

资源一号 02D 卫星高光谱数据黄河三角洲湿地景观分类

韩月^{1,2,3}, 柯樱海^{1,2,3}, 王展鹏^{1,2,3}, 梁德印⁴, 周德民^{1,2,3}

1. 城市环境过程与数字模拟国家重点实验室培育基地, 北京 100048;

2. 水资源安全北京实验室, 北京 100048;

3. 首都师范大学 资源环境与旅游学院, 北京 100048;

4. 中国空间技术研究院遥感卫星总体部, 北京 100086

摘要: 资源一号 02D 卫星 (ZY-1 02D) 于 2019 年成功发射, 2020 年 10 月正式投入使用, 是中国自主建造并成功运行的首颗民用高光谱业务卫星, 具有广泛的应用前景。本研究以黄河三角洲湿地为研究区, 以 ZY-1 02D 高光谱 (AHSI) 影像为数据源, 结合无人机和地面调查数据, 开展湿地景观分类研究。首先通过 ZY-1 02D AHSI 获取地物反射率波谱曲线, 分析不同地物波谱曲线的差异, 作为地物识别和分类的依据; 充分考虑研究区植被覆盖度的差异, 结合无人机影像制定研究区 7 类基本地物和 9 类精细地物两种湿地景观分类体系; 利用随机森林算法进行分类, 并引入 Tree SHAP 方法进行波段重要性排序和选择; 探究影响 ZY-1 02D AHSI 分类的重要波段, 选取与 Landsat 8 OLI 多光谱波段相重叠的波段进行分类, 并与 Landsat 8 OLI 分类结果进行比较。结果表明: (1) ZY-1 02D AHSI 数据能够较好地反映不同地物类型光谱曲线的差异; (2) 对于两种分类体系, 仅用前 40 个重要波段的总体分类精度达到最高, 7 类基本地物分类和 9 类精细地物分类的分类精度分别为 92.18% 和 90.76%, 这 40 个波段大多位于可见光、近红外波段; (3) 对于两种分类体系, 分别选取与 Landsat 8 OLI 多光谱波段重叠的 24 个和 29 个波段进行分类, 分类精度达到 90.01% 和 89.76%, 均明显高于 Landsat 8 OLI 的分类精度; (4) 蓝、绿波段对于识别高、低密度互花米草和芦苇较为重要, 短波红外波段对于芦苇的识别较为重要, 红波段对于识别高、低密度碱蓬较为重要。ZY-1 02D AHSI 数据的波谱范围较窄, 波谱连续, 能够较好地体现地物光谱曲线的细微变化, 在区分不同地物以及植被覆盖度差异上具有明显优势。本研究有利于及时有效地监测黄河三角洲湿地资源现状, 为 ZY-1 02D 高光谱数据在湿地生态监测应用提供科学参考依据。

关键词: ZY-1 02D, AHSI 影像, 高光谱数据, 植被覆盖度, 随机森林, Tree SHAP

中图分类号: P2

引用格式: 韩月, 柯樱海, 王展鹏, 梁德印, 周德民. 2023. 资源一号 02D 卫星高光谱数据黄河三角洲湿地景观分类. 遥感学报, 27 (6): 1387-1399

Han Y, Ke Y H, Wang Z P, Liang D Y and Zhou D M. 2023. Classification of the Yellow River Delta wetland landscape based on ZY-1 02D hyperspectral imagery. National Remote Sensing Bulletin, 27 (6): 1387-1399 [DOI: 10.11834/jrs.20211071]

1 引言

河口湿地位于海洋、淡水、陆地之间的过渡地带, 是海生和陆生生态系统的交错区, 生物资源丰富, 在维持生物多样性调节气候变化中发挥着重要的作用。河口湿地又属于生态脆弱带, 抵御外部干扰能力和生态系统稳定性较差。黄河三

角洲湿地是中国北方暖温带最年轻、保存最完整、总面积最大的湿地生态系统, 植被类型多样, 生物多样性丰富, 是中国国家级自然保护区 (李永涛等, 2019)。随着城市化加剧、农田开垦、油气工业发展、物种入侵等, 黄河三角洲湿地已成为全球受威胁最严重的生态系统之一。及时监测黄河三角洲湿地植被现状, 了解该区域不同类型湿

收稿日期: 2021-03-09; 预印本: 2021-09-09

基金项目: 国家自然科学基金 (编号: 42071396); 国家重点研发计划 (编号: 2017YFC0505903)

第一作者简介: 韩月, 研究方向为多源遥感数据融合、生态水文遥感应用。E-mail: hy1602311920@163.com

通信作者简介: 柯樱海, 研究方向为多源遥感数据融合、人工智能及数据挖掘方法、生态水文遥感应用。E-mail: yke@cnu.edu.cn

地植被的空间分布格局与特征,可为今后研究区湿地资源的保护与合理利用提供科学依据。

目前,有学者利用 Landsat、GF-1、Sentinel-2 等多光谱数据,对黄河三角洲湿地进行分类研究。如利用 GF-3 雷达数据和 Sentinel-2 多光谱数据协同对黄河三角洲地区进行分类(李鹏等,2019);利用 Landsat MSS/TM/OLI 数据研究黄河三角洲湿地近 30 年土地覆盖的时空分布和动态变化特征(徐振田,2020);基于多时相 Sentinel-2 影像提取黄河三角洲湿地信息(张磊等,2019)。河口湿地景观斑块破碎,目前基于多光谱影像数据进行河口湿地植被分类大多是对研究区基本地物类别进行分类(王霄鹏等,2014),难以区分植被覆盖度的差别。另一方面,由于土壤性状和水分梯度的高空间异质性,植被群落“同物异谱”和“异物同谱”现象严重,多光谱影像数据波段数目较少,无法保证植被分类的精度。高光谱影像具有更好的波段连续性,数百个波段蕴含着详细的光谱特征信息,有利于区分不同地物在光谱曲线上的细微差异,在地物景观分类中发挥了重要的作用(张良培和张立福,2005)。有学者利用国产高分五号(GF-5)高光谱影像对红树林物种进行分类,提高红树林物种的分类精度(Wan等,2020);也有学者以环境一号(HJ-1)卫星高光谱影像为数据源,实现了试验区扎龙湿地的景观地物分类(张雅春等,2018)。

由于卫星传感器在获取影像数据时往往不能兼顾空间和光谱分辨率,获得较高的光谱分辨率是以牺牲空间分辨率为代价的。目前大多数高光谱卫星传感器空间分辨率一般为 30 m 或者更低,河口湿地植被分布零散且覆盖度变化明显,在这样的空间分辨率下必然出现较为明显的混合像元现象。而前人基于高光谱数据的湿地分类研究多将像元分为单一地物类别,没有考虑同一种植被类型覆盖度的差异(Slagter等,2020)。有学者采用混合像元分解技术估算高光谱影像像元中各地物类别的丰度(刘帅和邢光龙,2020;Cheng等,2021;Jiao等,2019)。然而,湿地景观类型复杂,植被类型丰富,目前的混合像元分解技术难以求解各精细地物类别,尤其是各植被类型的丰度。

“资源一号 02D 卫星”(简称 ZY-1 02D)于 2019 年 9 月 12 日发射,2020 年完成在轨测试并正式投入使用,是中国自主建造并成功运行的首颗

民用高光谱业务卫星。该卫星运行于太阳同步轨道,搭载有可见短波红外高光谱相机 AHSI (Advanced Hyper Spectral Instrument) 和可见光/近红外相机 VNIC (Visible NearInfrared Camera),轨道重复周期为 55 d。AHSI 相机可获取 60 km 幅宽,空间分辨率为 30 m,共计 166 个波段的高光谱影像,其高光谱性能有利于实现地物的光谱信息调查和精细化制图,提高山水林田湖草等自然资源的调查监测能力。ZY-1 02D 作为最新发射的高光谱卫星,其应用研究鲜有报道。

