



DOI:10.12404/j.issn.1671-1815.2403753

引用格式:李宣,宋开山,刘吉平,等.基于遥感融合数据的玉米叶片叶绿素含量反演方法[J].科学技术与工程,2025,25(11):4428-4437.

Li Xuan, Song Kaishan, Liu Jiping, et al. Chlorophyll content inversion method in maize leaves based on remote sensing fusion data[J]. Science Technology and Engineering, 2025, 25(11): 4428-4437.

天文学、地球科学

基于遥感融合数据的玉米叶片叶绿素含量反演方法

李宣^{1,2}, 宋开山², 刘吉平¹, 朱冰雪^{2*}

(1. 吉林师范大学地理科学与旅游学院, 四平 136000; 2. 中国科学院东北地理与农业生态研究所, 长春 130102)

摘要 玉米是中国重要的粮食储备作物之一,其产量直接影响着国家的粮食安全。玉米的叶绿素含量与光合能力密切相关,且显著影响叶片的光合速率和植被生产力,是作物生长监测、病虫害监测和成熟度预测中重要的作物参数。实时、准确监测对于玉米参数和产量预测具有重要意义。以吉林省四平市梨树县典型黑土区为研究区,为解决 Sentinel-2 卫星重访周期间可能会出现的有效影像缺失问题,提出一种基于 Sentinel-2 与 MODIS 影像融合数据的玉米叶片叶绿素反演方法。基于融合影像提取叶绿素敏感特征波段,通过 3 种机器学习算法:随机森林(random forest, RF),梯度提升树(gradient boosting decision tree, GBDT)和极限梯度提升(extreme gradient boosting, XGBOOST)构建玉米叶片叶绿素含量估测模型,估测玉米叶片叶绿素含量并验证模型精度。得到结论如下:①使用 ESTARFM 数据融合算法模拟的数据与真实影像保持了高度的相关性;②影像缺失日期叶片叶绿素反演模型构建中,以融合影像波段反射率和植被指数为输入变量,XGBOOST 模型拟合精度有较好的效果。研究表明,结合融合影像特征波段和机器学习算法可以在影像缺失日期叶片叶绿素含量精准估算,这有效提升了玉米叶绿素含量获取的时间精度,为影像缺失等情况的逐日或大范围叶片叶绿素含量反演研究提供了新的方法,同时为时间间隔更短、更多种类作物生理生化参数的精细化监测提供新思路。

关键词 玉米; 叶绿素含量; 遥感反演; 机器学习; Sentinel2-MODIS 融合; ESTARFM

中图分类号 P237;

文献标志码 A

Chlorophyll Content Inversion Method in Maize Leaves Based on Remote Sensing Fusion Data

LI Xuan^{1,2}, SONG Kai-shan², LIU Ji-ping¹, ZHU Bing-xue^{2*}

(1. College of Geographic Science and Tourism, Jilin Normal University, Siping 136000, China;

2. Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Changchun 130102, China)

[Abstract] Corn is one of the important grain reserve crops in China, and its yield directly impacts national food security. The chlorophyll content of corn is closely related to its photosynthetic capacity and significantly affects the photosynthetic rate of the leaves and vegetation productivity. It is an important crop parameter for monitoring crop growth, pest and disease surveillance, and maturity prediction. Real-time and accurate monitoring is of great significance for corn parameters and yield prediction. This study was conducted in the typical black soil area of Lishu County, Siping City, Jilin Province. To solve the problem of missing effective images that may occur during the revisit period of Sentinel-2 satellites, a method for retrieving corn leaf chlorophyll based on the fusion data of Sentinel-2 and MODIS images was proposed. Using fused imagery, three machine learning algorithms were employed: random forest (RF), gradient boosting decision tree (GBDT), and extreme gradient boosting (XGBOOST) to construct a model for estimating corn leaf chlorophyll content, and the accuracy of the model was verified. The conclusions obtained were as follows. The data simulated using the ESTARFM data fusion algorithm maintained a high correlation with the real imagery. Among the leaf chlorophyll inversion models for missing image dates, where input variables included fused image band reflectance and vegetation index, the XGBOOST model showed good fitting accuracy. The research demonstrates that accurate estimation of leaf chlorophyll content can be achieved even on days with missing imagery, when fusion image feature bands are integrated with machine learning algorithms. This notably improves the

收稿日期: 2024-05-21 修订日期: 2024-08-10

基金项目: 中国科学院战略重点研究项目(XDA28050400); 吉林省科技发展计划重点研发项目(20230202040NC); 国家民用空间基础设施陆地观测卫星共性应用支撑平台项目(2017-000052-73-01-001735)

第一作者: 李宣(2000—), 男, 汉族, 吉林四平人, 硕士研究生。研究方向: 植被定量遥感。E-mail: lixuan@iga.ac.cn。

* 通信作者: 朱冰雪(1993—), 女, 汉族, 吉林辽源人, 博士, 助理研究员。研究方向: 遥感估产及农作物生理生化参数反演。E-mail: zhubingxue@iga.ac.cn。

temporal precision of corn chlorophyll content measurement, presenting a novel method for daily or large-scale inversion studies of leaf chlorophyll content, particularly in scenarios involving image gaps. Furthermore, it illuminates the potential for refined monitoring of physiological and biochemical parameters across a wider range of crops, with shortened time intervals.

[**Keywords**] corn; chlorophyll content; remote sensing inversion; machine-learning; Sentinel2-MODIS fusion; ESTARFM

叶绿素是作物体内重要的生化参数,是植物营养胁迫、光合能力和发育以及衰老各个阶段的关键指标^[1-2]。叶绿素含量直接影响作物氮素含量和生理状况^[3-4],对动态监测作物生长、病虫害、作物产量、预测作物成熟度等具有重要意义^[5]。传统的叶片叶绿素含量可以在实验室中通过化学溶剂萃取加以测量,该方法的精度较高,但是费时费力,仅能获得有限站点的观测数据^[6]。20世纪70年代,遥感技术进入了迅速发展阶段,具有覆盖面广、获取信息快、效率高、方便快捷等特点^[7],让区域作物叶绿素反演成为可能。

基于叶绿素特有的生化结构和光谱吸收特征,利用RGB相机、多光谱相机和高光谱成像仪等实现叶绿素含量的估算已较为普遍。其中普通高清相机虽便于携带,但仅能从较少的波段中提取信息,信息局限^[8];高光谱卫星影像可以获取到较多的波段信息,但影像空间分辨率低,时效性较差等缺点会导致模型精度偏低^[9];多光谱相机是指具有2个以上波段通道的相机,通常包括红、绿、红边和近红外波段,其数据有分辨率较高且易获取等优点^[10],韩茜等^[11]研究表明:利用光谱植被指数可实现作物参量的有效反演。与传统的多光谱数据(Landsat和MODIS)相比,从欧空局获取的Sentinel-2数据具有更高的空间分辨率(10 m)和光谱分辨率,在大面积作物理化参数的精细估算上有更大的优势,吴保升^[12]研究验证理论基于哨兵二号卫星数据冬小麦叶绿素含量相关反演研究的可信度,杨旭等^[13]研究验证了基于哨兵二号卫星数据水稻叶绿素含量反演的可行性。Sentinel-2卫星的重访周期(5 d),但在作物生长旺盛时期,数据质量容易受到天气、云量等因素的影响,导致无法获取指定时间的作物叶绿素含量分布结果,对作物叶绿素含量实时监测造成阻碍。

