

引文格式:王宇,杨丽萍,任杰,等.联合ALOS-2和Landsat 8的绿洲土壤水分反演模型研究[J].武汉大学学报(信息科学版),
2024,49(9):1630-1638.DOI:10.13203/j.whugis20220008



Citation: WANG Yu, YANG Liping, REN Jie, et al. An Oasis Soil Moisture Inversion Model Using ALOS-2 and Landsat 8 Data [J]. Geomatics and Information Science of Wuhan University, 2024, 49(9): 1630-1638. DOI: 10.13203/j.whugis20220008

联合ALOS-2和Landsat 8的绿洲土壤水分反演模型研究

王宇¹ 杨丽萍¹ 任杰¹ 张静¹ 孔金玲¹ 侯成磊²

¹ 长安大学地质工程与测绘学院, 陕西 西安, 710054

² 山东农业工程学院国土资源与测绘工程学院, 山东 济南, 250100

摘要: 机器学习和多源数据融合是土壤水分反演研究的热点方向, 但对L波段合成孔径雷达(synthetic aperture radar, SAR)数据的研究较少。以额济纳绿洲为研究区, 利用ALOS-2 PALSAR-2和Landsat 8影像提取雷达和光学特征参数, 通过参数重要性评分进行特征筛选, 采用随机森林方法建立基于雷达、光学以及雷达-光学特征参数协同的土壤水分反演模型, 对比模型精度, 反演绿洲土壤水分。结果表明, 与C波段相比, L波段SAR数据对干旱荒漠绿洲区土壤水分含量敏感性更高; 雷达特征参数中重要性较高的为表面散射和体散射分量, 二面角散射和螺旋体散射分量相对偏低; 光学特征参数中植被供水指数重要性最高, 增强型植被指数重要性最低。雷达特征参数方案最优模型决定系数 R^2 、均方根误差(root mean square error, RMSE)分别为0.67、2.16%, 光学特征参数方案模型精度普遍较低且精度相当, R^2 、RMSE分别为0.5、2.47%; 雷达-光学参数协同反演的最优模型 R^2 、RMSE分别为0.72、1.99%, 相比单一数据源, R^2 分别提升7.46%、38.4%, RMSE分别降低8.54%、22.6%。研究证明, 基于多源数据融合的随机森林模型在干旱荒漠绿洲区具有较高的预测精度和良好的适用性。

关键词: ALOS-2 PALSAR-2; Landsat 8; 土壤水分; 随机森林; 特征参数

中图分类号: P237

文献标识码: A

收稿日期: 2022-12-08

DOI: 10.13203/j.whugis20220008

文章编号: 1671-8860(2024)09-1630-09

An Oasis Soil Moisture Inversion Model Using ALOS-2 and Landsat 8 Data

WANG Yu¹ YANG Liping¹ REN Jie¹ ZHANG Jing¹ KONG Jinling¹ HOU Chenglei²

¹ School of Geological Engineering and Geomatics, Chang'an University, Xi'an 710054, China

² College of Land Resources and Surveying & Mapping Engineering, Shandong Agriculture and Engineering University, Ji'nan 250100, China

Abstract: Objectives: Integration of machine learning and multi-source data becomes a hot topic in soil moisture inversion, where relatively few studies are performed on L-band synthetic aperture radar (SAR) imagery. **Methods:** ALOS-2 PALSAR-2 and Landsat 8 images of Ejina Oasis are used to extract the radar and optical characteristic parameters which are then screened according to the importance score. Random forest is adopted to establish different soil moisture inversion models based on radar, optical, and radar-optical integrated parameters. Model accuracies are evaluated and soil moisture content in Ejina Oasis is inverted. **Results:** Compared with C-band, L-band SAR data is more sensitive to soil moisture content in arid desert oasis. With regard to radar characteristic parameters, surface and volume scattering components have higher important scores, while dihedral and helix scattering component are less important. As for optical characteristic parameters, vegetation water supply index takes the most important place while the enhanced vegetation index is the least important one. The determination coefficient R^2 and root mean square error (RMSE) of radar characteristic parameter scheme are 0.67 and 2.16%, respectively. The accu-

基金项目: 国家自然科学基金(41371220, 42071345); 中央高校基本科研业务费专项资金(300102269112)。

第一作者: 王宇, 硕士生, 主要从事土壤水分遥感研究。wangyu_9711@163.com

通讯作者: 杨丽萍, 博士, 副教授。zylpyang@chd.edu.cn

racy of optical characteristic parameter scheme model is generally low and the accuracy is equivalent, with R^2 and RMSE about 0.5 and 2.47%, respectively. R^2 and RMSE of the optimal radar-optical integrated parameter inversion model are 0.72 and 1.99%, respectively. Compared with either single data source, R^2 is increased by 7.46% and 38.4%, while RMSE is decreased by 8.54% and 22.6%. **Conclusions:** The research proves that the random forest model based on multi-source data fusion has higher prediction accuracy and better applicability in arid desert oasis area.

Key words: ALOS-2 PALSAR-2; Landsat 8; soil moisture; random forest; characteristic parameter

土壤水分是影响生态、水文和气候等环境过程的重要参数,土壤水分的准确获取和长期监测对于作物生长、旱情监测与预报等具有积极作用。重量法和时域反射法等传统的土壤水分测量方法需要定时定点测量,无法进行大面积长时效的实时观测。遥感技术和理论不断发展为大范围土壤水分的获取提供了有效途径^[1]。

