

# Sentinel-2影像多特征优选的黄河三角洲湿地信息提取

张磊, 宫兆宁, 王启为, 金点点, 汪星

1. 首都师范大学资源环境与旅游学院, 北京 100048;
2. 三维信息获取与应用教育部重点实验室, 北京 100048;
3. 资源环境与地理信息系统北京市重点实验室, 北京 100048

**摘要:** 以北方典型河口湿地—黄河三角洲湿地为研究区, 采用在特征选择和分类提取等方面具有明显优势的随机森林算法, 对研究区内的湿地信息进行提取。首先基于多时相、光谱信息丰富的Sentinel-2数据生成4类不同的特征变量, 包括光谱特征、植被指数和水体指数、红边指数、纹理特征; 再根据以上特征构建6种不同的提取方案, 对黄河三角洲湿地信息进行提取并验证不同方案的提取精度, 旨在选择最佳方案改善湿地信息提取的效果。结果表明: (1) 有效地使用多种特征变量是提高湿地信息提取的关键, 就不同特征对湿地信息提取的贡献率而言, 红边指数>植被指数和水体指数>光谱特征>纹理特征; (2) 基于随机森林算法优选的特征变量提取效果最佳, 总体精度高达90.93%, Kappa系数为0.90, 表明随机森林算法可以有效地进行特征选择, 在特征变量数据挖掘的同时, 仍能保证湿地信息提取的精度, 提高运行效率。本研究为湿地信息提取在数据源选择、特征选择和方法选择方面提供了一种新思路、方法和技术手段。

**关键词:** 河口湿地, 信息提取, Sentinel-2, 随机森林, 特征选择, 红边指数, 多时相数据

**引用格式:** 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

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 [DOI:10.11834/jrs.20198083]

## 1 引言

河口湿地位于江河入海的海陆交界处, 是两种截然不同的生态系统强烈作用而形成的生态边缘区, 为海生和陆生生态系统界面相互延展的重叠区域。河口湿地得天独厚的地理位置往往决定了该区域具有优越的区位优势、丰富的水资源、动植物资源、油气资源以及港口资源等。黄河三角洲湿地作为中国北方典型河口湿地, 在2013年被国际湿地公约秘书处列入“国际重要湿地名录”, 是世界范围内极具代表性的河口湿地之一, 也是中国暖温带最完整、最广阔、最年轻的湿地生态系统(温庆可等, 2011; 宫宁等, 2016; 刘润等, 2017; 宋红丽, 2015)。但是随着快速城市化

的加剧、大面积农田的开垦、工业的发展和港口的建设, 黄河河口湿地也已成为全球受威胁最严重的生态系统之一。因此, 急需采用先进的技术手段及时地监测黄河三角洲湿地的状况, 了解该区域不同类型湿地的空间分布格局与特征, 为今后保护区湿地资源的保护与合理利用提供科学依据。

相比于野外实地调查研究, 遥感技术具有观测范围广、监测周期短、获取信息量大等优点, 在湿地信息提取及其变化监测方面已经发挥了重要的作用。在湿地信息提取方法方面, 早期采取地目视解译方法对解译人员和工作量具有极高的要求, 使其不适宜作为一种独立的方法进行湿地信息提取研究(刘红等, 2004)。随着计算机自动分类技术逐步应用于遥感图像分类, 国内外学者一

收稿日期: 2018-03-08; 预印本: 2018-06-06

基金项目: 国家重点研发计划资助(编号: 2017YFC0505903)

第一作者简介: 张磊, 1993年生, 男, 硕士研究生, 研究方向为遥感技术及地学分析应用。E-mail: zhangleigis@outlook.com

通信作者简介: 宫兆宁, 1976年生, 女, 副教授, 研究方向为遥感技术及地学分析应用。E-mail: gongzhn@163.com

直在探索更高效、更精确提取影像信息的途径。利用多光谱数据对湿地信息进行提取时,最大似然分类算法是最常用的方法,但是该方法只能提供较低的分类精度(30%—60%)(Özesmi等, 2000; Na等, 2009; 黄进良, 1999);后来研究者发现,分类回归树CART(Classification And Regression Tree)(Davranche等, 2010; Na等, 2009; 那晓东等, 2008; Baker等, 2006)、支持向量机SVM(Support Vector Machines)(Wang等, 2012; 张策等, 2011)和人工神经网络(Artificial Neural Networks)(Bao和Ren, 2011; Abuelgasim等, 1996; 肖锦成等, 2013)等可以在一定程度上提高湿地信息提取的精度; Khatami等人(2016)统计了15年间基于像素的分类算法,研究得出支持向量机和随机森林算法分类精度较高,其次是神经网络,最大似然分类则是普遍使用的基准算法,提供精度较低。

利用以上方法进行湿地信息提取时,综合多源遥感信息的多特征变量提取与优化是今后湿地信息智能化提取的重点和难点之一,具有极大的研究空间和研究价值(张树文等, 2013)。目前,湿地信息提取所采用的主要特征变量为光谱特征及其衍生因子、纹理特征、时相特征以及DEM等其它辅助地学特征。黄进良(1999)在提取洞庭湖湿地信息时,只采用了Landsat数据的光谱反射率等信息,但提取精度有待商榷。Na等人(2009)采用Landsat-5 TM数据提取三江平原的湿地信息时,在原始光谱反射率的基础上加入植被指数(NDVI、EVI)、纹理特征和地形地貌特征之后,提取精度提高了6.53%;张猛等人(2017)采用时间序列的MODIS数据对洞庭湖流域大范围的湿地信息进行提取,发现多时相数据所反应的地物特征更加明显,能够有效地解决“异物同谱”、“同物异谱”的现象。近年来,一些与植被红边指数和叶绿素相关的光谱衍生因子已经被证实与地物信息息息相关(Delegido等, 2011; Shoko和Mutanga, 2017)。但是由于之前大部分传感器缺少相应光谱区域的波段,导致这些指数在实际应用中难以实现。

21世纪以来,随着各个国家和地区对地观测系统的完善,越来越多的遥感影像突破了波段数量、空间分辨率和重访周期的限制(郑阳等,

2017)。Sentinel-2对地观测卫星是欧盟委员会(European Commission)和欧洲航天局ESA(European Space Agency)共同倡议地哥白尼计划的重要组成部分,目的是帮助欧洲乃至全球监测陆地和海洋环境并满足其应对自然灾害和环境变化等安全需求(Drusch等, 2012)。Sentinel-2两颗卫星均搭载多光谱传感器MSI(Multispectral Instrument),可提供从可见光、近红外到短波红外(13个波段)多尺度、中高空间分辨率(10 m、20 m和60 m)的遥感影像;而且该系列卫星特意在植被光谱的红边区域(670—760 nm)设置了3个波段,这对植被生长信息及其健康状况的监测非常有效(Shoko和Mutanga, 2017; Delegido等, 2011; Korhonen等, 2017);同时,双星运行的重访周期缩短至5 d,从而极大地增强了对地观测能力(郑阳等, 2017)。虽然Sentinel-2影像具有时空分辨率高、光谱信息丰富,且获取成本低廉的优点,但是目前该数据在湿地信息提取方面的应用价值尚未挖掘,尤其是通过植被红边波段及其衍生因子对湿地信息进行更加精细化、准确化的提取。

本文主要完成以下两方面工作:首先,采用Sentinel-2系列影像为基础数据源,基于预处理的影像数据构建多种特征变量和湿地信息提取方案。之后,采用随机森林的算法对不同方案的湿地信息进行提取并使用该算法对所有特征变量进行优化选择,在确定不同特征变量重要性程度的同时,提高湿地信息提取精度。

## 2 研究区与数据

### 2.1 研究区概况

黄河三角洲位于黄河入海口处,是中国4大河流三角洲之一,现为国家级自然保护区。三角洲北临渤海,东靠莱州湾,地处118°32.981'E—119°20.450'E,北纬37°34.768'N—38°12.310'N之间。研究区位于北半球暖温带欧亚大陆东部,受海陆热力性质差异影响形成温带季风气候,四季分明,夏季高温多雨,冬季寒冷干燥(刘莉等, 2017)。自然保护区由两部分组成,分别为1976年之前黄河刁口河流路入海口部分和现行黄河入海口部分(黄河口、大汶流)。由于受影像质量的限制,本文的研究区主要分布在现行黄河入海口地区,即黄河口和大汶流地区(图1)。

