

引用格式: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.]  
doi:10.11873/j.issn.1004-0323.2021.2.0324

## 基于多时相 Sentinel-2 影像的黑河中游玉米种植面积提取研究

陈彦四<sup>1,2</sup>, 黄春林<sup>1</sup>, 侯金亮<sup>1</sup>, 韩伟孝<sup>1,2</sup>, 冯娅娅<sup>1,2</sup>, 李翔华<sup>1,2</sup>, 王 静<sup>3</sup>

(1. 中国科学院西北生态环境资源研究院 甘肃省遥感重点实验室, 甘肃 兰州 730000;

2. 中国科学院大学, 北京 100049;

3. 甘肃省食品检验研究院, 甘肃 兰州 730000)

**摘要:**玉米是黑河中游种植面积最大的农作物, 生长期需水量大、蒸散量高。准确获取玉米种植面积对该区域农作物种植结构调整、水资源合理规划有重要参考意义。基于2019年4月至9月 Sentinel-2 多时相影像, 采用随机森林算法开展了黑河中游玉米种植面积提取研究。研究方法分为两类——直接提取法和两步提取法。进一步探讨了多时间信息量对玉米种植面积提取精度的影响以及各输入特征参数在玉米面积提取过程中的重要性表现。结果表明: 基于 Sentinel-2 多时相影像, 直接提取法和两步提取法均可高精度地提取研究区玉米种植面积, 特别是两步提取法, 玉米分类总体精度可达 85.03%, F1\_Score 为 0.70, Kappa 系数为 0.83; 与单幅影像相比, 多时相影像可获取不同作物的物候信息, 有效减少作物错分/漏分, 提高作物分类精度。该方法对基于高分辨率光学影像结合机器学习方法获取具有高度异质性的作物信息具有重要的参考价值。

**关键词:**玉米种植面积; 多时相卫星影像; Sentinel-2; 随机森林

**中图分类号:** TP79 **文献标志码:** A **文章编号:** 1004-0323(2021)02-0324-08

### 1 引言

黑河中游是甘肃省中部河西走廊平原的主要部分<sup>[1]</sup>, 位于祁连山区的山前平原, 地形平坦、阳光充足、引水方便<sup>[2]</sup>。该区域气候干旱、降水量少、蒸散量大<sup>[3]</sup>。因此, 黑河中游绿洲以人工绿洲为主, 农作物种植结构具有高度异质性和多样性<sup>[4]</sup>。2002年开始玉米成为黑河中游甘州区、临泽县、高台县种植面积最大的农作物, 生长期需水量最大、蒸散量最高<sup>[5]</sup>, 准确获取玉米种植面积信息对该区域作物种植结构调整、水资源合理规划具有重要参考意义。

随着卫星遥感的快速发展, 遥感技术以其独特

的大面积观测能力在植被面积监测方面取得了巨大成就, 利用可监测农业信息的遥感卫星(例如 Landsat 卫星、SPOT 卫星、Terra 和 Aqua 卫星、高分 1 号卫星以及 Sentinel-2 卫星等)数据获取大范围农作物种植面积信息已成为一种重要手段<sup>[6-7]</sup>。过去几年农作物分类提取研究中, 卫星影像数据源以 Landsat 数据和 MODIS 数据为主<sup>[8-9]</sup>, 但是受限于数据空间分辨率较低, 即便 Landsat 数据具有较高的空间分辨率 30 m, 也存在混合像元的问题, 仍难以满足复杂异质性条件下的农作物种植面积制图精细制图的需求。Sentinel-2 是高分辨率多光谱成像卫星, 包括在同一轨道的两颗极轨卫星 S2A 和

收稿日期: 2020-06-18; 修订日期: 2021-03-19

基金项目: 中国科学院战略性先导科技专项(A类)(XDA19040500), 甘肃省重点研发计划项目(17YF1FA134)。

作者简介: 陈彦四(1992—), 女, 山西吕梁人, 硕士研究生, 主要从事农业遥感研究。E-mail: cheniansi@lzb.ac.cn

通讯作者: 黄春林(1979—), 男, 宁夏青铜峡人, 研究员/博导, 主要从事多源遥感数据同化及系统集成研究。E-mail: huangcl@lzb.ac.cn

S2B,两颗卫星互补,具有宽幅盖(290 km)、高重访周期(5天)、高空间分辨率(10 m)优势,可支持监测植被生长季节内变化<sup>[10]</sup>,大大减小混淆像元问题。因此本研究探讨高度异质性植被景观条件下,高分辨率的 Sentinel-2 影像在农作物种植面积提取研究中的应用潜力。

近几年,利用机器学习方法提取农作物空间分布的研究成为了热点。其中,基于逐像元的监督分类方法应用最为广泛,如决策树算法、最大似然分类、支持向量机以及随机森林(Random Forest, RF)等<sup>[11-12]</sup>,同时 K-means 聚类、迭代自组织非监督分类算法在无标签样本的大尺度区域分类中也具有一定优势<sup>[13-14]</sup>。已有研究表明,与最大似然分类器、支持向量机等作物类型提取方法相比,随机森林算法的精度更高<sup>[15]</sup>。故本文选择利用随机森林算法对具有高度异质性的黑河中游绿洲开展玉米种植面积提取研究。

