

基于FY-3D卫星和机器学习的辽宁省土壤水分监测技术研究

刘琳^{1,2,3}, 鹿罡⁵, 王井利¹, 张淑杰^{2,3,4}, 李荣平^{2,3,4*}

1. 沈阳建筑大学交通与测绘工程学院, 辽宁 沈阳 110168

2. 沈阳农业与生态气象研究院, 辽宁 沈阳 110166

3. 中国气象局沈阳大气环境研究所, 辽宁 沈阳 110166

4. 辽宁省农业气象灾害重点实验室, 辽宁 沈阳 110166

5. 辽宁省交通规划设计院有限责任公司, 辽宁 沈阳 110166

摘要: 干旱是国家粮食安全的重要威胁之一, 土壤水分精准监测技术既对抗旱减灾至关重要, 同时也是高标准农田节水灌溉的重要支撑。基于FY-3D遥感数据, 采用统计分析方法筛选与土壤水分变化显著相关的植被遥感指数和土壤参数等关键特征因子, 利用随机森林、XGBoost以及ConvLSTM深度学习模型构建了0~20 cm深度的土壤体积含水量反演模型。研究表明: 在裸土-播种期与成熟期, 土壤参数与地表温度(LST)成为主要特征因子; 出苗-拔节期与抽穗-灌浆期, 植被指数对土壤水分反演的影响明显增强。双时间节点较单时间节点建模的反演结果 R^2 提升了0.45以上, 显著提升了模型精度。ConvLSTM模型具有较强的稳定性与泛化能力, 全时期模型反演结果 R^2 均值达到0.71, RMSE低于0.046(cm^3/cm^3), MAE低于0.032(cm^3/cm^3)。基于ConvLSTM模型的双时间节点建模能够高精度反演辽宁省土壤水分, 实现了逐日的大面积土壤水分监测, 可有效地支撑气象抗旱减灾业务。

关键词: FY-3D; ConvLSTM; 土壤水分反演

中图法分类号: P237

文献标识码: A

文章编号: 1000-2324(2026)05-0918-14

Soil Moisture Monitoring Technology in Liaoning Province Based on FY-3D Satellite and Machine Learning

LIU Lin^{1,2,3}, LU Gang⁵, WANG Jing-li¹, ZHANG Shu-jie^{2,3,4}, LI Rong-ping^{2,3,4*}

1. School of Transportation and Geomatics Engineering/Shenyang Jianzhu University, Shenyang 110168, China

2. Shenyang Academy of Agriculture and Ecological Meteorology, Shenyang 110166, China

3. Shenyang Institute of Atmospheric Environment/China Meteorological Administration, Shenyang 110166, China

4. Key Laboratory of Agricultural Meteorological Disasters of Liaoning Province, Shenyang 110166, China

5. Liaoning Provincial Transportation Planning And Design Institute Co., LTD, Shenyang 110166, China

Abstract: Drought is one of the major threats to national food security. Accurate soil moisture monitoring technology is not only essential for drought mitigation and disaster reduction but also serves as a critical foundation for water-saving irrigation in high-standard farmland. Based on FY-3D remote sensing data, statistical analysis methods are used to screen key feature variables such as vegetation indices and soil parameters that are significantly correlated with soil moisture variations. These features are then used to develop Random Forest, XGBoost, and ConvLSTM deep learning models for retrieving soil volumetric water content at a depth of 0-20 cm. The results indicate that during the Bare soil - Sowing period and Maturity period, soil parameters and land surface temperature (LST) are the dominant factors for soil moisture retrieval. In contrast, during Emergence - jointing stage and Tasseling - Grouting stage, the impact of vegetation indices on soil moisture inversion is significantly enhanced. Compared with single-time-point modeling, the R^2 of dual-time-point modeling increases by more than 0.45, significantly enhancing model accuracy. The ConvLSTM model demonstrates strong stability and generalization ability, with mean R^2 values of 0.71 across different periods, RMSE below 0.046 (cm^3/cm^3), and MAE below 0.032 (cm^3/cm^3). The dual-time-point ConvLSTM modeling enables high-accuracy retrieval of soil moisture in Liaoning Province, achieving large-scale, daily soil moisture monitoring, and providing effective support for meteorological drought mitigation and disaster reduction operations.

