

# 基于 Sentinel-2 破碎化地块灌区作物种植结构的提取

吴迪<sup>1</sup>, 杨鹏<sup>2</sup>, 周黎勇<sup>3</sup>, 李芳松<sup>3</sup>, 李凌锋<sup>3</sup>, 张旭东<sup>2\*</sup>

(1.中国灌溉排水发展中心, 北京 100054; 2.沈阳农业大学 水利学院, 沈阳 110866;

3.新疆水利水电科学研究院, 乌鲁木齐 830049)

**摘要:**【目的】探究基于 Sentinel-2 遥感影像的决策树分类模型提取破碎化地块灌区作物种植结构的适用性。【方法】选取新疆阿拉沟灌区为研究区,以 2021 年覆盖作物全生育期的 Sentinel-2 遥感影像为数据源,结合田间调查和 Google 高清影像目视解译采样,基于主要作物物候信息、NDVI 时序特征等分析确定作物识别的关键期阈值,构建决策树模型进行灌区主要作物分类,并对分类结果精度验证。【结果】基于 Sentinel-2 提取的灌区种植结构分布图地块纹理清晰,能够满足灌区用水管理需要;构建的决策树分类模型可在灌区尺度实现作物分类,方法简便易行,总体精度达到 81.56%, Kappa 系数为 0.716 6。【结论】采用 Sentinel-2 遥感影像和决策树分类方法识别破碎化地块灌区复杂作物分类是可行的,可为灌区输配水决策和农业用水精细化管理提供基础信息。

**关键词:** Sentinel-2; 灌区作物分类; NDVI 时间序列; 决策树; 破碎化地块

中图分类号: TP79

文献标志码: A

doi: 10.13522/j.cnki.ggps.2022368

OSID:



吴迪, 杨鹏, 周黎勇, 等. 基于 Sentinel-2 破碎化地块灌区作物种植结构的提取[J]. 灌溉排水学报, 2023, 42(4): 74-80.

WU Di, YANG Peng, ZHOU Liyong, et al. Using Sentinel-2 Sensing Imagery to Estimate Planting Structure in Fragmented

Irrigated Lands[J]. Journal of Irrigation and Drainage, 2023, 42(4): 74-80.

## 0 引言

【研究意义】新疆地域广阔、水资源问题突出,灌区类型多样,其作物种植信息获取对灌区乃至区域水资源优化配置尤为重要。当前,灌区作物种类、种植面积等信息主要依靠统计和抽样调查获取,无法获取其空间分布,难以反映灌区实际情况,不能很好地满足用水管理需求。遥感技术作为灌区基础信息采集的一种有效手段,可弥补传统方法中数据获取客观性、时效性、经济性等方面不足,为灌区配水计划制定、水权分配、用水总量控制和定额管理等提供有力支撑<sup>[1]</sup>。【研究进展】基于光谱的影像分类可分为监督与非监督两类,目前监督分类方法较多,如决策树(Decision Tree, DT)、最大似然(Maximum Likelihood, ML)、神经网络(Neural Network, NN)、支持向量机(Support Vector Machine, SVM)等<sup>[2-5]</sup>。在众多分类方法中,决策树凭借其简单高效、逻辑性强,且能用于无规则、无次序的样本数据集等优点而成为一种较为常用的分类方法。目前常用的决策树算法有 ID3、CART、C4.5、C5.0、随机森林等,它们的学习和容错能力不同,对数据源及区域的适用性也

存在差异<sup>[6]</sup>。李鑫川等<sup>[7]</sup>使用决策树、支持向量机和最大似然分类法提取了黑龙江友谊农场大豆、玉米和水稻的种植信息,结果表明决策树分类效果优于支持向量机和最大似然分类法。刘淑琴等<sup>[8]</sup>使用面向对象的 C5.0 决策树算法提取江西南昌县土地利用类型信息,结果表明采用 C5.0 决策树算法在提取分类精度上明显优于最大似然法和 CART 决策树方法。在数据源方面,国内外常用的有 AVHRR、SPOT、MODIS、MERIS、Landsat、WorldView-2、HJ、GF、Sentinel 等,其时空分辨率、光谱波段信息、原始影像预处理、获取难易程度等各不相同,已应用于不同尺度和区域的作物面积提取研究中<sup>[9]</sup>。由于作物分布提取模型通常基于单个像元的光谱特征对像元所属作物类型进行分类,分类模型未考虑同一地块内邻近像元的关联关系,也未考虑遥感像元与实际地物间的对应关系<sup>[10]</sup>,尽管有些中低分辨率影像(AVHRR、SPOT、MODIS 和 MERIS 等)时间分辨率较高,但受到空间分辨率的限制,中低分辨率遥感数据的混合像元问题在地块边界尤为突出,给分类结果验证、种植面积核算带来较大不确定性,导致中低分辨率像元尺度的作物分布制图产品难以满足破碎化地块灌区用水管理实际需求。为解决目前卫星传感器不能同时获取高时间、高空间分辨率影像的缺陷,集成不同时空分辨率的多源遥感数据能够增强作物在物候和空间等特征上的差异信息,一定程度上克服了作物分类时“同物异谱,

收稿日期: 2022-06-27

基金项目: 新疆维吾尔自治区水利科技专项资金项目(XSKJ-2022-12)