微波遥感具有全天时、全天候工作的特点,且具有一定的穿透能力,已成为土壤水分反演的重要手段。被动微波遥感起步较早、算法成熟,但空间分辨率较低,在较小尺度研究中具有一定的局限性。主动微波遥感中的极化合成孔径雷达(polarimetric synthetic aperture radar, PolSAR)空间分辨率较高,同时可提供对土壤水分反演有利的极化特征参数,国内外学者已对此进行了大量研究。文献[2]提出了一种双分量模型(表面散射和体散射)的极化分解方法,忽略了二面角散射分量,在农田土壤水分反演中取得了较好的效果。文献[3]基于 Radarsat-2 数据,利用多种分解方法提取多种极化特征参数,对美国黑云杉林的土壤水分进行反演,效果良好。上述研究表明,PolSAR 极化特征参数在土壤水分反演中潜力巨大。然而,土壤水分和极化特征参数之间并非简单的线性关系,通常为了提高反演精度,往往考虑较多参数。简单的线性回归分析方法无法准确、全面地衡量大量特征参数在土壤水分反演中的作用,而机器学习方法对于参数的数量和类型没有限制,能够较好地弥补简单回归分析方法在土壤水分反演中的不足。因此,当考虑多极化特征参数时,常常与机器学习方法相结合。

随机森林(random forest, RF)^[4]和支持向量回归机(support vector regression, SVR)^[5]是目前较为常用的机器学习方法。文献[6]将后向散射系数作为输入变量、土壤水分作为输出变量,对 RF 和 SVR 等 12 种机器学习方法的性能进行了综合评估,认为 RF 性能最佳,适用于土壤水分反演。文献[7]基于 Radarsat-2 数据提取多种极化

特征参数并进行特征筛选,评价了 RF 在土壤水分反演中的适用性。结果表明,RF 能较好地模拟冬小麦各生长阶段的土壤水分含量,其性能优于人工神经网络(artificial neural network, ANN)和 SVR。综上,机器学习与 PolSAR 数据相结合能够较好地实现土壤水分模拟。然而,在植被覆盖地区,土壤水分的反演精度常常受到植被覆盖和地表粗糙度的影响,仅考虑极化特征具有一定局限性。

因此,不少学者协同利用 PolSAR 数据和光学数据,以提升土壤水分反演精度。文献[8]基于 Sentinel-1/2 数据,结合水云模型对冬小麦区域的土壤水分进行反演,结果表明,融合植被指数(fusion vegetation index, FVI)的反演效果最优,且能在一定程度上去除冬小麦对后向散射系数的影响。文献[9]基于 Sentinel-1/2 数据提取后向散射系数和植被指数,利用 SVR 和广义回归神经网络(generalized regression neural network, GRNN)反演了生长初期冬小麦区的土壤水分,认为加入植被指数的 SVR 模型效果最优。文献[10]基于 Sentinel-1 和 Landsat 8 影像,建立 SVR 模型,将后向散射系数、归一化植被指数(normalized difference vegetation index, NDVI)、地形因子等作为特征变量,在洪泛平原的水分反演中取得了较好效果,并认为 NDVI 能够在一定程度上减少植被覆盖的影响。上述研究表明,协同利用 PolSAR 和光学数据能够在一定程度上减少反演的不确定性,使反演结果更接近真实情况。然而,已有研究多基于 C 波段 SAR 数据,对于 L 波段在土壤水分反演中的效果有待深入^[11]。

本文以内蒙古额济纳绿洲为研究区,利用 L 波段 ALOS-2 PALSAR-2 影像和 Landsat 8 影像提取雷达、光学特征参数,通过平均精度下降(mean decrease accuracy, MDA)评估各参数重要性,建立并对比雷达、光学、雷达-光学协同的多参数组合方案的随机森林土壤水分反演模型,对额济纳绿洲区进行土壤水分反演。

1 研究区与数据

1.1 研究区概况

额济纳绿洲位于内蒙古自治区西部,气候干旱,降水稀少,年均气温 8.3°C ,年均降水量约 34 mm ,潜在蒸发量约 $3\ 700\sim 4\ 000\text{ mm}$,属温带大陆性气候。2019-09-24-2019-09-29,在研究区共布设土壤采样点111个,采集表层 $0\sim 10\text{ cm}$ 的土壤样品,在实验室通过烘干称重法^[12]获取各采样点的土壤水分含量。土壤采样点覆盖林地、草地、耕地和裸地等地物类型,实测土壤水分含量为 $0.02\%\sim 15.34\%$ 。研究区位置和采样点分布如图1所示。图1采用2019-09-24的Landsat 8影像,波段组合为7、5、4。

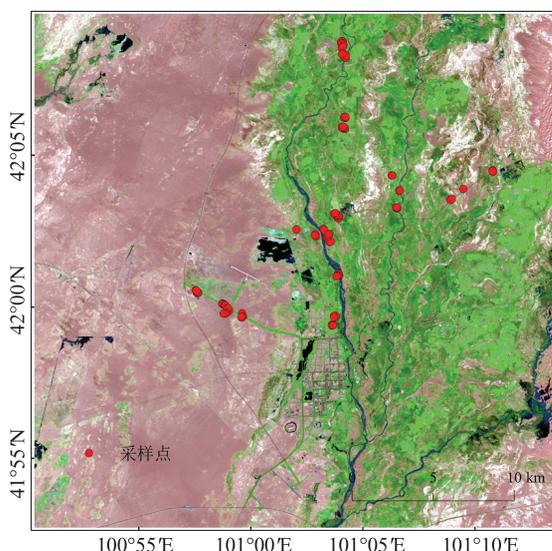


图1 研究区位置及采样点分布图

Fig. 1 Location of the Study Area and Distribution of the Sampling Points

1.2 遥感数据及预处理

2014年日本宇航局发射ALOS-2卫星,搭载L波段全极化合成孔径雷达传感器PALSAR-2,卫星轨道高度为 628 km ,重访周期为 14 d 。本文采用2019-09-29的ALOS-2 PALSAR-2升轨影像,包含水平极化(horizontal polarization, HH)、垂直极化(vertical polarization, VV)、水平-垂直极化(horizontal-vertical polarization, HV)、垂直-水平极化(vertical-horizontal polarization, VH)这4种极化方式,视向为右视,产品等级为Level 1.1级,中心入射角为 33.87° ,空间分辨率为 6 m 。利用SARscape软件首先对ALOS-2影像进行多视、滤波和地理编码,提取后向散射系数 σ_{VV} 、 σ_{VH} 、 σ_{HV} 、 σ_{HH} ;然后利用PolSARpro软件对ALOS-2影

