

引用格式: Yang Xianbao, Zhang Wangfei, Sun Bin, *et al.* Recognition of Vegetation Types in Hulunbuir Sandy Land and Its Surrounding Areas based on GEE Cloud Platform and Sentinel-2 Time Series Data[J]. Remote Sensing Technology and Application, 2022, 37(4): 982-992. [杨仙保, 张王菲, 孙斌, 等. 基于GEE和Sentinel-2时序数据的呼伦贝尔沙地及其周边植被类型识别研究[J]. 遥感技术与应用, 2022, 37(4): 982-992.]

DOI: 10.11873/j.issn.1004-0323.2022.4.0982

基于GEE和Sentinel-2时序数据的呼伦贝尔沙地及其周边植被类型识别研究

杨仙保^{1,2}, 张王菲¹, 孙斌^{2,3}, 高志海^{2,3}, 李毅夫^{2,3}, 王晗^{2,3}

(1. 西南林业大学 地理与生态旅游学院, 云南 昆明 650224;

2. 中国林业科学研究院资源信息研究所, 北京 100091;

3. 国家林业和草原局林业遥感与信息技术重点实验室, 北京 100091)

摘要: 沙地及其周边植被对固定沙丘、防止水土流失和环境治理等方面具有重要作用, 开展沙地及其周边植被类型识别研究对于客观地反映沙地及其周边的生态环境, 进而为沙地恢复治理政策制定具有重要意义。GEE云平台丰富的长时间序列遥感数据和强大的云计算能力, 为开展大区域植被类型识别提供了便捷。本研究基于GEE云平台存储的2019年Sentinel-2时序数据, 采用RF算法开展呼伦贝尔沙地及其周边主要植被类型的空间判识研究, 探索了GEE平台下顾及植被物候信息的植被类型识别效果。研究发现: ① Sentinel-2影像的光谱信息和近红外波段的纹理信息对研究区的主要植被类型识别能力有限, 而物候特征有效地弥补了原始光谱特征等对研究区不同植被类型区分能力的不足; ② 基于RF算法顾及物候特征的植被类型识别精度达到84.37%, Kappa系数为0.8, 比单一时相数据的识别精度提高了10.01%; ③ 呼伦贝尔沙地及其周边主要植被类型的物候特征差异明显, 有助于不同类型植被的空间识别, 特别是提高了灌草丛和草原的识别精度。研究表明利用Sentinel-2数据和GEE云平台对沙地等大区域植被类型的识别具有较大的潜力和广阔的应用前景。

关键词: GEE; Sentinel-2; 时序数据; 呼伦贝尔沙地; 植被类型识别

中图分类号: K901.4; TP79 **文献标志码:** E **文章编号:** 1004-0323(2022)04-0982-11

1 引言

根据第五次全国荒漠化和沙化监测结果中显示, 我国沙地面积达17 211.75万 hm^2 , 约占国土总面积的17.93%^[1], 受沙化土地危害严重。而沙地及周边的植被对固定沙丘、防止水土流失、环境治理等方面具有重要作用^[2]。开展沙地及周边植被类型识别研究对于客观地反映沙地及其周边的植被环境, 进

而为沙地恢复治理政策的制定具有重要意义。

遥感以大范围观测的特点成为大区域植被类型识别的主要技术手段^[3], 但遥感影像具有的海量空间信息目视解译难以分析, 计算机自动分类技术逐渐被运用, 其中根据分类过程中是否借助先验知识, 分为监督分类和非监督分类^[4-6]。正常情况下, 监督分类的准确性要优于非监督分类, 但监督分类需要获取大量有代表性的训练样本, 工作量较大^[7]。

收稿日期: 2021-02-11; 修订日期: 2022-06-14

基金项目: “中央级公益性科研院所基本科研业务费专项”(CAFYBB2019ZB004)和“国家高分辨率对地观测系统重大专项”(21-Y20A06-9001-17/18)资助。

作者简介: 杨仙保(1995—), 男, 云南德宏人, 硕士研究生, 主要从事区域地理研究。E-mail: 2296857887@qq.com

通讯作者: 孙斌(1987—), 男, 山东淄博人, 副研究员, 主要从事荒漠化和草原遥感监测与评价研究。E-mail: sunbin@ifrit.ac.cn

而机器学习算法则可在有少量训练样本的情况下,获得较高的正确率^[8]。此外,机器学习算法还可以从海量的数据中提取有用的信息,运算速度快和稳定性高,常作为植被类型识别的首选^[9-11]。在遥感影像时相的选择方面,相对于单一时相的遥感影像,长时间序列遥感影像因可表征植被生长时间差异性而被更多的研究人员关注。如梁爽等^[12]运用多时相的 Sentinel-2 影像进行湿地信息的提取,发现多季相影像的组合能够显著提高分类精度,相对单季影像总体分类精度提高了 10.9%—25.5%;Wakuliński 等^[13]基于 Sentinel-2 时间序列数据,采用支持向量机法,开展山区植被类型识别研究,也获得了较好的分类结果;Rapinel 等^[14]等基于 Sentinel-2 时间序列数据,采用支持向量机(SVM)和随机森林(RF)等分类方法,对草原植物群落进行了准确识别,分类精度得到有效的提高。

长时间序列遥感数据蕴藏了植被生长过程中大量的物候信息,但受研究范围和数据量等因素限制,导致传统的大区域植被类型识别效率较低。Google Earth Engine(GEE)云平台以丰富的长时间序列遥感数据和强大的云计算能力,在运用于大区域内植被长时间的演变分析时可突破以上限制,有效提高识别效率^[15-18],为开展大区域主要植被类型识别提供了有效的支撑。鉴于此,本研究以呼伦贝尔沙地及其周边的海拉尔市、陈巴尔虎旗、新陈巴尔虎左旗、鄂温克族自治旗为研究区,以 GEE 平台存

储的 Sentinel-2A 为数据源,采用 RF 算法开展呼伦贝尔沙地及其周边主要植被类型的空间判识研究,探索 GEE 平台下顾及植被物候信息的沙地及其周边植被类型识别的提取效果,以期为当地沙地恢复治理政策的制定提供科学数据支撑。

2 研究区与数据

2.1 研究区概况

研究区地处内蒙古东北部的呼伦贝尔高原之间,行政区包含海拉尔市、陈巴尔虎旗、新陈巴尔虎左旗、鄂温克族自治旗(47°20′—50°11′N,117°33′—121°09′E;图1)。研究区东部靠近黑龙江,南部与蒙古国接壤,西部至呼伦湖东岸,北部与俄罗斯相邻,地势由东向西逐渐降低,南部高于北部^{[19][20]}。研究区为大陆性季风气候,年平均气温介于-2.5—0℃之间,无霜期为90—100 d。年降水量270—370 mm,年蒸发量1 400—1 900 mm,干燥度为1.2—1.5,相对湿度为60%至70%^[21-22]。森林为针叶林、阔叶林和针阔混交林,树种以樟子松、落叶松(*Larix gmelinii*)、白桦(*Betula platyphylla*)、山杨(*Populus davidiana*)等;灌草丛以冷蒿(*Artemisia frigida* Willd)、差巴嘎蒿(*Artemisia halodendron Turcz. et Bess*)、小叶锦鸡儿(*Caragana microphylla* Lam)、黄柳(*Salix gordejewii* Chang et Skv)、榆树(*Ulmus Bergmanniana* Schneid)为主;草原以线叶菊(*Filifolium sibiricum*)、贝加尔针茅(*Stipa grandis*)、羊草