作者简介: 吴迪(1978-),男,正高级工程师,主要从事全国大型灌区规划和项目管理、水资源高效利用研究。E-mail: wudisyau075@163.com

通信作者: 张旭东(1979-),男,副教授,博士,主要从事节水灌溉理论与技术和农业节水研究。E-mail: zxxddd@126.com

异物同谱”现象发生，但目前尚无一种普适性的时空融合算法，且模型比较复杂，不利于推广应用<sup>[11]</sup>。因此，综合考虑遥感数据时空分辨率、数据质量、数据预处理和获取便捷程度等因素，采用免费且时空分辨率较好的 Sentinel-2 遥感影像作为灌区尺度种植结构提取的数据源是较为理想的选择。【切入点】对于沙漠绿洲灌区破碎化地块种植结构提取的相关研究较少且尚未有一种高效、精准的作物种植结构识别方法，制约着灌区农业用水的精细化管理。【拟解决的关键问题】本研究以阿拉沟灌区为研究区，基于 Sentinel-2 遥感影像的时序植被指数数据集，构建决策树分类模型并提取灌区种植结构，探究时序植被指数结合作物关键生育期特征在破碎化地块灌区作物识别中的优势，以期构建一种高效、准确的作物种植结构识别方法，为灌区输配水决策和农业用水精细化管理提供客观、有效的作物空间分布信息。

## 1 材料与方法

### 1.1 研究区概况

阿拉沟灌区位于新疆吐鲁番盆地托克逊县，为自流引水大型灌区，设计灌溉面积 25 440 hm<sup>2</sup>，涉及博斯坦镇、托克逊镇、夏镇、郭勒布依乡、伊拉湖镇和阿乐惠镇等 6 个乡镇，如图 1 所示。灌区属典型的大陆性暖温带荒漠气候，多年平均降水量 6.4 mm，且主要集中在夏季，年蒸发量 3 744 mm，年平均气温 13.8 ℃，年平均日照时间 3 043~3 224 h，非常适合

喜温作物生长<sup>[12]</sup>。阿拉沟灌区属于典型的绿洲灌溉农业区，地块破碎化较为严重，以经济作物为主且种类丰富，其中葡萄、玉米、高粱、棉花、蔬菜、瓜果等种植面积较大。

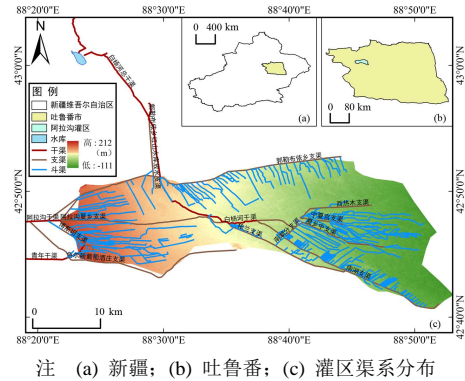


图1 阿拉沟灌区渠系分布

Fig.1 Distribution of canal system in Alagou irrigation area

### 1.2 数据来源及预处理

#### 1.2.1 数据来源

影像数据为 2021 年 1—10 月共 31 景 Sentinel-2 影像，下载于欧空局官网 (<https://scihub.copernicus.eu/>)。Sentinel-2 是欧空局“全球环境与安全管理”计划发射的卫星，Sentinel-2 系列卫星是高分辨率多光谱成像卫星，携带了多光谱成像仪 (Multispectral Imager, MSI)，飞行高度 786 km，覆盖 13 个光谱波段，幅宽达 290 km。地面分辨率分别为 10、20 m 和 60 m，1 颗卫星重访周期 10 d，2 颗卫星 (2A 和 2B) 互补，重访周期 5 d。

表 1 哨兵 2 号遥感影像信息

Table 1 Information of Sentinel-2 remote sensing images

序号	拍摄时间	传感器	云覆盖率/%	序号	拍摄时间	传感器	云覆盖率/%
1	20210105	S2B	9.29	17	20210520	S2A	0.42
2	20210110	S2A	4.58	18	20210525	S2B	0.00
3	20210115	S2B	5.23	19	20210604	S2B	15.14
4	20210120	S2A	17.63	20	20210624	S2B	0.00
5	20210125	S2B	4.38	21	20210629	S2A	0.49
6	20210130	S2A	0.34	22	20210704	S2B	0.00
7	20210219	S2A	2.01	23	20210724	S2B	0.26
8	20210301	S2A	5.50	24	20210818	S2A	4.58
9	20210321	S2A	9.93	25	20210823	S2B	0.19
10	20210331	S2A	6.53	26	20210907	S2A	0.22
11	20210405	S2B	9.07	27	20210912	S2B	10.33
12	20210410	S2A	13.22	28	20210927	S2A	2.44
13	20210415	S2B	1.06	29	20211012	S2B	1.77
14	20210420	S2A	12.08	30	20211022	S2B	0.17
15	20210430	S2A	0.33	31	20211027	S2A	0.49
16	20210505	S2B	1.69	-	-	-	-

本研究采用的 Sentinel-2 影像拍摄时间、传感器类型和云覆盖率详见表 1，该影像是 UTM/WGS84 投影下的 100 km×100 km 正射影像，其 L1C 级产品为经过正射校正和几何精校正的大气表观反射率产

品，并没有进行大气校正，因此，本文利用欧空局提供的 Sen2cor 插件对该产品进行大气校正和辐射定标。经处理后 L2A 产品包含经过大气校正的大气底层反射率数据，主要使用其蓝、绿、红以及近红

外 4 个波段, 其中蓝、绿、红波段合成真彩色影像用来检查研究区云的覆盖程度, 红与近红外波段用来合成归一化植被指数 (Normalized Difference Vegetation Index, *NDVI*) 影像<sup>[13]</sup>。由表 1 可知, 在 31 景影像中, 单景云覆盖率小于 20%, 其中 2021 年 1 月 21 日、6 月 4 日云覆盖率稍大, 分别为 17.63% 和 15.14%, 26 景影像云覆盖率均在 10% 以下。单景影像覆盖整个灌区, 裁剪后的灌区部分均为晴空状态, 遥感影像质量满足作物分类需要。