本文利用 ZY-1 02D AHSI 影像,结合野外实地调查和无人机遥感数据,充分考虑河口湿地植被覆盖度的空间差异,进行黄河口湿地景观分类。本研究首先通过 ZY-1 02D 获取不同地物反射率波谱曲线,分析不同地物波谱曲线的差异;结合无人机遥感数据和野外踏勘数据,确定研究区基本景观分类体系以及考虑植被覆盖度差异的精细分类体系,分别进行湿地景观制图。通过波段重要性评价进行特征波段选择,并与空间分辨率相同的 Landsat 8 OLI 多光谱影像数据分类结果进行比较,进而探究和评价 ZY-1 02D 卫星高光谱影像数据在河口湿地破碎景观的地物精细分类效果。

2 研究区和数据

2.1 研究区概况

黄河三角洲位于山东省东营市东北部($37^{\circ}34.768'N$ — $38^{\circ}12.310'N$, $118^{\circ}32.981'E$ — $119^{\circ}20.450'E$),渤海湾南岸和莱州湾西部,呈扇形分布,是世界上海生—陆生生态系统相互延展交叠的重要区域,总面积为 8100 km^2 ,其中湿地面积占 4500 km^2 。黄河三角洲属暖温带半湿润大陆性季风气候,四季分明,夏季多雨。年平均日照时间为 $2590\text{—}2830\text{ h}$,年均气温为 $11.5\text{ }^{\circ}\text{C}$ — $12.9\text{ }^{\circ}\text{C}$,年均降水量为 $530\text{—}660\text{ mm}$ 。研究区选择黄河东段,即黄河入海口河道两侧的湿地区域,面积为 893.95 km^2 (图 1)。此地区包含了黄河三角洲湿地生态系统中绝大部分湿地植物类别,包括互花米草、芦苇、碱蓬和柽柳。其中芦苇主要分布在灌溉沟渠旁、河堤沼泽地等区域,呈现聚集状,植被覆盖度较高;互花米草主要分布在中低潮位区域,新生互花米草呈现点簇状分布,植被覆盖度

较低, 发育后期逐渐形成密集群落, 植被覆盖度较高; 碱蓬主要分布在中高潮位区, 以盐地碱蓬为主, 潮沟附近密度较大; 柽柳多与芦苇混生, 植被覆盖度较小。

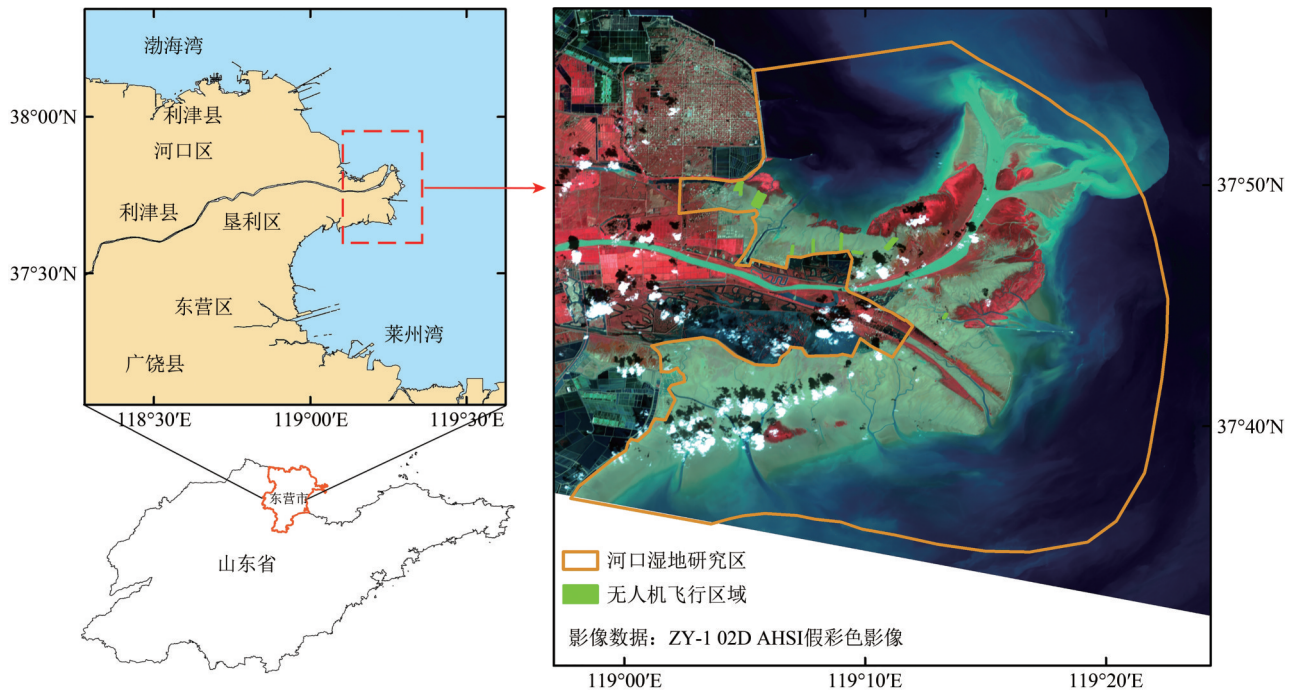


图1 研究区位置

Fig. 1 Location of study area

2.2 数据及预处理

2.2.1 遥感数据及预处理

研究数据源为 2020 年 9 月 26 日获取的 ZY-1 02D AHSI 数据, 卫星过境时间为北京时间 11:05。为与多光谱影像数据进行对比分析, 同时选取邻近时间 2020 年 10 月 24 日的 Landsat 8 OLI 多光谱影像数据, 数据参数见表 1。

ZY-1 02D AHSI 传感器光谱分辨率 10—20 nm, 空间分辨率 30 m, 幅宽 60 km, 可以同时获取 400—2500 nm 范围内共 166 个谱段的数据。该数据谱段主要由可见近红外 (VNIR) 和短波红外 (SWIR) 两个部分组成, 其中 VNIR 有 76 个波段, 光谱范围 400—1040 nm, 光谱分辨率为 10 nm; SWIR 有 90 个波段, 光谱范围 1005—2501 nm, 光谱分辨率为 20 nm。高光谱数据主要的预处理流程包括: 波段合成—去噪处理—辐射校正—几何校正。首先将 VNIR 和 SWIR 波段进行合成, 得到由 166 个波段组成的影像数据。随后, 对影像进行去噪处理, 剔除 15 条严重的噪声波段, 这些波段全部位于短波红外处, B98—B102 (1350—1429 nm), B125—

B133 (1804—1949 nm), B90 (2493—2505 nm)。处理后得到由 151 个有效波段组成的影像数据参与后续研究。其次, 依据每个波段对应的辐射定标系数进行辐射定标, 并利用 FLAASH 软件实现大气校正, 生成反射率数据。以 Landsat 8 多光谱数据为参照均匀布设控制点, 进行几何配准。Landsat 8 OLI 二级反射率产品从美国 USGS 网站获取, 多光谱数据空间分辨率 30 m, 波段数为 8 个 (表 1)。

2.2.2 野外踏勘和无人机数据

2020 年 9 月, 实验人员对研究区进行了详细调研。采用高精度手持 GPS 对不同地表覆被类型进行定位和相关信息记录。选取植被覆盖度具有差异性的 10 块区域 (图 1), 进行实地无人机航拍实验。本研究使用大疆精灵 4 多光谱无人机 P4M (Phantom4-M), 其搭载 6 个 CMOS 传感器, 包含 1 个用于可见光成像的彩色传感器和 5 个用于多光谱成像的单色传感器, 多光谱传感器包含蓝波段 (450±16 nm), 绿波段 (560±16 nm)、红波段 (650±16 nm)、红边波段 (730±16 nm) 和近红外

