

# 基于随机森林特征变量优化的湿地植物分类与密度反演

刘曙光<sup>1,2</sup>, 董行<sup>1</sup>, 娄厦<sup>1,2</sup>, DORZHIEVNA Radnaeva Larisa<sup>3</sup>,  
NIKITINA Elena<sup>3</sup>

(1. 同济大学土木工程学院, 上海200092; 2. 同济大学长江水环境教育部重点实验室, 上海200092; 3. 俄罗斯科学院西伯利亚分院 贝加尔湖自然管理研究所, 乌兰乌德 670047, 俄罗斯联邦布里亚特共和国)

**摘要:** 以长江口滨海湿地为研究区域, 采用随机森林算法对滨海湿地植被进行分类。在提取 Landsat-8 OLI 影像植被指数和水体指数的基础上, 提出利用植被指数季节差值对模型进行特征变量优化, 分析了长江口滨海湿地植物群落分布的空间特征。以所占面积最大的互花米草(入侵物种)为例, 采用多元线性回归模型结合实地测量数据, 估算了秋季的互花米草植物密度的空间特征。提出的多时相遥感数据结合随机森林特征变量优化方法, 可以较为便捷地提取长江口湿地3种优势物种的空间分布特征, 与最大似然法相比, 分类精度有较大提高, 总体分类精度由78.35%提高至87.55%, Kappa系数由0.72提高至0.84。该方法适用于存在“异物同谱”问题的湿地植物群落研究。

**关键词:** 长江口湿地; 湿地植被分类; 植物密度; 随机森林模型; 特征变量优化

中图分类号: X87

文献标志码: A

## Classification and Density Inversion of Wetland Vegetation Based on the Feature Variables Optimization of Random Forest Model

LIU Shuguang<sup>1,2</sup>, DONG Hang<sup>1</sup>, LOU Sha<sup>1,2</sup>,  
DORZHIEVNA Radnaeva Larisa<sup>3</sup>, NIKITINA Elena<sup>3</sup>

(1. College of Civil Engineering, Tongji University, Shanghai 200092, China; 2. Key Laboratory of Yangtze Water Environment of the Ministry of Education, Tongji University, Shanghai 200092, China; 3. Laboratory of Chemistry of Natural Systems, Baikal Institute of Nature Management of Siberian Branch of the Russian Academy of Sciences, Ulan-Ude, 670047, Republic of Buryatia, Russian)

**Abstract:** Taking the coastal wetland of the Yangtze River Estuary as the research area, the random forest model was used to classify the vegetation of wetland. Besides the vegetation index and water index extracted from Landsat-8 OLI image, the seasonal difference value of vegetation index based on plant phenology characteristics was proposed in this paper as the optimization of feature variables to analyze the spatial characteristics of vegetation distribution in the coastal wetland of the Yangtze River Estuary. Based on the vegetation classifications, the multiple linear regression model combined with the field data was used to estimate the vegetation density of the *Spartina alterniflora*, which was an invasive species occupying the largest area. It is indicated that the proposed multi-temporal data combined with the optimization of feature variable of random forest model can be used to conveniently analyze the spatially vegetation distributions in wetland. Compared with the maximum likelihood classification method, the proposed method in this paper greatly enhances the classification accuracy with an overall accuracy and Kappa coefficient of the classification results increasing from 78.35% and 0.72 to 86.02% and 0.82, respectively. The proposed method is proved to be applicable for solving the problem of “same spectrum for different surface features” in the study of wetland plant community distributions.

**Key words:** Yangtze River Estuary wetland; vegetation classification; vegetation density; random forest model; feature variables optimization

收稿日期: 2020-11-30

基金项目: 国家自然科学基金(42072281); 上海市科技创新行动计划(20230742500); 上海市土木工程高峰学科建设项目(2019010207)

第一作者: 刘曙光(1962—), 男, 教授, 博士生导师, 工学博士, 主要研究方向为海洋科学、资源与环境、水利工程、岩土工程。E-mail: liushglu@tongji.edu.cn

通信作者: 娄厦(1986—), 女, 副教授, 博士生导师, 工学博士, 主要研究方向为近岸水动力水环境、滨海湿地污染物输运。E-mail: lousha@tongji.edu.cn



论文  
拓展  
介绍

湿地是地球生态系统的重要组成部分,在净化环境、保护生物多样性、蓄滞洪水以及调节气候等方面发挥着不可替代的作用<sup>[1]</sup>。植物群落分布是湿地植物的主要特征之一,在维护湿地生态稳定、保护湿地环境等方面均具有重要的意义。20世纪90年代,大部分学者采用实地调查的方法分析湿地植物群落特征的分布规律,如张利权等<sup>[2]</sup>通过实地调查,分析了上海市南汇海堤外海三棱藨草的分布格局。但实地调查往往受时间、天气、地形限制,且需要耗费大量的时间和财力。

与实地调查相比,遥感技术具有经济、便捷等优势。在利用遥感数据进行植物及地物分类的研究中,监督分类方法是较为常用的方法,其中最大似然法、支持向量机、决策树算法以及基于决策树的随机森林算法均属于应用较为广泛的监督分类方法。孙楠<sup>[3]</sup>利用 Landsat-8 影像对比了最大似然法、支持向量机、CART(classification and regression tree)等方法对长江口湿地的分类结果,发现 CART 算法精度最优。但 CART 算法存在过度拟合和非全局的局部最优解等问题<sup>[4]</sup>。基于上述问题,Breiman<sup>[5]</sup>提出一种基于 CART 算法的机器学习算法,即随机森林算法,该算法通过集成学习有效解决了上述问题。Amani 等<sup>[6]</sup>利用多源多时相遥感影像对加拿大 5 个试点的湿地进行分类,并对比了最大似然法、支持向量机、CART 算法和随机森林算法的分类结果,指出随机森林算法的性能最好。

已有研究表明,随机森林算法与其他方法相比,在湿地植物分类中具有相对较高的精度。然而由于地物光谱反射率相近或地物反射能量衰减导致光谱差异减弱,湿地植物存在“异物同谱”的问题<sup>[7]</sup>。因此,本研究选取多时相遥感影像及特征变量,并根据当地湿地植物物候特性进行特征优化,用以提高“异物同谱”像元的识别精度,分析滨海湿地植物群落空间分布特征。最后,通过多元线性回归模型反演外来入侵植物互花米草的植物密度,实现对互花米草的精确检测、识别与密度反演。