研究中利用 BIGEMAP 软件下载了 Google Earth 的 16 级高清影像数据, 该影像的像元分辨率为 2.39 m, 用于辅助提取感兴趣区。

### 1.2.2 地面调查数据

田间作物样本采集使用手机新知卫星地图进行实地定位并标记作物类型。尽可能选择作物类型一致、长势均匀的较大地块 (宽度 10 m 以上) 的作物类型作为样本点, 采集地块中心点坐标, 灌区地面调查样本点分布见图 2。

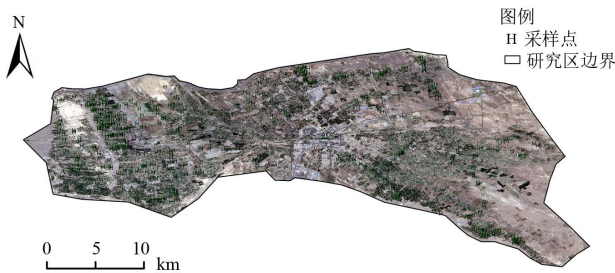


图2 灌区样本点分布

Fig.2 Spatial distribution of sample points in irrigated area

根据采样结果, 整理 10 类样本为复播 (瓜类复播玉米或高粱)、林果 (主要包括枣树、杏树和桃树等乔木类林果)、葡萄、棉花、蔬菜、设施农业、裸地-村庄、水体、草场、其他 (主要为未包含到以上地物类别中的, 包括地块边缘较窄的道路, 小片苜蓿、孜然、正播玉米、花生、芝麻、墓地及晒场等), 共计 588 个样本, 对 10 类样本进行可分离度检验, 属合格样本。为了对分类结果进行精度验证和评价, 将数量较多的 5 类样本 (复播、棉花、林果、葡萄和裸地-村庄) 拆分为训练样本和验证样本 (由于其余 5 类样本数量相对较少, 故不进行拆分), 随机选取训练样本, 采用训练样本作为掩膜提取验证样本, 使训练样本和验证样本相互独立, 拆分后的 5 类主要样本数量详见表 2。

表 2 训练样本和验证样本

Table 2 Training samples and verification samples

样本类型	复播	棉花	林果	葡萄	裸地-村庄
训练样本/个	75	17	70	20	75
验证样本/个	70	15	65	18	66

## 1.3 研究方法

决策树分类主要基于遥感影像等空间数据, 通过对其进行逻辑推理、数理统计、综合分析和归纳总结, 采用自上而下的递归方式, 形成分类规则进行分类。决策树分类采用一种二分递归分割技术, 将当前样本集分为 2 个子样本集, 使得生成的每个非叶子节点都有 2 个分支。决策树分类一般基于基尼系数不停地二分, 且一个特征可能会参与多次节点的建立<sup>[14]</sup>。实现面向对象决策树分类包括分类对象识别、决策树规则建立、规则输入、运行和精度评价 5 个步骤, 其中最关键环节是分类对象识别和决策树规则建立。决策树规则建立有多种方法, 如专家经验总结、传统统计方法等<sup>[15-17]</sup>。

阿拉沟灌区地块破碎化程度高、作物类型复杂, 针对这些特点, 利用不同作物生长的物候特征, 或不同时期的特征参数之间的差异, 通过适宜的分类规则和阈值设定, 能够进行作物类型识别, 因此采用决策树进行分类。主要过程为采用多时相的 Sentinel-2 遥感影像构建 *NDVI* 时序数据集, 根据实地调查作物分布及结构结合目视解译选取分类样本, 分析各类样本的 *NDVI* 特征及阈值, 结合作物关键期物候特征和解译人员对地物识别经验制定决策树规则, 根据建立的规则对遥感影像进行分类和精度评价, 其流程详见图 3。

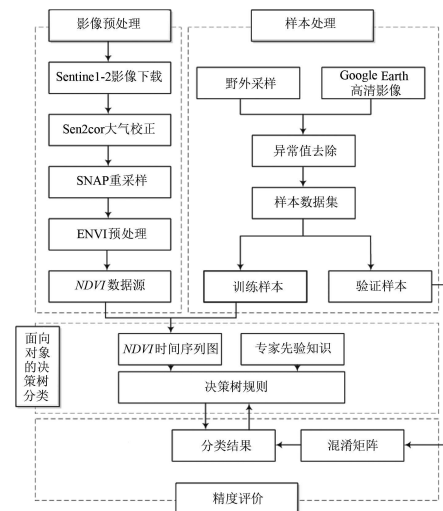


图3 种植结构识别流程

Fig.3 Flowchart of crop planting structure extraction

### 1.3.1 遥感影像预处理

Sentinel-2 的 L1C 文件经 Sen2cor 大气校正处理为 L2A 文件, 将 L2A 文件通过 SNAP 重采样为 ENVI 格式, 重采样分辨率设置为 10 m, 经过 ENVI 裁剪、波段合成和波段运算以备后用。

### 1.3.2 样本点转换 ROI

由于采集的样本为点矢量, 为了能够统计和提取各类样本在地块上的特征, 基于 Google Earth 的 16 级影像, 将作物样本点展布其上, 根据地块大小和纹

