Index (MNDWI). *Journal of Remote Sensing*, 9(5): 589-595 (徐涵秋. 2005. 利用改进的归一化差异水体指数(MNDWI)提取水体信息的研究. *遥感学报*, 9(5): 589-595) [DOI: 10.11834/jrs.20050586]
- Yan X and Niu Z G. 2021. Reliability evaluation and migration of wetland samples. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 14: 8089-8099 [DOI: 10.1109/jstars.2021.3102866]
- Yang J W, Zhang J S, Zhu X F and Yuan Z M Q. 2015. Random forest applied for dimension reduction and classification in hyperspectral data. *Journal of Beijing Normal University (Natural Science)*, 51(S1): 82-88 (杨珺雯, 张锦水, 朱秀芳, 谢登峰, 袁周米琪. 2015. 随机森林在高光谱遥感数据中降维与分类的应用. *北京师范大学学报(自然科学版)*, 51(S1): 82-88) [DOI: 10.16360/j.cnki.jbnuns.2015.s1.013]
- Zha Y, Ni S X and Yang S. 2003. An effective approach to automatically extract urban land-use from TM imagery. *Journal of Remote Sensing*, 7(1): 37-40 (查勇, 倪绍祥, 杨山. 2003. 一种利用TM图像自动提取城镇用地信息的有效方法. *遥感学报*, 7(1): 37-40) [DOI: 10.11834/jrs.20030107]
- Zhang H Y, Niu Z G, Xu P P, Chen Y F, Hu S J and Gong N. 2017. The boundaries and remote sensing classification datasets on large wetlands of international importance in 2001 and 2013. *Journal of Global Change Data and Discovery*, 1(2): 230-238 (张海英, 牛振国, 许盼盼, 陈燕芬, 胡胜杰, 宫宁. 2017. 大型国际重要湿地边界及遥感分类数据集(2001、2013). *全球变化数据学报*, 1(2): 230-238) [DOI: 10.3974/geodp.2017.02.15]
- Zhang L, Gong Z N, Wang Q W, Jin D D and Wang X. 2019. Wetland mapping of Yellow River Delta wetlands based on multi-feature optimization of Sentinel-2 images. *Journal of Remote Sensing*, 23(2): 313-326 (张磊, 宫兆宁, 王启为, 金点点, 汪星. 2019. Sentinel-2影像多特征优选的黄河三角洲湿地信息提取. *遥感学报*, 23(2): 313-326) [DOI: 10.11834/jrs.20190803]
- Zheng L L, Xu J Y and Wang X L. 2019. Application of random forests algorithm in researches on wetlands. *Wetland Science*, 17(1): 16-24 (郑利林, 徐金英, 王晓龙. 2019. 随机森林算法在湿地研究中的应用. *湿地科学*, 17(1): 16-24) [DOI: 10.13248/j.cnki.wetlandsci.2019.01.003]
- Zhou X C, Zheng L and Huang H Y. 2021. Classification of forest stand based on multi-feature optimization of UAV visible light remote sensing. *Scientia Silvae Sinicae*, 57(6): 24-36 (周小成, 郑磊, 黄洪宇. 2021. 基于多特征优选的无人机可见光遥感森林分类. *林业科学*, 57(6): 24-36) [DOI: 10.11707/j.1001-7488.20210603]

Wetlands mapping in typical regions of South America with multi-source and multi-feature integration

HUANG Yuling¹, YANG Gang¹, SUN Weiwei¹, ZHU Lin¹, HUANG Ke¹, MENG Xiangchao²

1. Department of Geography and Spatial Information Techniques, Ningbo University, Ningbo 315211, China;

2. Faculty of Electrical Engineering and Computer Science, Ningbo University, Ningbo 315211, China

Abstract: Wetlands play an important role in maintaining ecological balance, conserving water resources, recharging groundwater, and controlling soil erosion. They are often called the “kidneys of the earth” because they help purify water by filtering out pollutants and sediments. South America has a vast area of wetlands, as well as a variety of wetlands types. While most of these wetlands were conserved

in a relatively good condition until a few decades ago, pressures brought about by land use and climate change have threatened their integrity in recent years. However, no complete and uniform wetland map has provided adequate information on the location, distribution, size, and changing status of wetlands in South America. Remote sensing has been an effective tool for characterizing, mapping, and monitoring the complexity and dynamics of large areas of wetlands. Although fine wetland mapping may be done by combining data from many sources, the following two issues persist. On the one hand, given the complicated temporal dynamics and spectral heterogeneity of wetlands, large-scale wetland mapping remains a challenging task. On the other hand, supervised classification is a widely used technique for multi-category wetland mapping. However, selecting training samples is time consuming and labor intensive. Moreover, finer and more precise wetland information is currently unavailable for reference. In the study, we selected four study areas of typical wetlands in South America. First, an effective wetland sample collection process was proposed by using the existing land cover dataset to ensure the sample quality. Second, a multi-source feature set was constructed by combining Sentinel-1, Sentinel-2, and SRTM data. Then, feature selection is carried out on the basis of the random forest recursive feature elimination method (RF_RFE). We constructed a multi-feature combination scheme to compare the influence of multi-source features on wetlands classification. Finally, the random forest algorithm is used to classify wetlands in the study area. The research results show that the process facilitates the sample collection and improves the sample quality. The combination of Sentinel-1 and Sentinel-2 data can improve the accuracy of land cover mapping, and terrain features help greatly improve the overall classification accuracy and the accuracy of various types of objects. For wetland categories, the addition of multi-source data features can improve the separability of wetland categories. The feature selection based on RF_RFE can reduce the feature redundancy and improve the classification accuracy. The feature optimization results show that SAR polarization features and derived texture features can be used as an effective supplement to optical features. However, the dominant features account for less. The overall classification accuracy of the study area was 85.62%, and the Kappa coefficient was 0.8333. The study proposed an effective classification process and sample collection scheme to large-scale wetlands in South America. This study integrated Sentinel-1 synthetic aperture radar data, Sentinel-2 optical data, and terrain data to explore their importance to the extraction of different wetlands at a large scale. This work also verified the feasibility of feature selection based on random forest recursive feature elimination method. The research results reveal that the sample collection process facilitates sample collection and improves sample quality. The combination of Sentinel-1 and Sentinel-2 data can improve the accuracy of land cover mapping, and terrain features help greatly improve the overall classification accuracy and the accuracy of various types of objects. For wetland categories, the addition of multi-source data features can improve the separability of wetland categories. The feature selection based on RF_RFE can reduce the feature redundancy and improve the classification accuracy.

Key words: remote sensing, Sentinel-1, Sentinel-2, Google Earth Engine, wetlands classification, feature selection, South America

Supported by National Natural Science Foundation of China (No. 42122009, 41971296); Ningbo Science and Technology Innovation 2025 Major Special Project (No. 2021Z107); Public Projects of Ningbo City (No. 2021S089); China Postdoctoral Science Foundation (No. 2020M670440); Fundamental Research Funds for the Provincial Universities of Zhejiang (No. SJLZ2022002); Zhejiang Provincial Natural Science Foundation of China (No. LR19D010001)