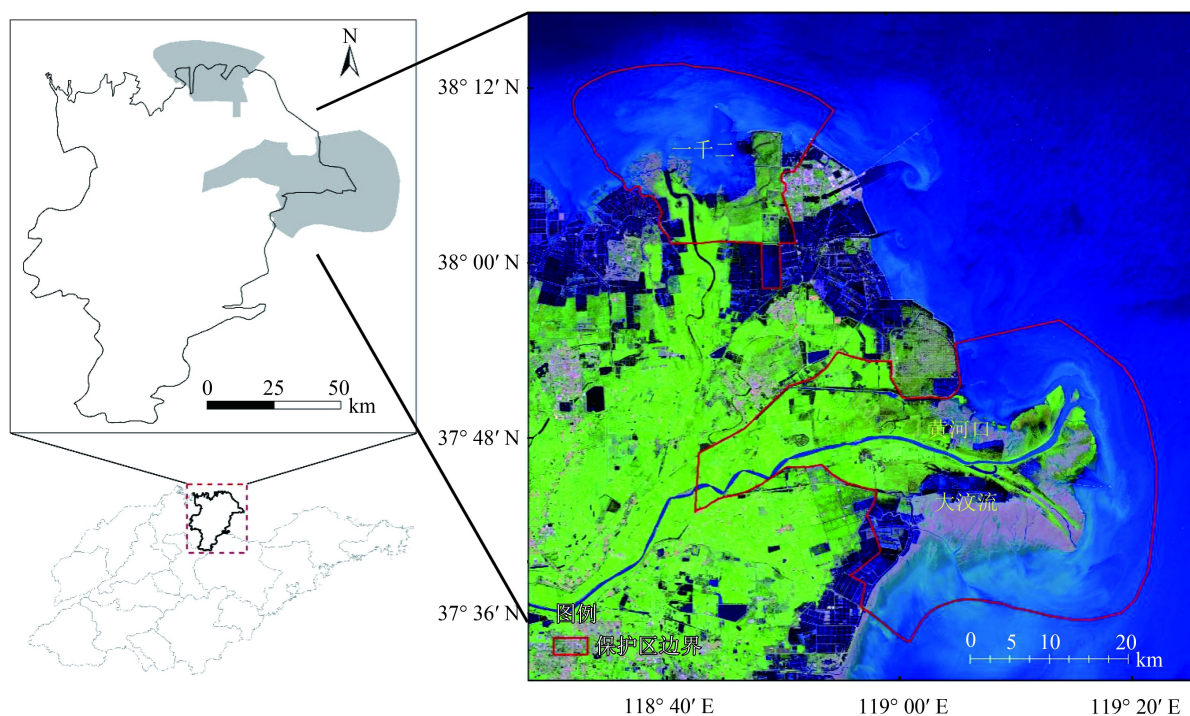


图1 研究区位置

Fig. 1 Location of the study area

现行黄河入海口地区独特的地理位置使其兼受河、海、陆交互作用, 导致该地区的陆地水文条件比较复杂, 主要分布地表水、地下水和海水3种水体类型。20世纪80年代以来, 黄河河口的来水来沙量大幅减小, 海水倒灌、海岸侵蚀严重, 导致研究区土壤盐渍化愈发明显。受研究区水、盐梯度分异性的影响, 黄河入海口地区自然植被可划分为盐生植被、水生植被和湿生植被。盐生植被主要分布于年高潮线上侧的微斜平地、潮间带滩涂的中上部, 潮上带分布的盐生植被通常和湿生植被呈复区分布; 湿地水生植被主要分布于距现代海岸线较远的微斜平地中上部、河间洼地, 地表经常或者较长周期有较深的积水, 土壤以含盐量较低的盐化潮土和潮盐土为主; 湿生植被集中分布于现代黄河入海口附近及古河道、决口扇形地间的积水洼地。(张绪良等, 2009; 黄葵等, 2012; 武吉华等, 2004)。

## 2.2 数据及预处理

### 2.2.1 遥感数据

为充分体现河口湿地复杂的生境信息及季相演替规律, 选取了涵盖2017年整个生长期的5景

Sentinel-2数据, 其中4景为Sentinel-2A(以下简称S2A)数据, 1景为Sentinel-2B(以下简称S2B)数据, 数据获取时相及其详细信息分别见表1和表2。

表1 S2A和S2B获取时间  
Table 1 Acquisition time of S2A and S2B

影像编号	获取日期	获取卫星	波段数	影像质量
A	2017-05-12	S2A	13	良好
B	2017-06-21	S2A	13	良好
C	2017-07-11	S2A	13	良好
D	2017-09-29	S2A	13	良好
E	2017-10-24	S2B	13	良好

本文所采用的Sentinel-2数据均来源于欧洲航天局的数据共享网站(<https://scihub.copernicus.eu/> [2018-03-08])。Sentinel-2发布的数据为已经进行过几何校正和辐射校正的大气上层表观反射率产品, 所以只需对该数据进行大气校正即可。采用ESA官方提供的SNAP软件对下载的数据进行大气校正。实验证明, SNAP软件中大气校正模型Sen2cor(Sentinel-2 Level-2A Atmospheric Correction Processor)与SMAC模型、6S模型相比较, 校正后获取的影像光谱曲线与地面实测的光谱曲线拟合

度、精度最高(苏伟等; 2018)。同时, 选取大气校正后波段数较多且空间分辨率较高的20 m遥感影像作为数据源, 大气校正之后影像输出的波段如表2中粗体字部分所示, 空间分辨率为10 m的波段会自动重采样成20 m。

表 2 S2A和S2B详细信息对比

Table 2 Spectral bands and resolutions of S2A and S2B sensors

波段	Sentinel-2A		Sentinel-2B		分辨率/ m
	中心波长/	波段宽度/	中心波长/	波段宽度/	
	nm	nm	nm	nm	
1 海岸波段	443.9	27	442.3	45	60
2 蓝波段	496.9	98	492.1	98	10
3 绿波段	560.0	45	559	46	10
4 红波段	664.5	38	665	39	10
5 植被红边1波段	703.9	19	703.8	20	20
6 植被红边2波段	740.2	18	739.1	18	20
7 植被红边3波段	782.5	28	779.7	28	20
8 近红外波段	835.1	145	833.0	133	10
8a 近红外波段	864.8	33	864.0	32	20
9 水汽波段	945.0	26	943.2	27	60
10 卷云波段	1373.5	75	1376.9	76	60
11 短波红外1	1613.7	143	1610.4	141	20
12 短波红外2	2202.4	242	2185.7	238	20

## 2.2.2 样本与验证数据

样本质量直接关系到湿地信息提取的精度, 应选择具有典型性、代表性的纯净像元作为样本(刘舒等, 2017)。本文的样本数据主要来自于现场实地踏勘和依据Google Earth高分辨率影像进行目视解译两种方式。2017年8月, 实验组对黄河三角洲自然保护区进行了详细的调研, 采用高精度GPS对不同湿地分布及其植被类型进行定位、记录并拍照。之后, 在室内通过Google Earth软件目视解译一部分湿地类型, 以上两部分数据构成样本数据和验证数据分别用于分类器的建立和精度验证。为了减少由于样本数据的比例不同而引起的误差, 本次实验设定各湿地类别样本点均不低于100(由于盐田面积较少, 样本点未达到100)。

## 3 湿地分类方案

参考湿地公约、《全国湿地资源调查与监测技术规程》以及相关文献资料(牟晓杰等, 2015; 王雪宏等, 2015), 结合研究区湿地分布的具体情况, 制订如下湿地分类方案(表3)。由于研究区经济发达, 人类活动频繁且剧烈, 湿地信息提取的同时不可避免的会受到人类活动的影响, 本文对处在现行黄河入海口地区, 但属于非湿地的类型也进行了提取, 主要包含建筑用地、旱地和林地。

表 3 黄河三角洲自然保护区湿地分类方案

Table 3 The categories plan of wetlands in the Yellow River Delta National Nature Reserve

一级分类	二级分类	具体说明
近海与海岸湿地	浅海水域	低潮时水深<6 m的浅海水域
	泥质海滩	淤泥质海滩、河滩
河流湿地	河流	常年有水或间歇性有水流动的河流, 包括河床部分
沼泽湿地	草本沼泽	以草本植物为主的永久或季节性咸淡水沼泽, 喜湿多年生草本和禾本科植物占优势, 研究区典型植被如芦苇、香蒲、盐地碱蓬、互花米草等
	灌木沼泽	以灌木为主的永久性或季节性沼泽, 如柽柳等
人工湿地	水库/坑塘	包括水库、坑塘、养殖池以及城市景观和娱乐水面等人工建造的静止水体
	水田	用于种植水稻田、水生作物的耕地, 如水稻田、藕池
	盐田	获取盐业资源而修建的晒盐场所或盐池