按照模型的理论基础,作物叶绿素含量的遥感反演方法可分为辐射传输模型方法和统计方法两大类。辐射传输模型能够较好地应用于作物叶绿素含量的估算,但是在建立模型时输入参数较为复杂^[14]。统计模型通过回归分析、主成分分析、神经网络分析、支持向量机等方法^[15],建立遥感数据(反射率或光谱指数)与作物叶绿素含量之间的关系,并以此来对作物的叶绿素含量进行估算,这类模型绕过作物生理机制,避免对作物复杂生理

过程的探索,在作物某些指标观测中更便捷^[16],Lee等^[17]研究表明对比线性回归、随机森林和支持向量机回归方法估算玉米冠层氮素浓度结果的精度,结果证明了两种机器学习模型均优于线性模型。

基于以上因素,通过Emelyanova等^[18]提出的基于混合算法ESTARFM的数据融合方法,将高分辨率空间数据(Sentinel-2)和高分辨率时间数据(moderate-resolution imaging spectroradiometer, MODIS)数据融合,模拟出高分辨率数据缺失日期的遥感影像,基于该影像采用3种机器学习算法:随机森林(random forest, RF)和梯度提升树(gradient boosting decision tree, GBDT)和极限梯度提升(extreme gradient boosting, XGBOOST)方法^[19]。进行玉米叶片叶绿素含量反演模型构建,并对比相近日期的Sentinel-2影像数据采用3种机器学习算法建模的精度,验证融合影像对玉米叶绿素含量估测的可行性,为今后遥感影像缺失日期的玉米叶片叶绿素含量估算提供一种新的研究方法。

1 研究区与数据

1.1 研究区概况

以位于吉林省四平市梨树县的国家百万亩绿色食品原料(玉米)标准化生产基地核心示范区,该示范区位于吉林省四平市西北方向约20 km的梨树县泉眼沟村(124°26.323'E, 43°17.563'N)(图1)。梨树县地处“三大黑土带”和“黄金玉米带”,东辽河、招苏台河横贯全境,是吉林省重要的黑土区,更是中国粮食生产先进县和国家商品粮基地县^[20]。属北温带半湿润大陆季风性气候,四季分明,雨热同季,作物生长期为每年的4—9月,日照、降水较充足^[21]。

1.2 样本数据

选取研究区玉米生长期有效数据三景(2022年6月19日、8月9日、8月24日)开展采样工作,每个生长期在47个样点进行叶片样品采集和玉米生长参数(图2):植株高度、叶片长、宽、叶绿素含量(C_{a+b})的测量和测试(表1)。为了有效减少测量误差,在5 m×5 m的样方选用三点取样法进行参数测量及叶绿素样品采样,将测量结果取均值代表该样方的实测结果,其中叶片叶绿素含量(C_{a+b})测试样本在采集后迅速装入自封袋放置恒温箱内,24 h内测试完毕。

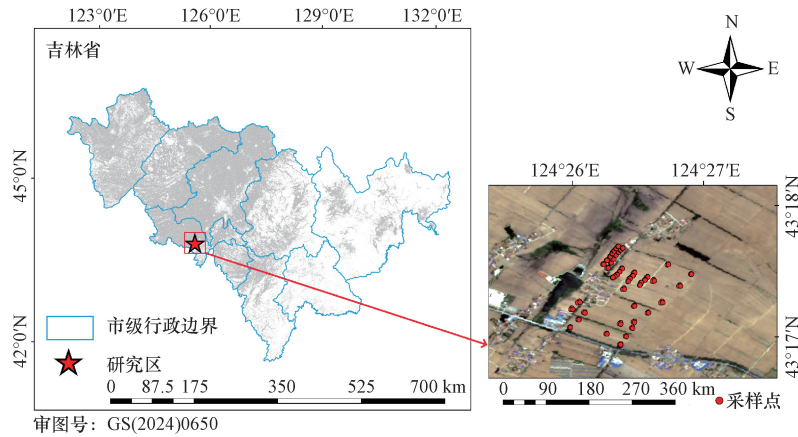


图1 研究区地理位置及采样点分布

Fig. 1 Location of study area and sites

表1 2022年研究区玉米生长季的参数观测数据集
Table 1 Parameter observation data set of maize growing season in the study area in 2022

参数	范围	样本数
株高/cm	26.4~332	179
叶片长/m	0.18~0.91	179
叶片宽/m	0.02~0.13	179
每平方米数	6~17	179
每株叶片数	4~14	179
页面积指数	0.04~9.91	179
叶片叶绿素含量/($\mu\text{g} \cdot \text{cm}^{-2}$)	1.49~145	179

示,使用打孔器在叶片上均匀打孔取样,称重后浸泡于0.01 L的丙酮-水混合液(80%丙酮,20%纯净水),避光静置24 h后放入离心机,得到叶绿素提取液。使用ShimadzuUV-2600分光光度计测量提取液固定波长300~800 nm处的吸光度,并换算成叶绿素含量,换算关系为

$$C_a = (12.25OD_{663.2 \text{ nm}} - 2.79OD_{646.8 \text{ nm}})V/W \quad (1)$$

$$C_b = (21.50OD_{646.8 \text{ nm}} - 5.10OD_{663.2 \text{ nm}})V/W \quad (2)$$

$$C_{a+b} = (7.15OD_{663.2 \text{ nm}} + 18.70OD_{646.8 \text{ nm}})V/W \quad (3)$$

式中: C_a 为叶片的叶绿素a含量,mg/g; C_b 为叶片的叶绿素b含量,mg/g; C_{a+b} 为叶片的叶绿素含量,mg/g; $OD_{646.8 \text{ nm}}$ 为提取液在646.8 nm处的吸光度; $OD_{663.2 \text{ nm}}$ 为提取液在663.2 nm处的吸光度; V 为叶绿素提取液的体积,L; W 为测试叶片样品的重量,g。