像进行滤波处理,通过极化分解提取极化特征参数;最后通过距离多普勒地形校正和重采样,得到与光学影像相匹配的后向散射系数和极化特征参数图像。

Landsat 8卫星于2013年由美国成功发射,卫星上携带运行陆地成像仪(operational land imager, OLI)和热红外传感器(thermal infrared sensor, TIRS)2种载荷。其中,OLI包括9个波段,除全色波段的空间分辨率为 15 m 外,其余波段的空间分辨率均为 30 m ;TIRS包括2个波段,空间分辨率为 100 m ,卫星重访周期为 16 d 。本文采用2019-09-24的Landsat 8影像,云量为 0.02% 。为了消除光照和大气等因素的影响,首先对Landsat 8影像进行辐射定标和大气校正,然后提取地表温度(land surface temperature, LST)和绿叶指数(green leaf index, GLI)等光学特征参数。

2 研究方法

2.1 雷达特征参数提取

极化分解是PolSAR影像目标特征描述与参数提取的主要方法^[13],其目的在于将复杂的散射机制表示为几种具有对应散射矩阵的单一散射分量。根据分解机制和针对目标的不同,已有学者提出了Freeman-Durden分解^[14](FD)、Cloude-Pottier分解^[15](CP)、Yamaguchi分解^[16](YM)、An & Yang分解^[17](AY)、H/A/ α 分解^[18]、van Zyl分解^[19](VZ)和Bhattacharya & Frery分解^[20](BF)等多种极化分解方法。基于ALOS-2 PALSAR影像和上述极化分解方法,提取36个雷达特征参数,具体见表1,其中 σ_{VV} 、 σ_{HH} 为同极化后向散射系数, σ_{HV} 、 σ_{VH} 为交叉极化后向散射系数,下标Vol、Odd、Surf、Dbl和Hlx分别代表体散射、奇次、表面、二面角和螺旋体散射分量。

2.2 光学特征参数提取

基于Landsat 8 OLI/TIRS数据,提取11个光学特征参数,其中植被指数通过波段反射率计算得到,LST通过ENVI扩展工具采用大气校正法提取。各参数计算公式参见文献^[21-31],包含NDVI、综合光谱响应指数(combined spectral response index, COSRI)、增强型植被指数(enhanced vegetation index, EVI)、GLI、全球植被水分指数(global vegetation moisture index, GVMI)、红外百分比植被指数(infrared percentage vegetation index, IPVI)、植被供水指数(vegetation supply water index, VSWI)、土壤改良植被指数

表 1 雷达特征参数

Tab. 1 Radar Characteristic Parameters

参数类型	处理方法	参数个数	雷达特征参数
后向散射系数	标准强度和相位处理	4	σ_{VV} 、 σ_{VH} 、 σ_{HV} 、 σ_{HH}
	Freeman-Durden 分解 ^[14]	3	FD _{Vol} 、FD _{Odd} 、FD _{Dbi}
	Cloude-Pottier 分解 ^[15]	3	CP _{Dbi} 、CP _{Surf} 、CP _{Vol}
极化特征参数	Yamaguchi 分解 ^[16]	3	YM3 _{Vol} 、YM3 _{Odd} 、YM3 _{Dbi}
		4	YM4 _{Vol} 、YM4 _{Odd} 、YM4 _{Dbi} 、YM4 _{Hlx}
	An & Yang 分解 ^[17]	3	AY3 _{Vol} 、AY3 _{Odd} 、AY3 _{Dbi}
		4	AY4 _{Vol} 、AY4 _{Odd} 、AY4 _{Dbi} 、AY4 _{Hlx}
	H/A/ α 分解 ^[18]	5	Alpha、Anisotropy、Entropy、Lambda、RVI
	van Zyl 分解 ^[19]	3	VZ _{Vol} 、VZ _{Odd} 、VZ _{Dbi}
	Bhattacharya & Frery 分解 ^[20]	4	BF _{Vol} 、BF _{Odd} 、BF _{Dbi} 、BF _{Hlx}

(modified soil adjusted vegetation index, MSA-VI)、温度植被干旱指数 (temperature vegetation drought index, TVDI)、比值植被指数 (ratio vegetation index, RVI) 和 LST。

2.3 随机森林建模与评价

随机森林是一种基于决策树和 Bagging 的集成学习方法。决策树是依据一定的划分规则对变量空间进行拆分,并根据相似实例对数据集分组的树形结构。Bagging 是一种从原样本集中随机抽取训练样本的技术,这意味着某些样本可能在训练中多次使用,而一些样本可能永远都不会被使用^[32],未被使用的样本称为袋外(out of bag, OOB)数据,可用来评估 RF 的预测效果。MDA 是 RF 中评估特征变量重要性的参数,通过随机打乱 OOB 数据中某一特征变量值,计算 OOB 误差,根据其下降程度评估特征变量重要性。通常,OOB 数据误差越大,参数越重要。

本文依据重要性评分结果选取最优特征参数方案建立 RF 模型,包括 2 个关键参数,即决策树数量 N 和节点分裂时输入的特征变量数 M 。采用 $N = \{400, 800, 1200, 1600, 2000, 2400\}$, $M = \{1, 2, 3, 4, 5, 6\}$ 的网格搜索,获得最优参数为 $N = 1600, M = 3$ 。模型精度评定采用 K 折交叉验证, $K = 10$, 精度评定指标采用决定系数 R^2 和均方根误差 (root mean square error, RMSE)。