波段 (840+26 nm) 5 个波段, 有效像素为 208 万。用千寻网络 RTK (Real-Time Kinematic) 进行实时动态定位, 对正射影像进行几何校正, 可保证厘米级位置精度。在 10:00—14:00 点获取无人机数据, 飞行时天气晴朗、风力较小。采用由陆向海的方式规划航线, 飞行高度约为 50 m, 航向和旁向重叠率均为 80%, 地面采样距离约为 3.2 cm。采用 DJI Terra 大疆同步处理软件对无人机多光谱数据进行预处理, 经过图像拼接和辐射校正, 最终得到研究区的多光谱正射影像图。

表 1 ZY-1 02D AHSI 和 Landsat 8 OLI 影像参数
Table 1 ZY-1 02D AHSI and Landsat 8 OLI image parameters

参数	ZY-1 02D AHSI	Landsat 8 OLI
谱段范围/nm	VNIR: 400—1040; SWIR: 1005—2501	Coastal: 433—453 Blue: 450—515 Green: 525—600 Red: 630—680 NIR: 845—885
		SWIR1: 1560—1660 SWIR2: 2100—2300 Cirrus: 1360—1390
光谱采样间隔/nm	VNIR: 10; SWIR: 20	20—60
波段数/个	166	8
空间分辨率/m	30	30

3 分类体系建立

3.1 湿地景观类别划分

3.1.1 7 类基本地物类别

由于湿地的地域性差异, 国内外对湿地景观分类的原则、方法及指标等并未统一 (王霄鹏, 2014; 宫宁 等, 2016; 李玉 等, 2020)。本文参考国际《湿地公约》、中国《国家湿地资源调查技术规程》中的分类体系和野外实地踏勘及研究区无人机影像目视解译的结果, 构建研究区河口湿地景观分类系统, 确定丰水潮滩、光滩、互花米草、芦苇、芦苇柎柳混生、水体和碱蓬共 7 类基本地物类别。

3.1.2 9 类精细地物类别

经实地调查发现, 研究区碱蓬和互花米草的覆盖度呈现较为明显的梯度变化。如图 2 无人机假

彩色影像所示, 互花米草植被覆盖度由低潮滩向内陆递减, 低潮滩处互花米草形成单一密集群落, 越靠近内陆, 呈现新生孤立的互花米草植株簇; 碱蓬的分布恰好相反, 越靠近内陆高潮位区域其生长越茂盛, 成片均匀分布, 靠近互花米草的区域碱蓬生长较为稀疏, 能够明显看到裸露的光滩 (Chen 等, 2017)。互花米草为入侵植物, 从低潮滩向内陆不断扩张, 破坏碱蓬生境。在互花米草入侵的前缘地带, 互花米草和碱蓬覆盖度较低。碱蓬的覆盖度还随着潮滩土壤水分、盐分条件的变化呈现较为明显的差异。土壤盐分含量越高, 中低潮滩的碱蓬种子的萌发率将会减小 (孙志高 等, 2017), 使得与靠近低潮滩盐分含量高的地方, 碱蓬的植被覆盖度越低, 在土壤盐分和水分适宜的中高潮位区域, 碱蓬的植被覆盖度相对较高。

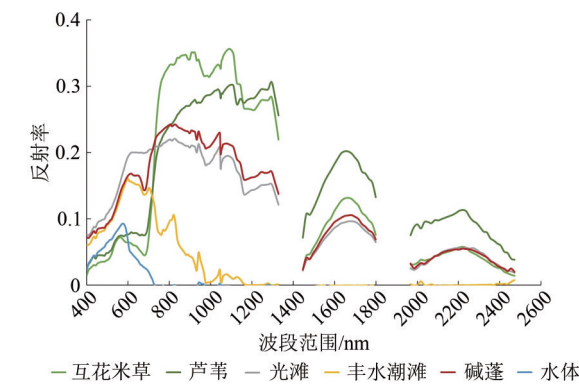


图 2 湿地典型地物光谱曲线图
Fig. 2 Reflectance curve of typical wetland features

因此, 本研究在 7 类基础地物类别的基础上, 增加不同植被密度的地物类别, 将研究区划分为 9 种地物类别: 丰水潮滩, 光滩, 高密度互花米草、低密度互花米草、芦苇、芦苇柎柳混生、水体、低密度碱蓬、高密度碱蓬。由于高光谱影像成像过程中有云的干扰, 对高光谱影像的分类中增加云和云阴影两个类别。

3.2 分类样本建立

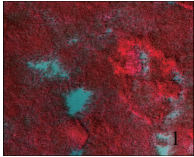
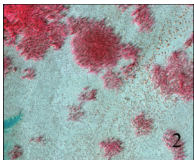
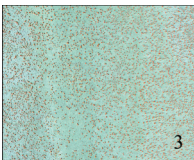
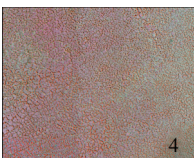
本研究通过实地踏勘记录、谷歌地球高清影像目视解译以及无人机影像数据解译结合的方法对两种分类体系分别建立样本。实地踏勘时, 利用手持 GPS 记录样本点位置以及地貌特征。借助谷歌地球高清影像对分布较为均一的地物, 如成片的芦苇、互花米草以及芦苇柎柳混生群落进行目视解译, 确定样本点。利用无人机影像, 结合

谷歌地球影像, 选取碱蓬密度高且较为纯净的像元, 确定碱蓬样本点。

对于9类精细地物分类体系, 利用无人机影像建立高、低密度互花米草和碱蓬样本。首先, 基于无人机影像的红和近红外波段计算归一化植被差异指数 NDVI (Normalized Difference Vegetation Index)。NDVI能够较好地反映植被分布情况, 将 NDVI>0 的像元定义为植被, 其余为非植被, 进

行数据二值化。为与卫星影像空间分辨率相匹配, 在无人机影像上建立与卫星影像像元对应的 30 m×30 m 的样方, 计算各样方内的植被覆盖度。表2示意了4个样方的植被覆盖情况。调查发现, 当植被覆盖度>0.5时, 植被分布较为密集, 划分为高密度植被; 覆盖度<0.1时, 基本为光滩; 覆盖度在 0.1—0.5 时, 植被分布相对稀疏, 划分为低密度植被。

表2 不同植被覆盖度无人机影像示意图
Table 2 Demonstration of UAV images of different vegetation coverage

无人机影像	样方中无人机影像	样方植被覆盖度	植被覆盖度范围	类别
		0.889	>0.5	高密度互花米草
		0.226	0.1—0.5	低密度互花米草
		0.180	0.1—0.5	低密度碱蓬
		0.869	>0.5	高密度碱蓬

本次实验进行地物分类时, Landsat 8 OLI影像上建立 2064 个样本点, 由于 ZY-1 02D 影像上存在云和云阴影, 共建立 2249 个样本点, 各类地物样本点数量都不低于 100。

4 研究方法

4.1 随机森林分类

随机森林算法是以分类回归树 CART (Classification and Regression Tree) 算法为基础, 通过集成学习的思想将多棵决策树集成的一种算法 (杜培军 等, 2016)。它以决策树为基本单元, 多个决策树共同对某个任务进行处理。随机森林算法支持高维输入样本, 能够运行在较大的数据

集上, 相较于其他分类算法具有较好的准确性和普适性, 因此被广泛应用到遥感地物分类中。

4.2 ZY-1 02D 波段重要性排序和选择

由于较高的光谱维数, 高光谱影像分类往往会遇到训练样本数量不足而导致的维数灾难问题 (Hughes, 1968; 杜培军 等, 2016)。为解决这一问题, 本文引入 Tree SHAP (Tree Shapley Additive exPlanations) 算法对 151 个波段进行特征重要性排序; 将波段按照特征重要性从高到底的顺序排列, 分析随机森林分类结果随波段数量增加的变化; 选择使分类精度达到最高的前 *m* 个特征波段, 实现分类和制图。