通过野外采集的叶片数据计算采样点的叶面积指数(leaf area index, LAI)公式为

$$LAI = 0.75LKmn \quad (4)$$

式(4)中: n 为采样点每平方米株数; m 为每株玉米叶片数量; L 为叶片长,m; K 为叶片宽,m。

将测量带回的叶片样品通过拍照识别的方式提取叶片实际面积[图2(c)],使用叶面积 S 和叶片的湿重 W_s ,计算单位面积的叶绿素含量(leaf chlorophyll content, CHL),公式为

$$CHL = W_s/S \times 1000C_{a+b}LAI \quad (5)$$

1.3 遥感数据

欧洲委员会和欧空局(European Space Agency, ESA)分别于2016年6月23日和2017年3月7日发射了Sentinel-2A和Sentinel-2B卫星,两颗卫星在轨运行的重访周期为5 d。本研究下载Sentinel-2的



图2 数据采集及测试

Fig. 2 Data acquisition and testing

马素霞等^[22]研究表明:利用叶绿素仪(SPAD-502)测定的叶绿素含量存在一定的超前或滞后,想要更好地了解叶片光合作用的内在变化,还需要研究由便携式叶绿素仪(如SPAD-502型)测定的指标SPAD(soil and plant analyzer development)值与各项叶绿素荧光参数的关系。因此,本文中采用化学测试的方法测定叶片的叶绿素的测量,如图[2(d)]所

L1C 级别产品,通过 Sen2cor 模型对 L1C 级别的产品进行大气校正,到 L2A 级别的遥感影像,使用 SNAP(Sentinel-2 Application Platform)软件对 L2A 级别的影像进行 10 m 分辨率的重采样,最后使用 Arcmap 软件提取 Sentinel-2 影像波段反射率。Sentinel-2 影像的下载地址为 <https://scihub.copernicus.eu/dhus/#/home>。

搭载在 Terra 和 Aqua 两颗卫星上的中分辨率成像光谱仪(MODIS),是美国地球观测系统(Earth Observing System,EOS)计划中用于观测全球生物和物理过程的重要仪器,1999 年 12 月 18 日发射的 Terra 和 2002 年 4 月 22 日发射的 Aqua 两颗卫星的交替运行,1、2 d 便可重复观测整个地球表面。MODIS 第 1~2 波段分辨率为 250 m,3~7 波段分辨率为 500 m,其他波段分辨率为 1 000 m,所以 MODIS 通常应用于中国或者全球尺度的长时间序列的影响研究。本研究通过谷歌地球引擎(Google Earth Engine,GEE)获取研究区采样当日的 MODIS 影像数据。

2 模型与方法

2.1 基于 ESTARFM 混合算法的影像融合方法

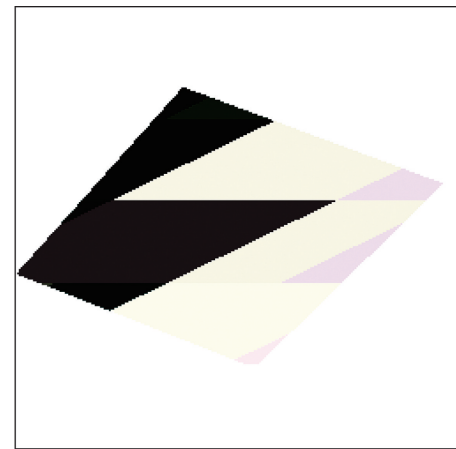
在本研究的卫星数据选取中,由于 Sentinel-2 重访周期相对较长且受采样当日天气和云量等因素的影响,导致无法获取采样当日的卫星影像,所以选择与采样当日最为接近日期的 Sentinel-2 影像进行建模,用于验证融合影像反演叶片和叶绿素含量的可靠性。从表 2 中可以看到采样日期与研究区 Sentinel-2 卫星过境时间。

在表 2 中可以得知,3 个采样时期都无法获取当日 Sentinel-2 的无云影像。选择 ESTARFM 数据融合方法,模拟出采样当日的 10 m 分辨率卫星影像。MODIS 卫星重访周期短(1、2 d),如图 3 所示,该影像的空间分辨率较低, Sentinel-2 卫星重访周期相对较长,但空间分辨率较高。两种卫星影像的波

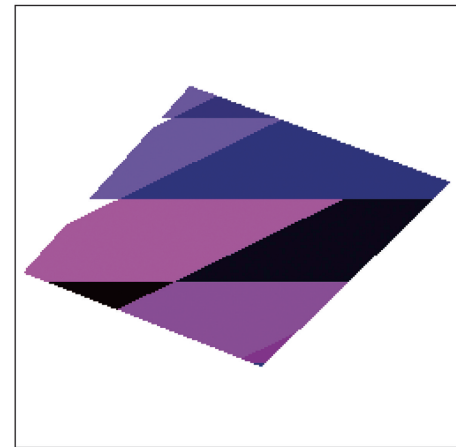
表 2 Sentinel-2 与 MODIS 卫星过境日期和采样日期

Table 2 Date of transit between Sentinel-2 and MODIS satellite and date of sampling

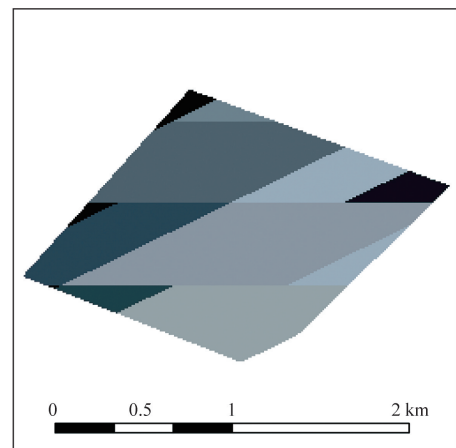
MODIS 过境日期	Sentinel-2 过境日期	采样日期
2022-06-02	2022-06-02	
2022-06-19		2022-06-19
2022-07-12	2022-07-12	
2022-08-06	2022-08-06	
2022-08-09		2022-08-09
2022-08-16	2022-08-16	
2022-08-24		2022-08-24
2022-08-31	2022-08-31	



(a)2022-06-19



(b)2022-08-09



(c)2022-08-24

图 3 采样当日 MODIS 影像

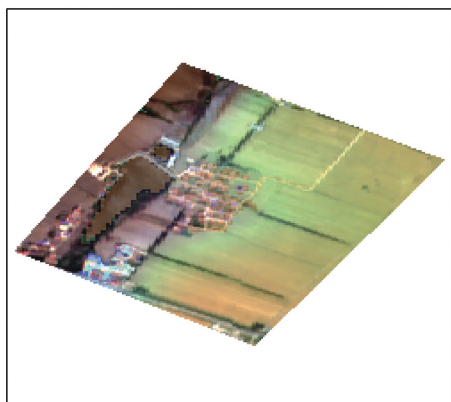
Fig. 3 MODIS image of sampling date

段设置不同, Sentinel-2 和 MODIS 共有 7 个相应波段,除去其中不适用于玉米叶片叶绿素含量的反演研究的沿海气溶胶波段,筛选 Sentinel-2 卫星和 MODIS 卫星相同的 6 个波段,如表 3 所示。通过 ESTARFM 的数据融合方法,模拟得到采样当日的 10 m 分辨率影像,如图 4 所示。