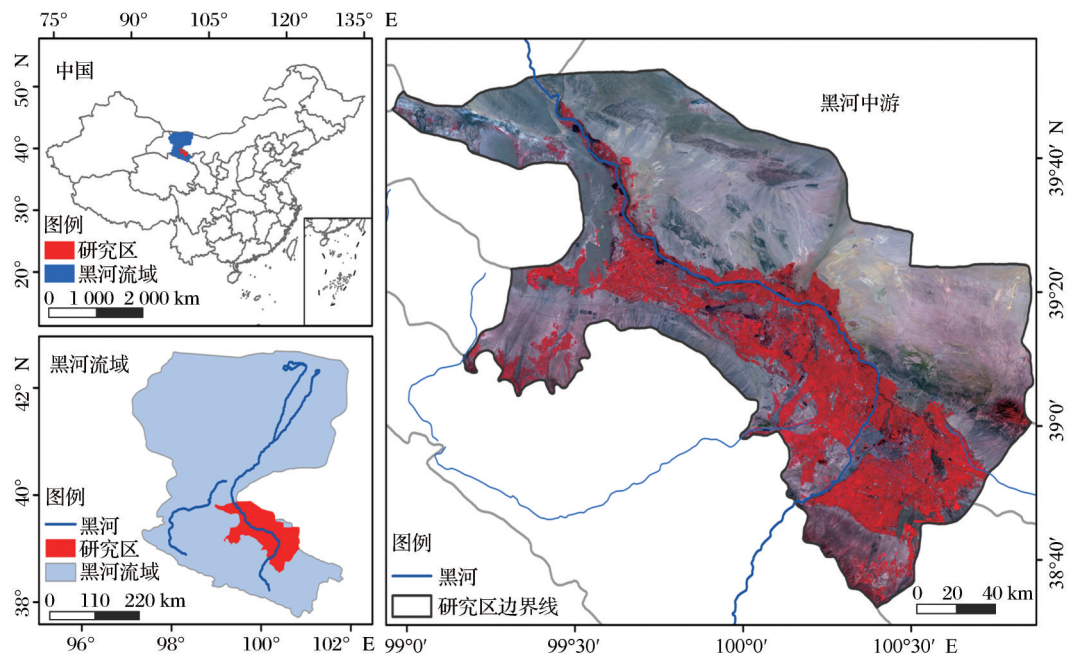
本文采用了两种不同的玉米种植面积提取方案,研究适合于复杂异质性植被景观条件下农作物种植面积提取的更优方法,并分析影响农作物种植

面积提取精度的主要因素。此外,还比较了基于单幅影像与多时相影像农作物种植面积提取结果差异,探讨多时间信息量对分类精度的影响。

## 2 研究区与数据

### 2.1 研究区概况

选取黑河中游张掖地区的甘州区、临泽县、高台县作为研究区(图1)。研究区是甘肃中部河西走廊平原的主要部分,地势平坦,为黑河绿洲农业集中区,气候干旱、年日照时数高、降水量少、蒸发量大的水热条件决定了该区域农业多为单季种植,农作物生长期主要集中在4~9月间。玉米是该区域种植面积最大、生长期需水量最大的农作物。本文选取该研究区开展玉米种植面积研究,主要基于以下3点考虑:①该绿洲农作物分布具有高度异质性和多样性,可为异质性区域作物提取提供参考;②该区域玉米种植收益为农民主要收入来源之一,可为种植结构提供参考;③该区域玉米种植生长期需水量最大,可为该区域水资源供应调整提供参考。



审图号:GS(2019)1823

图1 研究区概况

Fig.1 Overview of the study area

### 2.2 Sentinel-2 数据及预处理

选取2019年4月1日~9月30日研究区一个完整生长季上的428幅 Sentinel-2 MSI(Multi-spectral instrument,多光谱成像仪)Level-1C产品,该产品

提供大气顶层反射率影像,已经过辐射校正和几何校正。预处理过程包括重投影、重采样以及云掩膜。首先,将影像数据集重投影为 WGS 1984 投影坐标系;其次, Sentinel-2 影像包括13个波段,其中

B2/B3/B4/B8 波段空间分辨率为 10 m, B5/B6/B7/B8A/B11/B12 波段空间分辨率为 20 m, B1/B9/B10 波段空间分辨率为 60 m, 利用三次卷积插值法对所有较低空间分辨率波段重采样, 重采样后的空间分辨率为 10 m。利用 ESA 提供的 Sentinel-2 影像 QA60 波段对 428 幅影像质量进行评估, 掩去影像中云和卷云像元, 最终获得 6 幅月度中值合成影像。