SHAP 由 Lundberg 和 Lee (2017) 提出, 该方

法源于博弈论,可估算每个特征对每个样本预测结果的贡献,提高机器学习模型的可解释性。Tree SHAP 适用于基于树的机器学习模型,如决策树,随机森林和 XGBoost 等 (Rodríguez-Pérez 和 Bajorath, 2020; Lundberg 等, 2018)。Tree SHAP 将训练集中所有样本的每一个特征都进行重要性计算,得到逐个特征的 Shapley 值。Shapley 值表示样本中每个特征对该样本的“贡献”。这个“贡献”有可能是正值,也有可能是负值,其最终对样本的影响取决于 Shapley 值的绝对值 (李超等, 2021)。相较于常用的基尼指数 (Gini Index) 和准确率降低指标 MDA (Mean Decrease Accuracy), Tree SHAP 不仅能衡量各特征变量对整体分类结果的重要性,还能得到它们对各类类别的特征重要性。

假设样本集的第 i 个样本为 x_i , 第 i 个样本的第 j 个特征为 x_{ij} , 模型对该样本的评估值为 y_i , 整个模型的基线 (所有样本的评估值的均值) 为 y_{base} , 那么 Shapley 值服从以下等式:

$$y_i = y_{\text{base}} + w(x_{i1}) + w(x_{i2}) + \cdots + w(x_{ij}) \quad (1)$$

$w(x_{ij})$ 为 x_{ij} 的 Shapley 值, 其计算公式如下:

$$w(x_{ij}) = \frac{1}{N!} \sum_{S \in N \setminus \{j\}} |S|!(|N| - |S| - 1)! [f(S \cup \{j\}) - f(S)] \quad (2)$$

$$w(x_j) = \sum_{i=1}^n w(x_{ij}) \quad (3)$$

式中, N 表示样本所有特征的全集, S 表示去掉第 j 个特征后得到的所有特征子集。对于任意的特征序列而言, $f(S)$ 表示在模型中所有特征子集共同作用产生的贡献值, $f(S \cup \{j\})$ 表示在模型中所有特征子集和第 i 个样本的第 j 个特征共同作用产生的贡献值, $f(S \cup \{j\}) - f(S)$ 为特征 j 的边际贡献。对所有情况累加求均值即为第 i 个样本的第 j 个特征 Shapley 值。 $w(x_j)$ 为对所有的样本的特征 j 累加取均值即为特征 j 的 Shapley 值, n 为模型中所有的样本数。

4.3 分类方案比较

为分析光谱细分对于湿地景观精细分类的有效性,除 4.2 所述分类方案外,对于两种分类体系分别采取以下两种分类方案进行对比研究:(1) 基于 Landsat 8 OLI 多光谱数据的分类;(2) 在上述

ZY-1 02D 数据 m 个重要特征波段范围内,选取与 Landsat 8 OLI 波段相重合的 l 个波段进行分类。对 Landsat 8 OLI 多光谱数据进行分类时,选用除卷云波段外的其余 7 个多光谱波段反射率构建特征数据集。

4.4 模型训练与精度评价

为保证分类结果评价的可靠性和准确性,上述所有分类方案都采用十折交叉检验法对模型进行训练和验证。交叉检验是把得到的样本数据进行切分,组合为不同的训练集和测试集,每次都使用不同的训练集来训练模型,用测试集来评估模型,避免模型过拟合而导致评价精度偏高。通过十折交叉验证得到平均总体精度 (OA)、平均 Kappa 系数、平均生产者精度 (PA) 和平均用户精度 (UA) 作为评价指标。

5 结果与分析

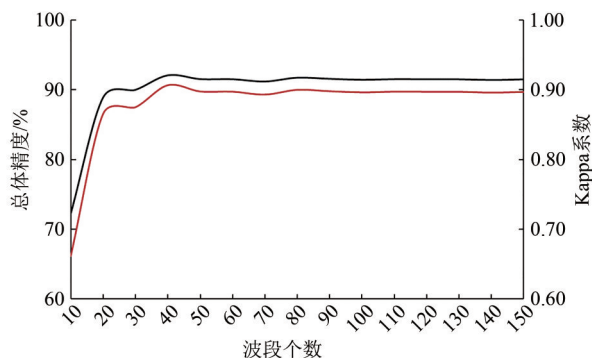
5.1 ZY-1 02D AHSI 地物光谱曲线分析

为分析各类地物在 ZY-1 02D AHSI 影像的光谱差异性,在影像中分别选取互花米草、芦苇、碱蓬、光滩、水体和丰水潮滩的 30 个均质纯像元点,对每类 30 个像元点的光谱曲线求均值,如图 2 所示。可以看出,不同地物的光谱曲线差异较为明显。丰水潮滩在 590 nm 处反射率达到顶峰,随后反射率逐渐下降,最后在短波红外处反射率基本为 0; 水体在 580 nm 处发射率达到顶峰,随后反射率开始下降,最后在 735 nm 处反射率基本为 0; 光滩在 600—1000 nm 反射率较高,在红波段和近红外波段整体较为平稳无较大波动,在 1100—1250 nm 的短波红外处开始反射率开始下降,整体反射率较低。互花米草、芦苇和碱蓬作为研究区分布较为广泛的典型湿地植被,符合绿色植物反射率特征; 芦苇在 750—1000 nm 近红外波段反射率低于互花米草,而在 1450—1800 nm 以及 1970—2450 nm 的短波红外波段高于互花米草; 碱蓬矮小,生物量较低,在可见光波段,尤其是红波段显著高于互花米草和芦苇,而在近红外波段反射率较低。不同地物独特的光谱曲线特征成为高光谱数据分类的重要依据。观察地物光谱曲线时发现有些波谱存在细微波动情况,这主要是由于影像成像中受到传感器及外界环境的影响,某些波

段中存在少许噪声导致。

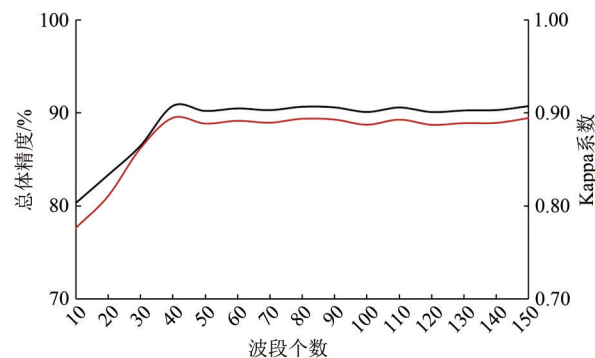
5.2 波段重要性排序和选择

利用 Tree SHAP 方法量化 ZY-1 02D AHSI 各波段反射率分别在 7 类基本地物和 9 类精细地物分类中的特征重要性, 根据重要性排序由高到低逐 10 个波段依次累加, 得到不同波段组合的随机森林分类十折交叉验证总体精度。如图 3 所示, 随着重要波段数目增加, 前期 (10—40 个波段特征) 分类精度迅速提高。基本地物分类的总体精度从 72.23% 提升到 92.18% (图 3(a)); 精细地物的总体精度从 80.35% 提高到 90.76% (图 3(b))。



(a) 7 类基本地物分类波段数量与分类精度

(a) The number of bands and classification accuracy for seven basic classes



(b) 9 类精细地物分类波段数量与分类精度

(b) The number of bands and classification accuracy for nine refined classes

— 总体精度 — Kappa 系数

图 3 波段数量与分类精度

Fig. 3 Number of bands and classification accuracy

如表 3 所示, 对于两种分类体系, ZY-1 02D AHSI 前 40 个重要波段绝大部分位于可见光近红外波段, 说明仅用 ZY-1 02D AHSI 的 VNIR 波段就能够达到较高的分类精度。进一步将前 40 个重要波段与 Landsat 8 多光谱影像数据的波谱范围进行匹配分析, 发现对于两种分类体系, 分别有 24 个波段和 29 个波段分布在 Landsat 8 OLI 的可见光、近红外波段 (表 3)。其中 16 个 (11 个) 波段位于 Landsat 8 OLI 海岸蓝、绿、红、近红外未涵盖的波谱范围中, 包含 2 个短波红外波段。这说明相较于 Landsat 8 OLI 多光谱数据, ZY-1 02D AHSI 数据的波谱范围较窄, 波谱连续, 能够更好的体现地物光谱曲线的细微变化和差异。随后, 选取与 Landsat 8 OLI 多光谱数据相匹配的 24 个以及 29 个重要波段分别作为一组新的分类特征, 分别对 7 类基本地物和 9 类精细地物进行分类。