Keywords: FY-3D; ConvLSTM; soil moisture inversion

收稿日期: 2025-09-24

修回日期: 2026-09-14

基金项目: 中国气象局人才项目“中国气象局农业气象重点创新团队”(CMA2024ZD02); 科院基本科研业务费“农田生态系统-气候-水资源相互作用与农田节水灌溉”(2024Z001); 国家自然科学基金面上项目“玉米开花后干旱胁迫导致叶片早衰的定量诊断及产量模拟研究”(42275202)

第1作者简介: 刘琳(1997-), 女, 硕士研究生, 研究方向: 土壤水分遥感监测研究。E-mail: Liulin4114@163.com

***通讯作者:** Author for correspondence. E-mail: rongpingli@163.com

土壤水分是全球水循环的重要组成部分,是植物吸收水分的主要来源,也是研究农业、生态、气候等的重要因素^[1]。因此,大范围、快速、准确的土壤水分反演监测,对防旱减灾、精准灌溉和种植规划具有重大意义^[2]。目前,土壤水分监测的方法主要为地面观测法和遥感监测法^[3]。地面观测法包括直接测量法、物理传感法,该方法既有准确度高的优点,也存在操作复杂、耗时长、投资大、范围小等缺点^[4]。遥感监测法包括微波遥感监测法和光学遥感监测法,因其具有监测范围广、成本低的优点而被广泛应用,但同时也存在受云层、植被覆盖影响等缺点^[5,6]。

随着卫星增多和技术发展,卫星遥感数据在土壤水分监测中已成为重要的数据源^[7]。研究表明,利用遥感指数能够有效实现土壤水分监测。蔡斌等^[8]利用归一化植被指数(NDVI)和植被条件指数(VCI)实现了对中国地区干旱和洪涝的实时监测。陈国茜等^[1]引入12个植被指数构建土壤水分反演模型,结果可以较好地实现土壤水分监测。鉴于单一遥感指数难以全面反映土壤水分动态,本研究整合多种遥感指数以实现更高精度的监测。除遥感指数外,土壤参数与地形因素对土壤水分的储存与分布也具有重要影响。土壤质地、密度、容重和持水能力等特性共同影响土壤的水分迁移与存储能力。同样数字高程模型(DEM)反映地形特征,对土壤水分的分布也有着重要的影响。张涛等^[9]发现同空间分布的裸土参数在被动微波遥感反演土壤水分中会造成不同程度的误差。土壤质地或土壤水力参数都可以有效地提高土壤水分的模拟精度^[10]。

近年来,随着计算能力的提升,各类机器学习模型被广泛应用于土壤水分和干旱监测研究并取得较好的效果^[11-14],其中长短期记忆网络(LSTM)、卷积神经网络(CNN)、卷积长短期记忆网络(ConvLSTM)等深度学习模型常用于高精度土壤水分预测^[15],ConvLSTM模型尤为突出。M ElSaadani等^[16]通过ConvLSTM模型实现36公里低分辨率土壤湿度(SM)预测。Yinglan A等^[17]研究结果显示ConvLSTM预测结果明显优于分辨率为0.25°的全球土地数据同化系统(GLDAS)产品。

随着遥感技术和机器学习的发展,两者的融

合将是未来干旱监测的关键手段,尽管学术界已经开始机器学习与遥感指数的探索性研究,尽管深度学习模型在土壤湿度估算中取得了显著进展,但目前多数研究仍局限于模型构建与精度验证阶段,对模型在实际业务化监测的适应性依然不足^[18]。针对当前土壤水分反演中存在分辨率受限与停留于实验阶段泛化性不足等问题,本文以FY-3D/MERSI数据为主要数据源,计算多种遥感指数并构建综合特征数据集,融合土壤参数数据,系统比较多种机器学习模型的反演性能。研究旨在提升土壤水分反演精度与空间分辨率,实现逐日0~20 cm土壤水分动态监测,并利用2021-2023年数据进行泛化性验证,系统评估其在不同时间尺度下的反演稳定性与适用性,为区域高精度土壤水分监测及农业干旱评估提供技术支持。