## 1 研究区域与数据

### 1.1 研究区域

长江三角洲位于 30°45′30″—32°11′56″N, 121°04′20″—122°28′40″E,北至江苏省启东市,南至杭州湾北部。长江口滨海湿地面积最大的 3 处湿地为崇明东滩湿地、九段沙湿地以及南汇促淤区湿地,如图 1 所示。

### 1.2 数据及预处理

Landsat-8 OLI 影像从美国地质调查局官网获取,重访周期 16 d,影像包含 9 个波段(海岸、蓝、绿、红、近红外、短波红外 1、短波红外 2、全色、卷云),空间分辨率 30 m,其中全色波段空间分辨率 15 m。本文选用 2018 年 12 月,2019 年 1 月、4 月、5 月、7 月共 5 景低云量退潮期影像,利用 ENVI 5.3 平台进行辐射定标和大气校正。遥感影像的处理及实地测量数据的收集均在 WGS-84 坐标系下进行。

### 1.3 训练样本及验证样本

样本的选取与分类精度直接相关。由于 Landsat-8 OLI 影像中生长期的芦苇与互花米草像元各波段反射率相近,本文通过实地调查结果结合当地湿地植物物候特性建立判读标志,通过目视解译 Google Earth 高空间分辨率影像进行训练样本和验证样本的选取。

湿地植物信息实地调查于 2019 年 5 月—10 月在上海南汇促淤工程附近开展。沿岸选取 18 个测点进行测量,测点位置见图 1a 和表 1。采用样方法测量,选取 100 cm×100 cm 样方计量植株数量,并用手持 GPS 记录样方坐标。将研究区域分为芦苇、互花米草、海三棱藨草、光滩和水体 5 种地物类型,生成训练样本时,选择具有代表性的纯净像元或实地考察结果所对应像元。生成验证样本时,利用 ArcMap 生成随机点 522 个,根据 Google Earth 高空间分辨率影像(按照 Google Earth 使用条款仅用其影像辨识边界信息)、RGB 假彩色组合和植物物候特性的夏、冬季影像对比判定样本点地物类型。根据前人研究<sup>[8]</sup>,红波段(B4)、近红外波段(B5)、短波红外 1(B6)3 个波段 RGB 假彩色组合能够体现的植物类型较丰富,可用于植物分类研究,本文获取了研究区域夏季冬季假彩色影像,可以较为清晰地辨识出冬季还未枯萎的互花米草植株范围。每种类型验证样本点均在 60~150 个之间。

## 2 研究方法

随机森林算法通过 Bagging 算法将多个 CART 决策树集成,根据所有决策树投票获得最终结果。基于随机森林模型的分类流程见图 2。该过程中设定决策树数量( $N$ )和生成决策树节点分裂时输入特征变量数量( $m$ )两个参数。研究表明<sup>[9]</sup>,决策树参数设置对分类精度影响不敏感,故本文选取参数默认值构建分类模型。

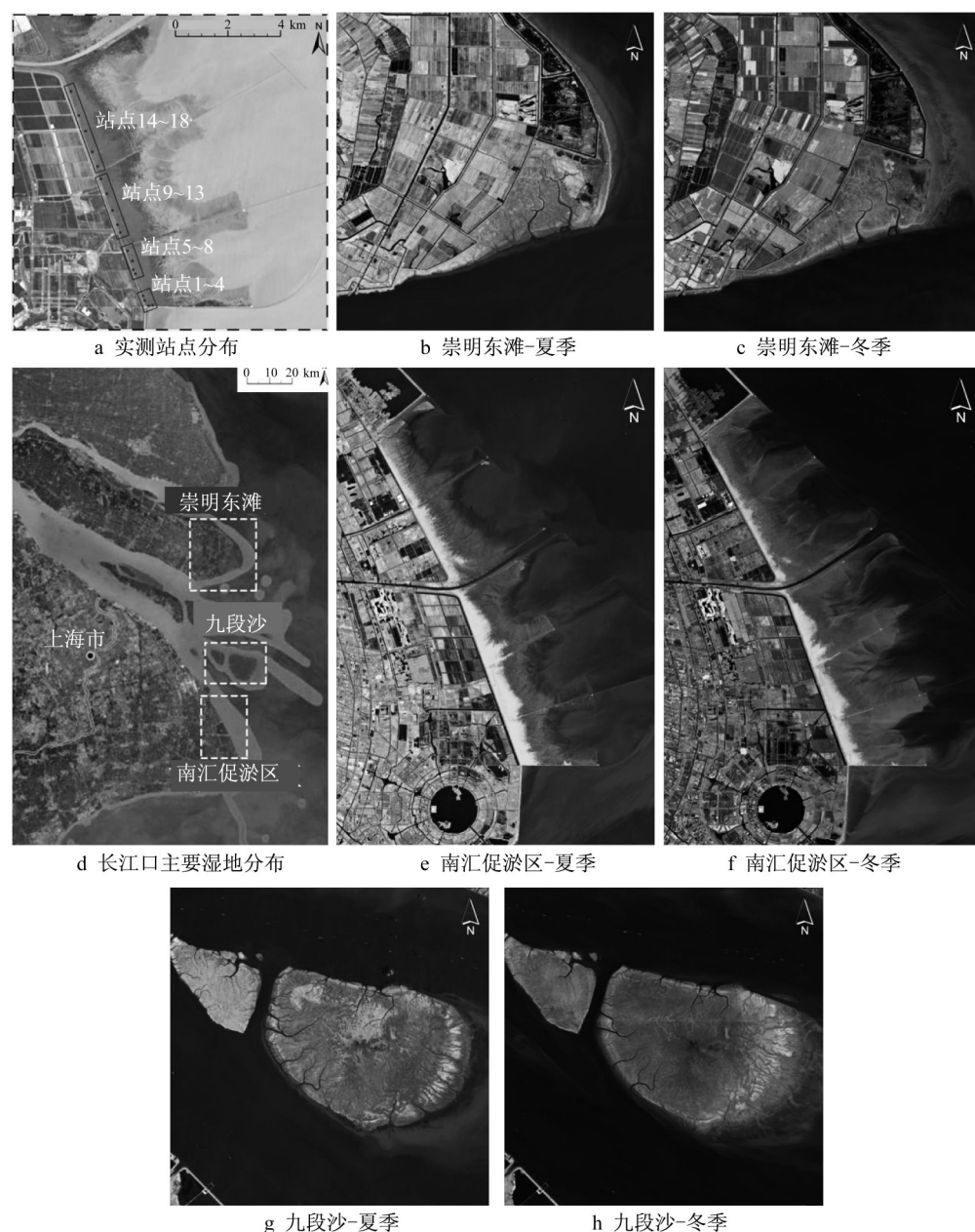


图1 研究区域

Fig. 1 Study regions

光谱特征及植被指数等指数是区分不同地物的最主要特征,本文利用ENVI 5.3中大气校正、波段计算及图层叠加工具,提取了遥感影像7个波段反射率及7种指数作为特征变量。波段反射率分别为:海岸波段(B1)、蓝波段(B2)、绿波段(B3)、红波段(B4)、近红外波段(B5)、短波红外1(B6)、短波红外2(B7)。指数特征包括:归一化差异植被指数(NDVI)<sup>[10]</sup>、比值植被指数(RVI)<sup>[10]</sup>、差值植被指数(DVI)<sup>[10]</sup>、改进的归一化差异水体指数(MNDWI)<sup>[11]</sup>以及缨帽变换中的3个分量(亮度BI、绿度GVI、湿度WI)<sup>[12]</sup>。

