

引用格式: Zhu Man, Zhang Lifu, Wan Nan, *et al.* Comparative Study on UNVI Vegetation Index and Performance based on Sentinel-2[J]. Remote Sensing Technology and Application, 2021, 36(4): 936-947. [朱曼, 张立福, 王楠, 等. 基于 Sentinel-2 的 UNVI 植被指数及性能对比研究[J]. 遥感技术与应用, 2021, 36(4): 936-947.]
DOI: 10.11873/j.issn.1004-0323.2021.4.0936

基于 Sentinel-2 的 UNVI 植被指数及性能对比研究

朱 曼^{1,2}, 张立福¹, 王 楠¹, 林昱坤^{1,2}, 张琳珊^{1,2}, 王 颀^{1,2}, 刘华亮³

(1. 中国科学院空天信息创新研究院, 北京 100094;

2. 中国科学院大学, 北京 100049;

3. 北京大学遥感与地理信息系统研究所, 北京 100871;)

摘要: 作物精准识别和分类是农业遥感检测的重要内容, 对作物长势监测以及估产十分重要。以美国混合农业带为研究区, 基于 Sentinel-2 时间序列影像, 根据其传感器响应函数计算了针对 Sentinel-2 的通用归一化植被指数 (Universal Normalized Vegetation Index, UNVI), 并通过两个对比实验, 分析 UNVI 等 6 个指数在作物精准分类中的性能。实验一以 JM (Jeffries-Matusita) 距离为指标对不同作物类别之间的可分性进行分析, 结果表明 UNVI 优于 NDVI、EVI、WDRVI、NDre1 和 NDWI 指数, 在玉米和棉花、玉米和水稻、玉米和水稻的区分上, UNVI 优于其他指数区分能力相当, 但在其余的作物组合上如棉花和水稻, NDVI 等指数则无法将其很好的区分, 此时 UNVI 指数依然可以表现出较好的区分能力; 实验二对 6 种时间序列指数特征分别使用随机森林和支持向量机进行作物分类, 结果表明 UNVI 指数的总体精度和 Kappa 系数最高, 其次是 NDre1 指数和 WDRVI 指数, EVI 的总体精度和 Kappa 系数最低, 这表明 UNVI 比其他 6 个指数更好地区分了研究区大豆、玉米、棉花和水稻等 4 种主要作物。综上, 基于 Sentinel-2 时间序列的 UNVI 指数在进行作物分类时与其他 5 种遥感植被指数相比, 具有较大的优势, UNVI 可为农作物长势分析和作物估产研究等农业研究和应用的可选植被指数。

关键词: Sentinel-2; 时间序列; UNVI 植被指数; 可分性; 作物识别

中图分类号: TP79 **文献标志码:** A **文章编号:** 1004-0323(2021)04-0936-12

1 引 言

农业遥感是遥感研究和应用的一个重要方向, 遥感技术在农作物的种类识别、种植面积估算、作物种植结构提取等方面发挥着重要作用。光谱特征是农作物遥感识别的基础特征, 诸多学者基于此识别了小麦、大豆、玉米和棉花等作物^[1]。但由于空间分辨率和光谱分辨率的限制而产生的“同物异谱”和“异物同谱”现象, 致使直接使用光谱特征进行作物识别的精度无法满足要求。为克服这一局

限性, 提高作物识别的精度, 时间序列光谱和指数影像被广泛使用, 其主要通过增加时间维度信息进行作物长势监测和遥感分类。目前, 已经定义的 40 多种植被指数, 被成功应用于生态环境监测、植被生物量估算、农作物长势评估以及种植面积估算等方面。其中, 归一化植被指数 (Normalized Difference Vegetation Index, NDVI)^[2-7] 和增强型植被指数 (Enhanced Vegetation Index, EVI)^[7-12] 应用较为广泛, 可有效指示作物生长状况和生育期, 且其多时相数据已广泛应用于作物分类^[6]、农业监测以及

收稿日期: 2020-06-12; 修订日期: 2021-06-28

基金项目: 国家自然科学基金重点基金项目 (41830108), 兵团重点领域创新团队建设计划 (2018CB004), 兵团重大科技课题 (2018A A00402), 中国科学院战略性先导科技专项 (XDA19080304) 资助。

作者简介: 朱 曼 (1994—), 女, 安徽宿州人, 硕士研究生, 主要从事高光谱遥感研究。E-mail: zhuman@radi.ac.cn

通讯作者: 张立福 (1967—), 男, 山东烟台人, 研究员, 主要从事高光谱遥感研究。E-mail: zhanglf@radi.ac.cn

作物估产中^[13]。然而,NDVI和EVI仅使用2~3个光谱波段,可能导致作物生长期某些独特特征的丢失,所以现在广泛使用的NDVI和EVI可能不是作物分类的最佳指数特征^[14]。除NDVI和EVI外,宽动态植被指数(Wide Dynamic Range Vegetation Index, WDRVI)、归一化红边指数(Normalized Difference red-edge 1, NDre1)和归一化水体指数(Normalized Difference Water Index, NDWI)等指数也被相继提出并在诸多场景得到应用。此外,有学者基于通用模式分解算法(Universal Pattern Decomposition Method, UPDM)建立了一种全谱段植被指数——通用归一化植被指数(Universal Normalized Vegetation Index, UNVI)。与其他只用少数波段信息进行计算的植被指数相比,UNVI具有较为明显的优势,能够更准确地反映植被生长过程中的细节变化^[15]。Liu等^[5]使用包含UNVI指数在内的4种指数来描述城市地表温度(Land Surface Temperature, LST)的变化,结果显示UNVI与LST变化具有最佳的相关性。Jiao等^[16]通过研究得出,基于UNVI的植被状况指数(Vegetation Condition Index, VCI)与长期干旱的相关性比基于NDVI的VCI更强,这意味着UNVI具有相当大的干旱监测潜力。UNVI还被应用于估算冬小麦的叶绿素含量,此外,与NDVI和三角植被指数(Triangle Vegetation Index, TVI)相比,UNVI能够有效提升冬小麦叶绿素含量估算的准确性和稳定性。对于叶面积指数(Leaf Area Index, LAI)的估计,UNVI具有比NDVI和EVI更高的饱和点,并且对更广泛的植被动态更为敏感^[16]。

然而,针对不同指数特征作物识别能力的研究,目前尚且较少。研究学者通常使用单一指数或多个指数特征组合进行作物识别,虽取得了一定的效果,但仍存在一定不足。当面向不同的作物或作物组合时,不同指数识别特征有效性缺乏统一和定量的描述,缺少最佳识别特征集,导致其在作物识别特征集的构建时存在盲目性。提取更高可分离性的特征进行作物识别尤为重要,可以有效解决特征冗余的问题,并提供鲁棒性更高的识别特征,以便提高作物识别的精度。

基于此,选择UNVI、NDVI、EVI、WDRVI、NDre1和NDWI 6个指数,通过设计两个对比实验评估不同指数的作物识别能力:第一个实验是使用单时相和多时相JM(Jeffries-Matusita)距离对比分

析各指数不同作物之间的类别可分性;第二个实验是利用这些时间序列指数特征分别使用随机森林和支持向量机来比较作物分类的准确性。

2 研究区与数据预处理

2.1 研究区概况

为了评估不同指数时间序列对作物分类的能力,选取美国中南部的混合农业带作为研究区。混合农业带分布在棉花带以北,玉米带和乳蓄带以南,如图1所示。该地区西部是密西西比河流域,中部为阿巴拉契亚山脉,东部为大西洋沿岸,亚热带湿润气候,由于多变的地理环境,该地区既适宜牧草生长,发展畜牧业;又适宜各种作物的生长,发展种植业。该区域种植作物类型丰富,包含大豆、玉米、水稻和棉花等多种作物,便于开展不同指数作物分类能力的研究。