## 4 研究方法

### 4.1 特征说明

本文选取多时相光谱特征、植被指数和水体

指数、红边指数(Gitelson和Merzlyak, 1994; Fernández-Manso等, 2016; Barnes等, 2000; Gitelson等, 2003)和纹理特征构建特征集, 表4详细描述了各种特征及其表达方式。

表 4 Sentinel-2特征集描述  
Table 4 Description of the feature set from Sentinel-2

特征变量	指数全称	指数简称	特征说明
光谱特征	Band	B	B2, B3, B4, B5, B6, B7, B8a, B11, B12
	Normalized Difference Vegetation Index	NDVI	(B8a-B4)/(B8a+B4)
	Ratio Vegetation Index	RVI	B8a/B4
植被指数水体指数	Difference Vegetation Index	DVI	B8a-B4
	Normalized Difference Water Index	NDWI	(B3-B8a)/(B3+B8a)
	Modified Normalized Difference Water Index	MNDWI	(B3-B11)/(B3+B11)
	Normalized Difference Vegetation Index red-edge 1	NDVIre1	(B8a-B5)/(B8a+B5)
	Normalized Difference Vegetation Index red-edge 2	NDVIre2	(B8a-B6)/(B8a+B6)
红边指数	Normalized Difference Vegetation Index red-edge 3	NDVIre3	(B8a-B7)/(B8a+B7)
	Normalized Difference red-edge 1	NDre1	(B6-B5)/(B6+B5)
	Normalized Difference red-edge 2	NDre2	(B7-B5)/(B7+B5)
	Chlorophyll Index red-edge	CIre	B7/B5-1
	Angular second moment	GLCM_A	角二阶矩
纹理特征	Correlation	GLCM_Cor	相关性
	Contrast	GLCM_Con	对比度
	Entropy	GLCM_E	熵
	Variance	GLCM_V	方差

实验选择5景影像的9个波段的反射率作为光谱特征;对影像进行预处理后提取植被指数和水体指数;由于植被红边波段作为Sentinel影像独特的波段特征,本文把红边指数单独作为一类特征,选取6种红边指数进行湿地信息提取(Hill, 2013; Shang 等, 2015)。

除光谱信息、植被指数、水体指数以及红边指数外,有研究者认为纹理信息可以提高湿地信息提取的精度(郑淑丹 等, 2014; 刘萌萌 等, 2014)。本文选取了经验精度较高的灰度共生矩阵GLCM (Grey-Level Co-occurrence Matrix),其数学表达模型如下

$$P(i, j; d, \theta) = \# \{ (x_1, y_1)(x_2, y_2) | f(x_1, y_1) = i, f(x_2, y_2) = j, |(x_1, y_1) - (x_2, y_2)| = d, \angle((x_1, y_1), (x_2, y_2)) = \theta \} \quad (1)$$

式中, #表示在该集合中的元素的数目,  $f(x_1, y_1)=i$ 表示图像 $(x_1, y_1)$ 位置像元的灰度级为*i*, *d*表示两个像元之间的距离,  $\theta$ 表示两个像元之间的方向角,通常 $d=\{1, 2, 3, 4\}$ ,  $\theta=\{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$ 。

在进行纹理分析时,选取角二阶矩、相关性、对比度、熵和方差5个冗余度较小的二阶统计量对影像进行定量分析。式2表明灰度共生矩阵与其统计量都是关于距离*d*和方向 $\theta$ 的函数,本文采取4个方向统计量叠加的均值消除角度的影响。为

避免纹理分析产生的统计分量较多且存在一定的信息交叉,本文参考前人研究经验,在对原始影像进行主成分分析之后选取第一主成分(第一个主成分方差占比均大于0.75)获取上述5个纹理特征(胡玉福 等, 2011; 侯群群 等, 2013; Pu和Landry, 2012)。

## 4.2 实验方案说明

本文共设计了以下6种方案进行对比研究(表5)。设置不同方案主要有以下两方面的目的:(1)研究不同特征变量对湿地信息提取的影响,确定不同特征变量的重要性程度;(2)通过方案之间的对比,探索提高湿地分类精度的最佳方法。

表 5 实验方案信息

Table 5 The information of experimental Programs

实验方案	特征组合
1	多时相光谱特征
2	多时相光谱特征+多时相植被指数和水体指数
3	多时相光谱特征+多时相红边指数
4	多时相光谱特征+多时相纹理特征
5	多时相光谱特征+多时相植被指数和水体指数+多时相红边指数+多时相纹理特征
6	所有特征优选组合

### 4.3 随机森林分类算法

随机森林作为一种比较新的机器学习模型,可以很好地预测多达几千个解释变量的作用,被誉为当前最好的机器学习算法之一。随机森林算法是以Breiman等人发明的分类回归树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,它的基本单元是决策树,如果把CART决策树看成分类任务中的一个专家,随机森林就是许多专家在一起对某种任务进行分类(Iverson等, 2008; Breiman, 2001)。

随机森林建立步骤如下:

(1) 在原始样本中,随机且有放回地抽取 $N$ 个训练样本(该采样方式称为bootstrap方法)组成训练样本集,每个训练样本集的数据约为原始样本数据集的63%;

(2) 基于抽取的训练样本集分别建立 $N$ 棵CART决策树组成随机森林,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;

(3) 将生成的多棵决策树组成随机森林分类器,使用该分类器对数据进行分类,采用投票的方式决定新样本的类别。

### 4.4 特征优选方法

随机森林算法不仅可以实现遥感影像的分类,而且在特征选择和降维等方面也发挥着重要的作用。由于在抽样的过程中约有37%的原始样本数据未被抽取,这部分数据被形象的称为袋外数据OOB(Out-Of-Bag)。通过OOB数据产生的袋外数据误差(Out-Of-Bag-Error)不仅可以评估分类精度,而且还可以计算不同特征变量的重要性VI(Variable Important),从而进行特征选择(Genuer等, 2010)。特征变量重要性评估模型如下

$$VI(M_A) = \frac{1}{N} \sum_{t=1}^N (B_{n_t}^{M_A} - B_{O_t}^{M_A}) \quad (2)$$

式中,  $VI$ 表示特征变量的重要性,  $M$ 为样本的全部特征数,  $N$ 为生成的决策树的棵数,  $B_{O_t}^{M_A}$ 为任意特征值 $M_A$ 未加入噪声干扰时第 $t$ 棵决策树的OOB误差,  $B_{n_t}^{M_A}$ 为任意特征值 $M_A$ 加入噪声干扰时第 $t$ 棵决

策树的OOB误差。若给某个特征 $M_A$ 随机加入噪声之后,袋外数据的准确率大幅度降低,则说明特征 $M_A$ 对分类结果影响很大,也就可以说明它的重要程度比较高。

基于表4的特征集和表5的研究方案,本文采用德国环境制图与分析计划(environment mapping and analysis program)项目组开发的EnMAP-BOX工具进行特征优选及湿地信息提取。在随机森林算法构建的过程中有两个十分重要的参数,分别是森林中决策树的个数 $N$ 和节点分裂过程中抽取的特征个数 $m$ 。在特征变量抽取时,本文默认选取EnMAP-BOX工具中总特征个数的算术平方根作为不同方案的特征个数;理论上,决策树 $N$ 的个数越多其分类准确率越高,但时间成本也越高。在抽取特征 $m$ 确定的基础上,本文通过大量的实验发现,决策树数量 $N \geq 100$ 时,所有方案的OOB误差逐渐收敛并趋于稳定。因此,本文选取 $N=100$ 作为生成决策树的数量。

### 4.5 精度评价

混淆矩阵(Confusion Matrix)也被称作误差矩阵,在精度评价中,主要用于比较分类结果和实际测得值之间的混淆程度进行精度评估。本文主要选取目前普遍采用的总体精度、Kappa系数、生产者精度和用户精度作为评价指标对各个方案进行评价。