自 40 个波段总体精度和 Kappa 系数达到最高之后, 后期 (50—151 个波段特征) 分类精度基本维持稳定且呈缓慢降低的趋势。当使用 151 个波段时, 基本地物分类精度为 91.60%, Kappa 系数为 0.90; 精细地物分类精度为 90.26%, Kappa 系数为 0.89。随着参与分类的波段数目增加, 特征维数增加, 逐渐造成分类过程中的维数灾难和过拟合的现象, 导致后期分类精度难以得到较大的提升, 基本维持稳定。故本文对于两个分类体系都选取前 40 个重要波段作为高光谱影像分类的特征波段, 降低高光谱数据特征维度, 减少数据冗余。

5.3 精度评价及分类结果

分别以 Landsat 8 OLI 7 个波段、ZY-1 02D AHSI 24 个 (或 29 个) 波段、ZY-1 02D AHSI 40 个波段作为分类特征, 利用随机森林算法分别进行 7 类基本地物和 9 类精细地物的分类, 十折交叉验证精度如表 4、表 5 所示。对于两种分类体系, ZY-1 02D 40 个波段得到的总体分类最高, Landsat 8 OLI 最低。对于 7 类基本地物分类, 由 ZY-1 02D 40 个波段得到的总体精度和 Kappa 系数分别为 92.18% 和 0.91, 而 Landsat 8 OLI 仅有 87.48% 和 0.84。对于 9 类精细地物分类, 由 ZY-1 02D 40 个波段得到的总体精度和 Kappa 系数分别为 92.18% 和 0.89, 而 Landsat 8 OLI 仅有 84.73% 和 0.82。相较于 Landsat 8 OLI, 当用与之相匹配的 24 个高光谱波段时, 7 类基本地物分类总体精度提高了 2.53% (90.01%)。对于 9 类精细地物分类而言, 使用 24 个

高光谱波段的总体精度提高了 5.06%（89.76%）。02D AHSI 连续的光谱波段能够更好的体现地物光谱曲线的细微变化，从而得到较高的分类精度。

表 3 ZY-1 02D 前 40 个重要波段与 Landsat 8 OLI 波段分布对应表
Table 3 Correspondence between the first 40 important bands of ZY-1 02D and Landsat 8 OLI bands

Landsat 8 OLI 波段	7 类基本地物分类		9 类精细地物分类	
	ZY-1 02D 波段	波段数目	ZY-1 02D 高光谱波段	波段数目
—	B1-B4(391—426 nm)	4	B2-B4(400—426 nm)	3
海岸蓝(433—453 nm)	B5-B7(426—460 nm)	3	B5-B7(426—460 nm)	3
蓝(450—515 nm)	B8-B15(451—520 nm)	8	B8-B15(451—520 nm)	8
绿(525—600 nm)	B16-B24(521—598 nm)	9	B16-B24(521—598 nm)	9
—	—		B25-B27(598—623 nm)	3
红(630—680 nm)	B32-B34(658—683 nm)	3	B28-B34(624—683 nm)	7
—	B 35-B36(684—709 nm)	2	B35-B36(684—709 nm)	2
近红外(845—885 nm)	B58(881—889 nm)	1	B56 B58(864—889 nm)	2
	B59(890—898 nm)			
—	B68-69(967—984 nm)	10	B59(890—898 nm)	2
	B72-76(1001—1044 nm)		B135-B136(1972—2002 nm)	
	B135-B136(1972—2002 nm)			
总计		40	总计	40

表 4 7 类基本地物分类精度统计表
Table 4 Statistical table of classification accuracy of seven types of basic features

分类类别	Landsat 8 OLI		ZY-1 02D AHSI 24 波段		ZY-1 02D AHSI 40 波段	
	平均 PA/%	平均 UA/%	平均 PA/%	平均 UA/%	平均 PA/%	平均 UA/%
丰水潮滩	79.13	84.14	83.36	80.58	85.77	85.16
光滩	87.93	78.39	92.79	85.58	93.34	88.14
互花米草	92.50	95.03	92.50	92.51	93.68	94.27
芦苇	83.94	86.79	85.88	91.84	90.21	91.04
芦苇怪柳混生	82.16	83.34	82.68	86.92	85.22	91.66
水体	95.32	95.12	89.60	96.04	93.71	97.09
碱蓬	80.89	84.98	87.15	88.84	90.44	90.55
云	—		99.38	96.33	99.29	98.57
阴影	—		96.07	100.00	99.00	100.00
平均总体精度/%	87.48		90.01		92.18	
平均 Kappa 系数	0.84		0.88		0.91	

对比不同影像各类别的平均 PA 和平均 UA 可知，对 7 类基本地物分类，利用 ZY-1 02D AHSI 40 个波段的 PA 和 UA 均达到 85% 以上（表 4）。对于 9 类精细地物分类而言，ZY-1 02D AHSI 40 个波段对于高、低密度互花米草以及高、低密度碱蓬的识别精度均高于 Landsat 8 OLI 7 个波段，平均 PA 和 UA 高出了 1%—13%。这说明相较于多光谱数据，ZY-1 02D AHSI 连续波段能够较好地反映植被覆盖度的差异。

图 4 为 7 类基本地物分类结果图，图 5 为 9 类

精细地物的分类结果图。可以看出，各类地物的空间分布较为一致，基于 Landsat 8 OLI 的高密度碱蓬的分布范围显著少于 ZY-1 02D 识别结果。进一步对比无人机影像（图 6(a)和图 6(d)），可看出 ZY-1 02D AHSI 前 40 个重要性波段的分类结果可以较好地区分不同密度互花米草和碱蓬以及光滩和水体的分布（图 6(c)和图 6(f)）；而在 Landsat 8 OLI 分类结果中，高、低密度互花米草和高、低密度碱蓬混淆较为严重（图 6(b)和图 6(e)）。

表 5 9类精细地物分类精度统计表

Table 5 Statistical table of classification accuracy of nine types of landscape features

分类类别	Landsat 8 OLI		ZY-1 02D AHSI 29波段		ZY-1 02D AHSI 40波段	
	平均PA/%	平均UA/%	平均PA/%	平均UA/%	平均PA/%	平均UA/%
丰水潮滩	78.05	78.70	84.36	82.29	88.22	84.58
光滩	85.33	77.92	92.54	87.61	92.91	89.34
高密度互花米草	94.42	93.18	94.36	95.88	95.43	96.25
低密度互花米草	74.21	83.83	87.27	83.45	88.76	89.01
芦苇	82.50	85.88	87.31	85.18	88.05	90.87
芦苇柆柳混生	81.55	81.88	84.65	87.05	84.61	88.43
水体	96.28	95.97	93.02	93.93	92.27	96.38
低密度碱蓬	74.89	80.01	88.05	85.42	87.67	85.39
高密度碱蓬	63.26	77.40	68.74	77.44	74.60	80.01
云	—	—	99.23	98.23	99.00	98.67
阴影	—	—	98.18	100.00	98.33	100.00
平均总体精度/%	84.73		89.76		90.76	
平均Kappa系数	0.82		0.88		0.89	