2.2 时间序列遥感数据

为完整覆盖研究区域作物的生育期,选取时间跨度为2018年4月至10月的20景Sentinel-2数据进行遥感时序分类实验。其中10景为Sentinel-2A数据(以下简称为S2A),10景为Sentinel-2B数据(以下简称为S2B)。为保证遥感时间序列影像质量,减少云层对作物识别的影响,只选择云量少于20%的影像。数据获取的时相和详细信息分别如表1和表2所示。

实验采用的Sentinel-2数据均来自于欧空局(ESA)官网,其数据下载地址为:<https://scihub.copernicus.eu/>。欧空局发布的Sentinel-2数据为L1C级数据,是经过正射校正和几何精校正的大气表观反射率产品,没有进行大气校正。为了方便用户进行下一步的处理,欧空局发布了专门生产L2A级数据的插件Sen2cor,用户可使用此插件进行大气校正生产Sentinel-2 L2A级数据,进而获得经过大气校正后的大气底层反射率数据^[17]。

2.3 CDL样本数据

样本数据使用2018年的CDL(The Cropland Data Layer)数据,该数据从美国农业部国家农业统计服务网站下载(<https://nassgeodata.gmu.edu/CropScape/>)^[18]。CDL数据是常见的用于作物分类监测训练和测试的数据^[19]。该数据产品使用中高分辨率的遥感影像和美国农业部收集的地面实测数据以及其他辅助数据,通过监督无决策分类的方法生成,分辨率为30 m,共定义了110个土地覆盖

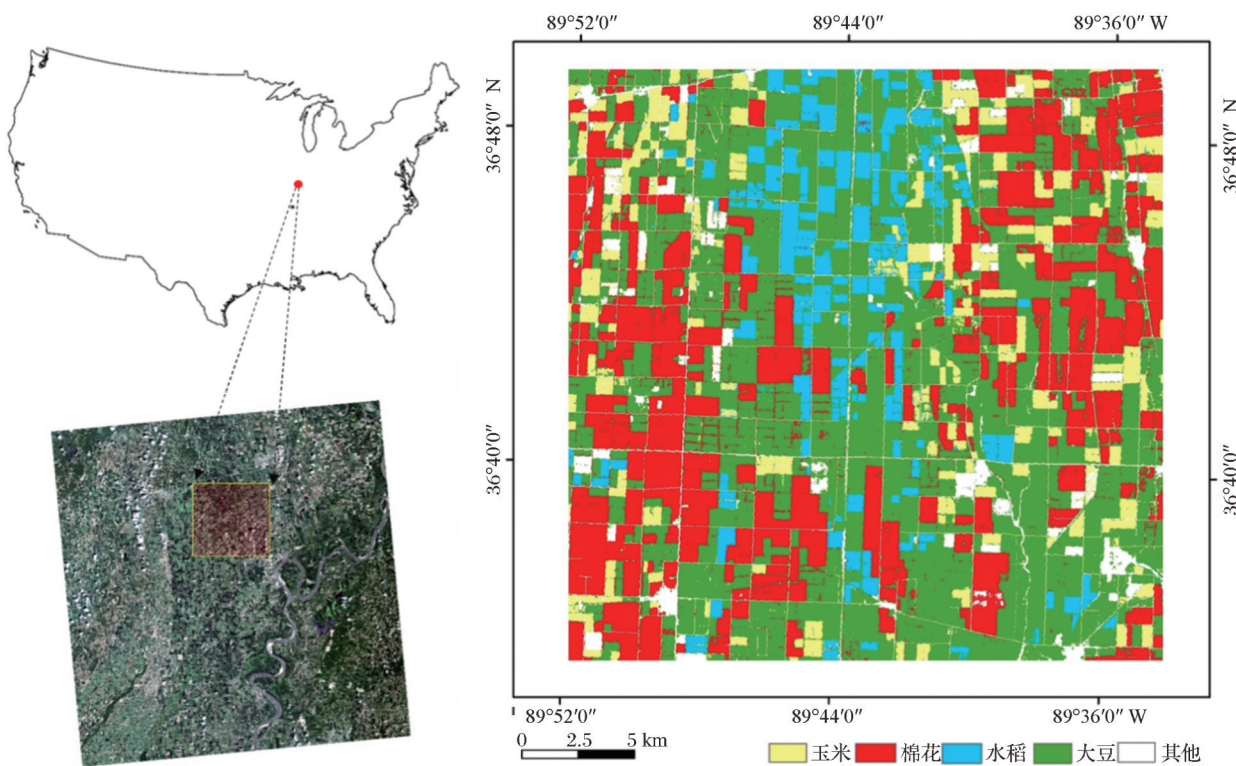


图 1 研究区位置

Fig.1 Location of the study area

表 1 S2A 和 S2B 获取时间

| Table 1 Acquisition time of S2A and S2B | | |
|---|-------------|------|
| 获取日期 | 获取卫星 | 影像质量 |
| 2018-04-21 | Sentinel-2B | 较好 |
| 2018-05-01 | Sentinel-2B | 较好 |
| 2018-05-11 | Sentinel-2B | 较好 |
| 2018-05-16 | Sentinel-2A | 良好 |
| 2018-06-05 | Sentinel-2A | 较好 |
| 2018-06-10 | Sentinel-2B | 较好 |
| 2018-06-15 | Sentinel-2A | 良好 |
| 2018-06-30 | Sentinel-2B | 良好 |
| 2018-07-05 | Sentinel-2A | 良好 |
| 2018-07-10 | Sentinel-2B | 良好 |
| 2018-07-20 | Sentinel-2B | 较好 |
| 2018-07-25 | Sentinel-2A | 良好 |
| 2018-08-04 | Sentinel-2A | 较好 |
| 2018-09-03 | Sentinel-2A | 良好 |
| 2018-09-13 | Sentinel-2A | 较好 |
| 2018-09-18 | Sentinel-2B | 良好 |
| 2018-09-28 | Sentinel-2B | 良好 |
| 2018-10-03 | Sentinel-2A | 良好 |
| 2018-10-18 | Sentinel-2B | 较好 |
| 2018-10-23 | Sentinel-2A | 较好 |

分类^[20]。

研究区所涉及的作物类型是混合农业带内的 4 种主要作物,分别是大豆、玉米、水稻和棉花。在使

表 2 S2A 和 S2B 详细信息对比

| Table 2 Spectral bands and resolutions of S2A and S2B | | | | | |
|---|-------------|-------------|-------------|-------------|-----------|
| 波段 | sensors | | | | 分辨率 /m |
| | Sentinel—2A | | Sentinel—2B | | |
| | 中心波长 /nm | 波段宽度 /nm | 中心波长 /nm | 波段宽度 /nm | |
| 1 海岸波段 | 443.9 | 27 | 442.3 | 45 | 60 |
| 2 蓝波段 | 496.6 | 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(宽) | 835.1 | 145 | 833 | 133 | 10 |
| 8a 近红外波段(窄) | 864.8 | 33 | 864 | 32 | 20 |
| 9 水汽波段 | 945.0 | 26 | 943.2 | 27 | 60 |
| 10 卷云波段 | 1 373.5 | 75 | 1 376.9 | 76 | 60 |
| 11 短波红外 1 | 1 613.7 | 143 | 1 610.4 | 141 | 20 |
| 12 短波红外 2 | 2 202.4 | 242 | 2 185.7 | 238 | 20 |

用 CDL 数据产品进行样本选取前需要将其重采样为 10m 分辨率,以使其和 Sentinel-2 影像分辨率保持一致,便于训练样本和测试样本的选取。

2.4 植被指数

通用归一化植被指数 UNVI 是基于通用模式分解算法(Universal Pattern Decomposition Method,

UPDM)建立的一种全谱段植被指数,它假设地物光谱可以表示为4个标准光谱的线性组合,即水、植被、土壤以及黄叶^[15],如公式(1)所示:

$$R(i) \rightarrow (C_w \cdot Pw(i) + C_v \cdot Pv(i) + C_s \cdot Ps(i) + C_4 \cdot P4(i)) \quad (1)$$