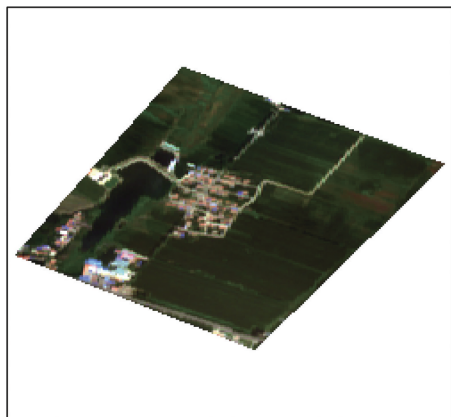
表3 Sentinel-2A 与 MODIS 相同的波段

Table 3 The same bands between Sentinel-2A and MODIS

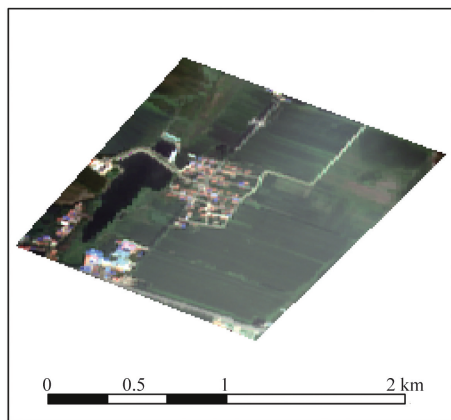
Sentinel-2A			MODIS		
波段	波长/nm	分辨率/m	波段	波长/nm	分辨率/m
B2	496.6	10	B03	459 ~ 479	500
B3	560	10	B04	545 ~ 565	500
B4	664.5	10	B01	620 ~ 670	500
B8	864.8	20	B02	841 ~ 876	500
B11	1 613.7	20	B05	1 230 ~ 1 250	500
B12	2 202.4	20	B06	1 628 ~ 1 652	500



(a)2022-06-19



(b)2022-08-09



(c)2022-08-24

图4 融合采样当日影像结果

Fig. 4 Fusion image of sampling date

2.2 模型构建

提取融合影像和 Sentinel-2 影像中的波段反射率并计算常见的 9 种植被指数(表 4), 采用 RF、GBDT 和 XGBOOST 构建研究区玉米叶片叶绿素含量的反演模型。其中, RF 是一种基于套袋集成学习算法和随机空间算法的分类和预测的机器学习算法^[23]。与其他机器学习算法相比, 它具有参数少, 操作效率高, 有较强的非线性仿真, 能够有效地处理与多变量和大量数据相关的问题^[24-25]。GBDT 属于一种有监督的集成学习算法, 较少参数的 GBDT 具有更高的准确率和更少的运算时间且 GBDT 模型在面对异常数据时具有更强的稳定性^[25]。XGBOOST 是梯度增强决策树算法的一个改进版本, 参看文献。该模型在保证精度的同时可以很好地控制过拟合, 同时也保留了树模型的高可解释性。通过选择自变量的数量生成重要性值, 便于模型分析, 因此它在定量遥感领域受到了越来越多的关注^[26]。为了更好地建立模型, 本研究利用 sklearn 中的网格搜索 (GridSearchcv) 方法^[27], 得到了 3 个模型主要参数的最优参数用于模型建立。

植被指数是一种简单、有效的经验指标。Hunt 等^[28]研究表明, 通过结合基于植被光谱特征的卫星可见光和近红外波段来反映地表植被的条件。它们还可以转化为冠层生物物理参数, 反映了植被的健康和生长状态, 是生态物理参数反演的重要输入参数。本研究选用几种常见的植被指数进行 Pearson 相关性分析, 筛选出相关性较高的植被指数作为叶绿素含量反演模型的输入变量。

表4 使用的植被指数及计算公式

Table 4 Vegetation index and sum calculation formula

植被指数	计算公式	参考文献
NDVI	$(NIR - RED) / (NIR + RED)$	[29]
DVI	$NIR - RED$	[30]
GRVI	$NIR / GREEN$	[31]
GDVI	$NIR - GREEN$	[32]
GNDVI	$(NIR - GREEN) / (NIR + GREEN)$	[33]
LICI	$RE2 / RE - (GREEN - RED) / (GREEN + RED)$	[34]
RECI	$(NIR / RED) - 1$	[35]
NDRE	$(NIR - RE) / (NIR + RE)$	[36]
GCI	$NIR / GREEN - 1$	[37]

注:①NDVI 为归一化植被指数 (normalized difference vegetation index), DVI 为差值植被指数 (difference vegetation index), GRVI 为绿色比值植被指数 (green ratio vegetation index), GDVI 为绿色差值植被指数 (green difference vegetation index), GNDVI 为绿色归一化植被指数 (green normalized difference vegetation index), LICI 为线性组合指数 (linear combinational index), RECI 为红边叶绿素指数 (red-edge chlorophyll index), NDRE 为归一化红边指数 (normalized difference red edge Index), GCI 为绿色叶绿素指数 (green chlorophyll index);②NIR 表示近红外波段反射率, RED 表示红光波段反射率, GREEN 表示绿光波段反射率, RE 和 RE2 表示两个红边波段反射率。

2.3 模型精度评价

共选用3个不同时期的174个叶绿素含量值用于数据分析和模型构建。以确定模型对时间变化的稳健性。将数据集划分为两个部分,其中2/3划分为训练集用来训练模型,剩余的1/3作为测试集用来评估模型。以决定性系数 R^2 、均方根误差(root mean squard error, RMSE)作为评价指标,其计算公式为

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (6)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (7)$$

式中: y_i 为第*i*个样本的叶绿素含量的真实值; \hat{y}_i 为第*i*个样本的叶绿素含量的预测值; \bar{y} 为叶绿素含量的真实值的平均值; n 为样本数。

对比3种模型对单位面积叶绿素总量的预测效果, R^2 越高, RMSE 越低表示该模型的拟合度越高,模型的模拟结果更接近实测值。

3 结果与分析

3.1 融合影像质量评价

本研究从定性和定量两个角度来对融合影像的质量进行评估。通过 ESTARFM 数据融合方法对采样日前后最近日期的 Sentinel-2 影像数据和 MODIS 影像数据使用采样当日的 MODIS 影像模拟采样当日的高分辨率融合影像(图4)。目视可见,相较于 MODIS 影像(图3),融合影像中的地物及植被都清晰可见,且拥有与 Sentinel-2 影像相同的分辨率。此外,选取有真实 Sentinel-2 影像的日期对比融合影像和真实 Sentinel-2 影像进行反射率的相关性分析,融合影像与 Sentinel-2 影像的反射率相关性分析结果表明(图5),融合影像与原始的 Sentinel-2 影