## 5 结果与分析

方案1—方案5作为对比实验,不需要进行特征选择,前5种方案旨在探究不同特征变量对湿地信息提取精度的影响。在方案5的基础上,采用随机森林算法对125个特征变量进行重要性评价,再通过随机森林分类器对逐一累加的湿地信息进行提取,最后选择分类精度最高的特征子集构成方案6。因此确定方案6的优选特征成为后续实验进行的关键。

### 5.1 方案6特征确定

为了更加直观清晰地呈现高重要性的特征变量,结合特征变量个数与分类精度和Kappa系数的关系(图3),选取前33个重要性得分较高的特征生成重要性分布图(图2)。

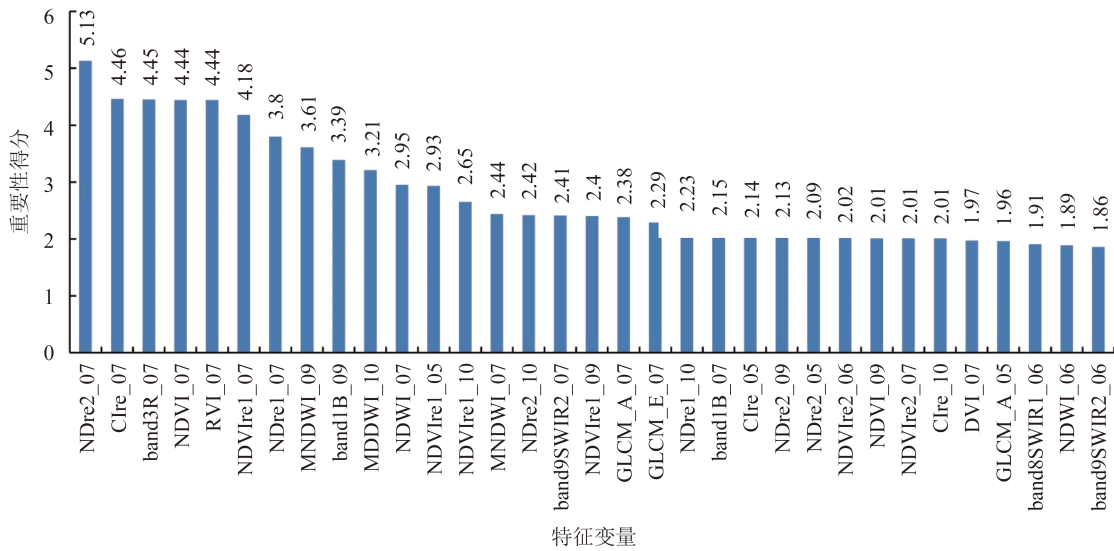


图2 特征重要性分布

Fig. 2 The distribution of characteristic importance

注: NDV Ire2\_07表示7月份的Normalized Difference Vegetation Index red-edge 2指数, 特征Clre\_07表示7月份的Chlorophyll Index red-edge指数, 特征band3R\_07表示7月份红波段, 以此类推, 表7如此。

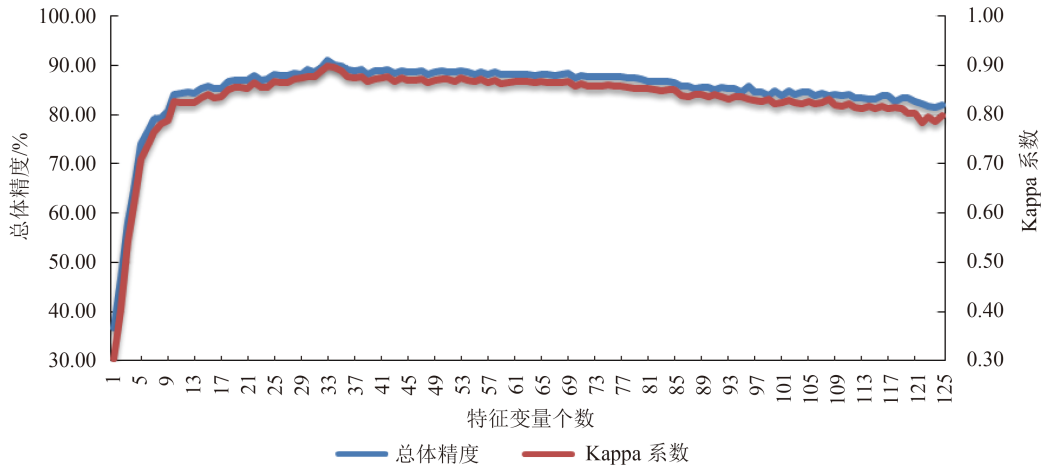


图3 特征变量个数与分类精度和Kappa系数关系

Fig. 3 The relation between the number of characteristic variables and the classification accuracy and the Kappa coefficient

统计不同特征变量的重要性得分发现: 不同特征变量的重要性得分差异较大, NDre2\_07特征的重要性得分最高, 高达5.13(图2), GLCM\_Cor\_09特征重要性得分最低(0.01), 几乎不产生影响。图3表明, 随着参与分类的特征变量的增加, 前期(前10个特征)分类精度呈现急速上升的趋势, 从单个特征分类精度为36.78%迅速达到84.15%, 这主要因为前期特征变量的重要性评分高, 特征之间相关性小和冗余特征少, 从而提高了分类器的性能; 中期(11—33个特征)的分类精度提升速度大幅降低, 但是仍然呈现平稳上升的趋势, 分类精度从84.23%逐步达到90.93%; 后期(34—125个特

征), 分类精度逐渐呈现下降的趋势, 而且总体下降速度逐渐加快。这是因为后期冗余特征和不相关特征增加, 降低了分类器的性能, 导致分类精度降低。图3同时表明, 当特征变量的个数达到33个时, 分类精度和Kappa均达到最大值, 分别为90.93%和0.90, 因此将前33个特征作为方案6的最终结果参与精度评价。

## 5.2 提取结果及精度评价分析

6种不同方案的提取结果如图4所示。通过目视解译的方法对比原始图像可以看出: 方案1、方案4、方案5的湿地信息提取不理想, 不同类别之

间存在较多错分的情况。如，在大汶流恢复区人工建造的水库/坑塘提取不完整或直接错分为河流类别，方案4中部分草本沼泽被错分为水田，方案5中部分泥质海滩被错分为灌木沼泽、建筑用