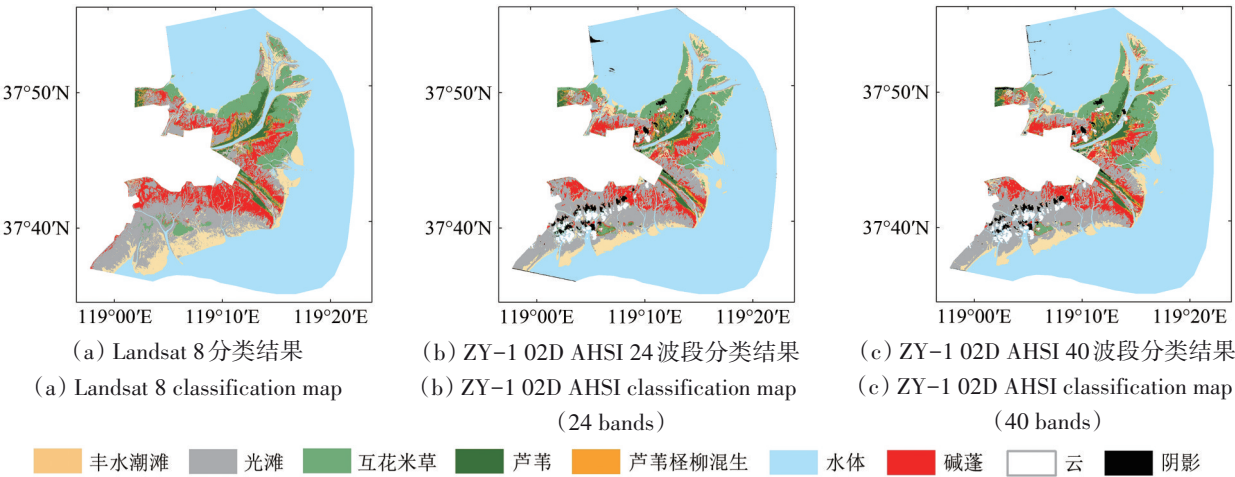


图 4 7类基本地物随机森林分类结果图

Fig. 4 Random forest classification maps for seven basic classes

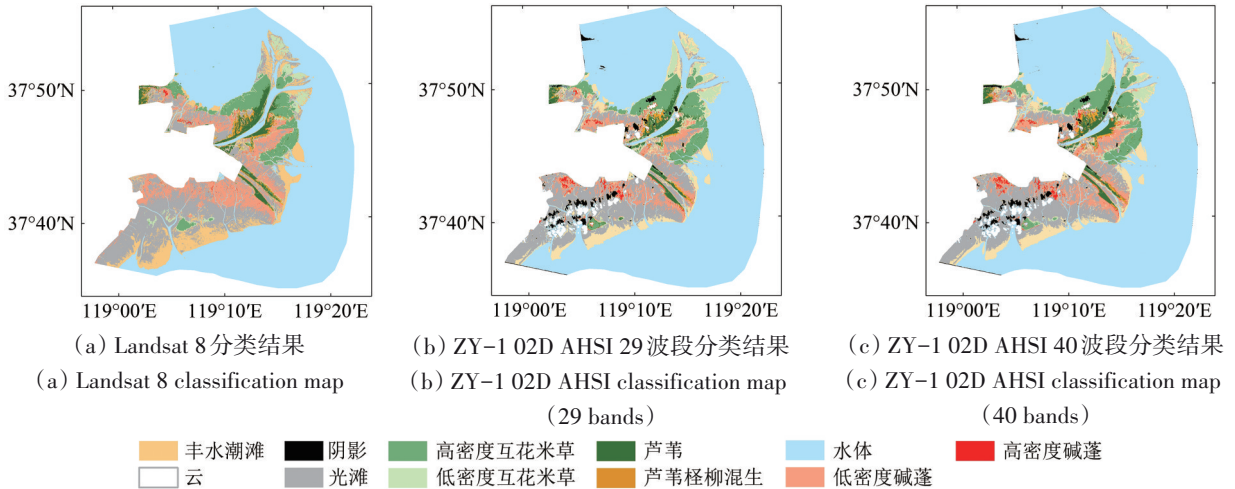


图 5 9类精细地物随机森林分类结果

Fig. 5 Results of random forest classification maps for nine refined classes

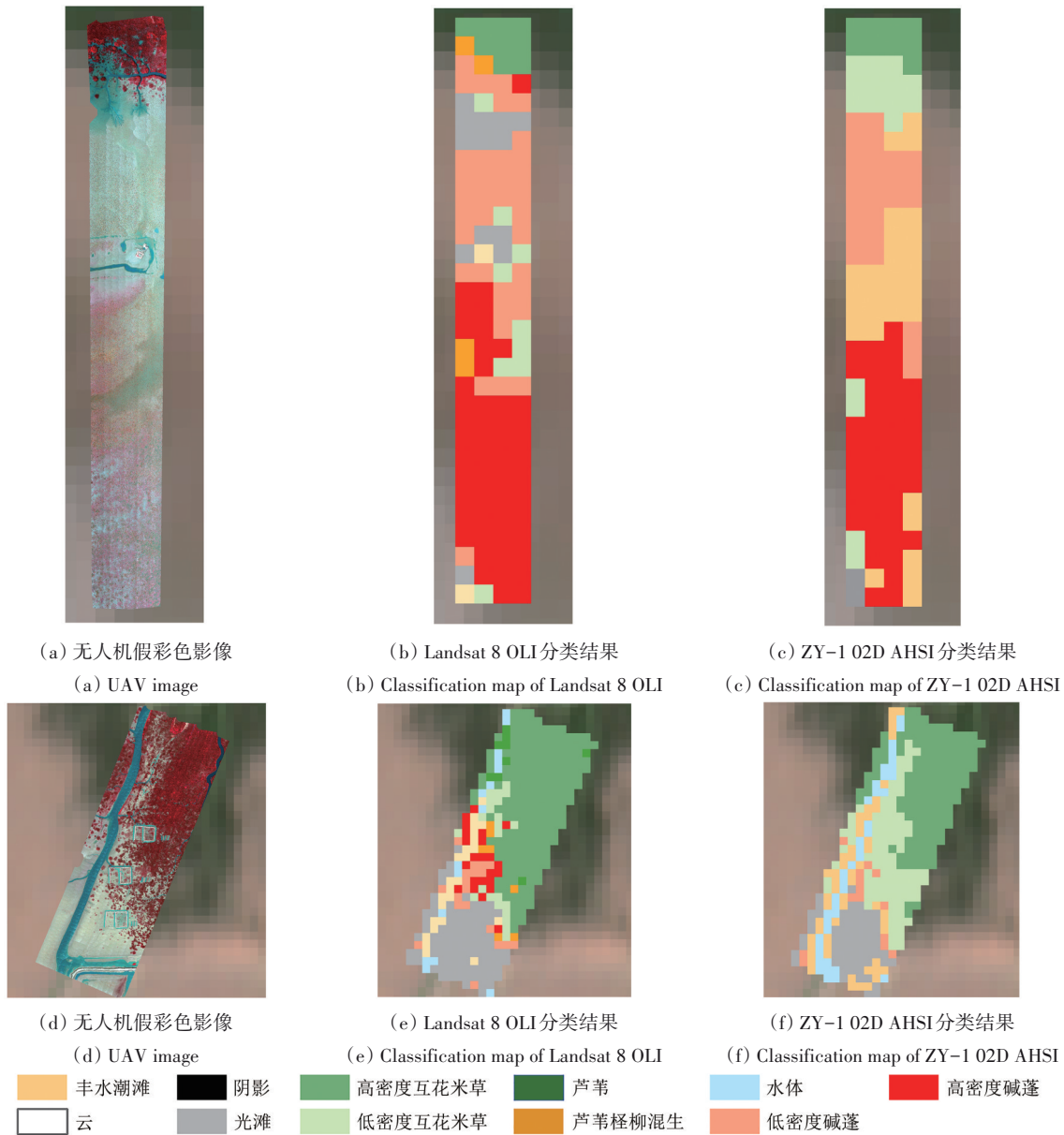


图6 无人机实测数据与分类结果对比图
Fig. 6 Comparison between the UAV image and the classification results

5.4 重要波段分析

Tree SHAP方法不仅可以计算不同波段对总体分类结果的重要性，还可以得到不同波段对各个地物类别分类的贡献。如图7所示，对于分类贡献最大的前20个波段全部位于可见光范围，其中位于蓝波段的B13、B14和B12对于总体分类精度的贡献最大；而对于各地物类别来说，其排序则有所不同，比如，位于红波段的B33、B34波段对于识别高密度碱蓬的贡献最大。表6总结了对于互花米草、碱蓬和芦苇3类盐沼植被分类的前10个重要波段与Landsat 8 OLI波段范围的匹配情况。可以看出，蓝、绿波段对高、低密度互花米草以及