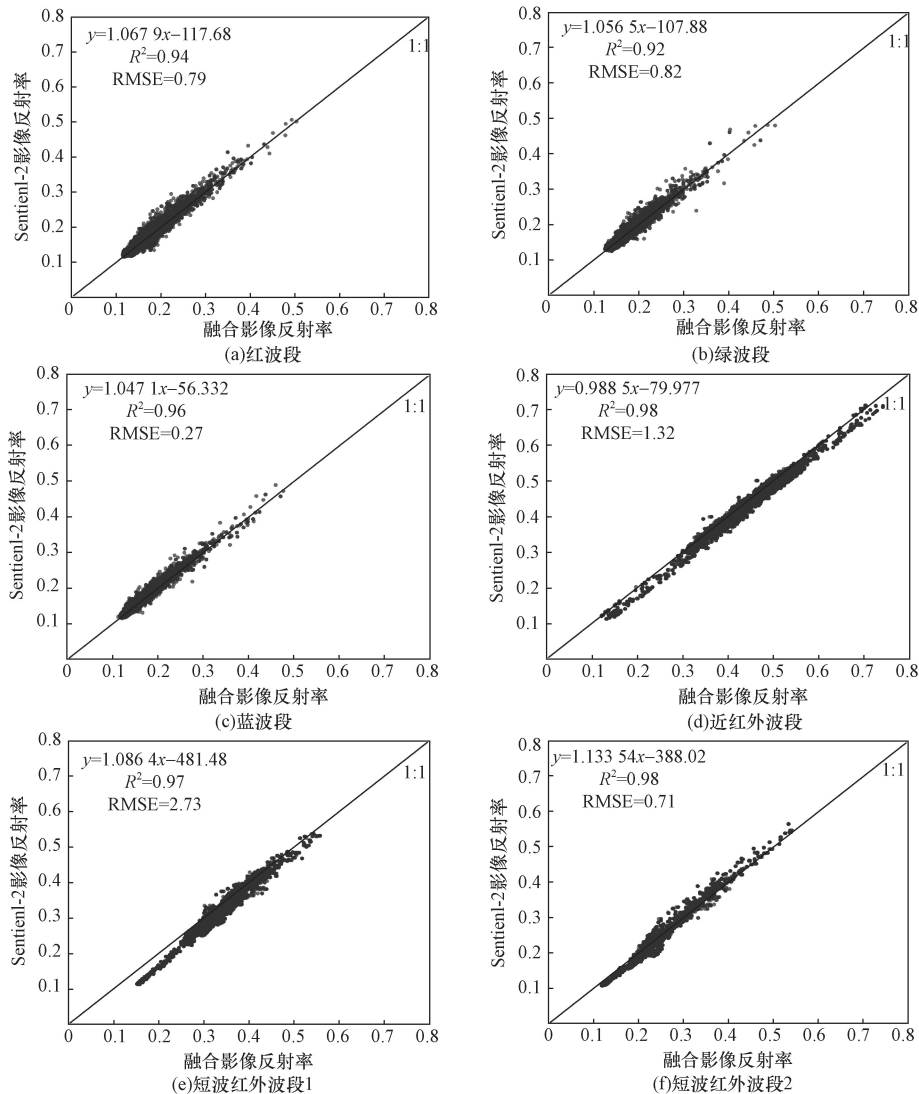


图5 融合影像与 Sentinel-2 影像的反射率相关性分析

Fig. 5 Fusion reflectivity correlation between fused image and Sentinel-2 image

像保持了高度的相关性(R^2 均在 0.92 以上)。以上分析说明,利用 ESTARFM 的数据融合方法来对 Sentinel-2 缺失日期的影像进行模拟,不仅提高了对研究区观测的连续性,也保持了影像数据的光谱一致性。

3.2 建模变量的相关性分析

过 Pearson 相关性分析可以得出 Sentinel-2 中的 12 个波段以及 9 个常见植被指数(表 4)与叶绿素总量的相关性的强弱,进一步剔除弱相关或可能干扰反演模型建立的变量,选择与叶绿素总量相关系数高的几个变量作为反演模型的输入变量。

图 6 中结果表明,Sentinel-2 影像中除了 B1 和 B6 波段反射率与叶绿素含量相关性较低外,波段 B4、B5、B11 和 B12 的反射率与叶绿素存在显著的负相关关系,并且相关系数明显高于其他波段。在图 7 中,植被指数 NDVI、NDRE、RECI、LICI 和 GNDVI 与叶绿素总量存在显著的正相关关系,并且相关系数也明显地高于其他的植被指数。在这些指数中,除 Sentinel-2 的 B1 和 B6 波段和植被指数中的 DVI 和 GDVI 之外,其他指数都与单位面积玉米植株的叶片叶绿素总量有较高的相关性,所以将 B1、B6、DVI 和 GDVI 从建模指数中剔除,其余指数都作为玉米叶片叶绿素含量遥感反演的敏感因子。