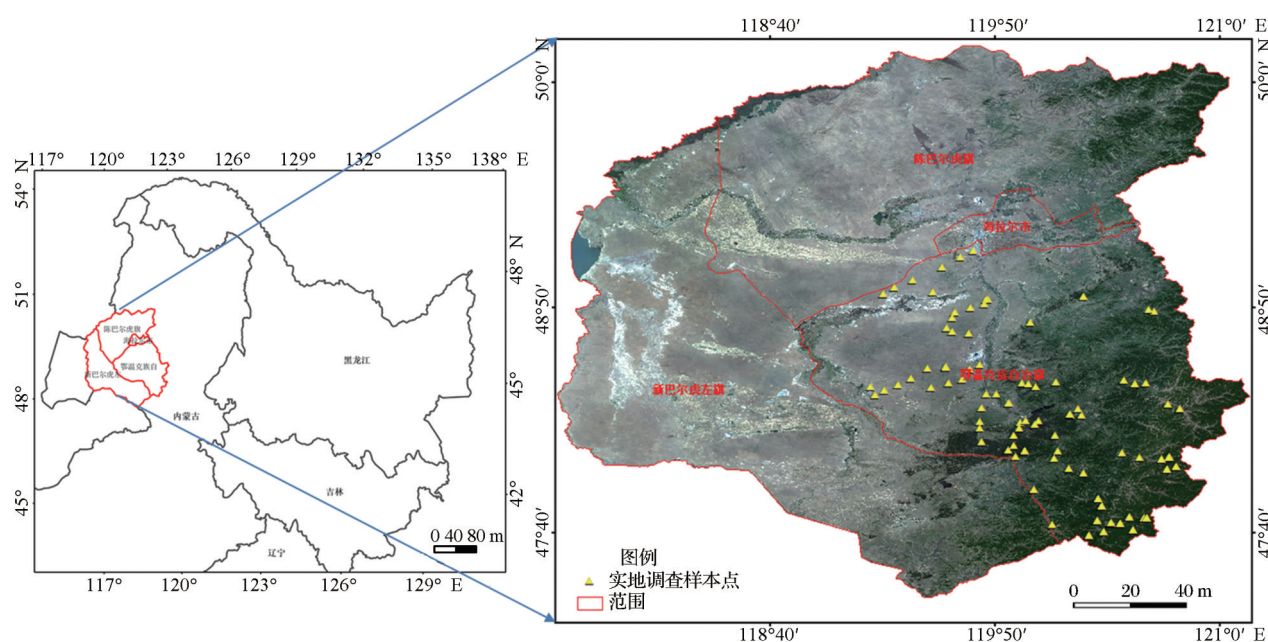


图1 研究区地理位置示意图

审图号:GS(2019)3333

Fig.1 The study area location diagram

(*Leymus chinensis*)、大针茅(*Stipa grandis*)、克氏针茅(*Stipa krylovii*)、隐子草(*Gramineae sp*)为主^[23-25]。农作物主要有大豆、玉米、大麦、小麦、油菜、蔬菜等,为一年一熟^[26]。作为呼伦贝尔重要的生态功能保护区和生态敏感的林草过渡带之一^[6],由于长期不合理开发利用,呼伦贝尔沙地植被退化和土地沙化现象严重。

2.2 数据获取及预处理

以存储于 GEE 云盘中 2019 年的 Sentinel-2A 无云数据作为数据源,同时为提取研究区不同植被类型的物候信息,数据获取时确保每个月都有 1 景影像覆盖。为验证时序数据中提取的物候特征对植被类型识别的有效性,减少因数据源不同而引起的误差,将 2019 年 6 和 7 月的 Sentinel-2A 数据作为单一时相数据用于对比分析。

Sentinel-2A 含有 13 个不同空间分辨率的波段,其中 B2~B4(绿、蓝、红)、B8(近红外)等波段的空間分辨率为 10 m;B5~B7(红边 1、2、3)、B8A(红边 4)、B11~B12(短波红外 1、2)等波段的空間分辨率为 20 m;B1(气溶胶)、B9(水蒸气)、B10(卷云)等波段的空間分辨率为 60 m^[27];而用于校正的 SRTM 数据的空間分辨率为 30 m,为保证分辨率的统一,本文中采用最邻近采样方法将不同分辨率的波段均重采样为 10 m。而研究使用的样本数据是通过实地调查和同期高分辨率遥感影像随机选取的方式获取。

研究区用于分类的样本于 2019 年和 2020 年实地调查获得,共获得了 105 个主要植被类型的样本。另外,为保证研究具有足够的样本量,我们又从高分辨率遥感影像随机选取了 1 665 个样本,最终获得不同植被类型的样本共计 1 770 个。其中,针叶

林样本 130 个,阔叶林样本 170 个,针阔混交林样本 90 个,灌草丛样本 300 个,草原样本 330 个,农作物样本 240 个,其他植被样本 70 个,非植被样本 440 个。

3 研究方法

研究首先进行分类体系构建,然后基于 GEE 云平台快速获取了研究区 2019 年的 Sentinel-2A 数据和 DEM 数据,并进行预处理和重采样,在此基础上进行光谱特征、植被指数、纹理信息、物候特征、植被覆盖度和地形特征等特征数据集的提取。结合所有样本数据通过 RF 特征选择法进行特征数据集的选择得到最优特征数据集,作为不同植被类型识别的特征,再利用 RF 分类方法实现研究区地物类型的识别。最后将时序数据与单一时相数据的植被类型识别结果进行精度评价和对比分析。本文详细技术流程见图 2。

3.1 分类体系构建

按《中国植被》^[28]和《中国土地覆盖遥感监测》^[29]对不同植被类型的定义,结合研究区植被类型分布概况和研究目的,将研究区植被类型分为针叶林、阔叶林、针阔混交林、灌草丛、草原、农作物、其他植被和非植被等 8 类。具体分类体系见表 1。

3.2 特征信息提取

为提高沙地及其周边植被类型的识别精度,研究提取不同植被类型的特征信息并进行有效组合表 2。首先选取不同植被类型在经过预处理的影像上 12 个原始波段的地表反射率作为光谱特征;选择归一化植被指数(Normalized Difference Vegetation Index,NDVI)、增强植被指数(Enhanced Vegetation Index,EVI)、归一化建筑指数(Normalized Differ-

表 1 分类体系
Table 1 Classification system

代码	植被类型	特征
1	针叶林	森林郁闭度≥60%,以针叶林为主要类型,以樟子松、落叶松等针叶树为建群种组成的森林群落。
2	阔叶林	森林郁闭度≥60%,以阔叶林为主要类型,以白桦和山杨等为建群种组成的森林群落。
3	针阔混交林	森林郁闭度≥60%,针叶林和阔叶林相互混生且每种类型面积均不超过 50%,是针叶林向阔叶林过渡的植被类型,是白桦、山杨和樟子松、落叶松相互混生的森林群落。
4	灌草丛	林木郁闭度在 10%—60%,高度在 2 m 以下的灌丛,以中生或旱中生多年生草本植物为主要建群种,其中散生灌木的植物群落。主要以小叶锦鸡、冷蒿、差巴嘎蒿、黄柳、榆树等植被为主。
5	草原	林灌郁闭度<10%,且覆盖度大于 5%,是一种以生长草本植物为主,由旱生或中旱生草本植物组成的草本植物群落,主要有线叶菊、贝加尔针茅、羊草、大针茅、克氏针茅、隐子草等。
6	农作物	农业上栽培的各种粮食和经济植物,包括大豆、玉米、小麦、油菜、蔬菜等植被。
7	其他植被	主要由天然草本植物为主的沼泽化低地草甸、沼泽植被和水生植被等。
8	非植被	植被覆盖度小于 5%,表层主要由水体、人造地表、盐碱地、裸沙、裸地、矿坑等覆盖。