地或旱地；所有方案在海陆交界地带都存在河流和浅海水域两种类别混分的现象。从目视解译对比的效果来看，方案2、方案3、方案6的提取结果较好。

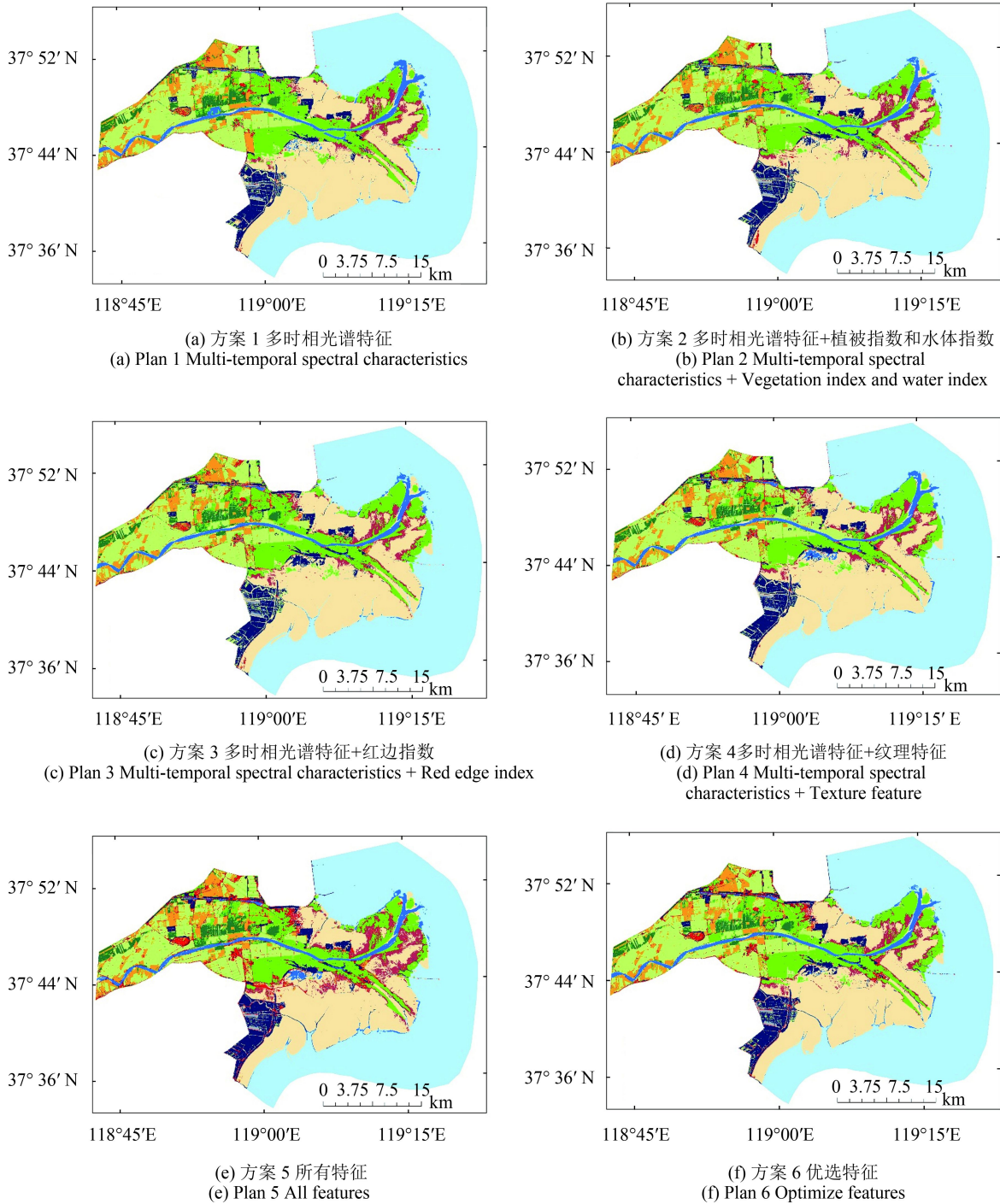


图 4 不同方案分类结果

Fig. 4 Classification results of different plans



为了定量准确地评价不同方案对湿地信息提取精度的影响,利用验证数据,对不同方案进行精

度评价分析。主要评价指标包括总体精度、Kappa系数、生产者精度以及用户精度,具体结果见表6。

表6 分类精度统计  
Table 6 The statistics of classification accuracy

类别	方案1		方案2		方案3		方案4		方案5		方案6	
	PA/%	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%	UA/%	PA/%	UA/%
浅海水域	90.20	86.79	90.20	86.79	88.68	88.68	86.54	84.91	86.27	83.02	94.00	88.68
泥质海滩	80.36	81.82	83.33	81.82	83.64	83.64	80.70	83.64	81.48	80.00	86.21	90.91
河流	82.35	84.00	85.71	84.00	86.00	86.00	78.85	82.00	80.39	82.00	90.00	90.00
草本沼泽	83.53	83.53	86.90	85.88	88.24	88.24	83.33	82.35	83.33	82.35	91.76	91.76
灌木沼泽	78.57	82.50	81.93	85.00	85.54	88.75	78.82	83.75	79.76	83.75	90.12	91.25
水库/坑塘	86.49	80.00	87.18	85.00	87.50	87.50	81.58	77.50	80.00	80.00	92.31	90.00
水田	83.95	85.00	85.19	86.25	88.61	87.50	81.48	82.50	80.25	81.25	90.12	91.25
盐田	100.00	100.00	100.00	100.00	100.00	100.00	96.55	93.33	96.43	90.00	100.00	100.00
建筑用地	90.14	85.33	90.14	85.33	91.67	88.00	91.30	84.00	89.86	82.67	94.29	88.00
旱地	75.71	81.54	76.39	84.62	81.16	86.15	73.91	78.46	75.36	80.00	84.29	90.77
林地	82.69	78.18	85.19	83.64	90.38	85.45	80.77	76.36	77.19	80.00	92.45	90.74
总体精度/%	83.82		85.66		87.67		82.37		82.01		90.93	
Kappa系数	0.83		0.84		0.86		0.81		0.80		0.90	

注:表中PA、UA分别代表各类的生产者精度和用户精度。

由表6可知,以多时相光谱特征为基础的方案1总体精度为83.82%, Kappa系数为0.83。在多时相光谱特征的基础之上加入不同的特征变量对湿地信息提取会产生不同的影响。分别加入植被指数和水体指数、红边指数之后,方案2、方案3分类精度略有上升,总体精度分别达到85.66%和87.67%, Kappa系数分别为0.84、0.86。但是加入纹理特征的方案4和所有特征都参与分类的方案5总体精度和Kappa系数均略有下降,总体精度分别为82.37%、82.01%, Kappa系数分别为0.81、0.80。这说明,植被指数和水体指数、红边指数会提高湿地信息提取的精度,但是20 m分辨率或中分辨率的影像的纹理特征不一定会提高分类精度,反而在一定程度上会导致分类精度降低。由随机森林算法进行特征筛选之后的方案6总体精度和Kappa系数都大幅度提高,相较于前5个方案中提取效果最好的方案3而言,总体精度提高了3.26%, Kappa系数增长了0.40。

单个类型的生产者精度和用户精度可以表

明:植被生化组分相关的特征(如植被指数、红边指数)会提高湿地信息提取的精度。例如,方案2和方案3中的草本沼泽、灌木沼泽、水田和旱地在加入以上两种特征之后精度都有所提高,而且红边指数对湿地信息提取的贡献要高于一般的植被指数。纹理特征只对泥质海滩和建筑用地的生产者精度略有改善,这说明利用中分辨率影像提取的纹理特征只适用于纹理信息较为明显、单一的类型。所有特征参与分类时,可能会由于信息冗余不仅导致总体精度降低,而且也会影响单个类型的精度。采用方案6中的优选特征进行湿地信息提取时,只保留了对分类起关键作用的特征信息,避免了交叉冗余信息对湿地信息提取的干扰,使所有类型的生产者精度和用户精度都有较大程度的改善。

### 5.3 特征优选分析

对方案6的优选特征进一步分析发现,不同类别的特征参与分类时所占比重存在差异。为了更加直观地表示这种差异,将特征变量按重要性分

段计算比重, 图5(a)截取了重要性得分较高的前50个特征变量, 以每10个特征为一组计算不同类别特征组内所占的比重, 从前30个特征变量的比重可以看出, 多时相光谱特征、植被指数和水体指数所占比重逐渐呈现下降趋势, 红边指数所占比重一直稳步提升; 30个特征变量之后, 多时相光谱特征、植被指数和水体指数、红边指数所占

比重与之前呈现相反的趋势; 纹理特征所占比重几乎保持不变。分类效果最佳时, 33个特征中不同类别特征的个数如下(图5(b)), 多时相光谱特征有6个, 植被指数和水体指数特征有9个, 红边指数特征15个, 纹理特征仅占3个, 这也证实了红边指数特征在分类中的作用较为显著, 纹理特征则不明显。

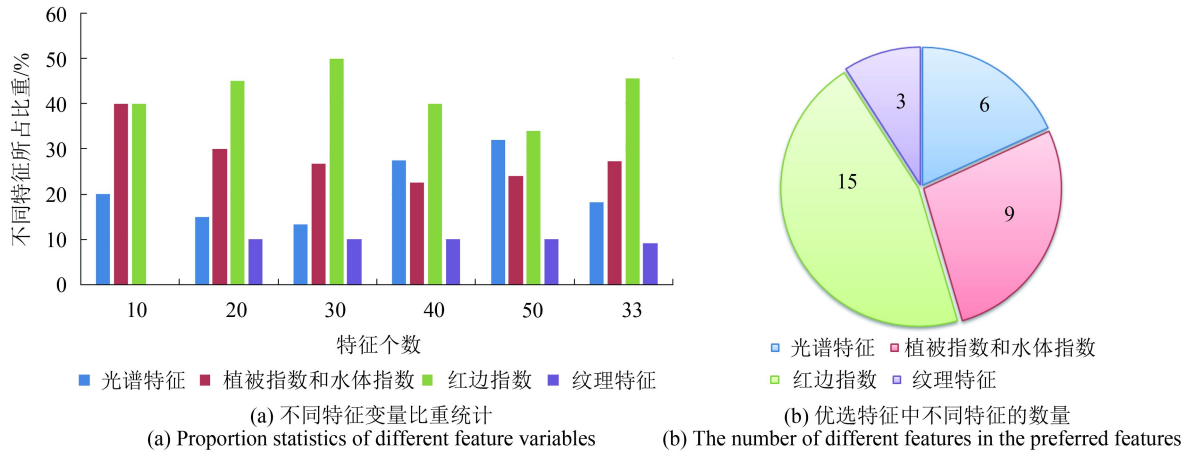


图5 特征变量统计图

Fig. 5 Statistical graph of feature variables

方案6的优选出的33个特征变量如表7所示, 从表中可以大致推断, 7月份是适宜分类提取的主要时相, 主要是因为研究区内人类活动频繁, 农作物种植范围较广, 黄河三角洲旱地作物主要以小麦、玉米为主, 水田主要为水稻、莲藕等, 7月