3 结果与分析

3.1 参数重要性评价

由于特征参数较多,因此本文采用 MDA 分别对雷达特征参数和光学特征参数的重要性进行评分,结果如图 2 所示。

从图 2(a)可以看出,从散射方式看,奇次散

射、表面散射和体散射在雷达特征参数中的重要性较高,而二面角散射和螺旋体散射的重要性相对较低。原因可能是由于研究区内部虽有胡杨、怪柳等植被覆盖,但大部分地物还是以低矮草地、戈壁、沙漠为主,因此主要散射机制仍为奇次散射、表面散射和体散射。(1)在奇次散射和表面散射分量中,FD、YM4、AY4、AY3 和 YM3 分解的重要性程度相对较高,在重要性排名中占据前 5;而 BF、VZ、CP 分解相对其他分解方式的重要性程度有所下降,分别占据 7~9 名。(2)在体散射分量中,YM3、AY3 和 FD 分解的重要性程度相对较高,分别排名 12~14 位,CP 分解的重要性偏低,对土壤水分含量变化不太敏感。(3)二面角散射分量的重要性程度大部分低于体散射分量,但也存在例外,如 VZ、AY3 分解的二面角散射分量分别优于 YM4、AY4 的体散射分量,重要性排名分别为 15、17 位。(4)螺旋体散射分量的重要性偏低,原因是由于该分量的引入是为了解决城区的人造地物不满足反射对称问题^[33],研究区内人造地物相对较少,因而其重要性偏低。(5)在 H/A/ α 分解中, Lambda 对土壤水分反演的贡献率较高,重要性排名第 6,仅次于表面散射分量;RVI 相对 Lambda 重要性程度有所下降,主要原因是 RVI 能够量化区域植被影响差异^[7],但研究区植被较为稀疏,因此重要性偏低;其他参数的重要性排序依次为 Entropy > Anisotropy > Alpha。(6)在 4 个后向散射系数中,同极化后向散射系数的重要性高于交叉极化的后向散射系数,说明同极化后向散射系数相较交叉极化后向散射系数更适合于稀疏植被,与文献[7]在冬小麦返青期的研究较为一致。

从图 2(b)可见,光学特征参数的重要性排序依次为 VSWI > COSRI > IPVI > LST > RVI > TVDI > MSAVI > NDVI > GVI > GLI > EVI。

VSWI重要性最高,EVI最低,这说明VSWI对于0~10 cm的土壤水分敏感度较高,能够较好地反映土壤水分的变化情况;EVI是为改善高植被区大气校正和土壤背景的影响而提出的^[34],研究区多为低矮植被和裸地,因此EVI重要性偏低。MSAVI、NDVI、GVMI和GLI相比EVI重要性程度有所提升,但总体依然较低,可

能是GVMI不太适合小尺度区域的干旱监测,NDVI和GLI适用于中等植被覆盖区,在稀疏植被和较高植被覆盖区,其土壤水分敏感性有所降低^[35]。相比NDVI,MSAVI能在一定程度上改善土壤背景影响,因此重要性略高于NDVI。其余光学特征参数均对土壤水分反演具有较高的重要性。

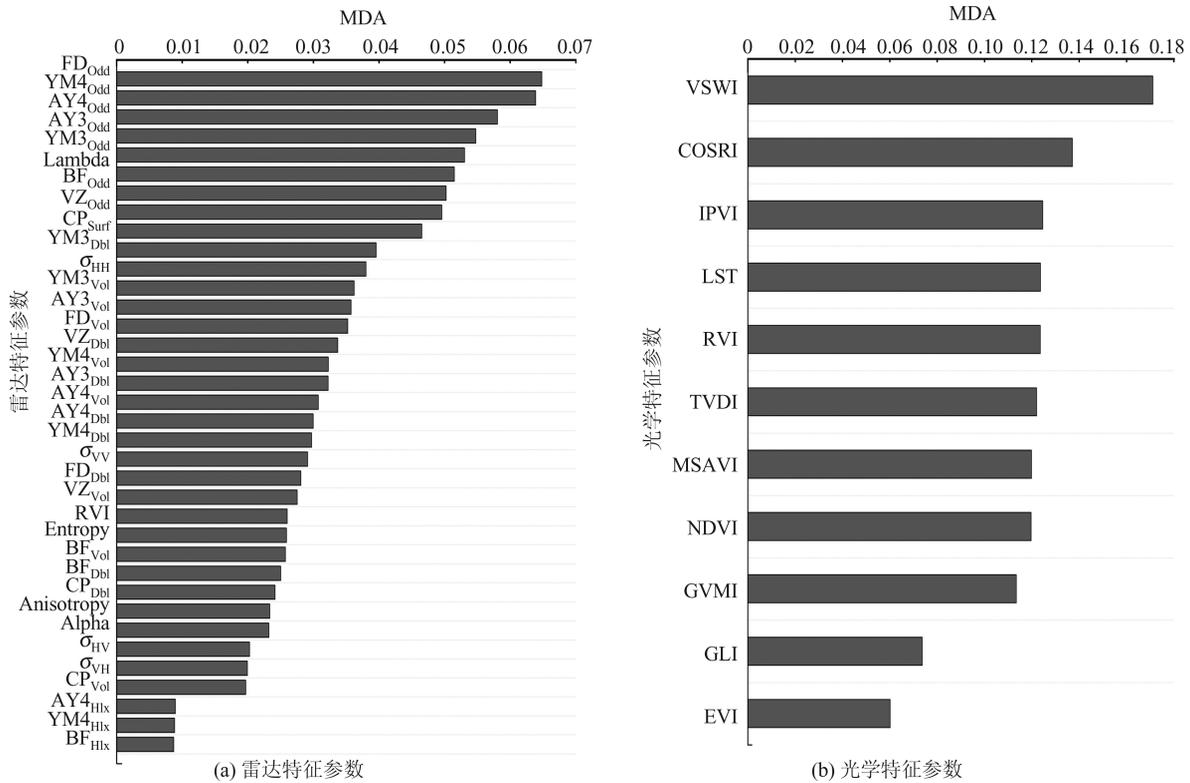


图2 重要性评分

Fig. 2 Importance Scores

3.2 模型精度对比

3.2.1 基于雷达特征参数的模型

在雷达特征参数中,剔除重要性排名靠后的6个参数,即对土壤水分含量变化敏感度较低的参数,将剩余的30个参数按重要性排序,从VIM_{R4}(VIM表示重要性,下标R表示雷达特征参数,4表示前4个参数)开始,依次放入2个特征变量,分别计算 R^2 和RMSE,选出最优特征组合方案,建立RF模型,各模型精度如图3所示。