1 研究区与数据源

1.1 研究区概况

辽宁省位于中国东北地区南部,地理位置介于东经118°53′—125°46′,北纬38°43′—43°26′之间(图1),属于温带大陆性季风气候,雨热同季,日照丰富,积温较高,冬长夏暖,春秋季节短,四季分明。全省年日照时数2 100~2 600 h,年降水量在600~1 100 mm。土壤以壤质土壤为主,大体呈现“东、西和北壤为主,南砂中黏占少许”的空间分布特征。良好的土地资源与光照条件为农业生产提供了有利的自然基础。全省耕地面积居全国第七,主要种植玉米、水稻、大豆等农作物,据统计玉米种植面积可达77%以上。农业生产对土壤水分变化高度敏感,对土壤水分的高精度、时空连续监测具有重要的科学与应用价值。

1.2 数据来源及预处理

1.2.1 遥感数据及预处理 FY-3D/MERSI卫星数据。数据来源于风云卫星遥感数据服务网(<http://data.nsmc.org.cn/DataPortal/cn/home/index.html>)。该星中分辨率光谱成像仪MERSI共有25个观测通道,其中,可见光—近红外通道16个、短波红外通道3个、中长波红外通道6个,地面分辨率通道有6个为250 m,19个为1 000 m^[19]。FY-3系列的(C、B、F、G)卫星均具备土壤水分产品数据,可获取时段上存在一定局限。FY-3B和FY-

表 1 遥感指数
Table 1 Rs index

监测参量 Monitoring parameters	遥感指数 RS index	指数算法 Index algorithm	来源 Origin
冠层温度 Canopy temperature	温度条件指数(TCI) Temperature condition index	$TCI = (LST_{max} - LST_i) / (LST_{max} - LST_{min})$	[21]
冠层含水量 Canopy water content	陆表水分指数(LSWI) Land surface water index	$LSWI = (R_{NIR} - R_{SWIR1}) / (R_{NIR} + R_{SWIR1})$	[22]
	红外指数归一化差值(NDIIB7) Normalized difference of infrared index	$NDIIB7 = (R_{NIR} - R_{SWIR2}) / (R_{NIR} + R_{SWIR2})$	[23]
	归一化干旱指数(NMDI) Normalized multi-band drought index	$NMDI = [R_{NIR} - (R_{SWIR1} + R_{SWIR2})] / [R_{NIR} + (R_{SWIR1} + R_{SWIR2})]$	[23]
	水分指数(WI) Water index	$WI = R_{905} / R_{940}$	[19]
	归一化植被指数(NDVI) Normalized difference	$NDVI = (R_{NIR} - R_{RED}) / (R_{NIR} + R_{RED})$	[24]
作物形态及绿度 Crop morphology and greenness	棕色植被指数(BVI) Brown vegetation index	$BVI = (R_{GREEN} - R_{SWIR2}) / (R_{GREEN} + R_{SWIR2})$	[19]
	植被条件指数(VCI) Vegetation condition index	$VCI = (NDVI - NDVI_{min}) / (NDVI_{max} - NDVI_{min})$	[19] [25]
	温度植被指数(TVI) Temperature Vegetation Index	$TVI = T_s / NDVI$	[26]
	植被健康指数(VHI) Vegetation Health Index	$VHI = 0.5 \times (VCI + TCI)$	[27]
	归一化差异水分指数(NDWI) Normalized difference water index	$NDWI = (R_{NIR} - R_{GREEN}) / (R_{NIR} + R_{GREEN})$	[28]
土壤含水量 Soil water content	植被供水指数(VSWI) Vegetation Supply Water Index	$VSWI = NDVI / LST$	[19]

注: max 和 min 分别表示最高和最低值, 其中 RNIR, RSWIR1, RSWIR2, R905, R940, RRED 和 RGREEN 分别对应的是 FY-3D/MERSI 数据中的 4、8、9、18、20、3 和 2 波段数据。

Note: "max" and "min" represent the highest and lowest values, respectively. RNIR, RSWIR1, RSWIR2, R905, R940, RRED and RGREEN correspond to bands 4, 8, 9, 18, 20, 3 and 2 in the FY-3D/MERSI data, respectively.