2.3 地面参考数据

于 2019 年 10 月 22 日~11 月 1 日对研究区农作物分布开展了野外实地考察, 主要记录了玉米和非玉米植被(森林、麦子、辣椒、燕麦、苜蓿、油白菜、洋葱、西兰花、笋、土豆、甜菜等)。受采样时间限制, 非玉米植被样本的采集只对植被类别显示明显的区域进行了记录。共采集了 780 个样本点, 其中玉米样本点 623 个, 非玉米植被样本点 157 个点。考虑到非玉米植被采样点数量较少, 且黑河中游大棚种植基本上均为蔬菜种植, 因此在 Google Earth 卫星影像手动采集了 241 个大棚种植点, 且假设这些点为非玉米植被样本。另外在 Google Earth 还采集了林地样本 38 个, 草地/林地样本 333 个, 建设用地样本 529 个以及其他土地覆盖类型(例如裸地、稀疏植被、水体等)样本 667 个。除此之外, 为避免在 Google Earth 上采集水体样本点造成地理位置偏差, 在 Sentinel-2 影像采集了 336 个水体样本点。因此研究区内总共采样点为 2 896 个(详见表 1), 其中玉米 623 个, 非玉米植被 731 个, 非植被 1 542 个。

表 1 野外采集数据以及其他辅助数据

Table1 Information of field collection data and other auxiliary data

auxiliary data				
地物类型		采样点	分类总计	来源
玉米	玉米	623	623	野外采集
非玉米植被	蔬菜、麦子等	157	731	野外采集
	大棚种植	241		Google Earth
	草地、林地	333		Google Earth
非植被	建设用地	529	1 542	Google Earth
	水体	333		Sentinel-2
	其他	677		Google Earth

3 方 法

3.1 随机森林分类算法

随机森林是 2001 年 Breiman 提出的一种集成学习方法, 被广泛应用于统计分类和非参数回归问题。该算法主要过程是: ①利用 bootstrap 重抽样方

法对样本数据集随机采样, 每次生成包含 m 个样本的训练样本集, 采样共进行 n 次; ②对每个训练样本集随机选择特征集或特征集的线性组合进行决策树建模, 生成 n 个分类器, 即 n 棵树; ③通过投票从 n 棵决策树中选择最优的分类结果。随机森林算法能评估特征变量的重要性, 具有快速、易于参数化和较强鲁棒性的特点, 因此在具有许多不同特性的大规模数据集的分类运用中精度较高且能处理具有高维和噪声的输入数据, 目前该方法已在土地覆被分类和农作物制图应用研究中已表现出卓越的性能<sup>[16]</sup>。

3.2 玉米种植面积提取方案

研究构建两种不同基于随机森林算法的玉米种植面积提取方案, 具体技术路线如图 2 所示。方案一, 直接提取法, 即基于多时相影像和 3 类样本(玉米、非玉米植被、非植被)构建 RF 分类模型(简称为 RF-0), 直接提取研究区玉米种植面积。方案二, 两步提取法, 首先基于多时相 S2 影像和训练样本, 构建提取植被区的 RF 分类模型(简称为 RF-1), 得到植被区月度时间序列影像; 然后, 基于植被区月度时间序列影像和玉米/非玉米植被样本, 构建提取玉米种植面积的 RF 分类模型(简称为 RF-2)。方案一是目前农作物分类提取中常用的方法<sup>[17]</sup>, 方