长江口湿地优势植物为芦苇、海三棱藨草以及互花米草3种,其中芦苇与互花米草的光谱特性相近,“同谱异物”现象明显,影响了两种植物的分类精度。芦苇与互花米草均在4月的返青期出现新旧植株交杂。7月芦苇种群生物量等各项生物指标达到峰值,而互花米草群落于9月达到最高峰。两种植物的各项生物指标均在冬季降至最低,芦苇群落地上部分于11月全部枯萎,而互花米草在冬季冠层仍呈现黄绿色<sup>[13]</sup>。通过实地测量记录坐标及部分目视解译,选取研究区域各类地物纯净像元,计算各地物的像元光谱反射率均值,获得两者光谱曲线,如图3



表1 测点位置及互花米草平均密度

Tab. 1 Locations of observation sites and average density of spartina alterniflora

站点编号	测点坐标/(°)	平均密度/(株·m <sup>-2</sup> )	站点编号	测点坐标/(°)	平均密度/(株·m <sup>-2</sup> )
1	30.9231N,121.9727E	144	10	30.9490N,121.9616E	175
2	30.9247N,121.9721E	177	11	30.9539N,121.9596E	178
3	30.9213N,121.9744E	182	12	30.9595N,121.9573E	145
4	30.9215N,121.9759E	186	13	30.9646N,121.9548E	179
5	30.9336N,121.9680E	191	14	30.9689N,121.9535E	162
6	30.9322N,121.9689E	174	15	30.9739,N121.9513E	151
7	30.9366N,121.9678E	165	16	30.9809N,121.9490E	148
8	30.9401N,121.9658E	172	17	30.9862N,121.9458E	160
9	30.9463N,121.9627E	185	18	30.9944N,121.9426E	179

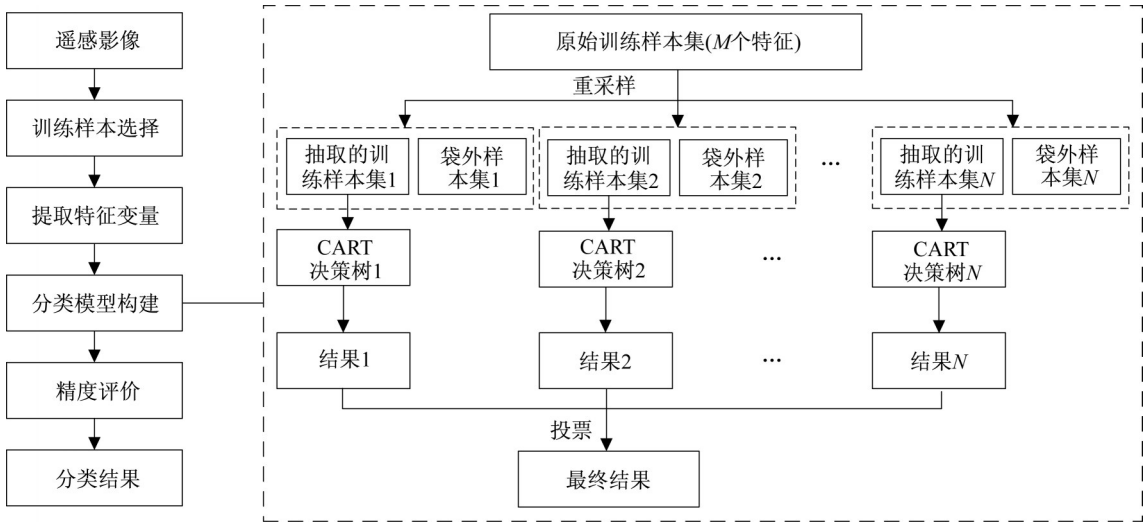


图2 基于随机森林模型的分类流程图

Fig. 2 Classification process based on random forest model

所示。通过对比夏、冬季两景遥感影像的假彩色组合(图1b、1c、1e、1f、1g、1h),冬季互花米草在假彩色组合影像中呈绿色,芦苇与海三棱藨草已枯萎,与裸地颜色相近,体现出两者的物候特性差异性,有助于目视解译植被的种类。

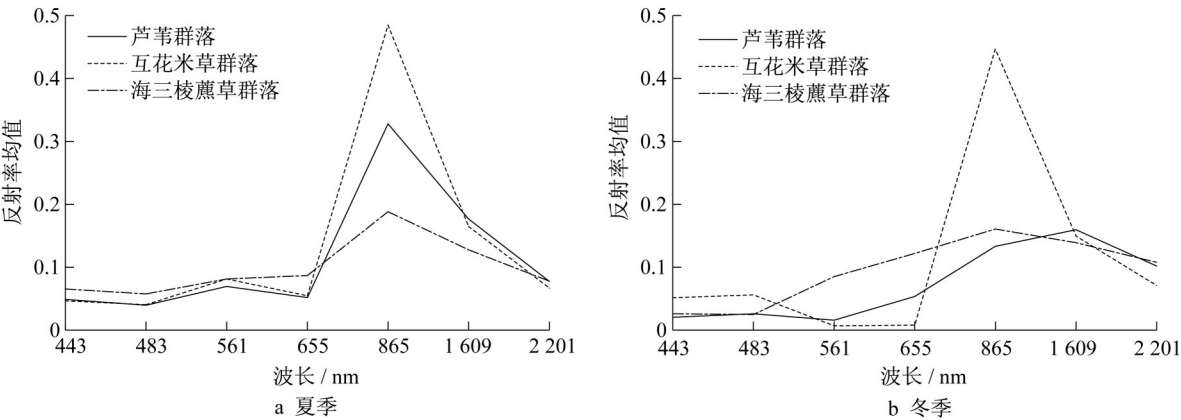


图3 长江口湿地植被光谱曲线

Fig. 3 Spectral curves of wetland vegetation in Yangtze River estuary

基于植物的物候特性差异性,本文对随机森林模型中的特征变量进行了优化。将夏季(7月)影像的植被指数与冬季(12月)影像的植被指数的差值定义为植被指数季节差值( $V_{SDI}$ ),将其作为为特征变量之一,以提高分类精度。相关计算公式如下:

$$V_{SDI(1)} = V_{NDVI(s)} - V_{NDVI(w)} \quad (1)$$

$$V_{SDI(2)} = V_{RVI(s)} - V_{RVI(w)} \quad (2)$$

$$V_{SDI(3)} = V_{DVI(s)} - V_{DVI(w)} \quad (3)$$

式中: $V_{SDI(1)}$ 为归一化植被指数季节差值,表示夏季归一化植被指数与冬季归一化植被指数差值; $V_{SDI(2)}$ 为比值植被指数季节差值,表示夏季比值植被指数与冬季比值植被指数差值; $V_{SDI(3)}$ 为差值植被指数季节差值,表示夏季差值植被指数与冬季差值植被指数差值; $V_{NDVI(s)}$ 、 $V_{NDVI(w)}$ 分别为夏季、冬季归一化植被指数; $V_{RVI(s)}$ 、 $V_{RVI(w)}$ 分别为夏季、冬季比值植被指数; $V_{DVI(s)}$ 、 $V_{DVI(w)}$ 分别为夏季、冬季差值植被指数。

本文采用随机森林算法和极大似然算法、单时相和多时相特征集合以及基于物候特性差异性优化的特征变量对长江口滨海湿地植物进行了分类,并采用混淆矩阵(confusion matrix)的方法,通过计算制图精度、用户精度、总体精度以及Kappa系数对分类精度进行评价<sup>[14]</sup>。植物密度是反映植物特征的重要参数之一,能够反映湿地生态环境的变化。目前对于植物密度的研究主要以实地测量为主,难以获取大范围植物密度数据。凌成星<sup>[15]</sup>采用遥感影像提取的光谱特征及植被指数特征18个自变量,通过相关性分析及逐步回归得到以B5、DVI、RVI、NDVI为自变量的最优多元线性回归模型<sup>[16]</sup>。该方法为湿地植物密度空间特征分析提供了新的思路。因此,本文在植物分类的基础上,以B5、DVI、RVI、NDVI

为自变量结合实测数据,构建长江口滨海湿地植物密度反演模型,分析长江口湿地植物密度空间分布的特征。模型回归系数的显著性通过决定系数 $R^2$ 和构造检验统计量 $F$ 检验评价<sup>[16]</sup>。

### 3 结果与讨论

#### 3.1 不同分类算法、特征集合及特征优化后分类结果与分析

本文通过对比选用不同分类算法、特征集合以及特征优化后的分类结果精度,分析了不同分类算法对所选取的特征集合的适用性及其产生精度差异的原因,并讨论了选用同种算法的情况下不同特征集合对分类结果精度的影响。通过提取Landsa-8 OLI 遥感影像的光谱特征(每景影像提取7个光谱特征:B1~B7)、指数特征(每景影像提取7个指数特征:NDVI、DVI、RVI、MNDWI、BI、GVI、WI)以及本文提出的优化特征变量( $V_{SDI(1)}$ 、 $V_{SDI(2)}$ 、 $V_{SDI(3)}$ ),构建3组特征集合。特征集合A为单时相光谱特征及单时相指数,即2019年7月遥感影像提取的7个光谱特征和7个指数特征;特征集合B为多时相光谱特征及多时相指数,即2018年12月,2019年1月、4月、5月、7月5景遥感影像提取的共35个光谱特征和35个指数;特征集合C为多时相光谱特征、多时相指数特征及优化特征变量(即2018年12月,2019年1月、4月、5月、7月5景遥感影像提取的共35个光谱特征和35个指数特征以及3个植被指数季节差值 $V_{SDI(1)}$ 、 $V_{SDI(2)}$ 、 $V_{SDI(3)}$ )。采用随机森林算法和最大似然法分别对长江口湿地植被进行分类,共设计5种实验方案进行对比研究(表2)。

表2 不同分类方案分类精度比较  
Tab. 2 Comparison of classification accuracies in different cases

实验 方案	分类 方法	特征 变量	总体 精度/%	制图精度/%			用户精度/%			Kappa 系数
				芦苇	互花米草	海三棱藨草	芦苇	互花米草	海三棱藨草	
方案1	随机森林算法	集合A	80.65	84.62	75.52	68.85	73.33	75.00	73.68	0.75
方案2	随机森林算法	集合B	85.82	89.23	87.41	79.51	77.63	86.21	82.20	0.82
方案3	最大似然法	集合A	78.35	78.46	77.62	81.15	80.95	80.43	58.24	0.72
方案4	最大似然法	集合B	75.67	75.38	86.71	90.98	87.50	86.71	52.36	0.68
方案5	随机森林算法	集合C	87.55	89.23	88.11	80.33	79.45	86.30	86.73	0.84

随机森林算法分类及最大似然法分类结果如图4所示。其中图4a~4d为崇明东滩湿地,图4e~4h为九段沙湿地,图4i~4l为南汇促淤区;图4a、4e、4i为方案1分类结果,图4b、4f、4j为方案3结果,图4c、4g、4k为方案2结果,图4d、4h、4l为方案4结果。由

图4可以看出,长江口滨海湿地植物呈明显的条带状或弧形空间分布格局,由低潮带至高潮带依次分布海三棱藨草、互花米草、芦苇群落。植物群落面积大小依次为互花米草、海三棱藨草、芦苇。其中,互花米草在九段沙下沙及南汇促淤工程区分布面