理信息绘制感兴趣区 (Region of interest, ROI)，将样本的点矢量转换为面矢量，如图 4 所示。

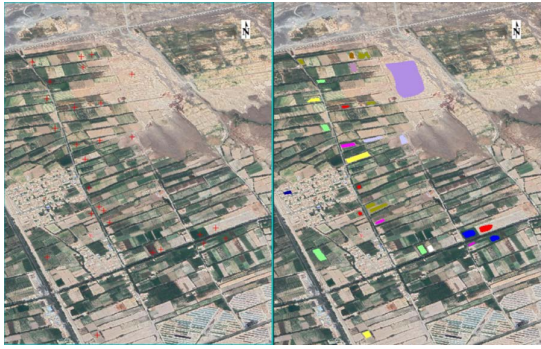


图4 基于样本点的感兴趣区

Fig.4 Region of interest based on sample points

### 1.3.3 NDVI 计算

NDVI 可用于检测植被生长状态、植被覆盖度和提取物候信息等，是目前应用最广泛的植被指数<sup>[18-21]</sup>，其计算式为：

$$NDVI = (R_{Nir} - R_{Red}) / (R_{Nir} + R_{Red}), \quad (1)$$

式中： $R_{Nir}$  和  $R_{Red}$  分别为近红外波段的反射值与红光波段的反射值。NDVI 值介于 -1 和 1 之间，负值表示地面覆盖为云、水、雪等；当  $R_{Nir}$  和  $R_{Red}$  近似相等时，NDVI 近似于 0，表示有岩石或裸土等；NDVI 为正值，表示有植被覆盖，且随覆盖度增大而增大。

### 1.3.4 分类精度评价方法

混淆矩阵是分类精度评价的一种常用方法，用  $n$  行  $n$  列的矩阵形式来表示， $n$  为验证样本的种类，列表示真实值，行表示预测值。本文采用的评价指标为总体精度 (Overall Accuracy, OA) 和 Kappa 系数<sup>[22]</sup>。

总体分类精度表征分类的整体准确率，即分类的结果与参考数据所对应区域的实际类型相一致的概率，计算式为：

$$OA = \sum_{i=1}^n X_{ii} / N, \quad (2)$$

式中： $N$  表示样本点总数； $X_{ii}$  表示被分到正确类别的样本数。

Kappa 系数表征分类结果的可信度，用来评定生产者精度和用户精度的稳定性，计算式为：

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

式中： $X_{i+}$  和  $X_{+i}$  分别表示第  $i$  行和第  $i$  列的总样本数量。

## 2 结果与分析

### 2.1 NDVI 时序特征分析及决策树构建

将裁剪好的灌区 31 个时相的 NDVI 波段合成得到一个 NDVI 时间序列数据集。采用 ENVI 对 10 类样本 ROI 不同时相的 NDVI 平均值和均方差进行统计，并绘制 NDVI 时序变化图，如图 5 所示。

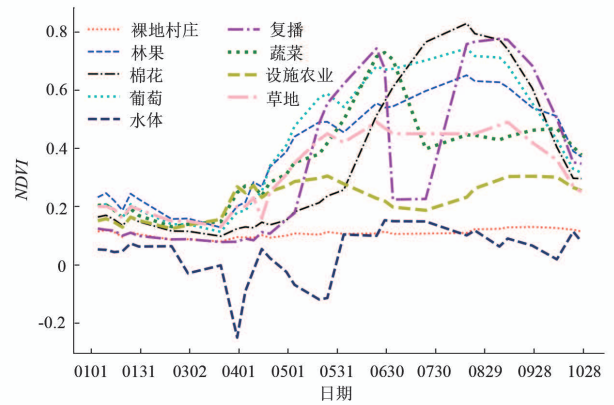


图5 各类样本 NDVI 时序变化曲线

Fig.5 Variation curve of NDVI covered by different crops and land surface

由图 5 可知，各类样本 NDVI 时序变化特征比较明显。其中，复播 NDVI 时间序列图呈“双峰”特征，在 2021 年 6 月 24 日达到第 1 个峰值 0.731 3，在 2021 年 9 月 7 日达到第 2 个峰值 0.778 3，而 7 月 4—19 日 NDVI 值均在 0.22 以下；裸地-村庄 NDVI 值全年均在 0~0.2 之间；设施农业 NDVI 值在 0.2~0.3 之间且呈“单谷”特征，在 5 月 25 日和 9 月 17 日超过了 0.3，而 7 月 4 日—8 月 8 日则  $< 0.2$ ；棉花 NDVI 值呈明显“单峰”曲线，在 8 月 13 日达到峰值 0.830 6；林果（含套种）和葡萄的 NDVI 动态变化整体趋势比较相似，在 3 月 16 日以后 NDVI 迅速上升，且在作物生长季 NDVI 有波动，可能是因为林果和葡萄在生长季节生长旺盛，田间管理中存在对树形和枝叶修剪情况，但它们之间仍然存在区别，林果 NDVI 值在 1—3 月较葡萄高，葡萄 NDVI 值 4—8 月较林果高；蔬菜 NDVI 值在 6 月末达到峰值 0.717 4，生长后期在 0.43 左右波动；草场 NDVI 值在 0.25~0.6 之间；除此之外，水体 NDVI 值  $< 0$ ；其他未分类数据为未包含到以上地物类别中的地表类型，包括地块边缘较窄的道路，苜蓿、孜然、玉米、花生、芝麻、墓地和晒场等。