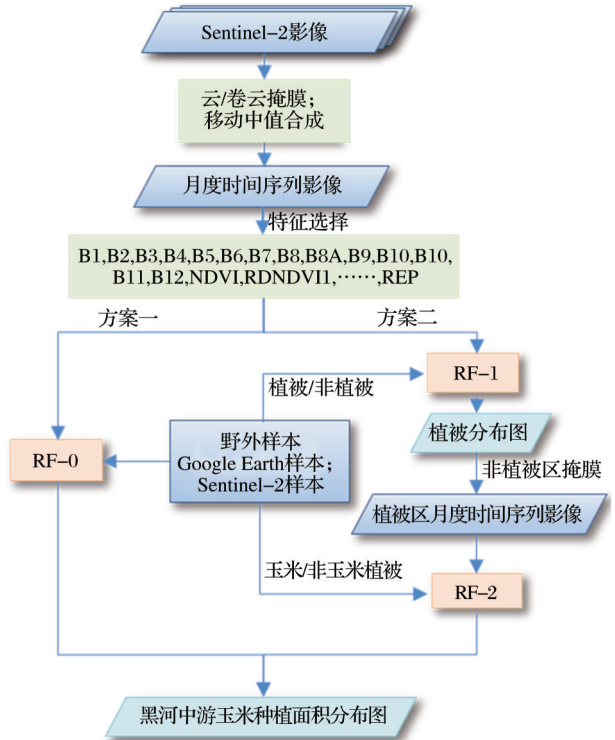


图 2 技术路线图

Fig.2 The flowchart of this study



案二将植被区独立后再对玉米种植面积进行提取。本文旨在通过比较两种方案得出适合于黑河中游玉米种植面积提取的方法。

### 3.2.1 特征选取

为了研究 Sentinel-2 影像中各个波段对植被信息提取的贡献性,选用了全部可用的 13 个光谱波段,包括可见光、近红外、短波红外波段。另外,还考虑了有利于植被区提取或与植被生长状况相关的 17 个植被指数。其中,NDVI(Normalized Difference Vegetation Index,归一化植被指数)用于定量光合能力、水分胁迫和植被生产力<sup>[18]</sup>,同时,时序 NDVI 能够观测不同土地覆被类型的物候信息<sup>[19]</sup>和植被特征<sup>[20]</sup>;NDWI(Normalized Difference Water Index,归一化水指数)是对地表水分含量敏感的植被指数<sup>[21]</sup>;EVI(Enhanced Vegetation Index,增强植被指数)可降低土壤反射率和大气污染,并增强对高覆盖植被区的敏感性<sup>[22]</sup>;GCVI(Green Chlorophyll Vegetation Index,叶绿素指数)的空间分布对研究区肥料(氮)等施用的响应具有重要意义<sup>[23]</sup>;STI(Soil Tillage Index,土壤耕作指数)和 NDTI(Normalized Difference Tillage Index,归一化耕作指数)是识别耕作土地的有效指数<sup>[24]</sup>,其他指数也经常被用于植被信息的提取。

本文共选取 30 个波段/指数,得到的月度合成影像共有 30×6 个波段/指数,因此,模型输入共有 180 个特征。

### 3.2.2 参数设定

在随机森林算法中,将一系列确定性伪随机数添加到样本数据集中,这些数字是双精度浮点数,范围是[0.0, 1.0),然后按照伪随机数小于 0.7 和大于等于 0.7 划分训练样本和验证样本,即 RF-0、RF-1、RF-2 模型的训练样本和验证样本个数分别为 2016/869、2069/816、900/394。对给定的训练数据集,使用字段随机分割方法对模型参数进行评估,在 RF-0/RF-1/RF-2 分类器中设置模型参数树的个数为 200,随机特征变量数为总特征数的平方根以及最小叶节点数为 10。

### 3.3 精度评估

构建混淆矩阵,定义总体精度(OA, Overall Accuracy)、F1\_Score 以及 Kappa 系数 3 个评价指数:

$$OA = \frac{\sum_{i=1}^n X_{ii}}{\sum_{i=1}^n \sum_{j=1}^n X_{ij}} \quad (1)$$

$$F1\_Score_i = \frac{2X_{ii}}{\sum_{i=1}^n X_{i+} + \sum_{i=1}^n X_{+i}} \quad (2)$$

$$Kappa = \frac{N \sum_{i=1}^n X_{ii} - \sum_{i=1}^n (X_{i+} \times X_{+i})}{N^2 - \sum_{i=1}^n (X_{i+} + X_{+i})} \quad (3)$$

其中:N 表示训练样本的个数,n 表示分类类别数,i 表示某一具体类别。OA 指在验证样本集上分类器预测正确的概率;F1\_Score 由用户精度和生产者精度计算得到,是用户精度和生产者精度调和均值,用于表征模型输出的好坏,取值范围位于 0~1 之间,其值越接近 1 表示模型输出越好;Kappa 系数是检测验证样本与预测结果吻合度的指标,其值的大小通常位于 0~1 之间,越接近 1 表示吻合度越高<sup>[25]</sup>。

## 4 结果

### 4.1 玉米种植面积提取评估

首先选择基于 Sentinel-2 的所有 180 个特征作为输入参数,分别训练两种研究方案下的随机森林分类模型,方案一中 RF-0 模型总体精度为 93.06%、Kappa 系数为 0.88;方案二中 RF-1 模型与 RF-2 模型的总体精度分别为 97.68% 和 89.78%,Kappa 系数分别为 0.95 和 0.79,表明通过两种方案构建的随机森林模型精度均较高。

进一步利用验证样本对所构建的模型进行精度验证,为了使两分类方案间具有可比性,对方案一中的混淆矩阵进行了处理,结果如表 2 所示。两种方案提取的植被区 OA、Kappa 和 F1\_Score 基本相同,OA 为 97%,Kappa 为 0.935 左右,以及植被类型 F1\_Score 为 0.97,表明两种研究方案均能较准确地提取植被区。在玉米种植面积提取中,方案一中 OA 为 84.48%,Kappa 为 0.68,玉米 F1\_Score 为 0.82;方案二中 OA 为 85.03%,Kappa 为 0.70,玉米 F1\_Score 为 0.83。与方案一相比,方案二中 RF-1 首先将植被区独立出来,RF-2 接着在植被区中分类出玉米作物,玉米错分为非植被类型的样本数明显减少,OA、Kappa 系数、玉米 F1\_Score 均明显高于方案一,此方案能有效降低分类过程中的总体类误差玉米种植面积提取研究。基于方案二的研究区玉米种植面积分布结果如图 3 所示。