其中: i 为波段编号; $R(i)$ 为地物 i 波段下的光谱; Pw 、 Pv 、 Ps 、 $P4$ 分别代表水、植被、土壤以及黄叶这4种标准光谱的归一化反射率值; C_w 、 C_v 、 C_s 、 C_4 则分别表示标准光谱对应的UPDM系数。

基于公式(1)需要计算复数系数矩阵 P ,为了使UPDM系数的计算简单化,Zhang等对UPDM系数矩阵计算公式进行优化后导出了一个简单的系数

$$M = \begin{bmatrix} 0.2159 & 0.1381 & -0.0137 & -0.0230 & -0.0008 & 0.0031 & 0.0028 & -0.0004 & -0.0821 & -0.08237 \\ -0.0288 & -0.0396 & -0.1077 & -0.0495 & 0.1012 & 0.1432 & 0.1500 & 0.1535 & -0.0988 & -0.18231 \\ -0.1057 & -0.0237 & 0.1990 & 0.1461 & -0.0258 & -0.0690 & -0.0778 & -0.0796 & 0.3557 & 0.4667 \\ -0.1126 & 0.0106 & 0.1442 & 0.1291 & 0.0259 & -0.0055 & -0.0074 & -0.0059 & -0.0590 & -0.1111 \end{bmatrix} \quad (3)$$

UNVI指数的计算公式为:

$$UNVI = \frac{C_v - a \cdot C_s - C_4}{C_w + C_v + C_s} \quad (4)$$

其中, $C=[C_w, C_v, C_s, C_4]^T$ 是UPDM中的系数,可以利用上面公式计算得到; a 为调节因子。为了使枯叶的UNVI值趋于0,茂密植被的UNVI值趋于1,本研究中 a 的值取为0.1^[15,21]。

本研究共选择了5种指数与UNVI进行比较,分别为归一化植被指数NDVI(Normalized Difference Vegetation Index)、增强型植被指数EVI(Enhanced Vegetation Index)、宽动态植被指数WDRVI(Wide Dynamic Range Vegetation Index)、归一化红边指数NDre1(Normalized Difference red-edge 1)、归一化水体指数NDWI(Normalized Difference Water Index)。表3显示了基于Sentinel-2 L2A地表反射率影像计算各指数的公式。

表3 用于对比分析的指数计算公式

Table 3 Indices used for comparison in this study

| 指数 | 计算公式 |
|----------------|--|
| 通用归一化植被指数 UNVI | $UNVI = \frac{C_v - a \cdot C_s - C_4}{C_w + C_v + C_s}$ |
| 归一化植被指数 NDVI | $NDVI = \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + \rho_{RED}}$ |
| 增强型植被指数 EVI | $EVI = 2.5 \times \frac{\rho_{NIR} - \rho_{RED}}{\rho_{NIR} + 6 \times \rho_{RED} - 7.5 \times \rho_{BLUE} + 1}$ |
| 宽动态植被指数 WDRVI | $WDRVI = \frac{0.2 \times \rho_{NIR} - \rho_{RED}}{0.2 \times \rho_{NIR} + \rho_{RED}}$ |
| 归一化红边指数 NDre1 | $NDre1 = \frac{\rho_{RED2} - \rho_{RED1}}{\rho_{RED2} + \rho_{RED1}}$ |
| 归一化水体指数 NDWI | $NDWI = \frac{\rho_{GREEN} - \rho_{NIR}}{\rho_{GREEN} + \rho_{NIR}}$ |

矩阵 M ,由此公式(1)中的复杂系数矩阵 P 可由矩阵 M 替换,优化后如公式(2)所示^[21]:

$$C = MR \quad (2)$$

其中: $R=[R_1, R_2, \dots, R_n]^T$ (T 代表矩阵转置)为原始遥感数据的反射率光谱向量; $M=[M_w, M_v, M_s, M_4]^T$ 是1个 $4 \times n$ 矩阵, n 是波段数量, M 的下标与上面公式中的含义相同。对于不同卫星传感器, M 值不同。

研究中基于Sentinel-2传感器,且由于海岸波段、水汽波段和卷云波段受云等的影响较大,所以在计算 M 矩阵时将这3个波段去除,使用剩余的10个波段计算 M 矩阵,如公式(3)所示:

其中NDVI是通过增加近红外(NIR)波段中叶绿素的强反射与红色波段中的强吸收之间的差异而被广泛使用;EVI通过加入蓝色波段来增强植被信号,进而矫正土壤背景和气溶胶散射的影响,在植被茂密时较为敏感^[22]。NDVI和EVI是进行时间序列分析时被广泛使用的植被指数,所以对任何其它指数进行评估应首先与NDVI和EVI进行比较。WDRVI是宽动态范围植被指数,是为了提高NDVI在高密度植被覆盖下的敏感性而建立的,与EVI类似,这个指数需要来自电磁波谱的红色和近红外波段的信息。NDre1利用红边波段5和波段6进行计算,是Sentinel-2独有的红边指数,它对植被叶片冠层的细小变化十分敏感,可用于精细农业、森林监测^[23]。此外由于研究区域中种植有大量水稻,NDWI水体指数用遥感影像的特定波段进行归一化处理,可以突显影像中的水体信息,对提取水稻较为有利。综上,实验选取以上6个指数进行对比分析,研究不同指数特征在识别作物方面的表现能力。

3 方 法

基于Sentinel-2的UNVI植被指数可分性对比研究总体流程如图2所示。

基于SG滤波后的不同指数时间序列影像和CDL样本数据,设计了两个对比实验来评估各个指数区分作物的能力,第一个实验是针对研究区内作物6个指数的可分性分析。每一类作物按照等比例随机抽样的方式选取整个研究区域的10%作为样

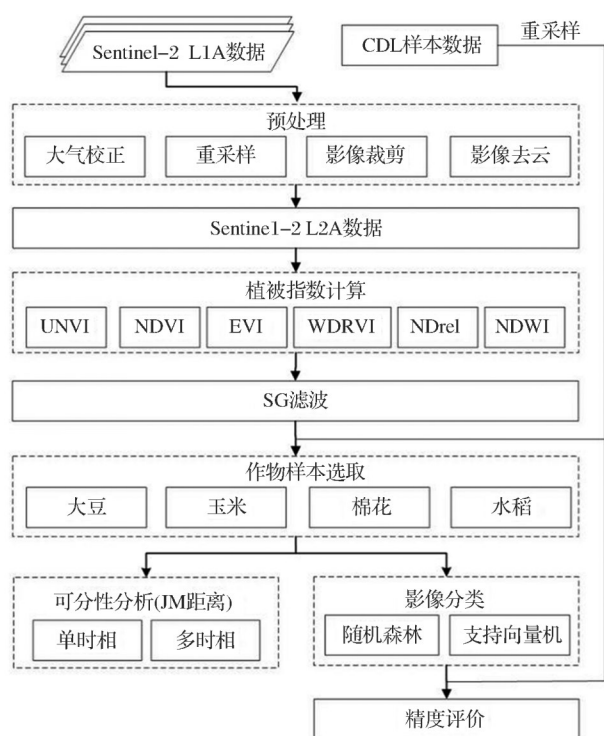


图2 技术路线图

Fig.2 Overall technology roadmap

本进行JM距离计算,选取的这些像素均结合Google Earth影像进行验证,以保证选取的每类作物样本像素的纯度、可靠性和分布均匀性^[24]。第二个实验是用不同的时间序列指数作为特征集,使用随机森林和支持向量机对研究区域的作物进行分类,比较6个指数在分类精度上的优势。