份冬小麦处于已经收获的时间, 玉米等作物刚刚栽种, 旱地植被生长稀疏, 覆盖度较小, 但是水稻、莲藕正处于抽穗、开花的时期, 生长较为茂盛, 所反映的植被光谱信息十分丰富, 两种地物类型在物候期上的差异也使其易于区分。

表7 优选特征分布表

Table 7 Distribution list of optimal feature

月份	优选特征
5月	NDVIre1_05、CIre_05、NDre2_05、GLCM_A_05
6月	NDVIre2_06、band8SWIR1_06、NDWI_06、band9SWIR2_06
7月	NDre2_07、CIre_07、band3R_07、NDVI_07、RVI_07、NDVIre1_07、NDre1_07、NDWI_07、MNDWI_07、band9SWIR2_07、GLCM_A_07、GLCM_E_07、band1B_07、NDVIre2_07、DVI_07
9月	MNDWI_09、band1B_09、NDVIre1_09、NDre2_09、NDVI_09
10月	MNDWI_10、NDre2_10、NDre1_10、CIre_10

## 6 结论

基于多时相Sentinel-2遥感数据, 提取多时相光谱特征、植被指数和水体指数、红边指数和纹理特征确定6种不同的实验方案, 结合随机森林算法进行特征优选和黄河三角洲湿地信息提取, 主要得到以下结论:

以多时相光谱数据为基础, 单独引入不同特征变量对湿地信息提取的精度影响不同。植被指数和水体指数、红边指数会对湿地信息提取的精度产生积极的影响, 相比较而言, 红边指数对分类精度的贡献更高; 单独加入纹理特征之后导致分类精度下降。

随机森林算法在进行特征优选时可以发挥很好的效果。对比方案3和方案6发现,采用随机森林算法进行特征优选之后的分类精度有了大幅度改善,而且参与分类的特征维数降低到33维,以上分析说明,基于随机森林算法的特征选择方法,能够保留地物最重要的特征信息,在降低数据维度的同时,分类精度仍能保持较高水平,进而缩短模型运算时间,提高分类效率。

本研究主要优势在于新数据源、新特征变量的引入以及对不同特征变量进行的评价分析,同时,本文也验证了随机森林算法对特征优选和信息提取的可行性。但是后续研究仍需要在以下两个方面进行补充:(1)分类所采用的随机森林算法是基于像素尺度进行的,不可避免地会存在“椒盐现象”,这种现象在旱地、水田分布区较为显著,主要是因为该区域的田埂上生长着草本植物容易造成混分,后续结合面向对象的方法,将两者进行比较研究,以期在一定程度上进一步提高湿地信息提取的精度。(2)重点探究了不同特征变量对湿地信息提取的影响,对于特征变量时相不同所带来的影响略有涉及,但缺乏深入思考,后续的工作重点将主要研究不同月份的影像数据对湿地信息提取的贡献,达到时相与特征统筹兼顾,以实现湿地信息自动化、智能化、精准化提取的目的。

## 参考文献(References)

- Abuelgasim A A, Gopal S, Irons J R and Strahler A H. 1996. Classification of ASAS multiangle and multispectral measurements using artificial neural networks. *Remote Sensing of Environment*, 57(2): 79–87 [DOI: [10.1016/0034-4257\(95\)00197-2](https://doi.org/10.1016/0034-4257(95)00197-2)]
- Baker C, Lawrence R, Montagne C and Patten D. 2006. Mapping wetlands and riparian areas using Landsat ETM+ imagery and decision-tree-based models. *Wetlands*, 26(2): 465–474 [DOI: [10.1672/0277-5212\(2006\)26\[465:MWARAU\]2.0.CO;2](https://doi.org/10.1672/0277-5212(2006)26[465:MWARAU]2.0.CO;2)]
- Bao Y H and Ren J B. 2011. Wetland landscape classification based on the BP neural network in DaLinor Lake area. *Procedia Environmental Sciences*, 10: 2360–2366 [DOI: [10.1016/j.proenv.2011.09.368](https://doi.org/10.1016/j.proenv.2011.09.368)]
- Barnes E M, Clarke T R, Richards S E, Colaizzi P D, Haberland J, Kostrzewski M, Waller P, Choi C, Riley E, Thompson T, Lascano R J, Li H and Moran M S. 2000. Coincident detection of crop water stress, nitrogen status and canopy density using ground-based multispectral data//Proceedings of the 5th International Conference on Precision Agriculture. South Segoe Road, Madison: American Society of Agronomy.
- Breiman L. 2001. Random forest. *Machine Learning*, 45(1): 5–32 [DOI: [10.1023/A:1010933404324](https://doi.org/10.1023/A:1010933404324)]
- Davranche A, Lefebvre G and Poulin B. 2010. Wetland monitoring using classification trees and SPOT-5 seasonal time series. *Remote Sensing of Environment*, 114(3): 552–562 [DOI: [10.1016/j.rse.2009.10.009](https://doi.org/10.1016/j.rse.2009.10.009)]
- Delegido J, Verrelst J, Alonso L and Moreno J. 2011. Evaluation of sentinel-2 red-edge bands for empirical estimation of green LAI and chlorophyll content. *Sensors*, 11(7): 7063–7081 [DOI: [10.3390/s110707063](https://doi.org/10.3390/s110707063)]
- Drusch M, Del Bello U, Carlier S, Colin O, Fernandez V, Gascon F, Hoersch B, Isola C, Laberinti P, Martimort P, Meygret A, Spoto F, Sy O, Marchese F and Bargellini P. 2012. Sentinel-2: ESA's optical high-resolution mission for GMES operational services. *Remote Sensing of Environment*, 120: 25–36 [DOI: [10.1016/j.rse.2011.11.026](https://doi.org/10.1016/j.rse.2011.11.026)]
- 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](https://doi.org/10.1016/j.jag.2016.03.005)]
- Genuer R, Poggi J M and Tuleau-Malot C. 2010. Variable selection using random forests. *Pattern Recognition Letters*, 31(14): 2225–2236 [DOI: [10.1016/j.patrec.2010.03.014](https://doi.org/10.1016/j.patrec.2010.03.014)]
- Gitelson A A, Gritz Y and Merzlyak M N. 2003. Relationships between leaf chlorophyll content and spectral reflectance and algorithms for non-destructive chlorophyll assessment in higher plant leaves. *Journal of Plant Physiology*, 160(3): 271–282 [DOI: [10.1078/0176-1617-00887](https://doi.org/10.1078/0176-1617-00887)]
- Gitelson A and Merzlyak M N. 1994. Spectral reflectance changes associated with autumn senescence of *Aesculus hippocastanum* L. and *Acer platanoides* L. leaves. Spectral features and relation to chlorophyll estimation. *Journal of Plant Physiology*, 143(3): 286–292 [DOI: [10.1016/S0176-1617\(11\)81633-0](https://doi.org/10.1016/S0176-1617(11)81633-0)]
- Gong N, Niu Z G, Qi W and Zhaug 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](https://doi.org/10.11834/jrs.20164210)]
- Hill M J. 2013. Vegetation index suites as indicators of vegetation state in grassland and savanna: an analysis with simulated SENTINEL 2 data for a North American transect. *Remote Sensing of Environment*, 137: 94–111 [DOI: [10.1016/j.rse.2013.06.004](https://doi.org/10.1016/j.rse.2013.06.004)]
- Hou Q Q, Wang F and Yan L. 2013. Extraction of color image texture feature based on gray-level co-occurrence matrix. *Remote Sensing for Land and Resources*, 25(4): 26–32 (侯群群, 王飞, 严丽. 2013. 基于灰度共生矩阵的彩色遥感图像纹理特征提取. *国土*