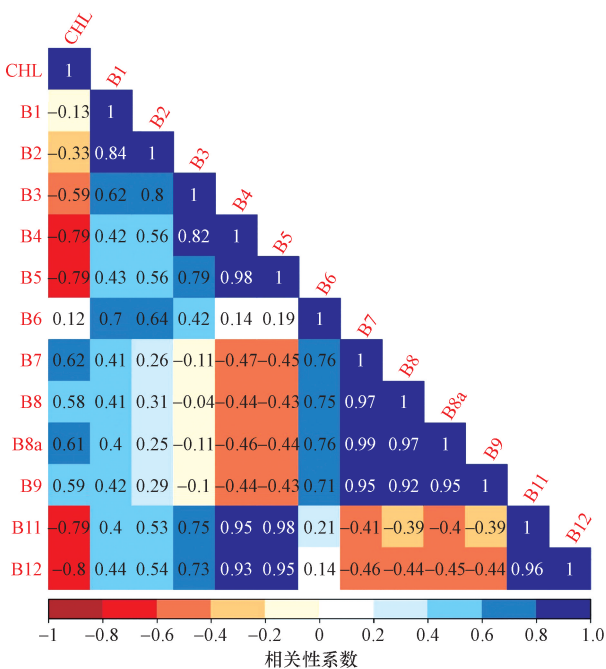


图 6 叶片叶绿素含量与波段的关系

Fig. 6 Relationship between chlorophyll content and wave band in leaves

3.3 利用融合影像变量估算叶片叶绿素含量

融合影像数据模型构建中,选择 6 个波段(表 3)和 6 个植被指数作为构建玉米叶片叶绿素含量模

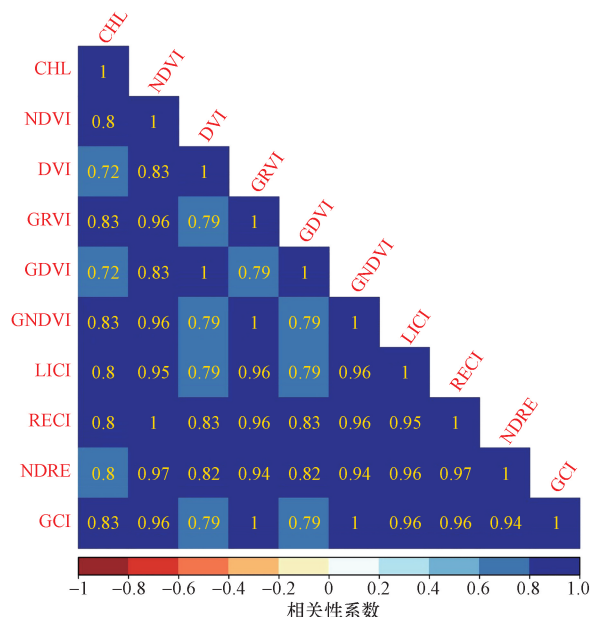


图 7 叶片叶绿素含量与植被指数的关系

Fig. 7 Relationship between chlorophyll content in leaves and vegetation index

型的输入变量。选用 3 种方法分别构建叶片叶绿素含量估测模型,将实测数据与预测数据进行线性拟合,3 个不同时期分 3 种颜色表示,拟合效果如图 8 所示,3 种模型里,XGBOOST 模型的拟合效果最好($R^2 = 0.82$, $RMSE = 19.34 \mu\text{g}/\text{cm}^2$)。

3.4 利用 Sentinel-2 影像变量估算叶片叶绿素含量

使用与采样日相近日期的 Sentinel-2 的 10 个波段反射率和 9 个植被指数,采用与融合数据叶片叶绿素含量估算相同的 3 种模型,构建叶片叶绿素含量的估算模型,将实测数据与预测数据进行线性拟合,3 个不同的采样时间时期分 3 种颜色表示,拟合结果如图 9 所示,3 个模型中,XGBOOST 模型的拟合效果最好($R^2 = 0.88$, $RMSE = 16.61 \mu\text{g}/\text{cm}^2$)。

3.5 分析与制图

融合影像模型的拟合效果($R^2 = 0.82$, $RMSE = 19.34 \mu\text{g}/\text{cm}^2$)较 Sentinel-2 影像数据建模的拟合效果($R^2 = 0.88$, $RMSE = 16.61 \mu\text{g}/\text{cm}^2$)略微逊色,但其优点在于可以对研究区域的叶片叶绿素含量进行逐日估测。从图 8 和图 9 中可以看出,3 个模型中拟合度最高的模型为 XGBOOST 模型。相较于 GBDT 模型只用到一阶导数信息,XGBoost 对代价函数进行了二阶泰勒展开,同时用到了一阶和二阶导数,所以 XGBOOST 模型对单位面积叶绿素含量的估算效果更好。通过融合影像数据进行 XGBOOST 回归为玉米植株单位面积叶绿素含量的反演提供了新的方法,而且本研究的训练数据和测试数据都

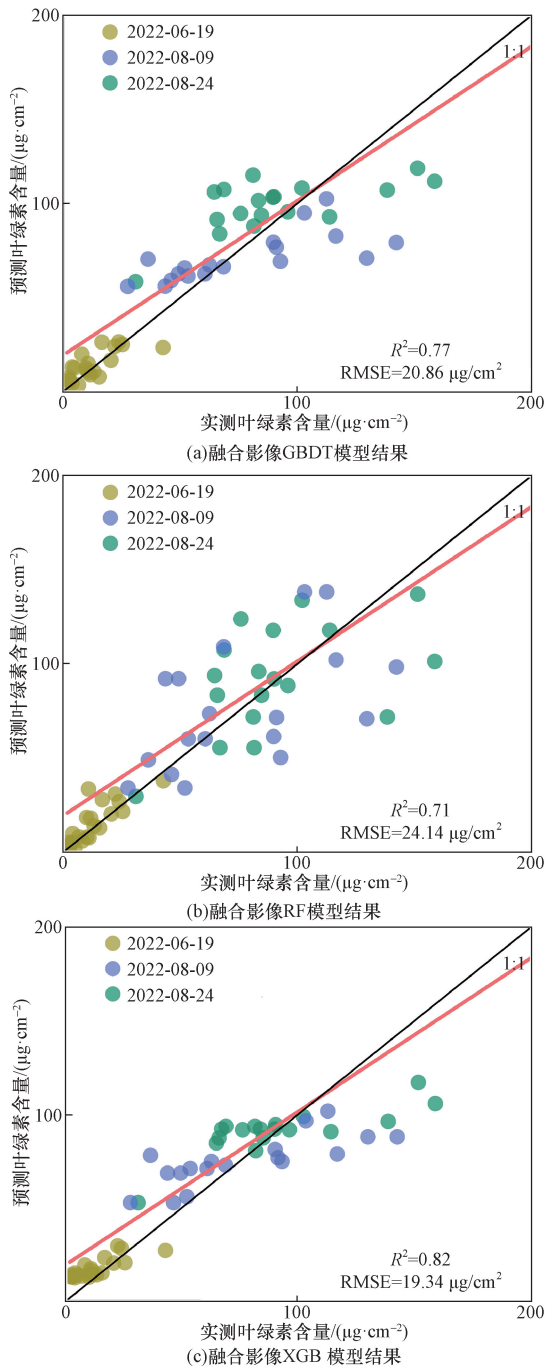


图8 融合影像3个模型的拟合结果对比
 Fig. 8 Comparison of fitting results of three models of fused images

选用了3个不同的玉米生长期的实测数据,这说明XGBOOST模型不受玉米生长期的时间限制,可以在整个玉米生长期对玉米植株单位面积叶绿素总量进行遥感监测,所以本研究选取拟合度最高的XGBOOST模型,使用融合影像估算叶片叶绿素含量模型进行遥感反演,反演结果如图10所示,从图10中可以看出3个不同生长期玉米植株单位面积叶绿素含量的变化与研究区实际情况基本相符,