芦苇识别的贡献较大。互花米草和芦苇在8—10月生长旺盛，在蓝色波段存在吸收峰、在绿色波段存在小的反射峰，从而与其他类型区分。同时，短波红外波段对于芦苇识别的重要性较高。如图2所示，互花米草和碱蓬短波红外波段的反射率均比芦苇低，这可归因于互花米草和碱蓬生长于潮间带，叶片含水量较高，因此短波红外波段反射率较低。红波段对于碱蓬，尤其是高密度碱蓬识别的贡献较大。研究区碱蓬在9—10月为结实期，生长旺盛，植株呈现紫红色。综合来看，ZY-1 02D AHSI能够捕捉不同盐沼植物在可见光范围的细微色调差异和理化特性，有利于分类。

表 6 各类典型植被前 10 个重要性波段分布表
Table 6 Distribution of top ten important bands for typical vegetation classes

Landsat8	ZY-1 02D AHSI 重要波段				
	高密度互花米草	低密度互花米草	低密度碱蓬	高密度碱蓬	芦苇
—	B4(417—426 nm)	—	—	B2-4(400—426 nm)	—
海岸蓝	B5(426—434 nm)	B5-6(426—442 nm)	B5-7(426—451 nm)	B5-6(426—443 nm)	B5(426—434 nm)
蓝	B8(451—456 nm) B10-14(469—512 nm)	B9-14(460—512 nm)	B9(460—468 nm) B11-14(477—512 nm)	B13(495—503 nm)	B11-14(477—512 nm)
绿	B22-23(572—589 nm)	B22—23(572—589 nm)	—	—	B22—24(572—598 nm)
红	—	—	B33-34(666—683 nm)	B32-34(658—683 nm)	—
—	—	—	—	B35(683—692 nm)	B135-136(1972—2002 nm)

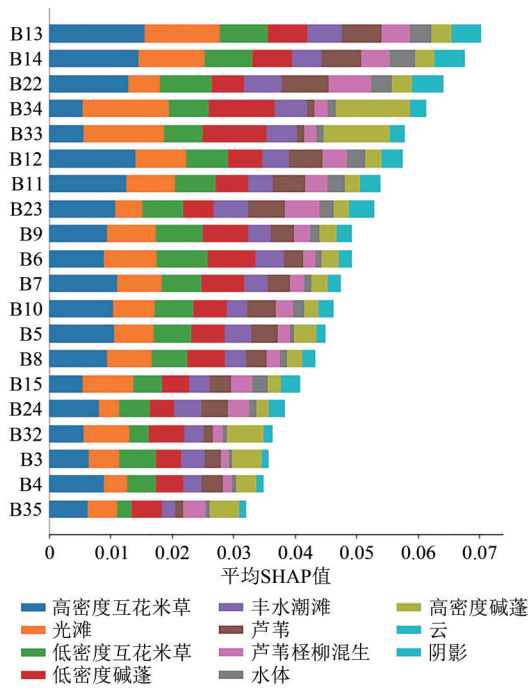


图 7 高光谱波段特征重要性权重分布图
Fig. 7 Hyperspectral band feature importance weight distribution map

6 结 论

本研究探索了新的高光谱数据源 ZY-1 02D AHSI 影像在河口三角洲湿地景观精细分类中的应用潜力。目前大部分湿地景观分类研究仍以植被群落尺度分类为主，未在分类体系中考虑植被群落覆盖度差异。针对黄河三角洲湿地斑块破碎、盐沼植被覆盖度差异明显的特点，本文利用 AHSI 数据进行了 7 类基本地物和考虑植被高、低覆盖度的 9 类精细地物分类。主要结论如下：

(1) ZY-1 02D AHSI 光谱曲线能够较好地反映研究区各类地物的光谱差异，可用于后续的景观分类研究。

(2) ZY-1 02D AHSI 对研究区 7 类基本地物的分类精度为 92.18%；9 类精细地物分类精度达到 90.97%。Tree SHAP 特征选择显示，前 40 个重要波段中，分别有 24 个和 29 个波段位于 Landsat 8 OLI 数据的可见光、近红外波段。

(3) 仅用与 Landsat 8 OLI 相重合的重要波段进行分类，对 9 类精细地物的分类精度也可达到 89.97%，远高于 Landsat 8 OLI 多光谱数据，表明 ZY-1 02D AHSI 数据连续的波谱曲线在识别植被覆盖度差异方面优势明显。

(4) 蓝、绿波段对于识别高、低密度互花米草较为重要，蓝、绿、短波红外波段对于芦苇的识别较为重要，红波段对于识别高、低密度碱蓬较为重要。

综上所述，ZY-1 02D AHSI 在黄河三角洲湿地景观分类中取得了良好的效果，在区分盐沼植被覆盖度方面发挥了多光谱数据不可替代的优势。今后研究将发展适宜于湿地植被识别和分类的高光谱指数，充分发挥 ZY-1 02D AHSI 波谱连续的特点，进行植被覆盖度定量反演研究。

参考文献(References)

Chen B Q, Xiao X M, Li X P, Pan L H, Doughty R, Ma J, Dong J W, Qin Y W, Zhao B, Wu Z X, Sun R, Lan G Y, Xie G S, Clinton N and Giri C. 2017. A mangrove forest map of China in 2015: analysis of time series Landsat 7/8 and Sentinel-1A imagery in Google Earth Engine cloud computing platform. ISPRS Journal of Photogrammetry and Remote Sensing, 131: 104-120. [DOI: 10.1016/j.isprsjprs.2017.07.011]

Cheng X Y, Cai Z Y, Li J, Wen M X, Wang Y M and Zeng D. 2021. A spatial-spectral clustering-based algorithm for endmember extraction and hyperspectral unmixing. International Journal of Remote Sensing, 42(5): 1948-1972. [DOI: 10.1080/01431161.2020.1849851]