研究区玉米种植面积提取过程受复杂因素影响,例如研究区作物灌溉方式大多为渠灌,受开渠

表 2 黑河中游玉米种植面积提取精度评估

Table 2 Evaluation of extraction precision of maize planting area in the middle reaches of Heihe River

植被区提取结果			OA	Kappa	F1_Score
方案一	植被	非植被	96.78%	0.94	0.97
植被	393	19			
非植被	9	448			
方案二	植被	非植被	96.32%	0.93	0.97
植被	252	15			
非植被	8	283			
玉米种植面积提取结果			OA	Kappa	F1_Score
方案一	玉米	非玉米	84.48%	0.68	0.82
玉米	190	34			
非玉米	27	142			
方案二	玉米	非玉米	85.03%	0.70	0.83
玉米	188	37			
非玉米	22	147			

时间等影响,地理位置相邻的大面积农作物灌溉时间相似,这部分农作物物候信息、生长健康状况相似,可能监测到的植被指数差异小,导致错分;玉米种植地边缘的杂草、套种玉米以及玉米地生长的其他果树、枣树等,这些因素均可能造成玉米与其他植被间光谱差异变小,导致分类困难;此外,山前平原土壤侵蚀严重,土质较差,而距离黑河较近的耕地土质较好,土质之间的差异导致玉米类内光谱差异变大从而导致玉米错分。

4.2 Sentinel-2 多时相影像分析

进一步采用增量式学习方法<sup>[26]</sup>分析多期月中值合成影像数据对分类准确性的影响。该方法在数据子集上训练初始模型,随后不断加入新的数据

更新该模型。首先以4月中值合成影像作为数据子集训练初始模型,并对模型进行验证,计算得出玉米的 Kappa;然后在此基础上加入5月份中值合成影像再次训练模型,并再次计算此模型分类中玉米的 Kappa。以此类推,不断加入新的数据直到新数据为9月份中值合成影像。基于单幅影像和多时相影像分类得到的总体分类精度如图4所示。基于单幅影像的分类中,4、5月和8月的玉米提取 Kappa 较高。在4月和5月该地区农田上基本压了薄膜,这些压薄膜的区域基本上种植玉米,因此玉米提取精度较高,但是由于薄膜和大棚信号很相似,导致大棚也被识别为玉米;在8月份,玉米处于抽穗期,生长旺盛,光谱特征与其它地物差异明显,因此提取精度较高。而6、7、9月精度较低,可能是由于6、7月份植被覆盖度整体较低,不同作物间光谱差异较小,存在同物异谱/异物同谱现象;9月部分成熟农作物被收割,导致不同作物间植被指数或波段值差异减小,玉米种植面积提取精度稍有下降。

基于多时相影像的分类,随模型输入影像数量的增加,玉米的 Kappa 明显提高。在不同月份中,基于多时相影像的提取精度均高于单幅影像的提取精度,主要是因为多时相影像能在很大程度上能克服基于单幅影像玉米提取过程中的外界因素影响,并提供更丰富的光谱信息,有效减少作物分类过程的错分、漏分现象,从而提高作物分类精度。在本研究中,到8月份为止,Kappa 值达到最高(0.70),比基于单幅影像的最佳分类精度高 2%。

4.3 随机森林分类中变量重要性评估

随机森林算法中共使用了 180 个变量,这些变

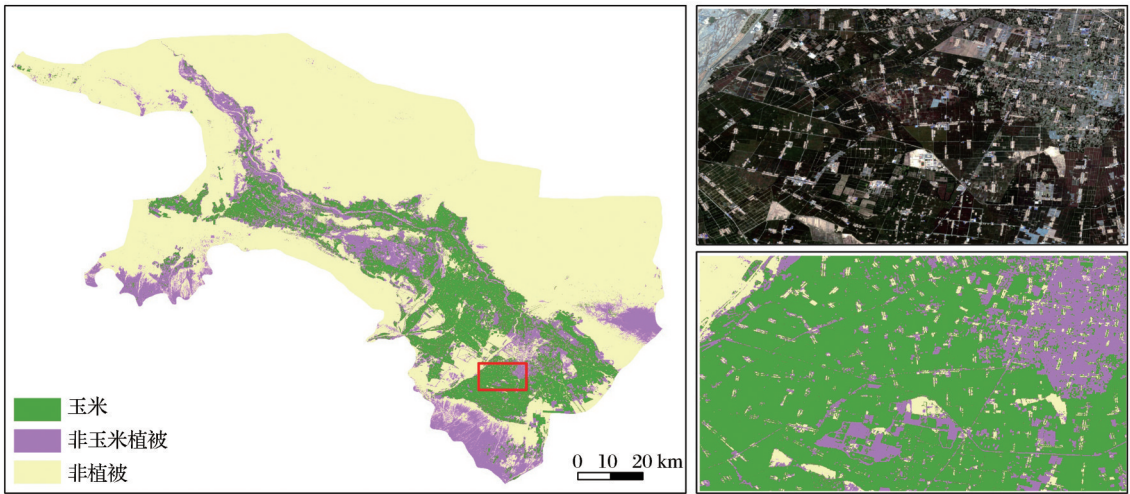


图 3 研究区玉米种植面积分布图(左图)以及局部放大 Sentinel-2 RGB 影像(右上)与分类结果比较(右下)