- 资源遥感, 25(4): 26–32 [DOI: [10.6046/gtzyyg.2013.04.05](https://doi.org/10.6046/gtzyyg.2013.04.05)]
- Hu Y F, Deng J J, Kuang X H, Wang P, He S and Xiong L. 2011. Study on land use classification of high resolution remote sensing image based on texture feature. *Geography and Geo-Information Science*, 27(5): 42–45, 68 (胡玉福, 邓良基, 匡先辉, 王鹏, 何莎, 熊玲. 2011. 基于纹理特征的高分辨率遥感图像土地利用分类研究. *地理与地理信息科学*, 27(5): 42–45, 68)
- Huang J L. 1999. The area change and succession of Dongtinghu wetland. *Geographical Research*, 18(3): 297–304 (黄进良. 1999. 洞庭湖湿地的面积变化与演替. *地理研究*, 18(3): 297–304) [DOI: [10.3321/j.issn:1000-0585.1999.03.011](https://doi.org/10.3321/j.issn:1000-0585.1999.03.011)]
- Huang K, Shida Z S, Deng C L and Du K. 2012. Response Relationship Between Ecosystem and Runoff into the Sea in the Yellow River Estuary. *Yellow River*, 34(11): 66–69 (黄葵, 石达扎实, 邓春蕾, 杜凯. 2012. 黄河口近海生态系统与入海径流的响应关系. *人民黄河*, 34(11): 66–69) [DOI: [10.3969/j.issn.1000-1379.2012.11.023](https://doi.org/10.3969/j.issn.1000-1379.2012.11.023)]
- Iverson L R, Prasad A M, Matthews S N and Peters M. 2008. Estimating potential habitat for 134 eastern US tree species under six climate scenarios. *Forest Ecology and Management*, 254(3): 390–406 [DOI: [10.1016/j.foreco.2007.07.023](https://doi.org/10.1016/j.foreco.2007.07.023)]
- Khatami R, Mountrakis G and Stehman S V. 2016. A meta-analysis of remote sensing research on supervised pixel-based land-cover image classification processes: general guidelines for practitioners and future research. *Remote Sensing of Environment*, 177: 89–100 [DOI: [10.1016/j.res.2016.02.028](https://doi.org/10.1016/j.res.2016.02.028)]
- Korhonen L, Hadi, Packalen P and Rautiainen M. 2017. Comparison of Sentinel-2 and Landsat 8 in the estimation of boreal forest canopy cover and leaf area index. *Remote Sensing of Environment*, 195: 259–274 [DOI: [10.1016/j.rse.2017.03.021](https://doi.org/10.1016/j.rse.2017.03.021)]
- Liu H Y, Lv X G and Zhang S K. 2004. Landscape biodiversity of wetlands and their changes in 50 years in watersheds of the Sanjiang Plain. *Acta Ecologica Sinica*, 24(7): 1472–1479 (刘红玉, 吕宪国, 张世奎. 2004. 三江平原流域湿地景观多样性及其50年变化研究. *生态学报*, 24(7): 1472–1479) [DOI: [10.3321/j.issn:1000-0933.2004.07.023](https://doi.org/10.3321/j.issn:1000-0933.2004.07.023)]
- Liu L, Han M, Liu Y B and Pan B. 2017. Spatial distribution of wetland vegetation biomass and its influencing factors in the Yellow River Delta Nature Reserve. *Acta Ecologica Sinica*, 37(13): 4346–4355 (刘莉, 韩美, 刘玉斌, 潘彬. 2017. 黄河三角洲自然保护区湿地植被生物量空间分布及其影响因素. *生态学报*, 37(13): 4346–4355) [DOI: [10.5846/stxb201508241763](https://doi.org/10.5846/stxb201508241763)]
- Liu M M, Liu Y L, Sun G Q and Peng L. 2014. SVM land cover classification based on spectral and textural features using stratified samples. *Remote Sensing Technology and Application*, 29(2): 315–323 (刘萌萌, 刘亚岚, 孙国庆, 彭立. 2014. 结合纹理特征的SVM样本分层土地覆盖分类. *遥感技术与应用*, 29(2): 315–323) [DOI: [10.11873/j.issn.1004-0323.2014.2.0315](https://doi.org/10.11873/j.issn.1004-0323.2014.2.0315)]
- Liu R H, Liang S C, Zhao H Y, Qi G C, Li L X, Jiang Y and Niu Z G. 2017. Progress of Chinese coastal wetland based on remote sensing. *Remote Sensing Technology and Application*, 32(6): 998–1011 (刘润红, 梁士楚, 赵红艳, 漆光超, 李丽香, 姜勇, 牛振国. 2017. 中国滨海湿地遥感研究进展. *遥感技术与应用*, 32(6): 998–1011) [DOI: [10.11873/j.issn.1004-0323.2017.6.0998](https://doi.org/10.11873/j.issn.1004-0323.2017.6.0998)]
- Liu S, Jiang Q G, Ma Y, Xiao Y, Li Y H and Cui C. 2017. Object-oriented wetland classification based on hybrid feature selection method combining with relief F, multi-objective genetic algorithm and random forest. *Transactions of the Chinese Society for Agricultural Machinery*, 48(1): 119–127 (刘舒, 姜琦刚, 马玥, 肖艳, 李远华, 崔璨. 2017. 基于多目标遗传随机森林特征选择的面向对象湿地分类. *农业机械学报*, 48(1): 119–127) [DOI: [10.6041/j.issn.1000-1298.2017.01.016](https://doi.org/10.6041/j.issn.1000-1298.2017.01.016)]
- Mou X J, Liu X T, Yan B X and Cui B S. 2015. Classification system of coastal wetlands in China. *Wetland Science*, 13(1): 19–26 (牟晓杰, 刘兴土, 阎百兴, 崔保山. 2015. 中国滨海湿地分类系统. *湿地科学*, 13(1): 19–26) [DOI: [10.13248/j.cnki.wetlandsci.2015.01.004](https://doi.org/10.13248/j.cnki.wetlandsci.2015.01.004)]
- Na X D, Zhang S Q, Kong B, Yu H, Li X F and Liu C Y. 2008. The extraction of freshwater marsh wetland information based on decision tree algorithm—A case study in the Northeast of the Sanjiang Plain. *Remote Sensing Technology and Application*, 23(4): 365–372 (那晓东, 张树清, 孔博, 于欢, 李晓峰, 刘春悦. 2008. 基于决策树方法的淡水沼泽湿地信息提取——以三江平原东北部为例. *遥感技术与应用*, 23(4): 365–372) [DOI: [10.11873/j.issn.1004-0323.2008.4.365](https://doi.org/10.11873/j.issn.1004-0323.2008.4.365)]
- Na X D, Zhang S Q, Zhang H Q, Li X F, Yu H and Liu C Y. 2009. Integrating TM and ancillary geographical data with classification trees for land cover classification of marsh area. *Chinese Geographical Science*, 19(2): 177–185 [DOI: [10.1007/s11769-009-0177-y](https://doi.org/10.1007/s11769-009-0177-y)]
- Özesmi S L. 2000. *Satellite Remote Sensing of Wetlands and a Comparison of Classification Techniques*. Twin Cities, MN: University of Minnesota
- Pu R L and Landry S. 2012. A comparative analysis of high spatial resolution IKONOS and WorldView-2 imagery for mapping urban tree species. *Remote Sensing of Environment*, 124: 516–533 [DOI: [10.1016/j.rse.2012.06.011](https://doi.org/10.1016/j.rse.2012.06.011)]
- Shang J L, Liu J G, Ma B L, Zhao T, Jiao X F, Geng X Y, Huffman T, Kovacs J M and Walters D. 2015. Mapping spatial variability of crop growth conditions using RapidEye data in Northern Ontario, Canada. *Remote Sensing of Environment*, 168: 113–125 [DOI: [10.1016/j.rse.2015.06.024](https://doi.org/10.1016/j.rse.2015.06.024)]
- Shoko C and Mutanga O. 2017. Examining the strength of the newly-launched Sentinel 2 MSI sensor in detecting and discriminating subtle differences between C3 and C4 grass species. *ISPRS Journal of Photogrammetry and Remote Sensing*, 129: 32–40 [DOI: [10.1016/j.isprsjprs.2017.05.011](https://doi.org/10.1016/j.isprsjprs.2017.05.011)]

[10.1016/j.isprsjprs.2017.04.016](https://doi.org/10.1016/j.isprsjprs.2017.04.016)