3.1 可分性分析

可分性分析是通过评估不同作物类别的UNVI、NDVI、EVI、WDRVI、NDre1、NDWI 6个指数的时间序列可分离性。用来评价可分离性的指标有Jeffries-Matusita距离(JM距离)^[8, 25-26]、可分性指数(SI)^[27-30]、变换后的散度(TD)^[25, 31]等。选择JM距离来表示地类之间的可分离性。两类之间的SI定义为通过VI标准偏差之和归一化的VI平均值之间的差异。均值的差异反映了类别间的变异性,而标准差的总和则代表类别内的变异性。SI的局限性在于,当两类均值相等时,SI始终为零,无法准确反映可分性^[32]。且SI只能测量一维特征空间(如单个时间点)中的可分离性,因此,它无法满足实验所需要的计算多时相VI指数的可分离性^[28]。与TD相比,JM距离用于测量可分性更可靠^[32-33],因此选择JM距离表示每对作物类型之间的可分性。

JM距离计算公式如公式(5)所示^[33]:

$$JM = 2(1 - e^{-B}) \quad (5)$$

其中:B为Bhattacharyya距离,如公式(6)所示:

$$B = \frac{1}{8}(\mu_i - \mu_j)^T \left(\frac{\Sigma_i + \Sigma_j}{2} \right)^{-1} (\mu_i - \mu_j) + \frac{1}{2} \ln \left(\left| \frac{\Sigma_i + \Sigma_j}{2} \right| / \sqrt{|\Sigma_i| \cdot |\Sigma_j|} \right) \quad (6)$$

其中: i 和 j 为两个不同的类; μ 为反射率的均值向量; Σ 为协方差矩阵。JM距离范围从0(完全不可分)到2(完全可分),数值越大,表示两个类之间的可分程度越高。

实验中分别计算了大豆、玉米、水稻和棉花这4种作物类别的6种可能组合的JM距离,分别为单时相和多时相的UNVI、NDVI、EVI、WDRVI、NDre1、NDWI的JM距离。

3.2 分类方法

为了确定时间序列6个指数区分作物的能力,同时为了验证各指数在不同分类器上的适用性,分别采用随机森林(Random Forest, RF)和支持向量机(Support Vector Machine, SVM)分类器对这6个指数分别进行训练和分类。

3.2.1 随机森林

随机森林(Random Forest, RF)是一种集成学习算法,它通过从随机选择的训练样本和特征中构造大量的决策树,并根据所有生成树的多数选票来决定最终的类^[34]。该方法具有训练速度快、参数化简单、精度高以及对噪声不敏感等优点^[10]。随机森林在遥感领域得到较为广泛的应用,特别适用于遥感时间序列数据的作物分类^[34-35]。在RF算法中需要设置两个参数:决策树的数量(ntree)和为每个决策拆分选择的特征的数量(mtry)。一般情况下,ntree可以尽可能地大,通常是几百个^[36]。在实验中,将UNVI和其他5个指数的ntree设置为100,这足以满足作物分类目的^[11, 34, 36]。mtry则设置为输入特征总数的平方根,这是常用的推荐默认值^[34]。

3.2.2 支持向量机

支持向量机(Support Vector Machine, SVM)是一种按照监督学习的方式对数据进行二元分类的广义线性分类器,其决策边界通过求解学习样本的最大边距超平面。SVM可以通过核方法进行非线性分类,是常见的核学习方法之一。SVM分类器具有较强的泛化能力,且其需要的样本数量不多,尤其适用于遥感影像的分类。

3.3 精度评价方法

分类性能的评估通常基于混淆矩阵(Confusion Matrix),也称为误差矩阵,是进行精度评价的一种标准格式,用 n 行 n 列的矩阵形式来表示^[37]。每一行代表参考数据中的实际类,每一列代表预测类。因此,对角线元素表示正确分类的像素个数。混淆矩阵中较常使用的具体评价指标有总体精度(Overall Accuracy, OA)、制图精度(Producer's Accuracy, PA)、用户精度(User's Accuracy, UA)以及 Kappa 系数(Kappa Coefficient, KC)等^[37],这些指标从不同的角度反映作物分类的精度。

总体精度(Overall Accuracy, OA),表明所有类别中正确分类的样本数量占总体检验样本总量的比例,如公式(7)所示:

$$OA = \frac{\sum_{i=1}^n c_{ii}}{\sum_{i=1}^n \sum_{j=1}^n c_{ij}} \quad (7)$$

制图精度(Producer's Accuracy, PA),表明不同类别的检验样本中被正确检出的数量占类别总检验样本的数量比例,如公式(8)所示:

$$PA = \frac{m_{ii}}{\sum_{j=1}^n c_{ij}} \quad (8)$$

用户精度(User's Accuracy, UA),表明某一类别中被正确分类的样本数量占被分类到该类别的所有样本数量总和的比例,如公式(9)所示:

$$UA = \frac{m_{ij}}{\sum_{i=1}^n c_{ij}} \quad (9)$$

Kappa 系数(Kappa Coefficient, KC),可以充分利用混淆矩阵的所有信息,并对分类精度进行全面衡量,如公式(10)所示:

$$Kappa = \frac{N \sum_{i=1}^n c_{ii} - \sum_{i=1}^n (c_{i+} c_{+i})}{N^2 - \sum_{i=1}^n (c_{i+} c_{+i})} \quad (10)$$

4 结果与讨论

4.1 时序指数分析

为定性了解大豆、玉米、棉花和水稻4种作物的生育期特征,计算了每个时间点所选各类作物样本像素的平均指数,并由SG滤波器重建时间序列曲线,如图3所示。同时为了更好地显示每个指数的制图效果,将6个指数的纵轴范围不同设置。本研究中,植被的NDVI、EVI和NDre1限制在0~1范围内,UNVI稍微超出这个范围^[21],WDRVI值的范围主要在-0.6~0.5范围内,而NDWI值的范围则在-0.8~-0.2范围内。对于每种作物类型,除玉米外,这6个指数随时间变化趋势类似。玉米的生育期早于其他作物,其指数值提前达到峰值,预计玉米可以和其他作物较好地地区分。水稻相较于棉花和大豆,其生育期亦略早,相对容易区分,但棉花和大豆生育期接近,各个指数峰值只有细微的差异。因此,棉花和大豆这两种作物类别区分较为困难。