Fig.3 Distribution map of maize planting area in the middle reaches of Heihe River (left) and comparison of local images (top right) and classification results(bottom right)

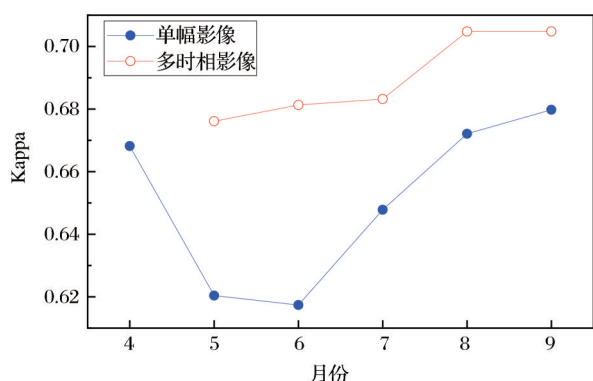
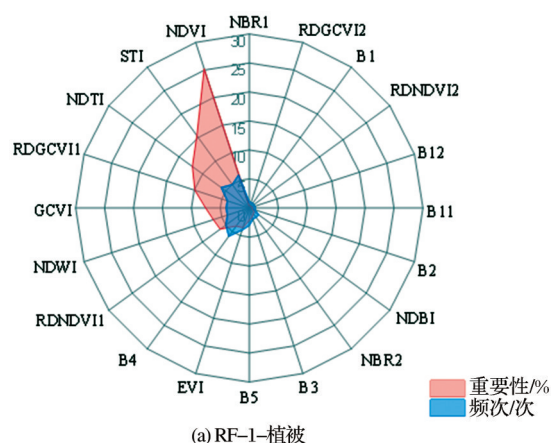


图4 玉米种植面积提取总体精度

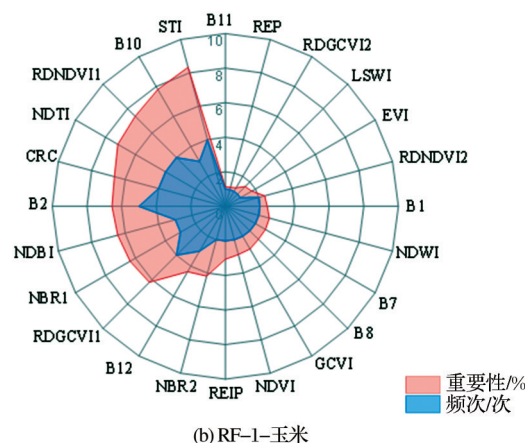
Fig.4 Overview of the overall accuracy of extraction of maize planting area

量在分类过程中的重要性各不相同。对 RF-1 和 RF-2 模型中 180 个变量进行特征重要性降序排序,并对排名前 60 变量统计其对应指数的重要性以及出现的频次(图 5)。RF-1 模型中排名前 4 的指数是 NDVI、STI、NDTI、RDGCVI1 (Red edge Green

Chlorophyll Vegetation Index 1, 红边叶绿素指数 1), 这 4 个指数与 B4、B5、B8、B11、B12 波段相关, 这些波段均可有效监测植被生长信息。其中 NDVI 指数排在第一位且重要性远高于其他指数或波段, 说明研究区植被提取在很大程度上依赖于 NDVI。与 RF-1 模型特征参数重要性排序相比, RF-2 模型使用指数数量明显增多, 这是由黑河中游复杂异质性植被结构决定的, 其中排名前 4 的指数为 STI、B10、NDT 1、RDNDVI1 (Red Edge Normalized Difference Vegetation Index 1, 红边归一化植被指数 1)。STI、NDTI 指数为 RF-1 和 RF-2 共有指数, 两指数均由 B11、B12 波段计算所得, 表明短波红外波段在植被信息提取中的重要性。另外, 值得注意的是, B10 波段属于“大气”短波红外波段, 具有水汽强吸收特性, 该波段对本研究中玉米识别具有重要作用, 这可能与玉米是研究区蒸散量最大的农作物相关。



(a) RF-1-植被



(b) RF-1-玉米

图5 随机森林模型输入指数的频次和重要性排序

Fig.5 Frequency and importance ranking of input index in Random Forest model

## 5 结 论

本文基于 Sentinel-2 多时相影像,通过搜集大量有利于研究的参数指数,采用随机森林算法开展了黑河中游玉米种植面积提取研究,得出以下结论:

(1) Sentinel-2 高分辨率多时相影像可用于复杂异质性景观条件下农作物种植面积精细提取,有效减少混淆像元的影响。本研究中玉米种植面积提取总体精度可达 85.03%, F1\_Score 为 0.70, Kappa 系数为 0.83, 可满足精细农业种植面积提取的需求,为该区域农作物种植结构调整、水资源合理规划提供数据支撑。