积占比较大,主要分布在中高潮滩,海三棱藨草在南汇促淤区的分布面积较大,主要分布在中潮滩,

芦苇主要分布于崇明东滩、九段沙中沙及下沙的高潮带。

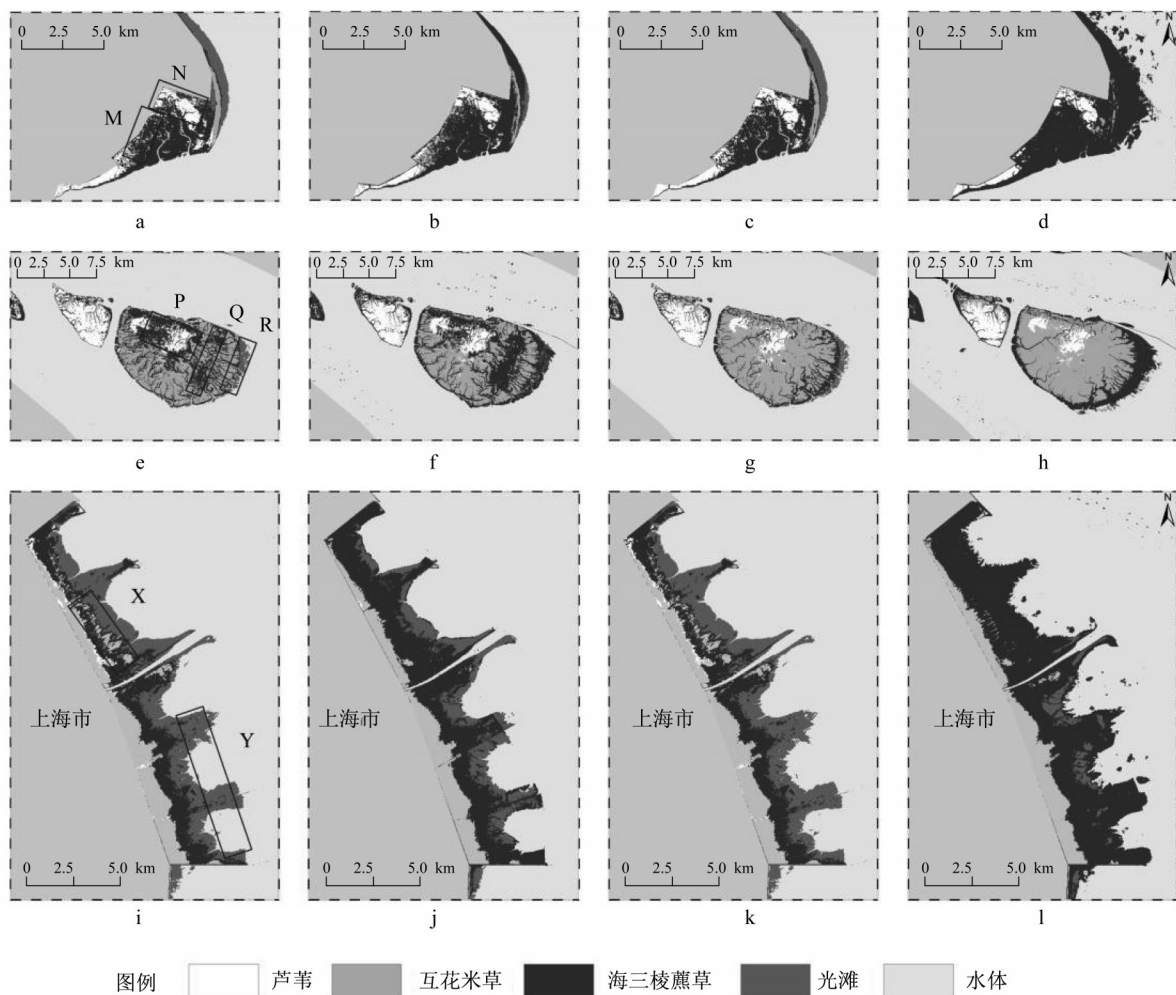


图4 不同分类方案分类结果

Fig. 4 Classification results of different cases

对比选用特征集合A条件下的随机森林算法和最大似然法分类结果(图4a、4b、4e、4f、4i、4j)发现,最大似然法分类结果中崇明东滩高潮带(M、N区域)有斑点状互花米草分布,九段沙下沙中潮带至高潮带间(Q区域)有大面积海三棱藨草分布,均为互花米草与海三棱藨草之间相互误分的像元。对比选用特征集合B时2种方法分类结果(图4c、4d、4g、4h、4k、4l)发现,最大似然法分类结果中崇明东滩低潮带、九段沙低潮带(R区域)有大面积海三棱藨草,均为裸地被误分为海三棱藨草的像元。对比采用随机森林算法选用不同特征集合的分类结果(图4a、4c、4e、4g、4i、4k)发现,选用特征集合A时九段沙中、高潮带(P、Q区域)有大面积海三棱藨草分布,均为互花米草被误分为海三棱藨草的像元。对比采用最

大似然法选用不同特征集合时的分类结果(图4b、4d、4f、4h、4j、4l)发现,选用特征集合A时Q区域有大量互花米草被误分为海三棱藨草的像元,选用特征集合B时R区域有大量光滩被误分为海三棱藨草的像元。基于以上分类结果,与实际目视情况(图1)相比,选用随机森林算法和特征集合B的误分、漏分最少。

由表2可知,在对特征集合A(低维特征集合)的处理上,随机森林算法和最大似然法的总体精度和Kappa系数相近。当选用特征集合B时,特征集合维数变高,使用随机森林算法总体精度及Kappa系数随之增加,而使用最大似然法的分类精度降低。产生这种现象的原因可能是,最大似然法将地物各特征变量的分布假定为正态分布,而由于各类地物的



光谱特性及植被指数等特征变量的分布具有高度的复杂性与随机性,很难获取完备的训练样本,选取的训练样本数据的统计信息与正态分布偏离,在特征集合维数较低时,误差并不明显,随着特征集合维数提高,无法准确获取各类别概率密度函数,所构造的多元正态分布模型误差增大,导致分类精度降低。而构建随机森林模型的CART决策树,其建立决策树过程中选取特征变量采用最小基尼系数的原则,即样本集合中随机抽取2个样本属同类别的概率最大,并且在以一定数量的CART决策树构建随机森林模型的过程中,采用Bagging原理进行集成,大大提高了模型的泛化能力,使得随机森林模型具有较高的稳定性和处理高维数据集的能力,因而特征集合维数变高,分类精度也随之提高。说明随机森林算法在处理高维数据集方面优势显著。

采用随机森林算法进行分类时,选取的特征集合对结果具有一定的影响。选用特征集合A的情

况下,3种植被的制图精度及用户精度较选用特征集合B的情况明显偏低。选用特征集合B的总体精度比选用特征集合A的提高5.2%,Kappa系数提高0.7,说明选取多时相特征变量能有效提高分类精度。

选用特征集合C的随机森林算法分类结果如图5a~5c所示。相比同样选用特征集合B的随机森林算法分类结果,M区域中部分误分为互花米草的芦苇像元被正确分类,R区域中部分被误分为互花米草的像元被正确分为海三棱藨草。受崇明东滩淤积变化影响,崇明东滩互花米草相较海三棱藨草处于潮滩的前沿。由于长江口水动力条件及含沙量改变,崇明东滩北部淤涨、南部侵蚀,导致东滩北部高程增加<sup>[17]</sup>,更适宜互花米草的生长。同时,崇明东滩湿地土壤盐度北部高于南部,也导致了耐盐性更强的互花米草在北部生长<sup>[17]</sup>。