由图3可见,基于雷达特征参数组合的14种模型训练集 R^2 普遍高于0.8, RMSE普遍低于1.5%,验证集 R^2 普遍高于0.5, RMSE普遍低于2.5%。在所有模型中,VIM_{R30}的 R^2 和RMSE并非最优,说明模型精度与输入特征数量并非简单的正相关关系,特征选择是提高模型精度的重要环节。从VIM_{R4}至VIM_{R12},随着输入特征增加,精度显著提升, R^2 从0.54提升至0.64, RMSE从

2.48%降低至2.26%;从VIM_{R12}至VIM_{R18},虽然精度有所提升,但提升速度相比之前有所减缓, R^2 由0.64提升至0.67, RMSE由2.26%降低至2.16%;从VIM_{R18}之后, R^2 和RMSE均呈现不同程度的下降,推测是前后因子间存在强相关性造成了数据冗余,使模型在训练时可能存在轻微过拟合现象,尽管如此,VIM_{R18}的 R^2 和RMSE分别达到0.67和2.16%。C波段极化雷达数据在植被覆盖地表土壤水分研究中应用较多,笔者团队前期对研究区Radarsat-2 C波段数据的研究表明,总体样本中各参数组合方案的RF模型验证集 R^2 均低于0.6, RMSE在3.5%~5%之间^[36]。本文基于L波段的参数方案相比C波段在精度上均有提升,尤其是VIM_{R18}效果更佳,说明相比C波段而言,L波段全极化SAR对土壤水分敏感性更高。文献[37-38]对比了不同NDVI状态下ALOS-2 PALSAR-2 L波段雷达数据和Sentinel-1 C波段

雷达数据在法国南部、突尼斯中部地区土壤水分反演中的敏感性,认为不同 NDVI 状态下 L 波段对土壤水分的敏感程度优于 C 波段,与本文的研究结果相吻合。因此,选取 VIM_R18 作为最优雷达特征参数组合模型反演土壤水分。

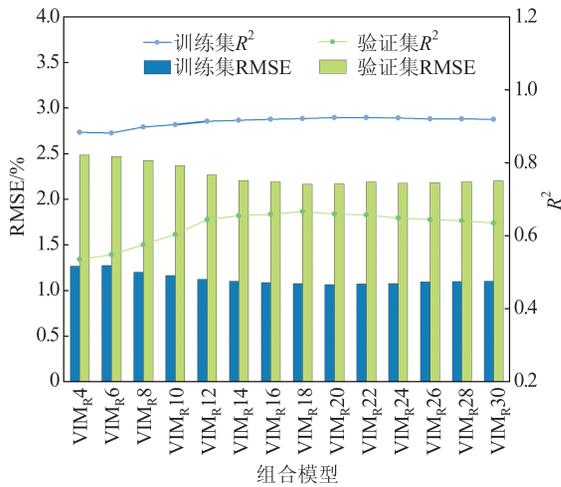


图 3 雷达特征参数组合模型精度对比

Fig. 3 Accuracy Comparison of Radar Characteristic Parameter Combination Models

3.2.2 基于光学特征参数的模型

在光学特征参数中,按重要性排名从 VIM_O4 (下标 O 代表光学特征参数)开始,每次放入 1 个特征参数,将其作为 RF 模型的输入变量,以土壤水分含量作为输出变量,建立 RF 模型,以 R^2 和 RMSE 评估模型精度,结果如图 4 所示。

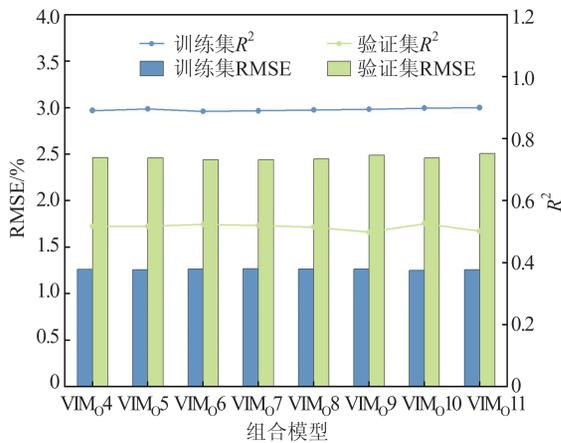


图 4 光学特征参数组合模型精度对比

Fig. 4 Accuracy Comparison of Optical Characteristic Parameter Combination Models

由图 4 可以看出,基于光学特征参数组合的 8 种模型训练集 R^2 均在 0.89 左右, RMSE 在 1.25% 左右,验证集 R^2 在 0.5 左右, RMSE 在 2.47% 左右。从 VIM_O4 至 VIM_O11,验证集 R^2 呈现先缓慢增长、后逐渐降低、再增长、又降低的趋势; RMSE

则呈现先降低、后增长、再降低、又增长的趋势;从 VIM_O6 至 VIM_O9,模型精度的第一次降低可能是由于因子间的强相关性导致模型训练集的过度拟合所致;从 VIM_O9 至 VIM_O11,验证集 R^2 呈现先增长、后下降的趋势,但对比 VIM_O9 和 VIM_O11 两组方案可知, VIM_O11 虽然 R^2 有所提升,但 RMSE 也呈现上升趋势,可能是由于采样点土壤水分低值数量较多、高值数量过少,导致出现低值点高估、高值点低估现象,最终导致验证集 RMSE 有所上升。由于在光学特征参数组合方案中综合使用 R^2 和 RMSE 无法确定最优组合方案,因此,选取验证集 R^2 位居前 3 位的 VIM_O6、VIM_O7、VIM_O10 光学特征参数组合方案进行土壤水分反演。