(2) 与单幅影像相比,多时相影像能监测不同

作物物候信息,获取更丰富的光谱信息,可有效减少作物分类过程中出现错分、漏分现象,提高作物分类精度。

(3) 玉米种植面积提取过程中,将非植被区独立出来可减小植被总体类内方差,有利于作物精准提取。作物灌溉的时空差异也可能导致作物类内光谱差异变大,而类间光谱差异减小,这些光谱上的差异使得玉米真实种植面积减小,同时伪种植面积增大。

## 参考文献 (References):

- [1] Chen Y, Lu D, Luo L, *et al.* Detecting Irrigation Extent, Frequency, and Timing in a Heterogeneous Arid Agricultural



- Region Using Modis Time Series, Landsat Imagery, and Ancillary Data [J]. *Remote Sensing of Environment*, 2018, 204: 197-211.
- [2] Lu Ling, Cheng Guodong, Li Xin. Study on the Landscape Pattern of Zhangye Oasis in the Middle Reaches of Heihe River Basin [J]. *Journal of Applied Ecology*, 2001, 12(1): 68-74. [卢玲, 程国栋, 李新. 黑河流域中游地区景观变化研究 [J]. *应用生态学报*, 2001, 12(1): 68-74.]
- [3] Wang Shuguo, Ma Chunfeng, Zhao Zebin, *et al.* Estimation of Soil Moisture of Agriculture Field in the Middle Reaches of the Heihe River Basin based on Sentinel-1 and Landsat 8 Imagery [J]. *Remote Sensing Technology and Application*, 2020, 35(1): 13-22. [王树果, 马春峰, 赵泽斌, 等. 基于 Sentinel-1 及 Landsat 8 数据的黑河中游农田土壤水分估算 [J]. *遥感技术与应用*, 2020, 35(1): 13-22.]
- [4] Jiao Yuanmei, Ma Mingguo, Xiao Duning. Study on the Landscape Pattern of Zhangye Oasis in the Middle Reaches of Heihe River Basin [J]. *Glacier and Frozen Soil*, 2003, 25(1): 94-99. [角媛梅, 马明国, 肖笃宁. 黑河流域中游张掖绿洲景观格局研究 [J]. *冰川冻土*, 2003, 25(1): 94-99.]
- [5] Zheng Luqian, Tan Minghong. A Comparative Study of Water Use Efficiency of Different Crops in the Middle Reaches of Heihe River and Its Implications for Planting Structure Adjustment [J]. *Journal of Geo-information Science*, 2016, 18(7): 977-986. [郑璐倩, 谈明洪. 黑河中游地区作物用水效率比较及种植结构调整方向研究 [J]. *地球信息科学学报*, 2016, 18(7): 977-986.]
- [6] Cabral A I R, Costa F L. Land Cover Changes and Landscape Pattern Dynamics in Senegal and Guinea Bissau Borderland [J]. *Applied Geography*, 2017, 82: 115-128.
- [7] Esmail M, Ali M, Negm A. Monitoring Land Use/Land Cover Changes Around Damietta Promontory, Egypt, Using RS/GIS [C] // 12th International Conference on Hydroinformatics (HIC)-Smart Water for the Future, South Korea, 2016, 154: 936-942.
- [8] Pea M A, Brenning A. Assessing Fruit-Tree Crop Classification from Landsat-8 Time Series for the Maipo Valley, Chile [J]. *Remote Sensing of Environment*, 2015, 171: 234-244.
- [9] Skakun S, Franch B, Vermote E, *et al.* Early Season Large-Area Winter Crop Mapping Using Modis NDVI Data, Growing Degree Days Information and a Gaussian Mixture Model [J]. *Remote Sensing of Environment*, 2017, 195 (15): 244-258.
- [10] Son N T, Chen C F, Chen C R, *et al.* Classification of Multi-temporal Sentinel-2 Data for Field-Level Monitoring of Rice Cropping Practices in Taiwan [J]. *Advances in Space Research*, 2020, 65(8): 1910-1921.
- [11] Kamthonkiat D, Honda K, Turrall H, *et al.* Discrimination of Irrigated and Rainfed Rice in a Tropical Agricultural System Using Spot Vegetation Ndvi and Rainfall Data [J]. *International Journal of Remote Sensing*, 2007, 26(12): 2527-2547.
- [12] Ozdogan M, Gutman G. A New Methodology to Map Irrigated Areas Using Multi-temporal Modis and Ancillary Data: An Application Example in the Continental Us [J]. *Remote Sensing of Environment*, 2008, 112(9): 3520-3537.
- [13] Basukala A K, Oldenburg C, Schellberg J, *et al.* Towards Improved Land Use Mapping of Irrigated Croplands: Performance Assessment of Different Image Classification Algorithms and Approaches [J]. *European Journal of Remote Sensing*, 2017, 50(1): 187-201.
- [14] Ragetti S, Herberz T, Siegfried T. An Unsupervised Classification Algorithm for Multi-temporal Irrigated Area Mapping in Central Asia [J]. *Remote Sensing*, 2018, 10(11).
- [15] Gomez C, White J C, Wulder M A. Optical Remotely Sensed Time Series Data for Land Cover Classification: A Review [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2016, 116: 55-72.
- [16] Mitsopoulos I, Chrysafi I, Bountis D, *et al.* Assessment of Factors Driving High Fire Severity Potential and Classification in a Mediterranean Pine Ecosystem [J]. *Journal of Environmental Management*, 2019, 235: 266-275.
- [17] Hao P, Zhan Y, Wang L, *et al.* Feature Selection of Time Series MODIS Data for Early Crop Classification Using Random Forest: A Case Study in Kansas, USA [J]. *Remote Sensing*, 2015, 7(5): 5347-5369.
- [18] Panda S S, Ames D P, Panigrahi S. Application of Vegetation Indices for Agricultural Crop Yield Prediction Using Neural Network Techniques [J]. *Remote Sensing*. 2010, 2(3): 673-696.
- [19] Reed B C, Brown J F, VanderZee D, *et al.* Measuring Phenological Variability from Satellite Imagery [J]. *Journal of Vegetation Science*, 1994, 5(5): 703-714.
- [20] Huemmrich K F, Privette J L, Mukelabai M, *et al.* Time-Series Validation of Modis Land Biophysical Products in a Kalahari Woodland, Africa [J]. *International Journal Of Remote Sensing*, 2005, 26(9): 4381-4398.
- [21] Jeong S, Kang S, Jang K, *et al.* Development of Variable Threshold Models for Detection of Irrigated Paddy Rice Fields and Irrigation Timing in Heterogeneous Land Cover [J]. *Agricultural Water Management*, 2012, 115: 83-91.
- [22] Huete A R, Liu H Q, Batchily K, *et al.* A Comparison of Vegetation Indices over a Global set of TM Images for EOS-MODIS [J]. *Remote Sensing of Environment*, 1997, 59(3): 440-451.
- [23] Huete A, Didan K, Miura T, *et al.* Overview of the Radiometric and Biophysical Performance of the Modis Vegetation Indices [J]. *Remote Sensing of Environment*, 2002, 83(1-2): 195-213.
- [24] Deventer A P, Ward A D, Gowda P H, *et al.* Using Thematic Mapper Data to Identify Contrasting Soil Plains and Tillage Practices [J]. *Photogrammetric Engineering & Remote Sensing*