根据灌区主要作物关键期物候和 NDVI 时间序列变化特征，对主要特征时段 NDVI 建立决策树分类规则，具体详见表 3。

### 2.2 主要作物种植面积提取

分类结果表明，灌区作物总播种面积 33 247  $hm^2$ ，其中棉花 4 913  $hm^2$ 、复播 2 460  $hm^2$ （瓜类与玉米、高粱复种）、林果 6 687  $hm^2$ （含果树套种）、葡萄 1 953  $hm^2$ 、蔬菜 1 627  $hm^2$ 、草场 2 620  $hm^2$ 、设施农业 1 760  $hm^2$ 、其他面积 11 227  $hm^2$ （包括油料、苜蓿、孜然、正播玉米等）；同时，结果显示灌区范围内的村庄-裸地 14 187  $hm^2$ 、水面 373  $hm^2$ （包括永久和临时性水体）。10 m 分辨率灌区主要作物分布如图 6 所示。

表 3 决策树规则表

Table 3 Rules of the decision tree

地物类型	NDVI 特征	决策树分类规则
复播	NDVI 时间序列图为双峰	$b_{10}<0.2 \& b_{11}<0.2 \& (b_{20}>0.5 b_{19} + 0.5  b_{17} + 0.5)&(b_{23}<0.4 b_{22}<0.4 b_{24}<0.4) \& (b_{27}>0.5 b_{28}>0.5)$
裸地-村庄	NDVI 值 0-0.2 间	$b_{10}>0 \& b_{10}<0.3 \& b_{11}>0 \& b_{11}<0.3 \& b_{20}>0 \& b_{20}<0.3 \& b_{27}>0 \& b_{27}<0.3$
设施农业	NDVI 值 0.2-0.3 间	$b_{10}>0.17 \& b_{10}<0.37 \& b_{28}>0.22 \& b_{28}<0.39 \& b_{23}>0.04 \& b_{23}<0.32 \& b_{17}>0.17 \& b_{17}<0.42$
棉花	NDVI 值在 7 月中旬达到峰值	$b_{24}>0.6 \& b_{24}<0.96 \& b_{19}>0.13 \& b_{19}<0.45 \& b_{28}>0.3 \& b_{28}<0.73 \& b_{10}>0.07 \& b_{10}<0.17$ $b_5>0.19 \& b_5<0.4 \& b_2>0.19 \& b_2<0.4 \& b_{24}>0.44 \& b_{24}<0.85 \& b_{19}>0.29 \& b_{19}<0.62 \& b_{28}>0.39 \& b_{28}<0.69$
林果 (包括套种)	NDVI 值在 1-3 月较葡萄高	$b_{24}>0.55 \& b_{24}<0.9 \& b_{28}>0.41 \& b_{28}<0.7 \& b_{19}>0.35 \& b_{19}<0.67 \& b_{11}>0.12 \& b_{11}<0.26$
葡萄	NDVI 值 4-8 月较林果高	$b_{17}>0.14 \& b_{17}<0.7 \& b_{20}>0.45 \& b_{20}<0.89 \& b_{21}>0.4 \& b_{21}<0.92 \& b_{23}>0.13 \& b_{23}<0.67 \& b_{10}>0.1 \& b_{10}<0.4 \& b_{27}>0.23 \& b_{27}<0.65$
蔬菜	NDVI 值在 6 月末达到峰值	$b_5>0.15 \& b_5<0.25 \& b_{17}>0.22 \& b_{17}<0.6 \& b_{20}>0.3 \& b_{20}<0.6 \& b_{27}>0.33 \& b_{27}<0.6$
草场	NDVI 值 0.25-0.6 间	$b_{10}<0 \& b_{11}<0 \& b_{20}<0 \& b_{27}<0 \& b_{17}<0 \& b_5<0$
水体	NDVI 值 <0	

注 其中  $b_1, b_2, \dots, b_{28}$  分别表示 28 个时相的 NDVI 值, 序号与表 1 中的序号相同。&表示和运算, |表示与运算。

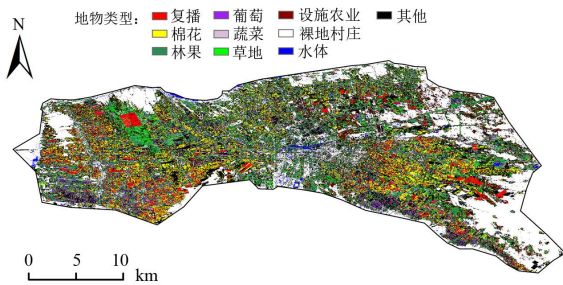


图 6 阿拉沟灌区主要作物分布

Fig.6 Crop planting structure based on remote sensing of the Alagou irrigated area

### 2.3 精度评价

基于感兴趣区的验证得到总体分类精度为 81.56%,  $Kappa$  系数为 0.716 6, 由图 7 可知, 主对角线的精度较高, 村庄-裸地正确分类达到了 99.78%, 其次是棉花 76.63%, 但是林果和葡萄容易错分, 这是因为它们的 NDVI 曲线比较接近, 区别特征不明显导致; 另外, 有较多的复播类别错分为棉花, 说明在决策树构建时, 棉花的规则相对较为松而复播的规则比较紧。

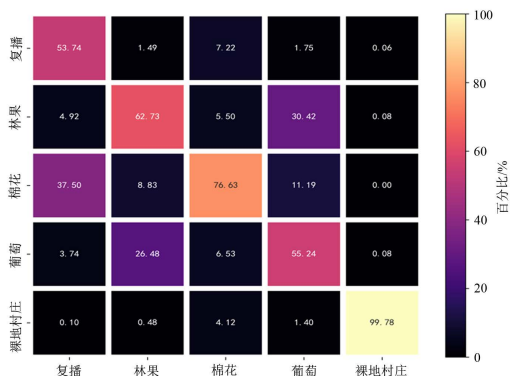


图 7 混淆矩阵热力图

Fig.7 Confusion matrix thermal diagram

由表 4 遥感分类结果与灌区年初灌溉用水计划统计数据对比可知, 棉花、复播、草场 3 类面积二者基本一致, 葡萄、蔬菜的遥感结果比统计分别多 620、627  $hm^2$ 。与统计结果相差比较大的是“其他”类别,