3.2.3 基于雷达-光学特征参数的协同模型

将上述优选的雷达和光学特征参数组合方案进一步组合,得到 3 种协同反演方案,将 3 种方案的反演值分别与验证点的土壤水分含量实测值进行对比分析,计算 R^2 和 RMSE,结果见表 2。

表 2 雷达-光学参数协同反演模型精度对比

Tab. 2 Accuracy Comparison of Radar-Optical Parameter Integrated Inversion Model

方案	组合类型	R^2	RMSE/%
方案 1	VIM _R 18+ VIM _O 6	0.72	1.99
方案 2	VIM _R 18+ VIM _O 7	0.71	2.00
方案 3	VIM _R 18+ VIM _O 10	0.70	2.02

由表 2 可知,3 种协同反演模型 R^2 均不低于 0.7, RMSE 均低于 2.05%。其中方案 1 的 R^2 为 0.72, RMSE 为 1.99%,在 3 种协同模型中精度最高。与仅使用雷达特征参数的最佳模型 VIM_R18 相比, R^2 提高 7.46%, RMSE 降低 8.54%,表明光学特征参数能够在一定程度上消除植被覆盖的影响,弥补雷达数据在土壤水分反演中的不足。与仅使用光学特征参数的模型 VIM_O6 相比, R^2 提升 38.4%, RMSE 降低 22.6%,说明相较于单一数据源而言,多源数据融合能够在一定程度上提升土壤水分反演精度,使预测结果更接近真实情况。相比方案 1,方案 2 和方案 3 总体精度虽然有所下降,但仍然较为理想,进一步说明融合多源数据的机器学习方法在土壤水分反演中的应用潜力,其适用范围相较单一数据源更为广泛。

3.3 多种模型土壤水分反演结果对比

图 5 为基于雷达、光学以及雷达-光学特征参数协同反演模型的土壤水分反演结果。由图 5 可以看出,仅使用光学特征参数的模型(图 5(b))存

在水分整体高估的现象,特别是在裸地和部分低植被覆盖区,这些区域易受土壤背景和天气的影响^[35],从而导致预测精度较低。相比图5(b),基于雷达特征参数(图5(a))和雷达-光学特征参数协同(图5(c))的模型对于裸地和低植被覆盖区的估算效果较好,与实测水分含量基本一致,表明雷达特征参数在裸地和低植被区土壤水分反演中的应用潜力。已有研究表明,对于植被覆盖较高地区而言,仅使用雷达特征参数难以完全消除植被覆盖和地表粗糙度的影响,由图5(a)可以

看出,东部植被覆盖区土壤水分含量相比实测值有所低估,加入光学特征参数后(图5(c)),低估现象有所改善,表明多源数据融合能在一定程度上克服因影响因子考虑不足导致的精度受限问题,这与文献[39]的研究结果较为吻合。但由于研究区土壤水分含量整体较低,导致土壤水分高值采样点较少,高值低估现象仍然存在。总体来说,图5能反映研究区土壤水分含量的整体趋势,其中雷达-光学协同模型反演效果最佳,能够更好地反映研究区土壤水分的空间分布格局。

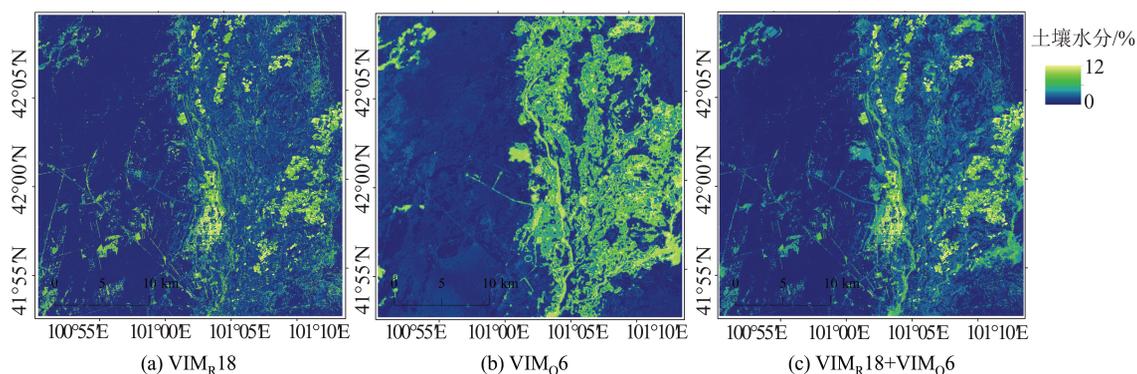


图5 多种模型土壤水分反演结果

Fig. 5 Soil Moisture Inversion Results of Various Models

4 结 语

本文基于L波段ALOS-2 PALSAR-2和Landsat 8影像提取雷达特征参数和光学特征参数,通过MDA优选特征参数,建立基于雷达、光学以及雷达-光学特征参数协同的RF土壤水分反演模型,定量评估模型精度,反演绿洲土壤水分。结论如下:(1)相比C波段而言,L波段对于干旱荒漠绿洲地区的土壤水分含量敏感性更高;(2)雷达特征参数中,表面散射分量重要性最高,体散射分量次之,二面角和螺旋体散射分量重要性较低,4个后向散射系数中,同极化后向散射系数的重要性优于交叉极化后向散射系数,光学特征参数中,VSWI重要性最高,EVI最低;(3)雷达特征参数方案最优模型 R^2 、RMSE分别为0.67、2.16%。光学特征参数方案模型精度普遍较低且精度相当, R^2 、RMSE分别为0.5、2.47%;(4)相比单一数据源而言,协同利用雷达和光学特征参数,模型适用范围更广,预测精度更高。最优协同反演模型的 R^2 、RMSE分别为0.72、1.99%。相比仅使用雷达特征参数模型 VIM_R18 , R^2 提升了7.46%,RMSE降低了8.54%,相比仅使用光学特征参数模型 VIM_O6 , R^2 提升了38.4%,RMSE降低了22.6%。