- ing, 1997, 63(1): 87-93.
- [25] Congalton R G. A Review of Assessing of the Accuracy of Classifications of Remotely Sensed Data[J]. Remote Sensing of Environment, 1991, 37(1): 35-46.
- [26] You N, Dong J. Examining Earliest Identifiable Timing of Crops Using all Available Sentinel 1/2 Imagery and Google Earth Engine[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2020, 161: 109-123.

## Extraction of Maize Planting Area based on Multi-temporal Sentinel-2 Imagery in the Middle Reaches of Heihe River

Chen Yansi<sup>1,2</sup>, Huang Chunlin<sup>1</sup>, Hou Jinliang<sup>1</sup>, Han Weixiao<sup>1,2</sup>, Feng Yaya<sup>1,2</sup>,  
Li Xianghua<sup>1,2</sup>, Wang Jing<sup>3</sup>

(1. *Key Laboratory of Remote Sensing of Gansu Province, Northwest Institute of Eco-Environment and Resources, Chinese Academy of Sciences, Lanzhou 730000, China;*  
2. *University of Chinese Academy of Sciences, Beijing 100049, China;*  
3. *Gansu Food Inspection and Research Institute, Lanzhou 730000, China*)

**Abstract:** Maize is the crop with the largest planting area in the middle reaches of the Heihe River, with large water requirements and high evapotranspiration during the growing period. Accurately obtaining the maize planting area has important significances for the adjustment of crop planting structure and reasonable planning of water resources in the region. The object of this paper is to assess the value of multi-temporal Sentinel-2 data for extraction of maize planting area in the middle reaches of the Heihe River from April to September 2019. The random forest algorithm was adopted in this work. The research methods were divided into two categories: extraction directly and two-step extraction. Further discussed the impact of multi-temporal information as input on the classification accuracy, and analyzed the importance of the input feature parameters of the model in the extraction process. The results showed that the two-step extraction method based on Sentinel-2 multi-temporal images could accurately extract the maize planting area in the study area with the overall classification accuracy of 85.03%, F1\_Score of 0.70, and Kappa coefficient of 0.83. Compared with single image, multi-temporal images could effectively improve the accuracy of crop classification, obtaining differently crop phenology information. The research demonstrates the value of obtaining highly heterogeneous crop information based on high-resolution optical image combined with machine learning method.

**Key words:** Maize planting area; Time series satellite imagery; Sentinel-2; Random forest