- Song H L. 2015. Influence of reclamation activities on ecosystem type and carbon sink function of the coastal wetland in the Yellow River estuary. Changchun: Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences. (宋红丽. 2015. 围填海活动对黄河三角洲滨海湿地生态系统类型变化和碳汇功能的影响. 长春: 中国科学院研究生院(东北地理与农业生态研究所))
- Su W, Zhang M Z, Jiang K P, Zhu D H, Huang J X and Wang P X. 2018. Atmospheric correction method for sentinel-2 satellite imagery. *Acta Optica Sinica*, 38(1): 128001 (苏伟, 张明政, 蒋坤萍, 朱德海, 黄健熙, 王鹏新. 2018. Sentinel-2卫星影像的大气校正方法. *光学学报*, 38(1): 128001) [DOI: [10.3788/aos201838.0128001](https://doi.org/10.3788/aos201838.0128001)]
- Wang L, Dronova I, Gong P, Yang W B, Li Y R and Liu Q. 2012. A new time series vegetation-water index of phenological-hydrological trait across species and functional types for Poyang Lake wetland ecosystem. *Remote Sensing of Environment*, 125: 49-63 [DOI: [10.1016/j.rse.2012.07.003](https://doi.org/10.1016/j.rse.2012.07.003)]
- Wang X H, Li Y Z, Meng H, Dong H F, Guo Y and Tong S Z. 2015. Distribution pattern of plant community in new-born coastal wetland in the Yellow River Delta. *Scientia Geographica Sinica*, 35(8): 1021-1026 (王雪宏, 栗云召, 孟焕, 董红芳, 郭岳, 佟守正. 2015. 黄河三角洲新生湿地植物群落分布格局. *地理科学*, 35(8): 1021-1026) [DOI: [10.13249/j.cnki.sgs.2015.08.012](https://doi.org/10.13249/j.cnki.sgs.2015.08.012)]
- Wen Q K, Zhang Z X, Xu J Y, Zuo L J, Wang X, Liu B, Zhao X L and Yi L. 2011. Spatial and temporal change of wetlands in Bohai rim during 2000-2008: An analysis based on satellite images. *Journal of Remote Sensing*, 15(1): 183-200 (温庆可, 张增祥, 徐进勇, 左丽君, 汪潇, 刘斌, 赵晓丽, 易玲. 2011. 环渤海滨海湿地时空格局变化遥感监测与分析. *遥感学报*, 15(1): 183-200) [DOI: [10.11834/jrs.20110115](https://doi.org/10.11834/jrs.20110115)]
- Wu J H, Zhang S, Jiang Y, Kang M Y and Qiu Y. 2004. *Plant Geography*. 4th ed. Beijing: Higher Education Press (武吉华, 张绅, 江源, 康慕谊, 邱扬. 2004. *植物地理学*. 4版. 北京: 高等教育出版社)
- Xiao J C, Ou W X and Fu H Y. 2013. Land cover classification of Yancheng Coastal Natural Wetlands based on BP neural network and ETM+ remote sensing data. *Acta Ecologica Sinica*, 33(23): 7496-7504 (肖锦成, 欧维新, 符海月. 2013. 基于BP神经网络与ETM+遥感数据的盐城滨海自然湿地植被分类. *生态学报*, 33(23): 7496-7504) [DOI: [10.5846/stxb201206300916](https://doi.org/10.5846/stxb201206300916)]
- Zhang C, Zang S Y, Jin Z and Zhang Y H. 2011. Remote sensing classification for Zhalong Wetlands based on support vector machine. *Wetland Science*, 9(3): 263-269 (张策, 臧淑英, 金竺, 张玉红. 2011. 基于支持向量机的扎龙湿地遥感分类研究. *湿地科学*, 9(3): 263-269) [DOI: [10.13248/j.cnki.wetlandsci.2011.03.004](https://doi.org/10.13248/j.cnki.wetlandsci.2011.03.004)]
- Zhang M, Zeng Y N and Zhu Y S. 2017. Wetland mapping of Dongting Lake Basin based on time-series MODIS data and object-oriented method. *Journal of Remote Sensing*, 21(3): 479-492 (张猛, 曾永年, 朱永森. 2017. 面向对象方法的时间序列MODIS数据湿地信息提取——以洞庭湖流域为例. *遥感学报*, 21(3): 479-492) [DOI: [10.11834/jrs.20176129](https://doi.org/10.11834/jrs.20176129)]
- Zhang S W, Yan F Q, Yu L X, Bu K, Yang J C and Chang L P. 2013. Application of remote sensing technology to wetland research. *Scientia Geographica Sinica*, 33(11): 1406-1412 (张树文, 颜凤芹, 于灵雪, 卜坤, 杨久春, 常丽萍. 2013. 湿地遥感研究进展. *地理科学*, 33(11): 1406-1412) [DOI: [10.13249/j.cnki.sgs.2013.011.1406](https://doi.org/10.13249/j.cnki.sgs.2013.011.1406)]
- Zhang X L, Ye S Y, Yin P and Chen D J. 2009. Characters and successions of natural wetland vegetation in Yellow River Delta. *Ecology and Environmental Sciences*, 18(1): 292-298 (张绪良, 叶思源, 印萍, 陈东景. 2009. 黄河三角洲自然湿地植被的特征及演化. *生态环境学报*, 18(1): 292-298) [DOI: [10.16258/j.cnki.1674-5906.2009.01.020](https://doi.org/10.16258/j.cnki.1674-5906.2009.01.020)]
- Zheng S D, Zheng J H, Shi M H, Guo B L, Sen B T, Sun Z Q, Jia X G and Li X J. 2014. Classification of cultivated Chinese medicinal plants based on fractal theory and gray level co-occurrence matrix textures. *Journal of Remote Sensing*, 18(4): 868-886 (郑淑丹, 郑江华, 石明辉, 郭宝林, 森巴提, 孙志群, 贾晓光, 李晓瑾. 2014. 基于分形和灰度共生矩阵纹理特征的种植型药用植物遥感分类. *遥感学报*, 18(4): 868-886) [DOI: [10.11834/jrs.20143282](https://doi.org/10.11834/jrs.20143282)]
- Zheng Y, Wu B F and Zhang M. 2017. Estimating the above ground biomass of winter wheat using the Sentinel-2 data. *Journal of Remote Sensing*, 21(2): 318-328 (郑阳, 吴炳方, 张森. 2017. Sentinel-2数据的冬小麦地上干生物量估算及评价. *遥感学报*, 21(2): 318-328) [DOI: [10.11834/jrs.20176269](https://doi.org/10.11834/jrs.20176269)]

# Wetland mapping of Yellow River Delta wetlands based on multi-feature optimization of Sentinel-2 images

ZHANG Lei, GONG Zhaoning, WANG Qiwei, JIN Diandian, WANG Xing

1. *College of Resources Environment & Tourism, Capital Normal University, Beijing 100048, China;*
2. *Key Laboratory of 3D Information Acquisition and Application of Ministry, Beijing 100048, China;*
3. *Key Laboratory of Resources Environment and GIS of Beijing Municipal, Beijing 100048, China*

**Abstract:** Estuary wetland is a special wetland type, and the extraction of estuary wetland information plays an important role in wetland conservation and scientific research. In this study, Yellow River Delta wetlands, as a typical estuary wetland in the north part of China, are considered the study area. The random forest method, which has evident advantages in feature selection and classification, was chosen to extract wetland information from the study area. First, five different characteristic variables, namely, spectral features, vegetation index, water index, red edge index, and texture features, were generated based on Sentinel-2 data with rich multi-temporal and spectral information. Then, six different classification schemes were constructed based on the preceding characteristic information. Finally, random forest classifier was used to extract the wetland information of the Yellow River Delta and verify the extraction accuracy of different results. The purpose is to select the best plan to improve the effect of wetland information extraction. Results are as follows: (1) The effective use of multiple feature variables is the key to improving the extraction of wetland information. The contribution of different characteristics to the wetland information extraction is described as follows: the red edge index > vegetation index and water index > spectral feature > texture feature. (2) The preferred features based on the random forest algorithm are crucial to extraction accuracy, with an overall accuracy of up to 90.93%, and Kappa coefficient of 0.90. This result shows that the random forest algorithm can effectively process feature selection. In feature variable data mining, the accuracy of the wetland information extraction can be guaranteed, and the operation efficiency can be improved. This study also provides a new idea, method, and technology for the selection of data sources and feature and method selections for wetland information extraction.

**Key words:** estuarine wetland, information extraction, Sentinel-2, random forest, feature selection, red edge index, multi-temporal data

**Supported by** National Key R&D Program of China (No. 2017YFC0505903)