本文研究表明,雷达-光学协同的随机森林模型在干旱荒漠绿洲区土壤水分反演中具有良好的适用性,能够较好地反映区域土壤水分空间分布格局。后期将采用更多的机器学习方法,探索协同利用L波段全极化雷达数据和光学数据在较深层次土壤水分含量反演中的效果和适用性。

参 考 文 献

- [1] Li Zhanjie, Chen Jipei, Liu Yanmin, et al. Soil Moisture Retrieval from Remote Sensing [J]. *Journal of Beijing Normal University (Natural Science)*, 2020, 56(3): 474-481. (李占杰, 陈基培, 刘艳民, 等. 土壤水分遥感反演研究进展[J]. 北京师范大学学报(自然科学版), 2020, 56(3): 474-481.)
- [2] Wang H Q, Magagi R, Goita K. Potential of a Two-Component Polarimetric Decomposition at C-band for Soil Moisture Retrieval over Agricultural Fields [J]. *Remote Sensing of Environment*, 2018, 217: 38-51.
- [3] Bourgeau-Chavez L L, Leblon B, Charbonneau F, et al. Evaluation of Polarimetric Radarsat-2 SAR Data for Development of Soil Moisture Retrieval Algorithms over a Chronosequence of Black Spruce Boreal Forests [J]. *Remote Sensing of Environment*,

- 2013, 132: 71-85.
- [4] Breiman L. Random Forest[J]. *Machine Learning*, 2001, 45(1): 5-32.
- [5] Ahmad S, Kalra A, Stephen H. Estimating Soil Moisture Using Remote Sensing Data: A Machine Learning Approach[J]. *Advances in Water Resources*, 2010, 33(1): 69-80.
- [6] Chaudhary S K, Srivastava P K, Gupta D K, et al. Machine Learning Algorithms for Soil Moisture Estimation Using Sentinel-1: Model Development and Implementation [J]. *Advances in Space Research*, 2022, 69(4): 1799-1812.
- [7] Li Pingxiang, Liu Zhiqu, Yang Jie, et al. Soil Moisture Retrieval of Winter Wheat Fields Based on Random Forest Regression Using Quad-Polarimetric SAR Images [J]. *Geomatics and Information Science of Wuhan University*, 2019, 44(3): 405-412. (李平湘, 刘致曲, 杨杰, 等. 利用随机森林回归进行极化 SAR 土壤水分反演[J]. 武汉大学学报(信息科学版), 2019, 44(3):405-412.)
- [8] Zhao Jianhui, Zhang Bei, Li Ning, et al. Cooperative Inversion of Winter Wheat Covered Surface Soil Moisture Based on Sentinel-1/2 Remote Sensing Data[J]. *Journal of Electronics & Information Technology*, 2021, 43(3): 692-699. (赵建辉, 张蓓, 李宁, 等. 基于 Sentinel-1/2 遥感数据的冬小麦覆盖地表土壤水分协同反演[J]. 电子与信息学报, 2021, 43(3):692-699.)
- [9] Guo Jiao, Liu Jian, Ning Jifeng, et al. Construction and Validation of Soil Moisture Retrieval Model in Farmland Based on Sentinel Multi-source Data[J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2019, 35(14): 71-78. (郭交, 刘健, 宁纪锋, 等. 基于 Sentinel 多源数据的农田地表土壤水分反演模型构建与验证[J]. 农业工程学报, 2019, 35(14):71-78.)
- [10] Holtgrave A K, Förster M, Greifeneder F, et al. Estimation of Soil Moisture in Vegetation-Covered Floodplains with Sentinel-1 SAR Data Using Support Vector Regression[J]. *ISPRS - Journal of Photogrammetry, Remote Sensing and Geoinformation Science*, 2018, 86(2): 85-101.
- [11] Liu C A, Chen Z X, Shao Y, et al. Research Advances of SAR Remote Sensing for Agriculture Applications: A Review [J]. *Journal of Integrative Agriculture*, 2019, 18(3): 506-525.
- [12] Robinson D A, Campbell C S, Hopmans J W, et al. Soil Moisture Measurement for Ecological and Hydrological Watershed-Scale Observatories: A Review [J]. *Vadose Zone Journal*, 2008, 7(1): 358-389.
- [13] Wang D, Liu C G, Zeng Y, et al. Dryland Crop Classification Combining Multitype Features and Multi-temporal Quad-Polarimetric Radarsat-2 Imagery in Hebei Plain, China[J]. *Sensors*, 2021, 21(2): 332.
- [14] Freeman A, Durden S L. A Three-Component Scattering Model for Polarimetric SAR Data [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 1998, 36(3): 963-973.
- [15] Cloude S R, Pottier E. A Review of Target Decomposition Theorems in Radar Polarimetry [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 1996, 34(2): 498-518.
- [16] Yamaguchi Y, Moriyama T, Ishido M, et al. Four-Component Scattering Model for Polarimetric SAR Image Decomposition [J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2005, 43(8): 1699-1706.
- [17] An W T, Cui Y, Yang J. Three-Component Model-Based Decomposition for Polarimetric SAR Data[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2010, 48(6): 2732-2739.
- [18] Cloude S R, Pottier E. An Entropy Based Classification Scheme for Land Applications of Polarimetric SAR[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 1997, 35(1): 68-78.
- [19] van Zyl J J. Application of Cloude's Target Decomposition Theorem to Polarimetric Imaging Radar Data [C]//*Radar Polarimetry*, San Diego, USA. 1993.
- [20] Bhattacharya A, Muhuri A, De S, et al. Modifying the Yamaguchi Four-Component Decomposition Scattering Powers Using a Stochastic Distance [J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2015, 8(7): 3497-3506.
- [21] Rouse J W, Haas R H, Schell J A, et al. Monitoring Vegetation Systems in the Great Plains with ERTS [J]. *NASA Special Publication*, 1974, 1: 309-351.
- [22] Fernández-Buces N, Siebe C, Cram S, et al. Mapping Soil Salinity Using a Combined Spectral Response Index for Bare Soil and Vegetation: A Case Study in the Former Lake Texcoco, Mexico [J]. *Journal of Arid Environments*, 2006, 65(4): 644-667.
- [23] Huete A, Justice C, Liu H. Development of Vegetation and Soil Indices for MODIS-EOS[J]. *Remote Sensing of Environment*, 1994, 49(3): 224-234.
- [24] Louhaichi M, Borman M M, Johnson D E. Spatial-