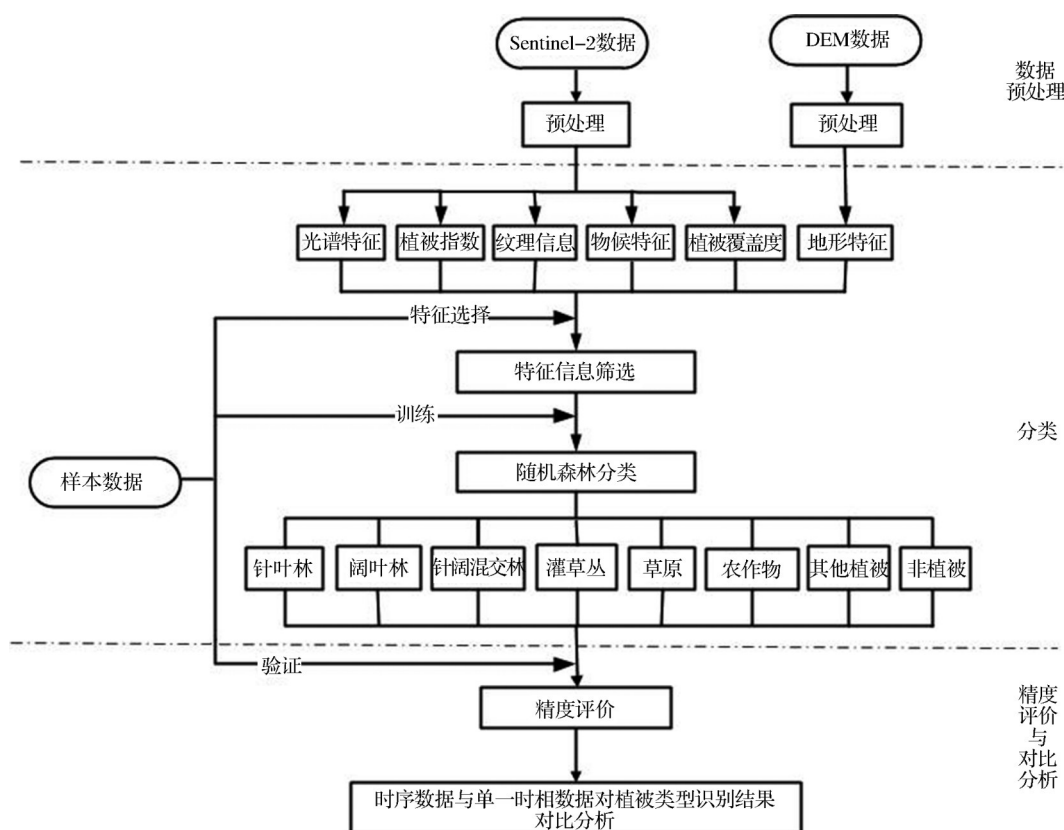


图2 技术路线

Fig.2 Technical route

ence Build-up Index, NDBI)、归一化差异水体指数 (Modified Normalized Difference Water Index, MNDWI) 等作为植被类型识别的植被指数。各植被指数的计算公式如下:

$$NDVI = \frac{\rho_{NIR} - \rho_{Red}}{\rho_{NIR} + \rho_{Red}} \quad (1)$$

$$EVI = 2.5 \times \frac{\rho_{NIR} - \rho_{Red}}{\rho_{NIR} + 6 \times \rho_{Red} - 7 \times \rho_{Blue} + 1} \quad (2)$$

$$NDBI = \frac{\rho_{MIR} - \rho_{NIR}}{\rho_{MIR} + \rho_{NIR}} \quad (3)$$

$$MNDWI = \frac{\rho_{Green} - \rho_{MIR}}{\rho_{Green} + \rho_{MIR}} \quad (4)$$

其中: ρ_{Red} 、 ρ_{Green} 、 ρ_{Blue} 、 ρ_{NIR} 、 ρ_{MIR} 分别代表红、绿、蓝、近红外、短波红外波段的地表反射率。

文中的纹理信息则使用GEE平台自带的灰度共生矩阵 (Gray-Level Co-occurrence Matrix, GLCM) 纹理特征函数 `glcmTexture(size, kernel, average)` 提取^[30]。纹理信息是一种结构特征,使用不同波段得到的纹理信息并无较大差异,但在近红外波段的植被反射率比较高,对植被的识别更有效^[31]。因此,将近红外B8波段的纹理信息,作为研究区内主要植被类型识别的纹理信息^[32]。

物候 (phenology) 是指自然界中植被受气候、水文和土壤等环境因素影响而发生周期性变化的生物学现象^[33],主要包括植被的发芽、展叶、抽穗及秋季叶变色及脱落过程等^[7,26,34]。但不同类型的植被在每个阶段的NDVI表现出不同的动态变化^[35]。因此物候特征的阶段性变化可通过年内月度NDVI变化来表征^[26]。植被覆盖度是基于GEE平台利用像元二分法进行计算,将累计频率为5%和95%的NDVI值作为 $NDVI_0$ 、 $NDVI_v$ 的值,当 $NDVI \leq NDVI_0$ 时取像元植被覆盖度为0, $NDVI \geq NDVI_v$ 时取像元植被覆盖度为1。将除去水体小于0的部分的NDVI值后进行统计的值作为NDVI值。然后利用植被覆盖度公式编写代码计算研究区植被覆盖度,并作为一个单独的波段,植被覆盖度公式如下:

$$f_v = \frac{NDVI - NDVI_0}{NDVI_v - NDVI_0} \quad (5)$$

其中: $NDVI_0$ 、 $NDVI_v$ 、 $NDVI$ 分别表示纯裸地、纯植被和混合像元的NDVI值; f_v 为植被覆盖度。并且通过实地调查发现,研究区的植被类型分布受坡度变化的影响较小,草原和森林在相对不同坡度的区域均有分布。但不同植被的分布容易受海拔的影响,森林主要分布在海拔较高的区域,而草原、灌草

丛和农作物等则分布在海拔相对较低的区域,不同的森林类型分布也容易受坡向的影响,阔叶林大部分分布在阴坡,针叶林主要分布在阳坡,并且山体阴影在乔灌木之间也存在差异,乔木因长得高大,山体阴影对它的影响也比较大,容易因山体阴影的原因而造成错分,但对于草原这样低矮的植被,山体阴影对它的影响就非常小。因此地形特征选择坡向、海拔、山体阴影等作为 3 个独立波段添加至影像中^[36]。