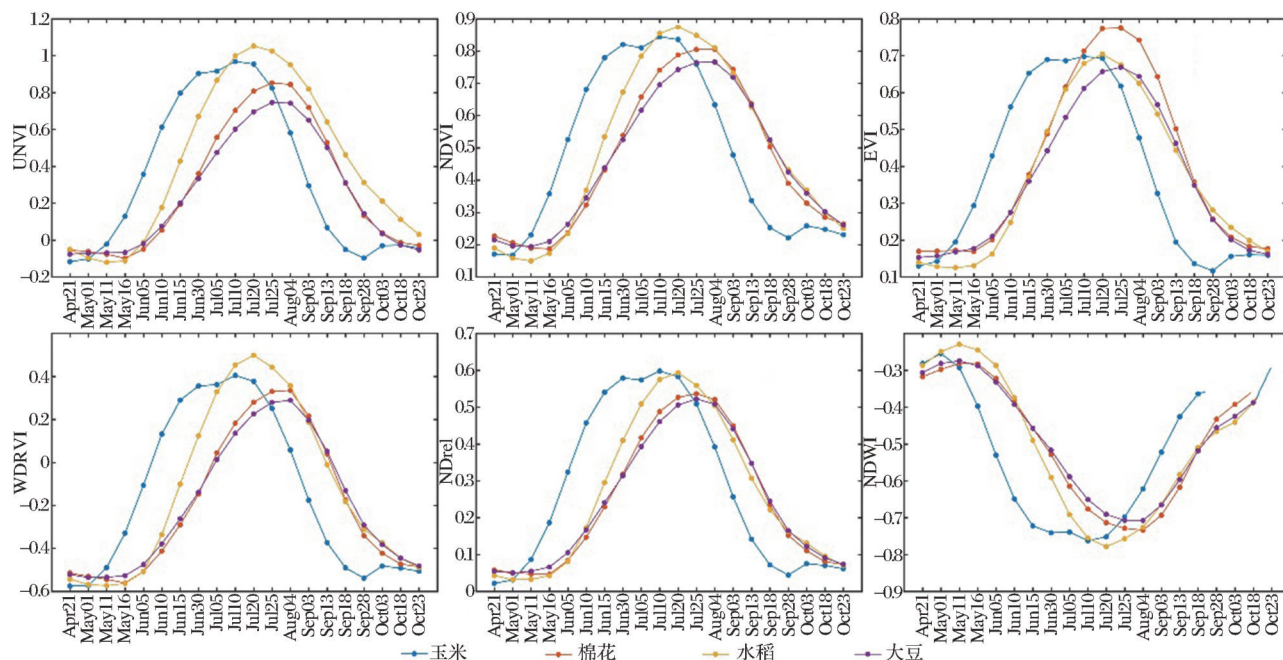


图3 4种作物的不同指数时间序列曲线

Fig.3 Exponential time series curves of four crops

4.2 可分性分析

上文根据时间序列指数对研究区内作物进行了定性分析,为定量了解大豆、玉米、棉花和水稻4种作物的可分性,分别计算每对类别作物单时相JM距离和多时相JM距离来进行展示。

4.2.1 单时相JM距离

对 UNVI、NDVI、EVI、WDRVI、NDre1 和 NDWI 指数,计算每对类别在单个时间节点上的 JM 距离,结果如图 4 所示。这 6 幅图分别代表了玉米和棉花、玉米和水稻、玉米和大豆、棉花和水稻、棉花和大豆以及水稻和大豆这 6 个类对的作物组合。横轴和纵轴分别代表各个指数和时相。每个网格单元的颜色表示对应的指数和时相的 JM 距离,网格单元越红,对应的 JM 距离值越大,两个类别的可分性越强。

从图中可以看出,由于作物的生育期不同,每类对作物之间区分度较高的时相亦存在差异,具体表现为 JM 距离在整个生育期内的变化。对于玉米和棉花来说,其较易区分的时间在 6 月,其次是 9 月。在较易区分的 6 月,UNVI 和 NDre1 的区分度最好;对于玉米和水稻,其区分的时间也在 6 月和 9 月,对于 NDVI、EVI 和 NDre1 等指数 6 月是区分玉米和水稻的最佳时间,但对于 UNVI 来说,9 月则可以较好地将玉米和水稻区分开,且区分度高于其他指数;对于玉米和大豆,UNVI 和其他指数在 6 月

9 月的区分度相当。在其余的棉花和水稻、棉花和大豆以及水稻和大豆这 3 种作物组合上,这 6 种作物均无法较好地将其区分,但 UNVI 对这 3 种作物组合的表现力仍然略优于 NDVI、EVI 等指数。

4.2.2 多时相JM距离

在实际的应用中,多时相指数组合是常用的农作物分类方法。因此,研究进一步计算了多时相指数的 JM 距离,结果如图 5 所示。横轴表示各个指数,纵轴表示时间序列长度。对于每个时间序列长度 $i(i=1,2,3,\dots,20)$,都有 C_{23}^i 个时相组合。

从图 5 中可以看出,随着时间序列长度的增加, JM 距离显著增大。当时间序列长度大于 8 时,玉米和棉花、玉米和水稻、玉米和大豆均可以较好地地区分,且 UNVI 指数相对于其他指数具有较好地表现力。在玉米和棉花的区分中,可以看出 UNVI 指数、WDRVI 指数以及红边指数 NDre1 随着时间序列长度的增加,均具有较好的表现力,而 NDVI 指数、EVI 指数和 NDWI 指数在同样的时间序列长度下对玉米和棉花的区分度低于前面 3 个指数;对于玉米和水稻,当时间序列长度相同时,UNVI 的表现要优于 EVI、NDre1 等指数。在玉米和大豆的区分上,UNVI 的表现力等同于 WDRVI,优于 NDre1 和 NDVI 等指数。