(4月1日—9月30日)自动土壤水分观测数据。考虑到土壤水分随时间变化,根据FY-3D/MERSI卫星过境时间与气象站点数据相差8 h,在卫星过境时间加8 h来选取对应的土壤水分数据,利用ArcMap10.8中的反距离权重法对土壤数据进行插值并保持与遥感数据统一的坐标系与分辨率。

1.3 中国气象局陆面数据同化系统土壤水分产品

中国气象局陆面数据同化系统CLDAS-V2.0产品,数据来源于国家气象科学数据中心(<http://data.cma.cn/>),该产品融合同化了149种多源地面观测数据、卫星遥感数据和数值模式

产品。基于多种陆面模式集成模拟,生成了空间分辨率为0.062 5°×0.062 5°(约6 km),覆盖东亚区域,多深度层次,逐日的土壤体积含水量产品^[29]。

1.4 划分研究时段

研究时段为4月1日—9月30日,覆盖了农作物从播种前期至成熟期。本研究以辽宁省农作物的生长状态为基础,并结合辽宁省春玉米发育期特点划分研究时段。共划分为4个时期,分别为裸土—播种期(4月1日—4月30日)、出苗—拔节期(5月1日—6月30日)、抽雄—灌浆期(7月1日—8月31日)、成熟期(9月1日—9月30日)。

1.5 土壤干旱等级划分

按照国标《农业干旱等级》(GB/T32136-2015)^[30],对土壤干旱等级进行划分,分为无旱、轻旱、中旱、重旱、特旱6个等级(表2)。

2 研究方法

研究步骤主要包括:(1)划分研究时段;(2)筛选不同研究时段相关的遥感指数,构成遥感指数

数据集;(3)构建机器学习模型;(4)分析、对比模型反演精度,选出最优模型;(5)干旱监测分析。研究技术路线如图2所示。

2.1 模型构建

本研究选取样本量82 000余个,其中80%的样本作为训练集,20%的样本作为验证集,用于模型训练与性能评估。

表2 基于土壤相对湿度指数(Rsm)划分的干旱等级
Table 2 Drought grades based on soil relative moisture index (Rsm)

等级 Grade	土壤相对湿度指数/% Soil relative moisture index		
	砂土 Sandy soil	壤土 Loam soil	黏土 Clay soil
	Rsm $\geq$ 55	Rsm $\geq$ 60	Rsm $\geq$ 65
无旱 Normal	Rsm $\geq$ 55	Rsm $\geq$ 60	Rsm $\geq$ 65
轻旱 Light drought	45 $\leq$ Rsm<55	50 $\leq$ Rsm<60	55 $\leq$ Rsm<65
中旱 Mid-drought	35 $\leq$ Rsm<45	40 $\leq$ Rsm<50	45 $\leq$ Rsm<55
重旱 Heavy drought	25 $\leq$ Rsm<35	30 $\leq$ Rsm<40	35 $\leq$ Rsm<45
特旱 Severe drought	Rsm<25	Rsm<30	Rsm<35