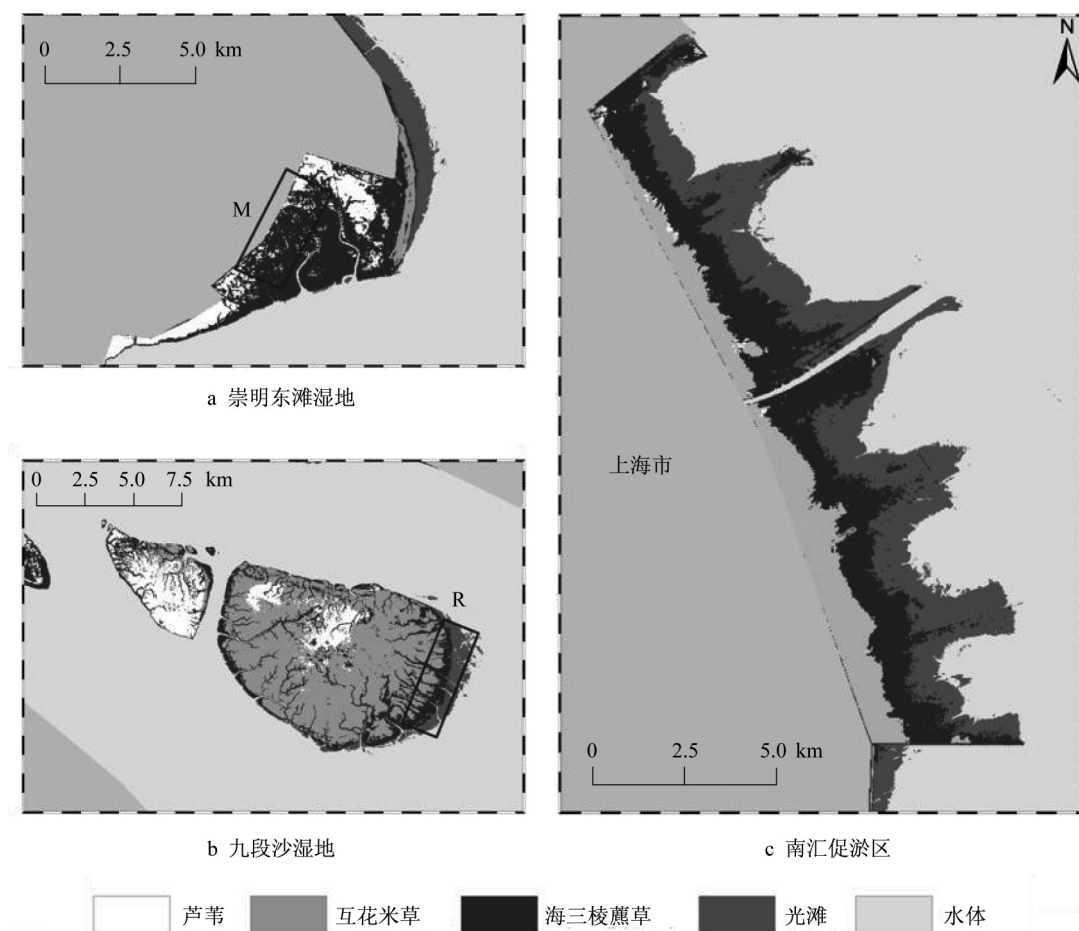


图5 方案5分类结果

Fig. 5 Classification results in Case 5

加入植被季节指数差值作为特征变量前后分类结果(图4c、4g、4k、图5a、5b、5c)及混淆矩阵如表2所示,可以看出,总体分类精度提高了1.73%,说明植被指数季节差值的计入,有利于分类精度的提高。除去水体与光滩,芦苇、互花米草与海三棱藨草的制图精度较高,均在80%以上。芦苇的用户精度较低,不足80%。互花米草与海三棱藨草的用户精度较高,均在85%以上。植被指数季节差值加入特征集合进行特征变量优化后,互花米草、海三棱藨草的制图精度均有不同程度提高,说明该2类漏分的像元变少。芦苇、互花米草和海三棱藨草的用户精度均有提高,说明3类植物的误分像元变少。3种地物

的制图精度及用户精度提高表明,植被指数季节差值加入有利于区分光谱特性相似但物候性有差异的地物。芦苇的制图精度与用户精度较低,主要原因可能是互花米草在长江口滨海湿地的竞争性大于芦苇,芦苇生长呈斑块状且面积较小,因此单个像元中可能存在混合生长的2种植被。

基于以上分析,最终选取总体精度最高的分类方法和特征集合(即随机森林算法和特征集合C),以其分类结果利用ENVI5.3进行后处理以减少图中的椒盐现象,最终得出长江口滨海湿地植物分类情况,如图6a~6c所示。