遥感提取比统计多 5 627  $hm^2$ , 主要原因为二者统计口径不完全一致。本研究中将灌区范围内除了复播、棉花、林果、葡萄、蔬菜、人工草场、裸地-村庄、设施农业、水体以外的土地面积全部归为其他, 包括了线状道路、渠道和部分不能够归类的苜蓿、孜然、玉米、花生、芝麻等小宗作物, 因此, 遥感提取所得的“其他”明显大于统计。

表 4 灌区主要作物遥感提取结果和统计面积对比

Table 4 Comparison of major crops area between statistics and remote sensing methods in irrigated area

类别/ $hm^2$	复播	棉花	林果	葡萄	蔬菜	草场	其他
遥感提取	2 460	4 913	6 687	1 953	1 627	2 620	11 227
统计	2 667	5 067	4 667	1 333	1 000	2 333	5 600
绝对误差	-207	-154	2 020	620	627	287	5 627

为进一步检验该方法的可靠性, 参考托克逊县“全国第三次国土调查”成果知灌区范围内的水浇地面积为 32 693  $hm^2$ 。鉴于阿拉沟灌区多年平均降水量只有 6.4 mm, 作物生长必须依靠灌溉, 且灌区内无水田, 因此灌区作物种植面积可认定为灌区范围内的水浇地面积, 遥感提取的作物分布面积为 33 246  $hm^2$ , 比全国第三次国土调查水浇地面积略大 1.02%, 说明该方法在提取破碎化地块作物种植结构方面具有较好的可靠性。

### 3 讨论

为探究 Sentinel-2 遥感影像对于破碎化地块灌区复杂作物类型提取的适用性, 本研究尝试了包括增强型植被指数 (Enhanced Vegetation Index,  $EVI$ )、陆表水分指数 (Land Surface Water Index,  $LSWI$ )、绿度归一化植被指数 (Green Normalized Difference Index,  $GNDVI$ )、红边拐点指数 (Red-Edge Inflection Point Index,  $REIP$ )、Sentinel-2 红边位置指数 (Sentinel-2 Red-Edge Position Index,  $S2REP$ ) 等 5 种遥感植被指数基于监督分类的效果, 其总体精度分别为 80.68%、42.08%、41.37%、52.09%和 51.47%,  $Kappa$  系数分别为 0.672 6、0.300 4、0.288 9、0.361 8 和 0.356 0, 均

劣于 *NDVI* 的结果（总体精度 81.56%，*Kappa* 系数 0.716 6），并且发现基于 *NDVI* 分类效果和真实地块的纹理较为符合。在构建 *NDVI* 时间序列数据库时，本研究共获得了 2021 年 1 月 5 日—10 月 27 日的 59 景时相遥感数据，由于质量不高等原因剔除了 47.46%，仅采用了剩余 31 景非等间隔时相数据。张馨予等<sup>[23]</sup>分析了 30、100、250、500、1 000 m 的空间分辨率，发现中高分辨率的影像经过时间滤波后会减小农作物的类内差异性，进而导致作物分类精度下降，本研究采用的影像空间分辨率为 10 m，没有进行滤波或插值处理仍得到较为满意的精度，进一步印证了该结论。

研究中采用植被指数时序曲线作为区分作物类别的主要特征，这与 Zheng<sup>[4]</sup>、李鑫川等<sup>[7]</sup>、张荣群等<sup>[17]</sup>、汪小钦等<sup>[22]</sup>和贾博中等<sup>[24]</sup>所采用的方法和原理类似。总体来看，采用高分辨率 Sentinel-2 遥感影像具有更高的总体精度与 *Kappa* 系数，但更高分辨率则意味着较大的遥感影像存储空间和处理时间，因此，针对区域多个灌区大面积作物种植结构提取，如何在满足精度的前提下减少存储和计算成本需要进一步探讨。

基于遥感数据的决策树作物分类方法具有建立分类规则灵活，分类速度快的优点。其分类结果的准确性取决于决策树的分类规则，而规则来源于分类样本的特征及其可分性，它同时具备客观因素和主观因素，该方法的不足之处在于当面对范围较大且复杂的研究区域时，受作物物候期年际变化的影响，决策树规则不易固定并进行业务化分类。因此，对于地块破碎化严重，作物种类随机的灌区，需规范田间采样步骤和流程，积累并形成不同作物的标准样本库，提高作物分类精度；同时，平衡构建分类规则过程中的主客观因素以及探索相同规则的适用范围仍需深入研究。

## 4 结论

1) 以 Sentinel-2 遥感影像数据为基础提取灌区种植结构分布图地块纹理清晰，可满足灌区用水管理的需求，为破碎化地块灌区作物种植结构提取提供了可行性。

2) 基于 Sentinel-2 卫星的多时相影像数据，结合作物关键物候期特征和 *NDVI* 时序特征构建决策树分类模型，分类总体精度达 81.56%，*Kappa* 系数达 0.716 6；遥感分类结果与灌区统计结果和“全国第三次国土调查”数据吻合度较高，提取精度满足灌区用水管理需求。