- ly Located Platform and Aerial Photography for Documentation of Grazing Impacts on Wheat [J]. *Geocarto International*, 2001, 16(1): 65-70.
- [25] Ceccato P, Gobron N, Flasse S, et al. Designing a Spectral Index to Estimate Vegetation Water Content from Remote Sensing Data: Part 1 [J]. *Remote Sensing of Environment*, 2002, 82(2/3): 188-197.
- [26] Crippen R. Calculating the Vegetation Index Faster [J]. *Remote Sensing of Environment*, 1990, 34(1): 71-73.
- [27] Feng Rui, Zhang Yushu, Li Danghong. Using AVHRR Data to Monitor Vegetation Soil Moisture Content [J]. *Journal of Meteorology and Environment*. 1998(4):30-31. (冯锐, 张玉书, 李党红. 利用 AVHRR 资料监测植被土壤含水量 [J]. 气象与环境学报, 1998(4):30-31)
- [28] Qi J, Chehbouni A, Huete A R, et al. A Modified Soil Adjusted Vegetation Index [J]. *Remote Sensing of Environment*, 1994, 48(2): 119-126.
- [29] Sandholt I, Rasmussen K, Andersen J. A Simple Interpretation of the Surface Temperature/Vegetation Index Space for Assessment of Surface Moisture Status [J]. *Remote Sensing of Environment*, 2002, 79(2/3): 213-224.
- [30] Luo Ling, Mao Dehua, Zhang Bai, et al. Remote Sensing Estimation for Light Use Efficiency of Phragmites Australis Based on Landsat OLI over Typical Wetlands [J]. *Geomatics and Information Science of Wuhan University*, 2020, 45(4): 524-533. (罗玲, 毛德华, 张柏, 等. 基于 Landsat OLI 影像的典型湿地芦苇植被光能利用率遥感反演方法初探 [J]. 武汉大学学报(信息科学版), 2020, 45(4): 524-533.)
- [31] Xu Huihui, Gao Meiling, Li Zhenhong, et al. Accuracy Comparison of near Surface Air Temperature Estimation Using Different Land Surface Temperature Sources [J]. *Geomatics and Information Science of Wuhan University*, 2023, 48(4): 568-578. (许慧慧, 高美玲, 李振洪, 等. 多源地表温度估算近地表气温的精度对比 [J]. 武汉大学学报(信息科学版), 2023, 48(4): 568-578.)
- [32] Yu Tingting, Dong Youfu. Correcting Elevation Error of ASTER GDEM Using Random Forest Regression Algorithm [J]. *Geomatics and Information Science of Wuhan University*, 2021, 46(7): 1098-1105. (余婷婷, 董有福. 利用随机森林回归算法校正 ASTER GDEM 高程误差 [J]. 武汉大学学报(信息科学版), 2021, 46(7): 1098-1105.)
- [33] Cheng Xiaoguang. Research of Model-Based Polarimetric SAR Decomposition Constrained for Nonnegative Eigenvalues [D]. Wuhan: Wuhan University, 2014. (程晓光. 特征值非负约束下的基于模型的极化 SAR 分解研究 [D]. 武汉: 武汉大学, 2014.)
- [34] Chen Liang, Wang Xuelei, Yang Chao, et al. Spatio-Temporal Variation Characteristics of Vegetation EVI and Their Topographic Effects in the West Mountain Regions of Hubei Province from 2000 to 2018 [J]. *Resources and Environment in the Yangtze Basin*, 2021, 30(2): 419-426. (陈亮, 王学雷, 杨超, 等. 2000—2018 年鄂西山区植被 EVI 时空变化特征及其地形效应 [J]. 长江流域资源与环境, 2021, 30(2):419-426.)
- [35] Huang Youxin, Liu Xiuguo, Shen Yonglin, et al. Advances in Remote Sensing Derived Agricultural Drought Monitoring Indices and Adaptability Evaluation Methods [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2015, 31(16): 186-195. (黄友昕, 刘修国, 沈永林, 等. 农业干旱遥感监测指标及其适应性评价方法研究进展 [J]. 农业工程学报, 2015, 31(16):186-195.)
- [36] Yang Liping, Wang Tong, Su Zhiqiang, et al. An Optimized Dimensionality Reduction Method of Characteristic Parameters for Retrieving Oasis Soil Moisture Using Random Forest [J]. *Journal of Northwest University (Natural Science Edition)*, 2022, 52(2): 181-191. (杨丽萍, 王彤, 苏志强, 等. 随机森林反演绿洲土壤水分的特征参数优化降维方法 [J]. 西北大学学报(自然科学版), 2022, 52(2):181-191.)
- [37] El Hajj M, Baghdadi N, Bazzi H, et al. Penetration Analysis of SAR Signals in the C and L Bands for Wheat, Maize, and Grasslands [J]. *Remote Sensing*, 2018, 11(1): 31.
- [38] Ayari E, Kassouk Z, Lili-Chabaane Z, et al. Cereal Crops Soil Parameters Retrieval Using L-Band ALOS-2 and C-Band Sentinel-1 Sensors [J]. *Remote Sensing*, 2021, 13(7): 1393.
- [39] Wang Hao, Luo Geping, Wang Weisheng, et al. Inversion of Soil Moisture Content in the Farmland in Middle and Lower Reaches of Syr Darya River Basin Based on Multi-source Remotely Sensed Data [J]. *Journal of Natural Resources*, 2019, 34(12): 2717-2731. (王浩, 罗格平, 王伟胜, 等. 基于多源遥感数据的锡尔河中下游农田土壤水分反演 [J]. 自然资源学报, 2019, 34(12): 2717-2731.)