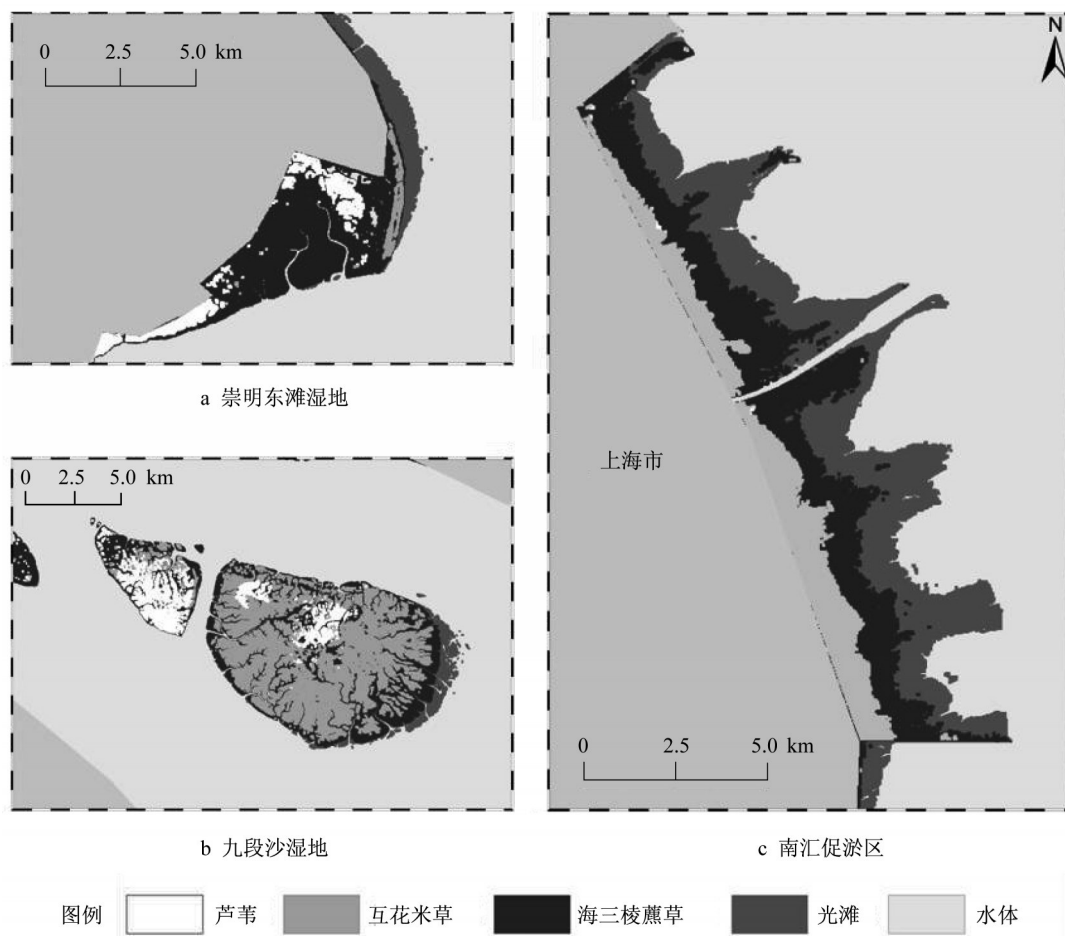


图6 长江口滨海湿地植物分布图

Fig. 6 Distributions of vegetation species in coastal wetlands of Yangtze River estuary

### 3.2 植物密度空间分布反演——以互花米草为例

利用2019年7月—9月实地调查获得的18个测站的互花米草密度以及遥感影像(2019年7月29日)提取的对应坐标点像元近红外波段反射率( $V_{B5}$ )、归一化植被指数( $V_{NDVI}$ )、差值植被指数( $V_{DVI}$ )、比值植被指数( $V_{RVI}$ )数据,进行多元线性回归参数计算及

回归结果分析,获得互花米草密度反演结果如下式所示:

$$V_{\text{dens}} = 1665.5 - 876.3V_{B5} - 2935.2V_{NDVI} + 1971.6V_{DVI} + 47.3V_{RVI} \quad (4)$$

根据上述多元线性回归反演模型,去除部分异常值,获得长江口滨海湿地互花米草密度的空间分布图,



如图7所示。

计算结果表明,长江口湿地互花米草的密度值多数在 $100\sim 300$ 株 $\cdot\text{m}^{-2}$ 之间,且沿低潮带至高潮带呈增

长趋势。模型决定系数 $R^2$ 为0.7928,经 $F$ 检验,自变量与因变量之间线性关系显著。该方法也可用于其他类似区域的植物特征反演分析。

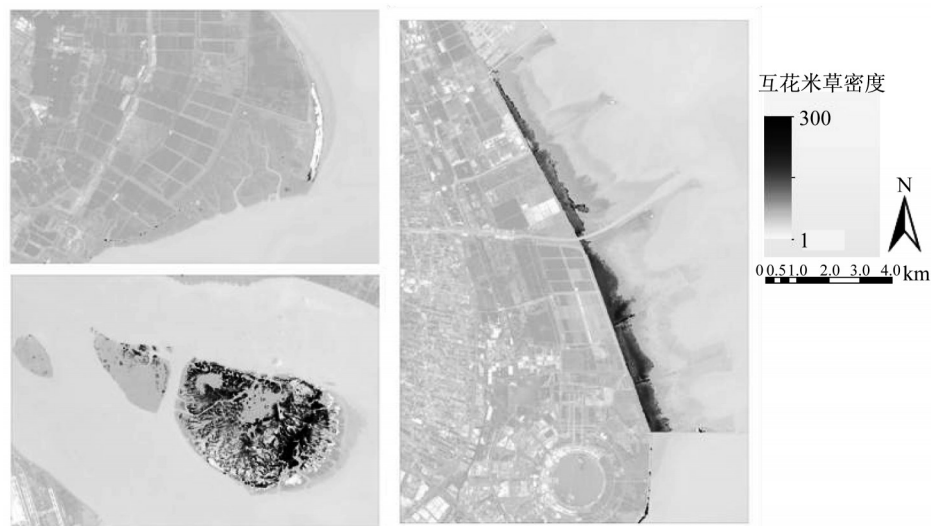


图7 长江口滨海湿地互花米草密度空间分布图

Fig. 7 Distributions of *Spartina alterniflora* density in coastal wetlands of Yangtze River estuary

## 4 结论

本文采用Landsat-8影像以及实地调查数据,利用随机森林算法,对长江口滨海湿地植被群落进行分类。在提取多时相光谱特征、植被指数和水体指数的基础上,基于植物物候特性进行了特征变量优化,获得了长江口芦苇、互花米草和海三棱藨草3种优势物种的空间分布。同时,以外来入侵物种互米花草为例,根据多元线性回归模型,以遥感影像提取的光谱特征及植被指数为因变量对滨海湿地植物密度进行反演,获得了长江口互米花草密度的空间分布特征。本文提出的分析方法和获得的分析结果,可用于长江口湿地生态环境的管理与保护,也可应用于其他类似地区的湿地环境研究。本文主要结论如下:

(1)为了更加精确识别长江口滨海湿地植物群落分布特征,提出了基于优化特征变量集合的随机森林模型,实现了长江口滨海湿地优势植被的精确分类,总体精度达到87.55%。基于植物分类算法比选,随机森林算法较传统最大似然法在处理高维特征集合时总体精度提高10%以上。本文通过随机森林模型的特征集合优化,构造了73个特征变量(35个光谱特征变量、35个指数特征变量及3个优化特征变量),与单时相特征集合分类结果相比,总体精度提高近7%。

(2)通过植物密度反演方法,获得互花米草的密度为 $100\sim 300$ 株 $\cdot\text{m}^{-2}$ ,所建立的多元线性回归模型,决定系数与 $F$ 值验证显示其线性关系显著,反演公式能较为精确地计算互花米草密度值。该研究成果可以为长江口入侵物种互花米草生态治理提供基础资料,该方法也可应用于类似地区湿地植物密度反演。