### 参考文献：

- [1] 胡琼, 吴文斌, 宋茜, 等. 农作物种植结构遥感提取研究进展[J]. 中国农业科学, 2015, 48(10): 1 900-1 914.  
HU Qiong, WU Wenbin, SONG Qian, et al. Recent progresses in
- [2] 陈智森, 司炳成. 基于高分辨率影像的河套灌区乡镇尺度作物种植面积监测[J]. 灌溉排水学报, 2018, 37(6): 124-128.  
CHEN Zhisen, SI Bingcheng. Monitoring cropland types at village-town scale in Hetao irrigation district using high-resolution satellite images[J]. Journal of Irrigation and Drainage, 2018, 37(6): 124-128.
- [3] 朱秀芳, 潘耀忠, 王双, 等. 训练样本对 TM 尺度小麦种植面积测量精度影响研究(II)-样本质量对小麦测量精度的影响分析[J]. 测绘科学, 2009, 34(5): 132-135.  
ZHU Xiufang, PAN Yaozhong, WANG Shuang, et al. The effects of training samples on the wheat planting area surveying accuracy in TM scale(II): The accuracy response of different classifiers to samples quality[J]. Science of Surveying and Mapping, 2009, 34(5): 132-135.
- [4] ZHENG B J, MYINT S W, THENKABAIL P S, et al. A support vector machine to identify irrigated crop types using time-series Landsat NDVI data[J]. International Journal of Applied Earth Observation and Geoinformation, 2015, 34: 103-112.
- [5] 刘昊. 基于 Sentinel-2 影像的河套灌区作物种植结构提取[J]. 干旱区资源与环境, 2021, 35(2): 88-95.  
LIU Hao. Extraction of crop planting structure in Hetao irrigated area based on Sentinel-2[J]. Journal of Arid Land Resources and Environment, 2021, 35(2): 88-95.
- [6] 边增淦, 王文, 江渊. 黑河流域中游地区作物种植结构的遥感提取[J]. 地球信息科学学报, 2019, 21(10): 1 629-1 641.  
BIAN Zenggan, WANG Wen, JIANG Yuan. Remote sensing of cropping structure in the middle reaches of the Heihe River Basin[J]. Journal of Geo-Information Science, 2019, 21(10): 1 629-1 641.
- [7] 李鑫川, 徐新刚, 王纪华, 等. 基于时间序列环境卫星影像的作物分类识别[J]. 农业工程学报, 2013, 29(2): 169-176, 298.  
LI Xinchuan, XU Xingang, WANG Jihua, et al. Crop classification recognition based on time-series images from HJ satellite[J]. Transactions of the Chinese Society of Agricultural Engineering, 2013, 29(2): 169-176, 298.
- [8] 刘淑琴. 基于面向对象 C5.0 决策树算法的南昌县土地遥感分类研究[D]. 银川: 宁夏大学, 2019.  
LIU Shuqin. Study on land remote sensing classification of Nanchang County based on object-oriented C5.0 decision tree[D]. Yinchuan: Ningxia University, 2019.
- [9] 靳镜宇, 白洁, 包安明, 等. 林农间作模式下和田绿洲特色林果结构遥感信息提取[J]. 农业工程学报, 2022, 38(3): 146-154.  
JIN Jingyu, BAI Jie, BAO Anming, et al. Extraction of the fruit trees based on satellite imagery under the pattern of forest and crops interplanting in Hotan oasis[J]. Transactions of the Chinese Society of Agricultural Engineering, 2022, 38(3): 146-154.
- [10] 宋茜, 胡琼, 陆苗, 等. 农作物空间分布遥感制图发展方向探讨[J]. 中国农业资源与区划, 2020, 41(6): 57-65.  
SONG Qian, HU Qiong, LU Miao, et al. Prospect of crop mapping[J]. Chinese Journal of Agricultural Resources and Regional Planning, 2020, 41(6): 57-65.
- [11] SHI Chenlie, WANG Xuhong, ZHANG Meng, et al. A comprehensive and automated fusion method: The enhanced flexible spatiotemporal DAta fusion model for monitoring dynamic changes of land surface[J]. Applied Sciences, 2019, 9(18): 3 693.
- [12] 赵鹏博, 吕昭, 买尼克·吾买尔, 等. 吐鲁番市灌溉水利用系数测定分析报告[R]. 吐鲁番: 吐鲁番市水利科学研究所, 2020.  
ZHAO Pengbo, LYU Zhao, MAINIKE Wumaier, et al. Analysis report on utilization coefficient of irrigation water in Turpan[R]. Turpan: Turpan Institute of Water Science, 2020.
- [13] 李中元, 吴炳方, 张淼, 等. 利用物候差异与面向对象决策树提取油