表 2 特征信息表

Table 2 Original features and optimization

特征类型	特征信息
光谱特征	B1、B2、B3、B4、B5、B6、B7、B8、B8A、B9、B11、B12
植被指数	NDVI、NDBI、MNDWI、EVI <i>B8_asm</i> 、 <i>B8_contrast</i> 、 <i>B8_corr</i> 、 <i>B8_var</i> 、 <i>B8_idm</i> 、 <i>B8_savg</i> 、 <i>B8_svar</i> 、 <i>B8_sent</i> 、 <i>B8_ent</i> 、 <i>B8_dvar</i> 、 <i>B8_dent</i> 、 <i>B8_imcorr1</i> 、 <i>B8_imcorr2</i> 、 <i>B8_maxcorr</i> 、 <i>B8_diss</i> 、 <i>B8_inertia</i> 、 <i>B8_shade</i> 、 <i>B8_prom</i>
纹理特征	<i>fv</i>
植被覆盖度	1 月—12 月 NDVI
物候特征	坡向、海拔、山体阴影
地形特征	

注：*B8_asm*、*B8_contrast*、*B8_corr*、*B8_var*、*B8_idm*、*B8_savg*、*B8_svar*、*B8_sent*、*B8_ent*、*B8_dvar*、*B8_dent*、*B8_imcorr1*、*B8_imcorr2*、*B8_maxcorr*、*B8_diss*、*B8_inertia*、*B8_shade*、*B8_prom* 分别表示角二阶矩、对比度、相关性、方差、逆差矩、和平均、和方差、和熵、差方差、相关信息测度 2、最大相关系数、差异性、惯性矩、聚类萌、聚类突等。

3.3 特征信息筛选

研究采用 RF 特征选取方法进行不同类型植被的特征信息优选,首先利用 RF 算法对特征变量的重要性进行排序,然后采用序列前项选择的方法,按特征重要性得分从高到低依次加入模型中训练分类器,通过逐次迭代并计算研究区植被类型识别的正确率。最终得到特征个数最少、分类准确率最高的特征数据集作为研究区植被类型识别的特征数据集^[37]。

3.4 沙地及其周边植被类型识别方法

采用 RF 算法作为研究区内主要植被类型的识别方法,利用 bootstrap 随机重采样技术从样本中随机抽取子样本,经过重复操作,构建多棵决策树,将决策树组成森林,对样本通过投票最终得到不同类型植被识别结果^[38-39]。通过两次随机性的引入,使 RF 算法不容易出现过拟合,具有较强的分析能力,对存在噪声的数据有很好的鲁棒性^[25]。研究选取 2019 年的 Sentinel-2A 时序数据和单一时相数据结

合 DEM 数据,通过 RF 特征选择获得最优特征数据集后,利用 GEE 云平台自带的 RF 分类器将样本随机切分成 70% 的训练样本和 30% 的验证样本进行不同植被类型的识别,最后对时序数据和单一时相数据的识别结果进行精度检验和对比分析。

3.5 精度评价和对比分析

精度评价主要通过混淆矩阵(confusion matrix)对研究区内主要植被类型的识别效果进行评价。将总体精度、制图精度、用户精度和 Kappa 系数等作为具体的精度评价指标^[40],然后对时序数据和单一时相数据的不同植被类型识别结果进行精度检验,并对比分析加入植被物候特征后的时序数据对沙地及其周边植被类型识别的效果。

4 结果与分析

4.1 沙地及周边主要类型植被的光谱和物候特征

基于不同植被类型的实测样本数据,在影像上提取不同波段的地表反射率,并生成光谱曲线(图 3)。从不同类型植被的光谱曲线(图 3)可以看出,针叶林在 B6—B11 波段的反射率明显低于其他的植被,易于提取;灌草丛、草原和其他植被除在 B12 波段的反射率具有一定的可分性外,在其他波段反射率相近可分性较差;农作物在 B2—B4 波段和 B12 波段与不同植被类型的反射率差异大,可分性较好;非植被区在 B1—B4 波段、B11 和 B12 波段的反射率高,便于区分。但阔叶林和针阔混交林二者之间因光谱曲线相近而难以区分。所以,仅通过光谱特征进行研究区植被类型识别,容易因不同类型植被之间的光谱分辨率相近而造成混分。

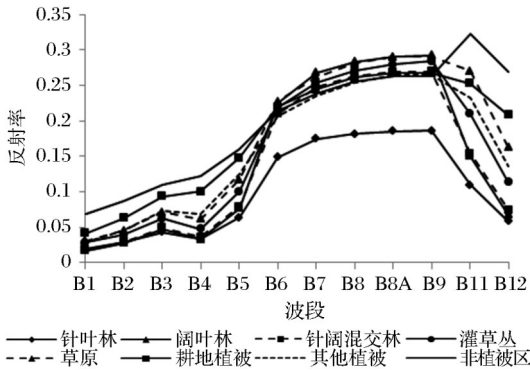


图 3 不同植被类型的光谱曲线
Fig.3 Spectral feature graph of different vegetation types

植被的 NDVI 变化能反映不同类型植被的生长情况,通过对比不同类型植被在 NDVI 曲线上的变化趋势,进行可识别性分析,结果如图 4 所示。

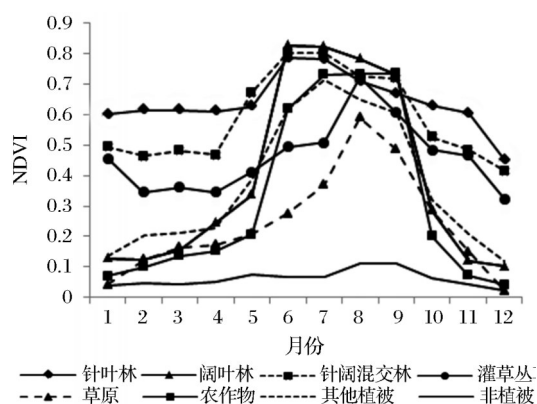


图4 不同植被类型的NDVI变化曲线

Fig.4 NDVI curve of different vegetation types

由图4可知,针叶林在1—5月的NDVI处于平稳状态,但NDVI与其他不同植被类型差异较大,便于针叶林的识别;针阔混交林在1—4月的NDVI波动变化较小,从4月开始上升,与针叶林二者在4—6月的NDVI变化差异较大,便于二者的区分;阔叶林在5月的NDVI急剧上升,上升速度比农作物的快,到6月达到峰值,9月急剧下降,与不同植被类型的NDVI变化差异较大,便于阔叶林的识别;灌草丛的NDVI在1月逐渐下降,到2—4月处于平稳状态,在4月以后缓慢上升,到7月急剧上升,8月达到峰值,然后开始快速下降,与不同类型的植被NDVI变化有较大差异;草原的NDVI从4月开始上升,但上升速度远比其他植被的小,8月达到峰值,之后便开始下降,与灌草丛的差异较大,便于二者的区分;农作物的NDVI从5月开始急剧上升,9月达到峰值,然后急剧下降,与其他植被类型变化差异较大,便于农作物的识别;其他植被的NDVI从4月开始上升,7月达到峰值,然后逐渐下降,到9月急剧下降,与草原和灌草丛的NDVI变化有较大差异,在8月与草原和灌草丛具有一定的可分性;而非植被的NDVI全年波动变化不大,易于区分。因此,从不同植被类型的NDVI变化曲线看,不同植被类型的NDVI变化趋势差异较大,便于识别。其中森林类型内部的针叶林、针阔混交林、阔叶林之间的NDVI变化也较大,便于森林类型的识别,可以有效弥补原始光谱特征对森林类型识别能力的不足,增加不同植被类型之间的可分性。