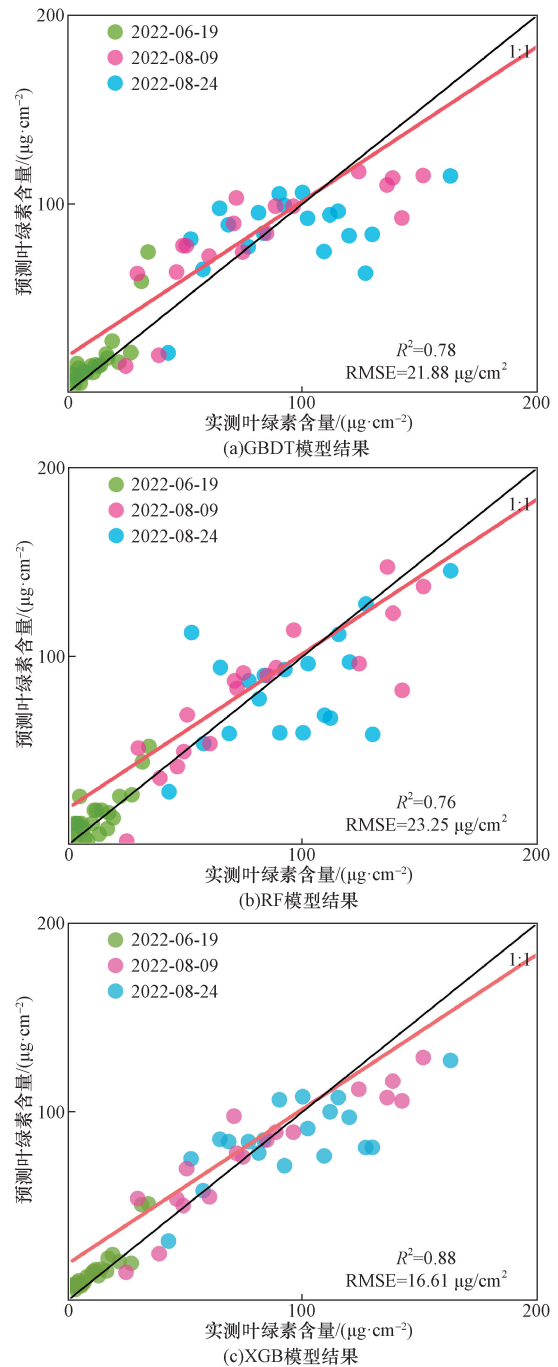


图9 Sentinel-2影像3种模型的拟合结果对比
 Fig. 9 Comparison of fitting results of three models by Sentinel-2 images

实测值与反演值间也有较好的一致性。

4 结论

在大面积和长时间序列反演过程中,云量等因素往往会影响Sentinel-2原始影像的数据质量从而导致无法估测当日作物叶片的叶绿素含量,本研究针对此类问题提出利用融合影像和机器学习算法结合,建立玉米叶片单位面积叶绿素含量的反演模



图10 XGBOOST模型反演结果图

Fig. 10 Inversion results of XGBOOST model

型的解决方案,利用 Sentinel-2 影像和 MODIS 数据成功反演示范区玉米叶片单位面积叶绿素含量,使农作物叶片叶绿素含量的反演不再受有效影像时间间隔的限制。结果表明:

(1) MODIS 数据与 Sentinel-2 的融合影像与 Sentinel-2 影像保持了高度的相关性 (R^2 均在 0.92 以上),很好地解决了在采样当日无法获取高分辨率卫星影像的问题。

(2) 分析 Sentinel-2 的 12 个波段和 9 个植被指数与 3 个玉米生长时期的单位面积叶绿素含量之间的相关性,发现 Sentinel-2 的 B2-B5, B7-B12 波段和植被指数 NDVI、LICI、RECI、NDRE、GRVI、GNDVI、GCI 都与玉米植株的叶片叶绿素含量有较高的相关性,将其用于叶绿素含量的建模,进而提高了建模精度。

(3) 在融合数据构建玉米叶片叶绿素含量反演模型时, XGBOOST 模型 ($R^2 = 0.82$, RMSE = 19.34 $\mu\text{g}/\text{cm}^2$) 拟合效果最优,可以实现玉米多生长期的单位面积叶绿素含量的遥感反演。

(4) 本文方法对于高空间分辨率影像缺失等情况的逐日或大范围叶片叶绿素含量反演研究提供了新的方法,同时为实现时间间隔更短,更多作物化学参数的精细化监测和农业农情的实时监控提供理论依据。

参 考 文 献

[1] Niinemets Ü, Tenhunen J. A model separating leaf structural and physiological effects on carbon gain along light gradients for the shade-tolerant species *Acer saccharum* [J]. *Plant Cell and Environment*, 1997, 20: 845-866.

[2] Peng Y, Gitelson A A. Remote estimation of gross primary productivity in soybean and maize based on total crop chlorophyll content [J]. *Remote Sensing of Environment*, 2012, 117: 440-488.

[3] Feng M, Yang W, Zhang D, et al. Monitoring planting area and

growth situation of irrigation-land and dry-land winter wheat based on TM and MODIS data [J]. *Nongye Gongcheng Xuebao/Transactions of the Chinese Society of Agricultural Engineering*, 2009, 25: 103-109.

- [4] 田静国, 王树东, 张立福, 等. 应用高光谱植被指数反演冬小麦叶绿素的光谱指标敏感性研究 [J]. *科学技术与工程*, 2016, 16(15): 1-8.
- Tian Jingguo, Wang Shudong, Zhang Lifu, et al. Spectral index sensitivity study of winter wheat chlorophyll inversion using hyperspectral remote sensing vegetation index [J]. *Science Technology and Engineering*, 2016, 16(15): 1-8.
- [5] Walburg G, Bauer M, Daughtry C. Effects of nitrogen nutrition on the growth, yield and reflectance characteristics of corn canopies [J]. *LARS Technical Reports*, 1982, 74: 677-681.
- [6] Xue L H, Yang L Z. Deriving leaf chlorophyll content of greenleafy vegetables from hyperspectral reflectance [J]. *ISPRS Journal of Photogrammetry and Remote Sensing*, 2009, 64(1): 97-106.
- [7] Huang Y B, Chen Z X, Yu T, et al. Agricultural remote sensing big data: management and applications [J]. *Journal of Integrative Agriculture*, 2018, 17(9): 1915-1931.
- [8] Ying H, Deng L, Mao Z H, et al. SPAD remote sensing estimation of corn canopy based on digital camera [J]. *Chinese Agricultural Sciences*, 2018, 51(15): 66-77.
- [9] Mao Z H, Deng L, Jie S, et al. Application of UAV multispectral remote sensing in chlorophyll prediction of maize canopy [J]. *Spectroscopy and Spectroscopy*, 2018, 38(9): 677-683.
- [10] Carmona F, Rivas R, Fonnegra D C, et al. Vegetation index to estimate chlorophyll content from multispectral remote sensing data [J]. *European Journal of Remote Sensing*, 2015, 48(1): 319-326.
- [11] 韩茜, 张潇元, 王树东, 等. 冬小麦典型多参量冠层高光谱反演的光谱指标敏感性研究 [J]. *科学技术与工程*, 2017, 17(25): 1701376.
- Han Xi, Zhang Xiaoyuan, Wang Shudong, et al. Sensitivity research on the canopy hyperspectral inversion for the typical multi-parameter of winter wheat [J]. *Science Technology and Engineering*, 2017, 17(25): 1701376.
- [12] 吴保升. 基于哨兵二号卫星影像的冬小麦叶绿素反演研究 [D]. 阿拉尔: 塔里木大学, 2022.
- Wu Baosheng. Study on winter wheat chlorophyll retrieval based