- 菜种植面积[J]. 地球信息科学学报, 2019, 21(5): 720-730.
- LI Zhongyuan, WU Bingfang, ZHANG Miao, et al. Identifying rapeseed planting area using an object-oriented method and crop phenology[J]. Journal of Geo-Information Science, 2019, 21(5): 720-730.
- [14] 谢鑫, 张贤勇, 杨霖琳. 融合信息增益与基尼指数的决策树算法[J]. 计算机工程与应用, 2022, 58(10): 139-144.
- XIE Xin, ZHANG Xianyong, YANG Jilin. Decision tree algorithm fusing information gain and Gini index[J]. Computer Engineering and Applications, 2022, 58(10): 139-144.
- [15] 王利军, 郭燕, 贺佳, 等. 基于决策树和 SVM 的 Sentinel-2A 影像作物提取方法[J]. 农业机械学报, 2018, 49(9): 146-153.
- WANG Lijun, GUO Yan, HE Jia, et al. Classification method by fusion of decision tree and SVM based on sentinel-2A image[J]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(9): 146-153.
- [16] 张旭东, 迟道才. 基于异源多时相遥感数据决策树的作物种植面积提取研究[J]. 沈阳农业大学学报, 2014, 45(4): 451-456.
- ZHANG Xudong, CHI Daochai. Mapping crop fields by using multi-sensor and multi-temporal remote sensing data with decision-tree[J]. Journal of Shenyang Agricultural University, 2014, 45(4): 451-456.
- [17] 张荣群, 王盛安, 高万林, 等. 基于时序植被指数的县域作物遥感分类方法研究[J]. 农业机械学报, 2015, 46(S1): 246-252.
- ZHANG Rongqun, WANG Sheng'an, GAO Wanlin, et al. Remote-sensing classification method of County-level agricultural crops using time-series NDVI[J]. Transactions of the Chinese Society for Agricultural Machinery, 2015, 46(S1): 246-252.
- [18] 贾云飞, 李云飞, 范天程, 等. 基于长时间序列 NDVI 的黄土高原延河流域及其沟壑区植被覆盖变化分析[J]. 水土保持研究, 2022, 29(4): 240-247.
- JIA Yunfei, LI Yunfei, FAN Tiancheng, et al. Analysis of vegetation cover change in Yanhe Basin and its gully area on the Chinese loess plateau based on long time series NDVI[J]. Research of Soil and Water Conservation, 2022, 29(4): 240-247.
- [19] 孔冬冬, 张强, 黄文琳, 等. 1982—2013 年青藏高原植被物候变化及气象因素影响[J]. 地理学报, 2017, 72(1): 39-52.
- KONG Dongdong, ZHANG Qiang, HUANG Wenlin, et al. Vegetation phenology change in Tibetan Plateau from 1982 to 2013 and its related meteorological factors[J]. Acta Geographica Sinica, 2017, 72(1): 39-52.
- [20] BUITENWERF R, ROSE L, HIGGINS S I. Three decades of multi-dimensional change in global leaf phenology[J]. Nature Climate Change, 2015, 5(4): 364-368.
- [21] 潘霞, 高永, 汪季, 等. 植被指数遥感演化研究进展[J]. 北方园艺, 2018(20): 162-169.
- PAN Xia, GAO Yong, WANG Ji, et al. Review on vegetation index using remote sensing evolution[J]. Northern Horticulture, 2018(20): 162-169.
- [22] 汪小钦, 邱鹏勋, 李娅丽, 等. 基于时序 Landsat 遥感数据的新疆开孔河流域农作物类型识别[J]. 农业工程学报, 2019, 35(16): 180-188.
- WANG Xiaoqin, QIU Pengxun, LI Yali, et al. Crops identification in Kaikong River Basin of Xinjiang based on time series Landsat remote sensing images[J]. Transactions of the Chinese Society of Agricultural Engineering, 2019, 35(16): 180-188.
- [23] 张馨予, 蔡志文, 杨靖雅, 等. 时序滤波对农作物遥感识别的影响[J]. 农业工程学报, 2022, 38(4): 215-224.
- ZHANG Xinyu, CAI Zhiwen, YANG Jingya, et al. Impacts of temporal smoothing methods on crop type identification[J]. Transactions of the Chinese Society of Agricultural Engineering, 2022, 38(4): 215-224.
- [24] 贾博中, 白燕英, 魏占民, 等. 基于 MODIS-EVI 的内蒙古沿黄平原区作物种植结构分析[J]. 灌溉排水学报, 2021, 40(4): 114-120.
- JIA Bozhong, BAI Yanying, WEI Zhanmin, et al. Using MODIS-EVI to identify cropping structure in Plains along the Yellow River in Inner Mongolia[J]. Journal of Irrigation and Drainage, 2021, 40(4): 114-120.

## Using Sentinel-2 Sensing Imagery to Estimate Planting Structure in Fragmented Irrigated Lands

WU Di<sup>1</sup>, YANG Peng<sup>2</sup>, ZHOU Liyong<sup>3</sup>, LI Fangsong<sup>3</sup>, LI Lingfeng<sup>3</sup>, ZHANG Xudong<sup>2\*</sup>

(1. China Irrigation and Drainage Development Center, Beijing 100054, China; 2. College of Water Conservancy, Shenyang Agricultural University, Shenyang 110866, China; 3. Xinjiang Institute of Water Resources and Hydropower Research, Urumqi 830049, China)

**Abstract: 【Objective】** Understanding planting structure and crop growth in a region is important to assess its food supply and security. The objective of this paper is to investigate the feasibility of a decision-tree model derived from the Sentinel-2 remote sensing imagery to map cropping structure in fragmented irrigation regions. **【Method】** The study site was Alagou irrigation area in Xinjiang. The planting areas of major crops in 2021 were estimated using the Sentinel-2 remote sensing imagery. We then compared these with both field investigation and visual interpretation from the Google HD images. Critical growth stage for identifying each crop was determined based on the phenological information and the *NDVI* time series, from which we derived a decision tree classification model. Accuracy of the model was verified against ground-truth data. **【Result】** The planting structure mapped from the Sentinel-2 remote sensing imagery had sharp textures, meeting the requirements for agricultural water management. The decision tree classification model can accurately classify crops at the scale required for irrigation management. The model is simple and feasible. Compared with ground-truth data, its average accuracy is 81.56% and the *Kappa* coefficient is 0.716 6. **【Conclusion】** The Sentinel-2 remote sensing imagery and the decision tree classification method derived from it can be used to accurately identify planting structure in fragmented lands. They can provide support information for decision-making in water management, and improve agricultural water usage in irrigation districts.

**Key words:** Sentinel-2; crops classification in irrigated areas; *NDVI* time series; decision tree; fragmented land parcel

责任编辑: 赵宇龙