4.2 分类结果与精度评价

通过RF特征选择算法对43个特征进行筛选,特征信息的筛选过程和结果如图5所示。从图5可以看出,随着特征维数的依次增加,沙地及其周边

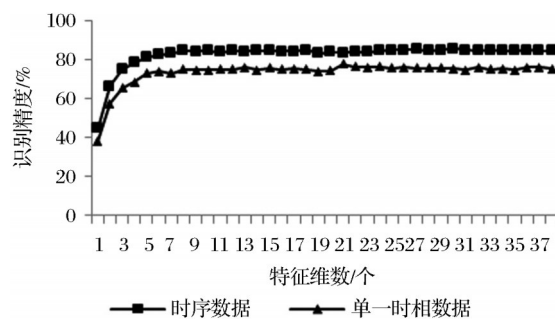


图5 特征维数与识别精度的关系

Fig.5 Feature dimension and the identification accuracy

的不同类型植被的识别精度逐渐提高,特征维数达到5个时,识别精度开始趋于稳定,其他特征的加入,对识别精度的影响非常小。因此将这5个特征作为研究区不同类型植被识别的特征。其中,获得光谱特征2个(B11和B12),植被指数1个(EVI),物候特征1个(全年NDVI变化),地形特征1个(海拔)。将通过特征选择后得到的5个特征,用于不同植被类型的识别,得到单一时相数据和时序数据的识别精度,结果如表3所示。由表3可以看出,利用单一时相数据对不同植被类型识别总体精度为74.36%,Kappa系数为0.66,整体精度较低。但加入物候特征后的时序数据总体精度达到84.37%,Kappa系数达到0.8,总体精度提高了10.01%,并且生产者精度和用户精度均得到提高,特别是针阔混交林与其他植被的生产者精度和用户精度提高超过了10%。时序数据加入物候特征可以有效提高不同植被类型的识别精度,改善沙地及其周边植被类型识别效果。

将单一时相数据的不同植被类型识别结果与时序数据的主要植被类型识别结果局部放大(图6)对比发现:单一时相数据中森林和农作物破碎度严重,边界比较模糊,大部分草原被错分为农作物和灌草丛,而农作物也被错分为草原和灌草丛。相对而言,加入物候信息的时序数据降低了森林和农作物类型识别结果的破碎度,使阔叶林和农作物的边界更加清晰,还减少草原与农作物之间的混分,有效改善了沙地及其周边植被类型的识别效果。

4.3 呼伦贝尔沙地及其周边植被类型空间分布格局

通过GEE平台自带的RF分类器进行呼伦贝尔沙地及其周边植被类型的识别,得到不同植被类型的空间分布格局,结果如图7所示,其中各类型的占比统计结果如图8所示。

由图7和图8可以看出,针叶林主要分布于鄂温克族自治旗,分布区域较少,约为研究区面积的

表 3 不同植被类型识别的精度评价表
Table 3 Different vegetation type recognition accuracy evaluation

数据	总体精度 /%	Kappa 系数	类别	生产精度 /%	用户精度 /%
单一时相 数据	74.36	0.66	针叶林	60.61	64.52
			阔叶林	71.79	51.85
			针阔混交林	31.82	50.00
			灌草丛	62.68	62.07
			草原	76.92	66.67
			农作物	72.86	79.69
			其他植被	12.5	50.00
			非植被	87.21	86.89
时序数据	84.37	0.80	针叶林	76.00	77.56
			阔叶林	85.19	67.65
			针阔混交林	43.75	77.78
			灌草丛	78.07	72.92
			草原	93.75	80.54
			农作物	84.78	86.67
			其他植被	31.82	87.50
			非植被	91.04	94.31

1.51%；针阔混交林分布于鄂温克族自治旗的针叶林和阔叶林的过渡区域,约为研究区面积的 1.7%；阔叶林则分布在陈巴尔虎旗和鄂温克族自治旗,是

3 种森林类型中分布最多的 1 种植被类型,约为研究区面积的 9.76%；灌草从主要分布于沙地及其周边地形较为低矮的区域,呈条状分布,占研究区的面积约为 11.63%；草原分布于研究区内的每个旗,分布面积大,约 55.30% 的区域属于草原；农作物则沿陈巴尔虎旗和鄂温克族自治旗的林草过渡带分布,但陈巴尔虎旗分布比较多,是呼伦贝尔主要的农作物种植区之一,约占研究区面积的 6.09%；其他植被在研究区内分布最少,主要分布于沼泽和湖泊周围,大约为研究区面积的 0.04%。

5 讨 论

以上研究表明,基于 GEE 云平台,利用机器学习算法结合植被的物候信息,可以有效进行呼伦贝尔沙地及其周边主要的植被类型识别。受数据源与研究方法的影响,导致研究仍然存在一些不足,还需在今后的研究中做进一步改进。

(1) 鉴于近红外波段对植被特征反映的敏感性,本研究仅使用了近红外 B8 波段的纹理信息,而其他波段的纹理信息没有得到充分利用。本研究发现红外和近红外波段信息用于研究区植被类型