对于棉花和水稻、棉花和大豆、水稻和棉花,其可分性均低于上述的 3 种作物组合,但 UNVI 在这 3

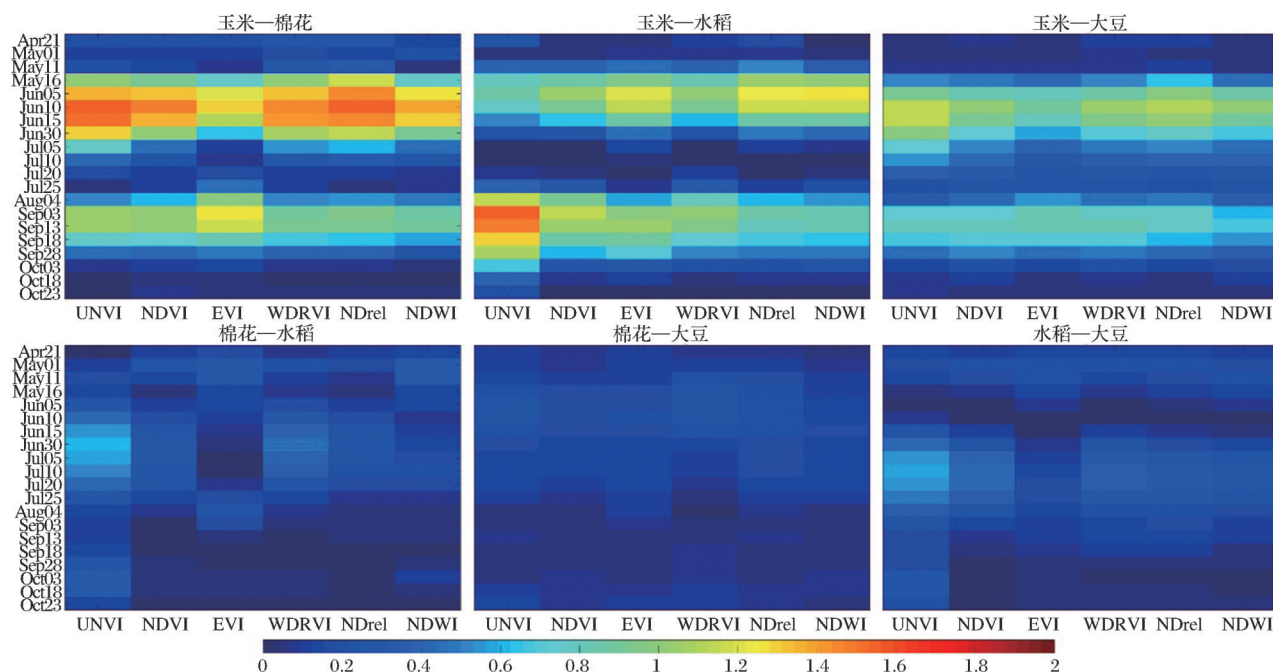


图4 单时相下每个类对之间的JM距离

Fig4 Pairwise JM distance charts for each single temporal VIs

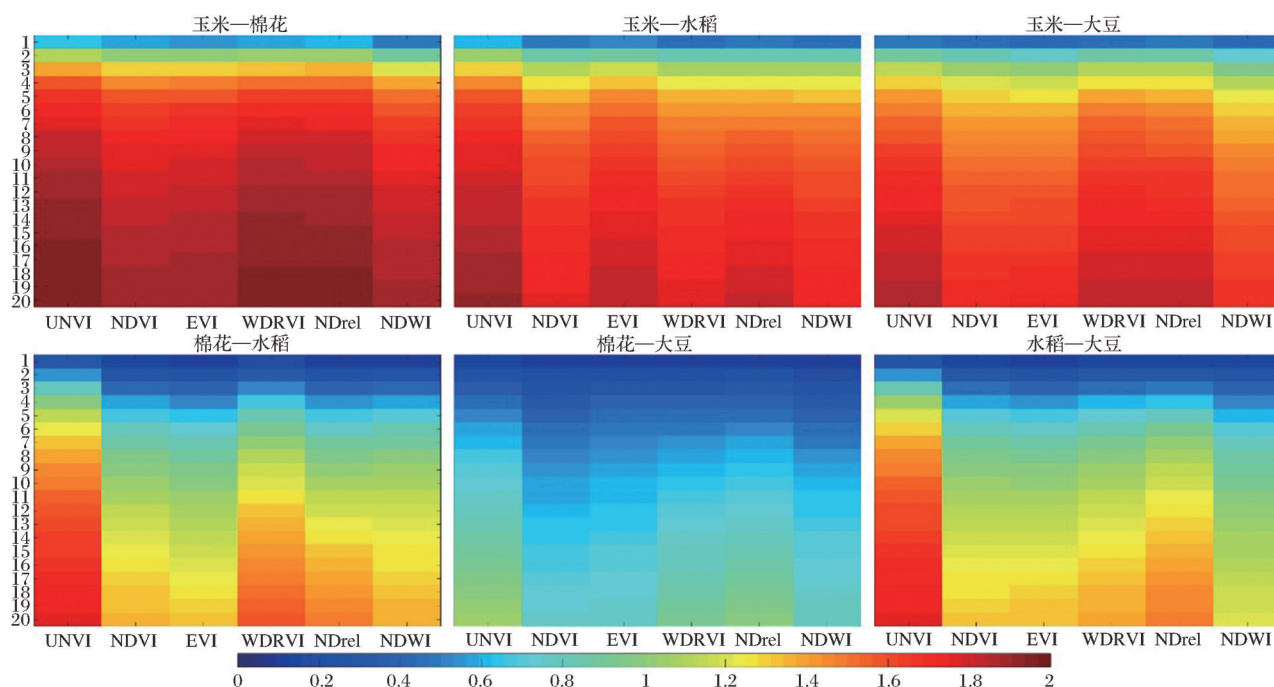


图5 多时相下每个类对之间的JM距离

Fig.5 Pairwise JM distance charts for all multi-temporal VI combinations

种作物组合的区分中仍具有较好的表现力。对于棉花和水稻,从图5中可以看出,随着时间序列长度的增加,UNVI指数的JM距离增加显著,其次是WDRVI指数和NDre1指数,但其JM距离明显低于UNVI指数。在棉花和大豆的区分中,UNVI和WDRVI、NDre1均不易将其区分。同样地,在水稻和大豆的区分中,随着时间序列长度的增加,当时相个数达到10时,UNVI指数已经可以将其很好地区分,但NDre1和WDRVI等指数区分度则较差。

综上所述,UNVI在每种作物组合上均具有较好的表现力,而NDre1、WDRVI等指数,只在某一种或某几种作物组合上略逊于UNVI,但在其他的作物组合上其区分度则较差。

4.3 分类结果与精度分析

通过对不同作物进行可分性分析,可以看出UNVI指数在区分不同作物时,其表现力明显优于其他指数。为了进一步验证时间序列UNVI指数区分不同作物的能力,同时为验证UNVI指数在不同分类器上的适用性,使用随机森林和支持向量机分别对UNVI指数以及其他5个指数对研究区域内的作物进行分类,并计算其混淆矩阵,使用总体精度(OA)、Kappa系数、用户精度(PA)和制图精度(UA)来评价分类精度。

4.3.1 随机森林(RF)

根据各个时间序列指数的特征,采用随机森林

分类器进行分类,并计算出混淆矩阵,图6显示了UNVI和其他5个指数的总体精度(OA)和Kappa系数。由图中可见,UNVI指数的总体精度和Kappa系数最高,其次是NDre1指数。WDRVI、NDWI和NDVI的总体精度和Kappa系数则低于UNVI和NDre1,且EVI的总体精度和Kappa系数最低。这表明使用随机森林进行分类时,UNVI在区分不同作物类别方面明显优于其他指数。

表4显示了4种作物类型的制图精度(PA)和用户精度(UA)。对于UNVI而言,玉米、棉花和水稻

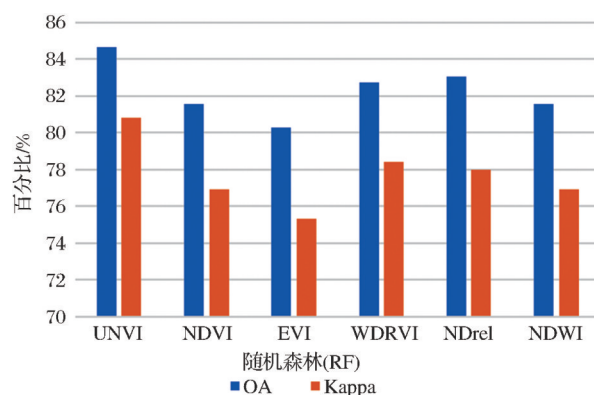


图6 不同指数下4种作物分类的总体精度(OA)和Kappa系数比较

Fig.6 Comparison of overall accuracy(OA) and Kappa coefficient of four crop classifications under different indices

的制图精度均为最高,大豆的制图精度为 94.1%,仅次于 NDWI。

表 4 不同指数下的每种作物分类的制图精度和用户精度(单位:%)

| | UNVI | | NDVI | | EVI | | WDRVI | | NDre1 | | NDWI | |
|----|------|-------|------|-------|------|-------|-------|-------|-------|-------|------|-------|
| | PA | UA | PA | UA | PA | UA | PA | UA | PA | UA | PA | UA |
| 玉米 | 89.8 | 94.03 | 88.2 | 94.23 | 87.8 | 92.91 | 88.6 | 94.36 | 89 | 94.38 | 88.1 | 94.63 |
| 棉花 | 91.3 | 85.33 | 86.7 | 83.93 | 85.7 | 84.6 | 89 | 85.49 | 89 | 85.29 | 87.3 | 85.93 |
| 水稻 | 92.2 | 97.46 | 87.2 | 96.67 | 85.1 | 96.7 | 88.2 | 97.67 | 88 | 97.13 | 86.3 | 97.4 |
| 大豆 | 94.1 | 65.17 | 93.1 | 59.41 | 92.7 | 56.98 | 93.9 | 60.74 | 93.3 | 61.83 | 94.9 | 58.36 |

4.3.2 支持向量机(SVM)

从随机森林分类结果中可以看出,UNVI的分类效果优于其他指数,为了进一步验证 UNVI在其他分类器上的适用性,本研究再次使用支持向量机对研究区作物进行分类,并计算出混淆矩阵。图 7 显示了 6 个指数的总体精度(OA)和 Kappa 系数,由图 7 可见,UNVI 指数的总体精度和 Kappa 系数最高,其次是 WDRVI 指数,NDre1 指数则略差于 WDRVI 指数。这表明使用支持向量机进行分类时,UNVI 仍具有优势,在区分不同作物类别方面优于其他指数。

表 5 显示了使用支持向量机分类时 4 种作物类型的制图精度(PA)和用户精度(UA)。对于 UNVI 而言,棉花、水稻和大豆的制图精度均为最高,玉米的制图精度为 80.2%,且使用 UNVI 指数分类时,其玉米、水稻和棉花的制图精度亦高于其他指数。这