- on Sentinel-2 image [D]. Alaer; TARIM University, 2022.
- [13] 杨旭, 卢学鹤, 石晶明, 等. 基于 Sentinel-2 卫星数据的水稻叶片叶绿素含量反演研究[J]. 光谱学与光谱分析, 2022, 42(3): 866-872.
Yang Xu, Lu Xuehe, Shi Jingming, et al. Satellite data inversion of rice leaf chlorophyll content based on Sentinel-2[J]. Spectroscopy and Spectral Analysis, 2022, 42(3): 866-872.
- [14] Jacquemoud S, Verhoef W, Baret F, et al. PROSPECT + SAIL models: a review of use for vegetation characterization[J]. Remote Sensing of Environment, 2009, 113: S56-S66.
- [15] Houborg R, McCabe M, Cescatti A, et al. Joint leaf chlorophyll content and leaf area index retrieval from Landsat data using a regularized model inversion system (REGFLEC)[J]. Remote Sensing of Environment, 2015, 159: 203-221.
- [16] Dong T, Liu J, Shang J, et al. Assessment of red-edge vegetation indices for crop leaf area index estimation[J]. Remote Sensing of Environment, 2019, 222: 133-143.
- [17] Lee H, Wang J, Leblon B. Using linear regression, random forests, and support vector machine with unmanned aerial vehicle multispectral images to predict canopy nitrogen weight in corn[J]. Remote Sensing, 2020, 12(13): 2071.
- [18] Emelyanova I V, Mccicar T R, Vanniel T G, et al. Assessing the accuracy of blending Landsat-MODIS surface reflectances in two landscapes with contrasting spatial and temporal dynamics: a framework for algorithm selection[J]. Remote Sensing of Environment, 2013, 133: 193-209.
- [19] 邱士其, 赵明松, 芦园园, 等. 基于 XGBoost 模型的安徽省土壤 pH 空间建模[J]. 科学技术与工程, 2023, 23(4): 1472-1480.
Qiu Shiqi, Zhao Mingsong, Lu Yuanyuan, et al. Spatial modeling of soil pH in Anhui Province based on XGBoost model[J]. Science Technology and Engineering, 2023, 23(4): 1472-1480.
- [20] 徐子曦, 唐友, 钟闻宇, 等. 基于 GBDT 算法的吉林省玉米产量预测模型研究[J]. 智慧农业导刊, 2024, 4(2): 15-18.
Xu Zixi, Tang You, Zhong Wenyu, et al. Research on the prediction model of corn yield in jilin Province based on GBDT algorithm [J]. Journal of Smart Agriculture, 2024, 4(2): 15-18.
- [21] 刘伟, 袁福香, 李忠辉, 等. 高光效对吉林省玉米气候资源利用效率的影响[J]. 科学技术与工程, 2022, 22(6): 2237-2244.
Liu Wei, Yuan Fuxiang, Li Zhonghui, et al. Effects of high light efficiency on utilization efficiency of maize climate resources in Jilin Province [J]. Science Technology and Engineering, 2022, 22(6): 2237-2244.
- [22] 马素霞, 王艳玲, 唐剑武. 水稻冠层叶片 SPAD 值与叶绿素、叶氮含量及净光合速率的变化特征[J]. 科学技术与工程, 2016, 16(24): 1601803.
Ma Suxia, Wang Yanling, Tang Jianwu. Variation of rice canopy SPAD value and chlorophyll, leaf nitrogen content and net photosynthetic rate[J]. Science Technology and Engineering, 2016, 16(24): 1601803.
- [23] Leo B. Random forests[J]. Machine Learning, 2001, 45: 5-32.
- [24] Han H, Wan R, Li B. Estimating forest aboveground biomass using Gaofen-1 images, Sentinel-1 images, and machine learning algorithms: a case study of the Dabie Mountain Region, China [J]. Remote Sensing, 2021, 14(1): 176.
- [25] Liang W Z, Luo S Z, Zhao G Y, et al. Predicting hard rock pillar stability using GBDT, XGBoost, and LightGBM algorithms [J]. Mathematics, 2020, 8(5): 765.
- [26] Zhang W G, Wu C Z, Zhong H Y, et al. Prediction of undrained shear strength using extreme gradient boosting and random forest based on Bayesian optimization[J]. Geoscience Frontiers, 2021, 12(1): 469-477.
- [27] Ahmad G N, Fatima H, Ullah S, et al. Efficient medical diagnosis of human heart diseases using machine learning techniques with and without grid search CV [J]. Ieee Access, 2022, 10: 80151-80173.
- [28] Hunt E R, Daughtry C S T, Eite L J U H, et al. Remote sensing leaf chlorophyll content using a visible band index[J]. Agronomy Journal, 2011, 103(4): 1090-1099.
- [29] Pettorelli N. The normalized difference vegetation index [M]. Bridge: Oxford University Press, 2013: 18-29.
- [30] Jiang Z, Huete A R, Chen J, et al. Analysis of NDVI and scaled difference vegetation index retrievals of vegetation fraction[J]. Remote Sensing of Environment, 2006, 101(3): 366-378.
- [31] Bendig J, Yu K, Aasen H, et al. Combining UAV-based plant height from crop surface models, visible, and near infrared vegetation indices for biomass monitoring in barley [J]. International Journal of Applied Earth Observation and Geoinformation, 2015, 39: 79-87.
- [32] Wu W. The Generalized difference vegetation index (GDVI) for dryland characterization [J]. Remote Sensing, 2014, 6(2): 1211-1233.
- [33] Moradi F, Darvishsefat A A, Pourrahmati M R, et al. Estimating aboveground biomass in dense hyrcanian forests by the use of Sentinel-2 data[J]. Forests, 2022, 13(1): 104.
- [34] Li D, Chen J M, Zhang X, et al. Improved estimation of leaf chlorophyll content of row crops from canopy reflectance spectra through minimizing canopy structural effects and optimizing off-noon observation time [J]. Remote Sensing of Environment, 2020, 248: 111985.
- [35] Fierro A, Angers D, Beauchamp C J, et al. Decomposition of paper de-inking sludge in a sandpit minesoil during its revegetation [J]. Soil Biology and Biochemistry, 2000, 32(2): 143-150.
- [36] Boiarskii B, Hasegawa S. Comparison of NDVI and NDRE indices to detect differences in vegetation and chlorophyll content [J]. Journal of Mechanics of Continua and Mathematical Sciences, 2019, 4: 20-29.
- [37] Wu C, Niu Z, Gao S J. The potential of the satellite derived green chlorophyll index for estimating midday light use efficiency in maize, coniferous forest and grassland[J]. Ecological Indicators, 2012, 14(1): 66-73.