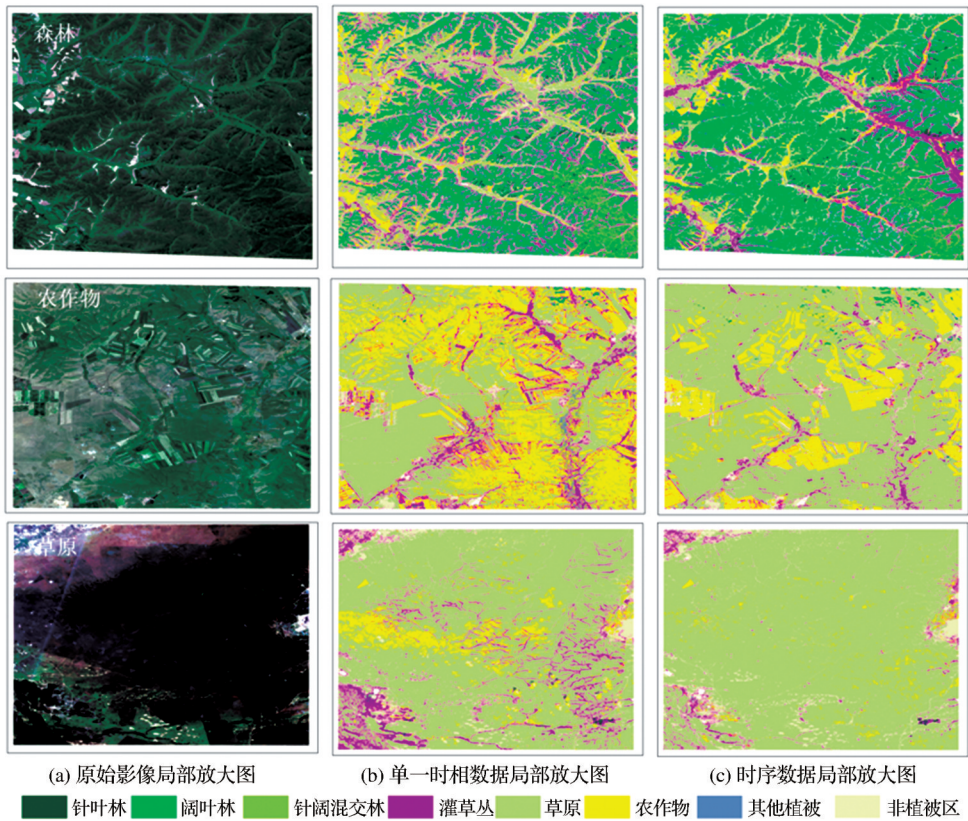


图 6 不同时相的识别结果局部放大图

Fig. 6 Partial enlargement diagram of recognition results in different temporal

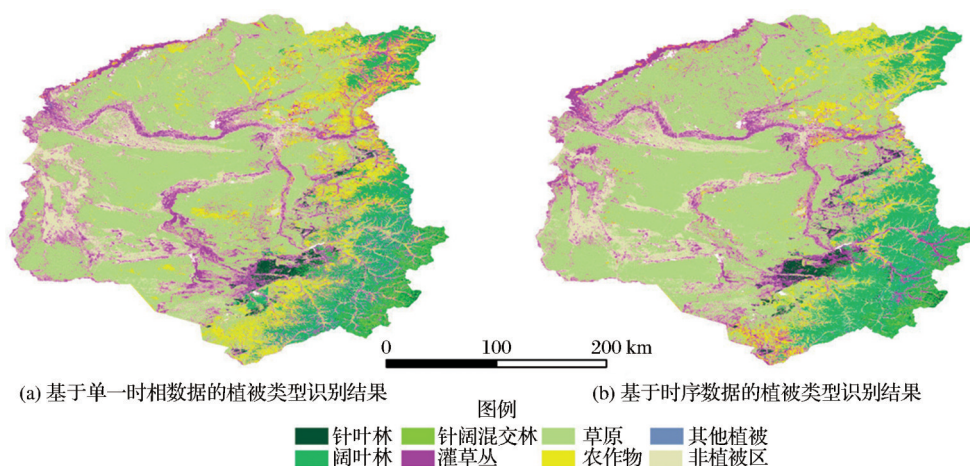


图 7 不同时相数据的植被类型识别结果

Fig.7 Vegetation type recognition results of different temporal data

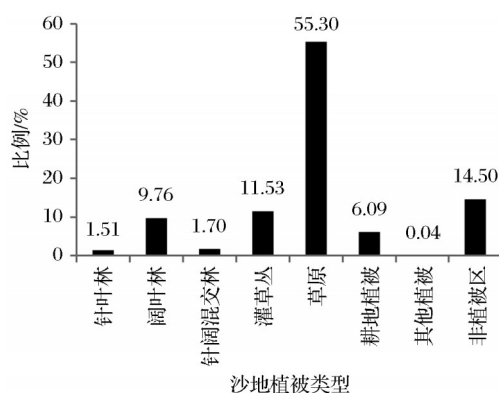


图 8 不同植被类型的比例

Fig.8 Proportion of different vegetation types

识别能力有限,仅在 EVI 中有所体现,可能其他波段的纹理信息对沙地及其周边植被类型具有好的效果,因此使用不同波段的纹理信息对研究区内主要植被类型识别还有待进一步的研究。

(2)不同植被类型的物候信息均是直接通过 NDVI 的变化趋势来得到,具有时间序列长的优势。有研究表明,在数学知识的基础上,根据整个 NDVI 曲线重新定义物候指标,提取了不同植被类型的生长季始期、生长季末期、生长季长度、NDVI 振幅、NDVI 左导数、NDVI 右导数、生长季期间的 NDVI 积分等物候参数,对植被的识别效果也有显著的提高^[40],下一步考虑使用更多物候参数进行沙地及其周边不同植被类型的识别。

(3)时序数据通过利用植被在不同时期变化的差异性,虽然能在很大程度上克服植被类型识别过程中外界因素的影响,提高了识别效果。但沙地及其周边植被交叉分布,容易产生较多的混合像元,导致草地、灌草丛、针叶林、阔叶林和针阔混交林仍

存在一定的错分和误分现象。此外,本研究获取的样本量虽然较大且具备代表性,但实地数据空间分布不够均匀,对植被类型识别结果存在一定的影响,今后通过增加研究区内均匀分布的样本,结合多源遥感数据融合,降低混合像元的影响,提高沙地及其周边植被类型的识别精度。

6 结 论

本研究以呼伦贝尔沙地及其周边植被类型为研究对象,基于 GEE 云平台 and 其中存储的 Sentinel-2 为数据源,采用 RF 分类方法并顾及物候信息差异,实现沙地及其周边不同植被类型的识别。对比不同时相数据的识别结果,探索了物候特征对于研究区植被类型识别的作用,得到以下结论:

(1)Sentinel-2 影像的光谱信息和近红外波段的纹理信息对研究区的主要植被类型识别能力有限,而物候特征有效弥补了原始光谱特征对研究区不同植被类型区分能力的不足。

(2)采用时间序列 Sentinel-2 影像和 GEE 平台可以有效地提高研究区沙地及其周边植被类型识别。

(3)呼伦贝尔沙地及其周边主要植被类型的物候特征差异明显,有助于不同类型植被的空间识别,特别是能够提高草原与农作物之间的混分,提高灌草丛、草原及其与农作物之间的识别精度。

尽管利用 Sentinel-2 数据的高时空分辨率和 GEE 云平台强大的数据处理能力对沙地等大区植被类型的识别具有较大的潜力,但本文中仅对 RF 方法进行了探索,采用其他机器学习方法的识别潜力仍然需要进一步研究;另外采用其他光学和

微波遥感数据源和 GEE 平台在沙地及其周边植被类型识别中的潜力在本文中并未涉及,未来可进一步探索不同时间序列数据源及 GEE 平台的潜力。

参考文献 (References):