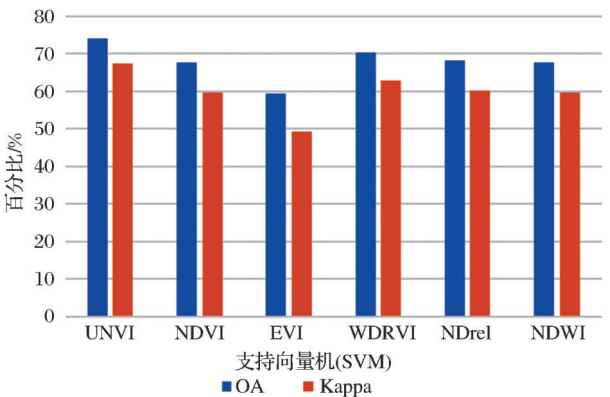


图 7 不同指数下四种作物分类的总体精度(OA)和 Kappa 系数比较

Fig.7 Comparison of overall accuracy (OA) and Kappa coefficient of four crop classifications under different indices

表明使用支持向量机进行作物分类时,UNVI 可以取得较好的精度。

表 5 不同指数下的每种作物分类的制图精度和用户精度(单位:%)

| | UNVI | | NDVI | | EVI | | WDRVI | | NDre1 | | NDWI | |
|----|------|-------|------|-------|------|-------|-------|-------|-------|-------|------|-------|
| | PA | UA | PA | UA | PA | UA | PA | UA | PA | UA | PA | UA |
| 玉米 | 80.2 | 89.91 | 80.5 | 88.75 | 79.2 | 85.81 | 81.3 | 89.54 | 80 | 86.67 | 81.3 | 84.07 |
| 棉花 | 82.6 | 70.42 | 71.5 | 64.36 | 60.2 | 70 | 77.1 | 68.11 | 78.1 | 64.07 | 74.1 | 72.01 |
| 水稻 | 85.1 | 96.7 | 69.5 | 92.91 | 45.3 | 94.18 | 76.2 | 94.07 | 68 | 90.91 | 67.5 | 93.75 |
| 大豆 | 87.3 | 52.31 | 84.7 | 45.13 | 85.5 | 35.37 | 86.3 | 47.6 | 83.7 | 47.42 | 85.9 | 44.19 |

5 结 论

计算了 Sentinel-2 植被指数 UNVI,并通过两个对比实验研究了时间序列 UNVI、NDVI、EVI、WDRVI、NDre1 和 NDWI 6 个指数用于作物种类识别的能力,评价了 UNVI 的应用潜力。实验一通过计算 JM 距离评价 6 个指数的类别可分性,结果显示,UNVI 优于 NDVI、EVI、WDRVI、NDre1、NDWI。在某些作物之间的区分上,WDRVI 和 NDre1 识别效果和 UNVI 相当,但在其他作物类对

的区分上,其效果则不显著,总体上,UNVI 在所有作物类对之间的可分性均较好。结果表明 UNVI 对作物生长动态的敏感性高于其他指数,使得作物之间可分性更高。实验二使用随机森林和支持向量机进行作物分类,分类结果表显示 UNVI 指数的总体精度和 Kappa 系数均居于首位,这表明使用不同的分类算法进行分类时,UNVI 均可以取得较好的作物分类效果,具有稳定性。在使用随机森林进行作物分类时,NDre1 指数的分类效果仅次于 UNVI 指数,但在使用支持向量机进行分类时,WDRVI 指

数的分类效果则优于 NDre1 指数,这表明 NDre1 指数和 WDRVI 指数不具备稳定性,会因分类算法的不同而出现不同的分类结果。

因此,与实验研究的其他植被指数相比,基于 Sentinel-2 的时间序列 UNVI 植被指数在进行作物分类时具有更高的识别精度,可作为农业遥感研究和应用的植被指数。

参考文献 (References):

- [1] Begue A, Arvor D, Bellon B, *et al.* Remote sensing and cropping practices: a review [J]. Remote Sensing, 2018, 10: 99. DOI:10.3390/rs10010099.
- [2] Jia K, Liang S, Wei X, *et al.* Land cover classification of landsat data with phenological features extracted from time series MODIS NDVI data [J]. Remote Sensing, 2014, 6(11): 11518-11532. DOI:10.3390/rs61111518.
- [3] Knight J F, Lunetta R S, Ediriwickrema J, *et al.* Regional scale land cover characterization using MODIS-NDVI 250 m Multi-Temporal imagery: a phenology-based approach [J]. Mapping Sciences & Remote Sensing, 2006, 43(1): 1-23. DOI:10.2747/1548-1603.43.1.1
- [4] Kong F, Li X, Hong W, *et al.* Land Cover classification based on fused data from GF-1 and MODIS NDVI time series [J]. Remote Sensing, 2016, 8(9): 741. DOI: 10.3390/rs8090741.
- [5] Liu K, Su H, Zhang L, *et al.* Analysis of the urban heat island effect in Shijiazhuang, China using satellite and airborne data [J]. Remote Sensing, 2015, 7(4): 4804-4833. DOI: https://doi.org/10.3390/rs70404804.
- [6] Yang Shao, Ross S, Lunetta Brandon, *et al.* An evaluation of time-series smoothing algorithms for land-cover classifications using MODIS-NDVI multi-temporal data [J]. Remote Sensing of Environment, 2016, 174: 258-265. DOI:10.1016/j.rse.2015.12.023.
- [7] Brown J C, Kastens J H, Coutinho A C, *et al.* Classifying Multiyear Agricultural land use data from Mato Grosso using time-series MODIS vegetation index data [J]. Remote Sensing of Environment, 2013, 130: 39-50. DOI: 10.1016/j.rse.2012.11.009.
- [8] Arovor D, Jonathan M, Meirelles M S P, *et al.* Classification of MODIS EVI time series for crop mapping in the state of Mato Grosso, Brazil [J]. International Journal of Remote Sensing, 2011, 32(22): 7847-7871. DOI: 10.1080/01431161.2010.531783.
- [9] Bendini H N, Fonseca L M G, Schwieder M, *et al.* Detailed agricultural land classification in the brazilian cerrado based on phenological information from dense satellite image time series [J]. International Journal of Applied Earth Observation and Geoinformation, 2019, 82: 101872. DOI: 10.1016/j.jag.2019.05.005.
- [10] Clark M L, Aide T M, Grau H R, *et al.* A Scalable approach to mapping annual land cover at 250 m using MODIS time series data: a case study in the dry chaco ecoregion of south America [J]. Remote Sensing of Environment, 2010, 114(11): 2816-2832. DOI:10.1016/j.rse.2010.07.001.
- [11] Peng G, Jie W, Le Y, *et al.* Finer resolution observation and monitoring of global land cover: first mapping results with Landsat TM and ETM+ data [J]. International Journal for Remote Sensing, 2013, 34(7-8): 2607-2654.
- [12] Senf C, Pflugmacher D, Der Linden, S V, *et al.* Mapping rubber plantations and natural forests in Sishuangbanna (south-west China) using multi-spectral phenological metrics from MODIS time series [J]. Remote Sensing, 2013, 5: 2795-2812. DOI:10.3390/rs5062795.
- [13] Wolter P T, Mladenoff D J, Host G E, *et al.* Improved forest classification in the Northern Lake States using Multi-Temporal Landsat imagery [J]. Photogrammetric Engineering and Remote Sensing, 1995, 61(9): 1129-1143.
- [14] Qiong H U, Wenbin W U, Song Q, *et al.* How do temporal and spectral features matter in crop classification in Heilongjiang province, China [J]. Journal of Integrative Agriculture, 2017, 16(2): 324-336. DOI: 10.1016/S2095-3119(15)61321-1.
- [15] Zhang L, Furumi S, Muramatsu K, *et al.* A new vegetation index based on the universal pattern decomposition method [J]. International Journal of Remote Sensing, 2007, 28(1-2): 107-124. DOI:10.1080/01431160600857402.
- [16] Jiao W, Zhang L, Chang Q, *et al.* Evaluating an enhanced Vegetation Condition Index (VCI) based on VIUPD for drought monitoring in the Continental United States [J]. Remote Sensing, 2016, 8(3): 224. DOI:10.3390/rs8030224.
- [17] Su Wei, Zhang Mingzheng, Jiang Kunping, *et al.* Atmospheric correction method for Sentinel-2 satellite imagery [J]. Acta Agronomica Sinica, 2018, 38(1): 0128001. [苏伟, 张明政, 蒋坤萍, 等. Sentinel-2 卫星影像的大气校正方法 [J]. 光学学报, 2018, 38(1): 0128001. DOI:10.1117/3.2533493.]
- [18] Han W, Yang Z, Di L, *et al.* Crop Scape: A web service based Application for exploring and disseminating US continuous geospatial dropland data products for secession support [J]. Computers & Electronics in Agriculture, 2012, 84: 111-123. DOI:10.1016/j.compag.2012.03.005.
- [19] Hansen M C, Loveland, T R. A review of large area monitor-