(3)受水动力条件间接影响,长江口滨海湿地植物分布呈明显条带状或弧形分布格局,沿低潮带至高潮带方向依次分布海三棱藨草、互花米草、芦苇。通过像元统计,研究区域中3种湿地优势植物总面积约为 $266.6\text{ km}^2$ 。互花米草面积最大,为 $145.4\text{ km}^2$ ,占湿地植物总面积54.5%,芦苇与海三棱藨草面积分别为 $46.7\text{ km}^2$ 、 $74.5\text{ km}^2$ ,占比分别为17.5%和27.9%。崇明东滩前缘出现弧形互花米草群落,该区域受水动力条件影响,高程与盐度增加,更适宜耐盐性强的互花米草生长。

本研究基于长江口滨海湿地植物物候特性提出了植被指数季节差值,探讨了其对分类精度的影响,但存在一些不足:所采用的分类方法是基于像元的分类方法,存在部分“椒盐现象”,后续研究可结合面向对象分类方法,在一定程度上提高分类精度。

### 作者贡献说明:

刘曙光:指导研究方案和论文撰写,全文审阅。

董 行:数据收集与处理,设计论文框架,撰写论文。

娄 厦:提出研究思路,审阅及修订论文。

DORZHEVNA Radnaeva Larisa:提供研究思路与技术指导。

NIKITINA Elena:提供研究思路与技术指导。

## 参考文献:

- [1] 周念清,王燕,钱家忠,等.湿地氮循环及其对环境变化影响研究进展[J]. 同济大学学报(自然科学版), 2010,38(6):865.  
ZHOU Nianqing, WANG Yan, QIAN Jiazhong, *et al.* Advances in research on nitrogen cycle in wetlands and its influence on environmental change [J]. Journal of Tongji University(Natural Science), 2010,38(6):865.
- [2] 张利权,雍学葵.海三棱藨草种群的物候与分布格局研究[J]. 植物生态学与地植物学学报, 1992(1):43.  
ZHANG Liqun, YONG Xuekui. Studies on phenology and spatial distribution pattern of scirpus mariqueter population [J]. Chinese Journal of Plant Ecology, 1992(1):43.
- [3] 孙楠.利用GF-1和Landsat遥感数据监测长江口海岸带湿地近40年变化及其驱动因子分析[D]. 杭州:浙江大学, 2018.  
SUN Nan. GF-1 and Landsat observed a 40-year wetland spatiotemporal variation and its coupled environmental factors in Yangtze River Estuary [D]. Hangzhou: Zhejiang University, 2018.
- [4] 曹正凤. 随机森林算法优化研究[D]. 北京:首都经济贸易大学, 2014.  
CAO Zhengfeng. Study on optimization of random forests algorithm [D]. Beijing: Capital University of Economics and Business, 2014.
- [5] BREIMAN L. Random forests [J]. Machine Learning, 2001, 45(1):5.
- [6] AMANI M, SALEHI B, MAHDAVI S, *et al.* Wetland classification using multi-source and multi-temporal optical remote sensing data in newfoundland and Labrador, Canada [J]. Canadian Journal of Remote Sensing, 2017,43(4):360.
- [7] 陈珂,彭志平,柯文德. 遥感图像异物同谱干扰消除技术与仿真[J]. 计算机仿真, 2013,30(1):423.  
CHEN Ke, PENG Zhiping, KE Wende. Remote sensing image classification based on improvement of neural network HSV model [J]. Computer Simulation, 2013,30(1):423.
- [8] 翟天林,金贵,邓祥征,等. 植被信息的Landsat8卫星影像提取方法[J]. 测绘科学, 2016, 41(10): 126.  
ZHAI Tianlin, JIN Gui, DENG Xiangzheng, *et al.* Study of vegetation information extraction method based on Landsat8 satellite images [J]. Science of Surveying and Mapping, 2016, 41(10): 126.
- [9] 张晓羽,李凤日,甄贞,等. 基于随机森林模型的陆地卫星-8遥感影像森林植被分类[J]. 东北林业大学学报, 2016, 44(6):53.  
ZHANG Xiaoyu, LI Fengri, ZHEN Zhen, *et al.* Forest vegetation classification of Landsat8 remote sensing image based on random forests model [J]. Journal of Northeast Forestry University, 2016,44(6):53.
- [10] 田庆久,闵祥军. 植被指数研究进展[J]. 地球科学进展, 1998(4):3.  
TIAN Qingjiu, MIN Xiangjun. Advances in study on vegetation indices [J]. Advances in Earth Science, 1998(4):3.
- [11] 徐涵秋. 利用改进的归一化差异水体指数(MNDWI)提取水体信息的研究[J]. 遥感学报, 2005(5):589.  
XU Hanqiu. A study on information extraction of water body with the modified normalized difference water index (MNDWI) [J]. Journal of Remote Sensing, 2005(5):589.
- [12] 李博伦,逆超普,颜晓元. Landsat 8陆地成像仪影像的缨帽变换推导[J]. 测绘科学, 2016,41(4):102.  
LI Bolun, TI Chaopu, YAN Xiaoyuan. Study of derivation of tasseled cap transformation for Landsat 8 OLI images [J]. Science of Surveying and Mapping, 2016,41(4):102.
- [13] 王卿. 长江口盐沼植物群落分布动态及互花米草入侵的影响[D]. 上海:复旦大学, 2007.  
WANG Qing. The dynamics of plant community distribution of the salt marshes in the Yangtze River Estuary as influenced by spartina alterniflora invasions [D]. Shanghai: Fudan University, 2007.
- [14] FOODY G M. Status of land cover classification accuracy assessment [J]. Remote Sensing of Environment, 2002, 80 (1):185.
- [15] 凌成星. Worldview-2八波段影像支持下的湿地信息提取与地上生物量估算研究[D]. 北京:中国林业科学研究院, 2013.  
LING Chengxing. Study on wetland information extraction and above-ground biomass estimation supported by Worldview-2 images [D]. Beijing: Chinese Academy of Forestry, 2013.
- [16] MAC N R. Multiple regression and inference in ecology and conservation biology: further comments on identifying important predictor variables [J]. Biodiversity & Conservation, 2002,11(8):1397.
- [17] 丁文慧. 互花米草生态治理工程对崇明东滩白头鹤潮滩生境的影响[D]. 上海:华东师范大学, 2016.  
DING Wenhui. The effect of spartina ecological management project on the tidal flat habitat of grus monacha in Chongming Dongtan [D]. Shanghai: East China Normal University, 2016.