- [1] Tu Zhifang, Li Mengxian, Sun Tao. The status and trend analysis of desertification and sandification[J]. *Forest Resources Management*, 2016(1):1-5,13.[屠志方,李梦先,孙涛.第五次全国荒漠化和沙化监测结果及分析[J].林业资源管理,2016(1):1-5,13.]
- [2] Yao Xueling, Jiang Lina, Li Long, *et al.* Biomass simulation of six shrub species in otindag sandy land[J]. *Acta Ecologica Sinica*, 2019, 39(3):905-912.[姚雪玲,姜丽娜,李龙,等.浑善达克沙地6种灌木生物量模拟[J].生态学报,2019,39(3):905-912.]
- [3] Jia Kun, Li Qiangzi, Tian Yichen, *et al.* A review of classification methods of remote sensing imagery[J]. *Spectroscopy and Spectral Analysis*, 2011,31(10):2618-2623.[贾坤,李强子,田亦陈,等.遥感影像分类方法研究进展[J].光谱学与光谱分析,2011,31(10):2618-2623.]
- [4] Li Shihua, Wang Jinliang, Bi Yan, *et al.* A review of methods for classification of remote sensing images[J]. *Remote Sensing for Land & Resources*, 2005,17(5):1-6.[李石华,王金亮,毕艳,等.遥感图像分类方法研究综述[J].国土资源遥感,2005,17(5):1-6.]
- [5] Yang Chao, Wu Guofeng, Li Qingquan, *et al.* Fast map rectification algorithm based on error homogeneous region[J]. *Geography and Geo-Information Science*, 2018, 34(4):24-32.[杨超,邹国锋,李清泉,等.植被遥感分类方法研究进展[J].地理与地理信息科学,2018,34(4):24-32.]
- [6] Xu Dawei. Dynamic change and analysis of different grassland types distribution in the Hulunber Grassland[D]. Beijing: Chinese Academy of Agricultural Sciences Dissertation, 2019.[徐大伟.呼伦贝尔草原区不同草地类型分布变化及分析[D].北京:中国农业科学院,2019.]
- [7] Zhang Ying. Classification of forest types based on multi-dimensional features using multi-seasonal Landsat-8 OLI remote sensing images[D]. Beijing: Beijing Forestry University, 2018.[张莹.基于多季相Landsat-8 OLI影像的多特征森林植被分类[D].北京:北京林业大学,2018.]
- [8] Zhang Yan. Study of remote sensing image classification based on machine learning[D]. Beijing: Beijing Forestry University, 2014.[张雁.基于机器学习的遥感图像分类研究[D].北京:北京林业大学,2014.]
- [9] Zhu Sha. Remote sensing image classification based on machine learning[D]. Ganzhou: Jiangxi University of Science and Technology, 2017.[朱莎.基于机器学习的遥感图像分类研究[D].赣州:江西理工大学,2017.]
- [10] Foody G M, Mathur A. Toward intelligent training of supervised image classifications: Directing training data acquisition for SVM classification[J]. *Remote Sensing Environment*, 2004(93):107-117. DOI: 10.1016/j.rse.2004.06.017.
- [11] Ding Xiangyuan, Gao Zhihai, Sun Bin, *et al.* Sandy lands classification using GF-1 time series NDVI data[J]. *Remote Sensing for Land and Resources*, 2017, 29(3):196-202.[丁相元,高志海,孙斌,等.基于高分一号时间序列数据的沙化土地分类[J].国土资源遥感,2017,29(3):196-202.]
- [12] Liang Shuang, Gong Zhaoning, Zhao Wenji, *et al.* Information extraction of Baiyangdian wetland based on multi-season Sentinel-2 Images[J]. *Remote Sensing Technology and Application*, 2021, 36(4):777-790.[梁爽,宫兆宁,赵文吉,等.基于多季相Sentinel-2影像的白洋淀湿地信息提取[J].遥感技术与应用,2021,36(4):777-790.]
- [13] Wakulińska M, Marcinkowska-Ochtyra A. Multi-temporal Sentinel-2 data in classification of mountain vegetation[J]. *Remote Sensing*, 2020,12(17):2696.DOI:10.3390/rs 12172696.
- [14] Sébastien Rapinel, Cendrine Mony, Lucie Lecoq, *et al.* Evaluation of Sentinel-2 time-series for mapping floodplain grassland plant communities[J]. *Remote Sensing of Environment*, 2019, 223(15):115-129. DOI: 10.1016/j.rse.2019.01.018.
- [15] Dong Jinwei, Li Shiwei, Zeng Yelu, *et al.* Remote sensing cloud computing and scientific analysis-applications and practices[M]. Beijing: Science Press, 2020:8-11.[董金玮,李世卫,曾也鲁,等.遥感云计算与科学分析一应用与实践[M].北京:科学出版社,2020:8-11.]
- [16] Chen Kangming, Zhu Xudong. Remote sensing of spatio-temporal dynamics of saltmarsh vegetation along south China coast based on Google Earth Engine[J]. *Remote Sensing Technology and Application*, 2021, 36(4):751-759.[陈康明,朱旭东.基于Google Earth Engine的南方滨海盐沼植被时空演变特征分析[J].遥感技术与应用,2021,36(4):751-759.]
- [17] Long Shuang, Guo Zhengfei, Xu Li, *et al.* Spatiotemporal variations of fractional vegetation coverage in China based on Google Earth Engine[J]. *Remote Sensing Technology and Application*, 2020, 35(2):326-334.[龙爽,郭正飞,徐粒,等.基于Google Earth Engine的中国植被覆盖度时空变化特征分析[J].遥感技术与应用,2020,35(2):326-334.]
- [18] Wang Xiaona, Tian Jinyan, Li Xiaojuan, *et al.* Benefits of Google Earth Engine in remote sensing[J]. *National Remote Sensing Bulletin*, 2022,26(2):299-309.[王小娜,田金炎,李小娟,等. Google Earth Engine云平台对遥感发展的改变.遥感学报,2022,26(2):299-309.]
- [19] Deng Fang, Chen Jie. Dynamic analysis and quality evaluation of Hulunbuir Sandy Land[J]. *Inner Mongolia Forestry Investigation and Design*, 1997, 11(1):34-37.[邓芳,陈杰.呼伦贝尔沙地动态分析与质量评价[J].内蒙古林业调查设计,1997,11(1):34-37.]
- [20] Bai Yan, An Huijun, Zhang Honglei, *et al.* Remote sensing monitoring of vegetation changes in Hulunbeier Sandy Land[J]. *Forest Resources Management*, 2013(3):134-137.[白彦,安慧君,张红蕾,等.呼伦贝尔沙地植被变化遥感监测[J].林业资源管理,2013(3):134-137.]