- ing of land cover change using Landsat data[J]. Remote Sensing of Environment, 2012, 122: 66-74. DOI: 10.1016/j.rse.2011.08.024.
- [20] Boryan C, Yang Z, Mueller R, *et al.* Monitoring US agriculture: the US department of agriculture, national agricultural statistics service, cropland data layer program[J]. Geocarto International, 2011, 26(5): 341-358. DOI: 10.1080/10106049.2011.562309.
- [21] Zhang L, Qiao N, Baig M H A, *et al.* Monitoring vegetation dynamics using the Universal Normalized Vegetation Index (UNVI): an optimized vegetation index-VIUPD[J]. Remote Sensing Letters, 2019, 10(7-9): 629-638. DOI: 10.1080/2150704X.2019.1597298.
- [22] Huete A. 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. DOI: 10.1016/S0034-4257(96)00112-5.
- [23] Fernandez-Manso, Oscar, Quintano, *et al.* Sentinel-2A Red-Edge spectral indices suitability for discriminating burn severity [J]. International Journal of Applied Earth Observation & Geoinformation, 2016, 50: 170-175. DOI: 10.1016/j.jag.2016.03.005.
- [24] Liu Shu, Jiang Qigang, Ma Yue, *et al.* Object-oriented wetland classification based on hybrid feature selection method combining with relief f, multi-objective genetic algorithm and random forest[J]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(1): 119-127. [刘舒, 姜琦刚, 马玥, 等. 基于多目标遗传随机森林特征选择的面向对象湿地分类[J]. 农业机械学报, 2017, 48(1): 119-127.]
- [25] Hofmann W. Remote sensing: the quantitative approach[J]. Earth Science Reviews, 1980, 16: 0-387. DOI: 10.1016/0012-8252(80)90089-6.
- [26] Yeom J, Han Y, Kim Y. Separability analysis and classification of rice fields using KOMPSAT-2 High resolution satellite imagery[J]. Research Journal of Chemistry & Environment, 2013, 17(12): 136-144.
- [27] Deng C, Wu C. BCI: A Biophysical composition index for remote sensing of urban environments[J]. Remote Sensing of Environment, 2012, 127: 247-259. DOI: 10.1016/j.rse.2012.09.009.
- [28] Hu Q, Sullamenashe D, Xu B. *et al.* A Phenology-based spectral and temporal feature selection Method for crop mapping from satellite time series[J]. International Journal of Applied Earth Observation and Geoinformation, 2019, 80: 218-229. DOI: 10.1016/j.jag.2019.04.014.
- [29] Kaufman Y J, Remer L A. Detection of forests using Mid-IR Reflectance: an application for aerosol studies [J]. IEEE Transactions on Geoscience and Remote Sensing, 1994, 32(3): 672-683. DOI: 10.1109/36.297984.
- [30] Somers B, Asner G P. Multi-Temporal hyperspectral mixture analysis and feature selection for invasive species mapping in rainforests [J]. Remote Sensing of Environment, 2013, 136: 14-27. DOI: 10.1016/j.rse.2013.04.006.
- [31] Chen L, Jin Z, Michishita R, *et al.* Dynamic monitoring of wetland cover changes using time-series remote sensing imagery [J]. Ecological Informatics, 2014, 24: 17-26. DOI: 10.1016/j.ecoinf.2014.06.007.
- [32] Thomas I L, Ching N P, Benning V M, *et al.* Review article a review of multi-channel indices of class separability [J]. International Journal of Remote Sensing, 1987, 8(3): 331-350. DOI: 10.1080/01431168708948645.
- [33] Auffarth B, López M, Cerquides J. Comparison of redundancy and relevance measures for feature selection in tissue classification of CT images [C]// Proceedings of the 10th Industrial Conference on Advances in Data Mining: Applications and Theoretical Aspects, 1970. DOI: 10.1007/978-3-642-14400-4_20.
- [34] Belgiu M, Dragut L. Random forest in remote sensing: a review of applications and future directions [J]. ISPRS Journal of Photogrammetry & Remote Sensing, 2016, 114(4): 24-31. DOI: 10.1016/j.isprsjprs.2016.01.011.
- [35] Nitze I, Barrett B, Cawkwell F. Temporal optimisation of image acquisition for land cover classification with random forest and MODIS time-series [J]. International Journal of Applied Earth Observations & Geoinformation, 2015, 34: 136-146. DOI: 10.1016/j.jag.2014.08.001.
- [36] Guan H, Li J, Chapman M. Integration of orthoimagery and lidar data for Object-based urban thematic mapping using random forests [J]. International Journal of Remote Sensing, 2013, 34(13-14): 5166-5186. DOI: 10.1080/01431161.2013.788261.
- [37] Foody G M. Status of land cover classification accuracy assessment [J]. Remote Sensing of Environment, 2002, 80(1): 185-201. DOI: 10.1016/S0034-4257(01)00295-4.

Comparative Study on UNVI Vegetation Index and Performance based on Sentinel-2

Zhu Man^{1,2}, Zhang Lifu¹, Wan Nan¹, Lin Yukun^{1,2}, Zhang Linshan^{1,2},
Wang Sa^{1,2}, Liu Hualiang³

(1. *Arerospace Information Research Institute, Chinese Academy of Sciences, Beijing 100094, China;*

2. *University of Chinese Academy of Sciences, Beijing 100049, China;*

3. *Institute of Remote Sensing and Geographic Information Systems, Peking University, Beijing 100871, China*)

Abstract: Accurate crop identification and classification is an important part of agricultural remote sensing detection, which is very important for crop growth monitoring and yield estimation. In this paper, based on the Sentinel-2 time series images of the United States mixed agricultural belt as the research area, the Universal Normalized Vegetation Index (UNVI) for Sentinel-2 is calculated according to its sensor response function, and two comparisons are made. Experiment to analyze the performance of UNVI and other six indexes in the accurate classification of crops. Experiment 1 uses the JM (Jeffries-Matusita) distance as an indicator to analyze the separability between different crop categories. The results show that UNVI is better than NDVI, EVI, WDRVI, NDre1 and NDWI index. In corn and cotton, corn and rice, In terms of distinguishing between corn and rice, UNVI is better than other indexes in distinguishing ability, but in other crop combinations such as cotton and rice, NDVI and other indexes cannot distinguish them well. At this time, UNVI index can still perform better Distinguishing ability of experiment; Experiment 6 uses random forests and support vector machines to classify crops of the six time series index features. The results show that the UNVI index has the highest overall accuracy and Kappa coefficient, followed by the NDre1 index and the WDRVI index, and the EVI overall accuracy and The Kappa coefficient is the lowest, which indicates that UNVI distinguishes the four main crops of soybean, corn, cotton and rice in the study area better than the other five indexes. In summary, the UNVI index based on the Sentinel-2 time series has greater advantages in crop classification than other remote sensing vegetation indexes studied in this paper. UNVI can be used for agricultural research and application such as crop growth analysis and crop yield research Optional vegetation index.

Key words: Sentinel-2; Time series; UNVI vegetation index; Separability; Crop identification