- Du P J, Xia J S, Xue Z H, Tan K, Su H J and Bao R. 2016. Review of hyperspectral remote sensing image classification. *Journal of Remote Sensing*, 20(2): 236-256 (杜培军, 夏俊士, 薛朝辉, 谭琨, 苏红军, 鲍蕊. 2016. 高光谱遥感影像分类研究进展. *遥感学报*, 20(2): 236-256) [DOI: 10.11834/jrs.20165022]
- Gong N, Niu Z G, Qi W and Zhang H Y. 2016. Driving forces of wetland change in China. *Journal of Remote Sensing*, 20(2): 172-183 (宫宁, 牛振国, 齐伟, 张海英. 2016. 中国湿地变化的驱动力分析. *遥感学报*, 20(2): 172-183) [DOI: 10.11834/jrs.20164210]
- Hughes G. 1968. On the mean accuracy of statistical pattern recognizers. *IEEE Transactions on Information Theory*, 14(1): 55-63 [DOI: 10.1109/TIT.1968.1054102]
- Jiao L L, Sun W W, Yang G, Ren G B and Liu Y N. 2019. A hierarchical classification framework of satellite multispectral/hyperspectral images for mapping coastal wetlands. *Remote Sensing*, 11(19): 2238 [DOI: 10.3390/rs11192238]
- Li C, Chen G, Chu W Q and Hu H H. 2021. Research on main influencing factors of bursting of urban water supply network based on improved SHAP. *Bulletin of Science and Technology*, 37(1): 79-84 (李超, 陈功, 储文强, 胡鸿昊. 2021. 基于改进 SHAP 的城市供水管网爆管主影响因素研究. *科技通报*, 37(1): 79-84) [DOI: 10.13774/j.cnki.kjtb.2021.01.014]
- Li P, Li D H, Li Z H and Wang H J. 2019. Wetland classification through integration of GF-3 SAR and sentinel-2B multispectral data over the Yellow River Delta. *Geomatics and Information Science of Wuhan University*, 44(11): 1641-1649 (李鹏, 黎达辉, 李振洪, 王厚杰. 2019. 黄河三角洲地区 GF-3 雷达数据与 Sentinel-2 多光谱数据湿地协同分类研究. *武汉大学学报(信息科学版)*, 44(11): 1641-1649) [DOI: 10.13203/j.whugis.20180258]
- Li Y, Li Y R, Wang G H and Shi X. 2020. A hyperspectral image classification algorithm based on the weighted exponential function model. *Journal of Geo-Information Science*, 22(8): 1642-1653 (李玉, 李奕燃, 王光辉, 石雪. 2020. 基于加权指数函数模型的高光谱图像分类方法. *地球信息科学学报*, 22(8): 1642-1653) [DOI: 10.12082/dqxxkx.2020.190383]
- Li Y T, Du Z Y, Wang X, Yang Q S, Chen Z Q, Sun Y Y and Liu D X. 2019. Evaluation of wetland ecosystem services in Yellow River delta nature reserve. *Marine Environmental Science*, 38(5): 761-768 (李永涛, 杜振宇, 王霞, 杨庆山, 陈占强, 孙燕燕, 刘德玺. 2019. 黄河三角洲自然保护区湿地生态服务功能价值评估. *海洋环境科学*, 38(5): 761-768) [DOI: 10.13634/j.cnki.mes.2019.05.017]
- Liu S and Xing G L. 2020. Multi-dimensional convolutional network collaborative unmixing method for hyperspectral image mixed pixels. *Acta Geodaetica et Cartographica Sinica*, 49(12): 1600-1608 (刘帅, 邢光龙. 2020. 高光谱图像混合像元多维卷积网络协同分解法. *测绘学报*, 49(12): 1600-1608) [DOI: 10.11947/j.AGCS.2020.20190461]
- Lundberg S M, Erion G G and Lee S I. 2018. Consistent individualized feature attribution for tree ensembles. *arXiv*: 1802.03888.
- Lundberg S M and Lee S I. 2017. A unified approach to interpreting model predictions//*Proceedings of the 31st International Conference on Neural Information Processing Systems*. Long Beach: ACM: 4768-4777 [DOI: 10.5555/3295222.3295230]
- Rodríguez-Pérez R and Bajorath J. 2020. Interpretation of compound activity predictions from complex machine learning models using local approximations and Shapley values. *Journal of Medicinal Chemistry*, 63(16): 8761-8777 [DOI: 10.1021/acs.jmedchem.9b01101]
- 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]
- Sun Z G, Song H L and Hu X Y. 2017. Response of germination and seeding growth of Suaeda salsa seeds from tidal marshes in the Yellow River estuary to N/P ratio of soil. *Wetland Science*, 15(1): 10-19 (孙志高, 宋红丽, 胡星云. 2017. 黄河口潮滩碱蓬种子萌发与幼苗生长对土壤中氮磷比的响应. *湿地科学*, 15(1): 10-19) [DOI: 10.13248/j.cnki.wetlandsci.2017.01.002]
- Wan L M, Lin Y Y, Zhang H S, Wang F, Liu M F and Lin H. 2020. GF-5 hyperspectral data for species mapping of mangrove in Mai Po, Hong Kong. *Remote Sensing*, 12(4): 656 [DOI: 10.3390/rs12040656]
- Wang X P. 2014. Study on the Yellow River Delta Wetland Typical Vegetation Using Hyperspectral Remote Sensing. Dalian: Dalian Maritime University (王霄鹏. 2014. 黄河三角洲湿地典型植被高光谱遥感研究. 大连: 大连海事大学)
- Wang X P, Zhang J, Ma Y and Ren G B. 2014. Comparison of coastal wetland classification accuracy using fused images of different polarization SAR and TM images. *Journal of Marine Sciences*, 32(1): 40-46 (王霄鹏, 张杰, 马毅, 任广波. 2014. 不同极化方式 SAR 与 TM 融合影像滨海湿地分类精度比较. *海洋学研究*, 32(1): 40-46) [DOI: 10.3969/j.issn.1001-909X.2014.01.005]
- Xu Z T. 2020. Information Extraction and Dynamic Change Analysis of Yellow River Delta Wetland Based on Landsat Data. Qingdao: Qingdao University (徐振田. 2020. 基于 Landsat 数据的黄河三角洲湿地信息提取及动态变化分析. 青岛: 青岛大学) [DOI: 10.27262/d.cnki.gqda.2020.002244]
- 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.20198083]
- Zhang L P and Zhang L F. 2005. *Hyperspectral Remote Sensing*. Wuhan: Wuhan University Press: 1-8 (张良培, 张立福. 2005. 高光谱遥感. 武汉: 武汉大学出版社: 1-8)
- Zhang Y C, Na X D and Zang S Y. 2018. Wetland high precision classification based on the HJ-1A hyperspectral image. *Progress in Geography* 37(12): 1705-1712 (张雅春, 那晓东, 臧淑英. 2018. 基于 HJ-1A 高光谱影像的湿地精细分类. *地理科学进展*, 37(12): 1705-1712) [DOI: 10.18306/dlkxjz.2018.12.012]

Classification of the Yellow River Delta wetland landscape based on ZY-1 02D hyperspectral imagery

HAN Yue^{1,2,3}, KE Yinghai^{1,2,3}, WANG Zhanpeng^{1,2,3}, LIANG Deyin⁴, ZHOU Demin^{1,2,3}

1. State Key Laboratory Cultivation Base of Urban Environment Process and Simulation, Beijing 100048, China;

2. Beijing Laboratory of Water Resources Security, Beijing 100048, China;

3. College of Resource Environment and Tourism, Capital Normal University, Beijing 100048, China;

4. Remote Sensing Satellite General Division, China Academy of Space Technology, Beijing 100086, China

Abstract: The ZY-1 02D satellite was successfully launched in 2019 and officially used in October 2020. It is the first civil hyperspectral satellite independently developed and operated by China and has a wide application prospect. Taking the Yellow River Delta wetland as study area, this work investigated the performance of ZY-1 02D Advanced Hyperspectral Instrument (AHSI) images in wetland landscape classification. First, the ZY-1 02D AHSI spectral curves of typical land cover types in the study area were evaluated. Second, two levels of wetland landscape classification systems, including seven basic classes and nine refined classes considering the difference of vegetation coverage, were developed with the assistance of field surveys and unmanned aerial vehicle images. Then, the random forest algorithm was used for classification, and the tree SHAP method was introduced to sort and select important bands. The most important ZY-1 02D AHSI bands overlapping with Landsat 8 OLI bands were selected for classification, and the classification results were compared. Our results showed the following: (1) ZY-1 02D AHSI data can represent the differences of spectral features of different landscape types. (2) For the two classification systems, the first 40 important bands achieved the highest overall classification accuracy. The classification accuracy of seven basic classes and nine refined classes were 92.18% and 90.76%, respectively. Most of the important bands were visible and near-infrared bands. (3) For the two classification systems, a total of 24 and 29 bands overlapping with Landsat 8 OLI multispectral bands were selected for classification. The classification accuracy reached 90.01% and 89.76%, which were remarkably higher than that of Landsat 8 OLI. (4) Blue and green bands are important for the identification of high- and low-density *Spartina alterniflora* and *Phragmites australis*; blue, green, and short-wave infrared bands are important for *Phragmites australis* identification, and the red band is important for the identification of high- and low-density *Suaeda salsa*. Results showed that ZY-1 02D AHSI data have advantages in distinguishing different landscape features and differences in vegetation coverage. Our study is expected to provide scientific basis for the application of ZY-1 02D hyperspectral data in wetland ecological resources.

Key words: ZY-1 02D, AHSI satellite image, hyperspectral data, vegetation coverage, random forest, Tree SHAP

Supported by National Natural Science Foundation of China (No. 42071396)