- [21] Dong Jianlin, Ya Jie. Analysis on the changes of land desertification in hulunbeier sandy land area over the last 10 years [J]. Forest Resources Management, 2002(4): 39-43. [董建林, 雅洁. 呼伦贝尔沙地近十年来土地沙漠化变化分析[J]. 林业资源管理, 2002(4): 39-43.]
- [22] Ki Hyung Park, Qu Z Q, Wan Q Q, *et al.* Effects of enclosures on vegetation recovery and succession in Hulunbeier steppe, China [J]. Forest Science and Technology, 2013, 9(1): 25-32. DOI: 10.1080/21580103.2013.774124.
- [23] Zhang Yujing. Remote sensing monitoring of Hulunbuir grassland phenology and the correlativity between phenology and climatic factors [D]. Beijing: Chinese Academy of Agricultural Sciences Dissertation, 2018. [张玉静. 呼伦贝尔草原物候遥感监测及其与气候因子的关系[D]. 北京: 中国农业科学院, 2018.]
- [24] Bai Yan. Study on vegetation coverage change of Hulunbeier sand land by remote sensing [D]. Hohhot: Inner Mongolia Agricultural University, 2013. [白彦. 呼伦贝尔沙地植被覆盖度变化遥感监测研究[D]. 呼和浩特: 内蒙古农业大学, 2013.]
- [25] Zhu Xiaoyu, Xu Dawei, Xin Xiaoping, *et al.* The spatial-temporal distribution of different grassland types in hulunber grassland based on remote sensing from 1992 to 2015 [J]. Scientia Agricultura Sinica, 2020, 53(13): 2715-2727. [朱晓昱, 徐大伟, 辛晓平, 等. 1992—2015年呼伦贝尔草原区不同草地类型分布时空变化遥感分析[J]. 中国农业科学, 2020, 53(13): 2715-2727.]
- [26] Kang Jun. Fine vegetation classification and fractional vegetation cover retrieval using time series remote sensing data [D]. Beijing: University of Chinese Academy of Sciences (Institute of Remote Sensing and Digital Earth, China), 2018. [康峻. 基于时间序列遥感数据的植被精细分类与覆盖度反演研究[D]. 北京: 中国科学院大学 (中国科学院遥感与数字地球研究所), 2018.]
- [27] Chen Yansi, Huang Chunlin, Hou Jinliang, *et al.* Extraction of maize planting area based on multi-temporal Sentinel-2 imagery in the middle reaches of Heihe river [J]. Remote Sensing Technology and Application, 2021, 36(2): 324-331. [陈彦四, 黄春林, 侯金亮, 等. 基于多时相 Sentinel-2 影像的黑河中游玉米种植面积提取研究[J]. 遥感技术与应用, 2021, 36(2): 324-331.]
- [28] Wu Zhengyi. Chinese vegetation [M]. Beijing: Science Press, 1980: 143-682. [吴征镒. 中国植被 [M]. 北京: 科学出版社, 1980: 143-682.]
- [29] Zhang Zengxiang, *et al.* Monitoring land cover in China [M]. Beijing: Planet Map Press, 2010: 5-6. [张增祥, 等. 中国土地覆盖监测 [M]. 北京: 星球地图出版社, 2010: 5-6.]
- [30] Zhang Bo. Texture feature extraction analysis of remote sensing image based on gray level co-occurrence matrix [J]. Cultural Geography, 2017(22): 190-200. [张博. 基于灰度共生矩阵的遥感影像纹理特征提取分析[J]. 城市地理, 2017(22): 190-200.]
- [31] He Zhaoxin, Zhang Miao, Wu Bingfang, *et al.* Extraction of summer crop in Jiangsu based on Google Earth Engine [J]. Journal of Geo-information Science, 2019, 21(5): 752-766. [何昭欣, 张淼, 吴炳方, 等. Google Earth Engine 支持下的江苏省夏收作物遥感提取[J]. 地球信息科学学报, 2019, 21(5): 752-766.]
- [32] Haralick R M, Shanmugam K, Dinstein I. Textural features for image classification [J]. IEEE Transactions on systems, man, and cybernetics, 1973, 3(6): 610-621. DOI: 10.1109/TSMC.1973.4309314.
- [33] Wang Minyu, Luo Yi, Zhang Zhengyang, *et al.* Recent advances in remote sensing of vegetation phenology: Retrieval algorithm and validation strategy [J]. National Remote Sensing Bulletin, 2022, 26(3): 431-455. [王敏钰, 罗毅, 张正阳, 等. 植被物候参数遥感提取与验证方法研究进展[J]. 遥感学报, 2022, 26(3): 431-455.]
- [34] Li Zhengguo, Tang Huajun, Yang Peng, *et al.* Progress in remote sensing of vegetation phenology and its application in agriculture [J]. Chinese Journal of Agricultural Resources and Regional Planning, 2012, 33: 20-28. [李正国, 唐华俊, 杨鹏, 等. 植被物候特征的遥感提取与农业应用综述[J]. 中国农业资源与区划, 2012, 33: 20-28.]
- [35] Wei X Q, Gu X F, Tao Y, *et al.* Land-cover classification using multi-temporal GF-1 wide field view data [J]. Taylor & Francis, 2018, 39(20): 6914-6930. DOI: 10.1080/01431161.2018.1468106.
- [36] Bi Kaiyi, Niu Zheng, Huang Ni, *et al.* Identifying vegetation with decision tree model based on object oriented method using multi-temporal Sentinel-2A images [J]. Geography and Geo-Information Science, 2017, 33(5): 16-20, 27, 127. [毕恺艺, 牛铮, 黄妮, 等. 基于 Sentinel-2A 时序数据和面向对象决策树方法的植被识别[J]. 地理与地理信息科学, 2017, 33(5): 16-20, 27, 127.]
- [37] Yao Dengju, Yangjing, Zhan Xiaojuan. Feature selection algorithm based on random forest [J]. Journal of Jilin University (Engineering and Technology Edition), 2014, 44(1): 137-141. [姚登举, 杨静, 詹晓娟. 基于随机森林的特征选择算法[J]. 吉林大学学报(工学版), 2014, 44(1): 137-141.]
- [38] Lu D, Weng Q. A survey of image classification methods and techniques for improving classification performance [J]. International Journal of Remote Sensing, 2007, 28(5): 823-870. DOI: 10.1080/01431160600746456.
- [39] Huang D M, Xu S J, Sun J Q, *et al.* Accuracy assessment model for classification result of remote sensing image based on spatial sampling [J]. Journal of Applied Remote Sensing, 2017, 11(4): 1-13. DOI: 10.1117/1.JRS.11.046023.
- [40] Ping Yuepeng. Crop classification based on analysis of phenological characteristics of MODIS times serie [D]. Harbin: Harbin Normal University, 2016. [平跃鹏. 基于 MODIS 时间序列地表物候特征分析及农作物分类[D]. 哈尔滨: 哈尔滨师范大学, 2016.]

Recognition of Vegetation Types in Hulunbuir Sandy Land and Its Surrounding Areas based on GEE Cloud Platform and Sentinel-2 Time Series Data

Yang Xianbao^{1,2}, Zhang Wangfei¹, Sun Bin^{2,3}, Gao Zhihai^{2,3},
Li Yifu^{2,3}, Wang Han^{2,3}

(1.College of Geography and Ecotourism, Southwest Forestry University, Kunming 650224, China;

2.Institute of Forest Resource Information Techniques, Chinese Academy of Forestry, Beijing 100091, China;

3.Key Laboratory of Forestry Remote Sensing and Information System, Beijing 100091, China)

Abstract: Sand and its surrounding vegetation types play an important role in fixing dunes, preventing soil erosion and environmental management for sandy land. Identification of Sand and its surrounding vegetation types can objectively reflect the vegetation growth environment of sandy land and its surrounding areas, so as to provide a valuable reference for ecological restoration and the control policies formulating of sandy land. With huge amount of long-term earth observation data and powerful cloud computing capabilities, Google Earth Engine (GEE) cloud platform provides a convenient way for the identification of vegetation types in a large areas. In this study, based on the Sentinel-2 time series data of 2019 stored in the GEE cloud platform, the applied potentialities of GEE cloud platform in vegetation types identification was explored by combining the RF algorithm and vegetation phenology information in Hulunbuir sandy land and its surroundings. Results showed that: ① The spectral information of Sentinel-2 image and the texture information obtained from the near-infrared band have limited ability to identify the main vegetation types in the study area, but the phenological characteristics effectively make up for this shortcoming; ② Accuracy of the vegetation types identification method achieved by the RF algorithm and considering the phenological characteristics extracted from the long time series remote sensing data is 84.37% (with the Kappa coefficient of 0.8), which is 10.01% higher than that identification result acquired based on single-phase data; ③ Phenological characteristics of the main vegetation types in the Hulunbuir sandy land and its surroundings show significant differences, which is helpful for the identification of the vegetation types, especially to improve the recognition accuracy of shrubs and grassland. The research shows that the use of Sentinel-2 data and GEE cloud platform to identify vegetation types in large areas such as sandy land has great potential and broad application prospects.

Key words: GEE; Sentinel-2; Time series data; Hulunbuir sandy land; Identification of vegetation types