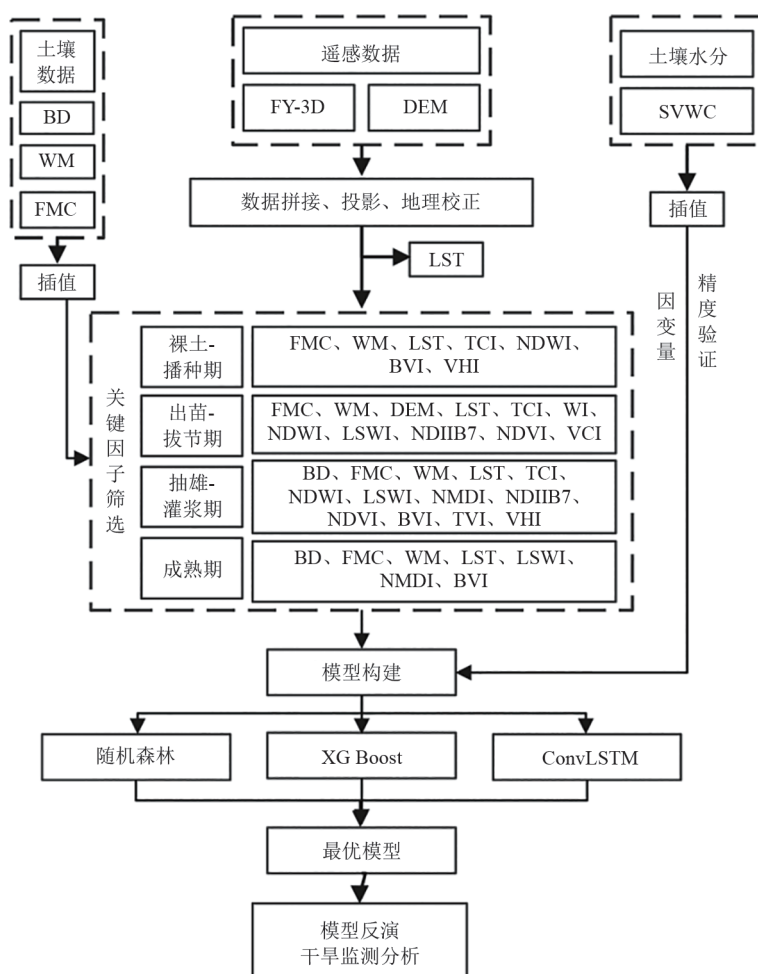


图2 技术路线

Fig. 2 Technical route

2.1.1 随机森林 随机森林(Random Forest, RF)是一种集成学习(Ensemble Learning)方法,由多个决策树组成^[31]。其核心思想是“集成多个弱学习器(回归树)来形成一个强学习器”,最终的反演结果通过多个树的平均值来计算。随机森林具有较强的稳定性,具有较高的反演精度和较好的泛化能力,被广泛应用于分类和回归问题^[32]。随机森林模型对数据的噪声和异常值并不敏感,个别异常值不会影响整体反演结果^[33]。因此,本研究采用默认参数构建随机森林土壤水分反演模型。

2.1.2 XGBoost XGBoost (Extreme Gradient Boosting)是梯度提升决策树(GBDT, Gradient Boosting Decision Tree)的改进版^[34]。其基本思想是,通过初始模型简单预测,计算真实值与预测值的残差作为新的目标,训练新的决策树来学习残差(误差),并迭代多个树,使误差逐步降低。利用加权求和,得到最终预测值,每棵树都会产生一个预测值,最终结果是所有树的加权和。利用梯度下降优化目标,计算损失函数的梯度,指导新树如何修正误差。

2.1.3 ConvLSTM ConvLSTM (Convolutional Long-Short-Term Memory)是一种将卷积神经网络(CNN)和长短期记忆网络(LSTM)相结合的深度学习模型架构^[35]。LSTM能够有效地捕捉时间序列中的长期依赖关系。而CNN擅长提取空间特征^[36]。在ConvLSTM中,卷积操作被整合到LSTM单元中。传统LSTM单元中的点乘和加法运算被卷积运算所替代^[37]。模型的结构与参数设置参考了已有研究,并结合数据规模与模型收敛情况进行经验性调整。ConvLSTM网络包含8层,两层卷积单元(64与128个滤波器),全连接层节点数为64,Dropout比例为0.2。优化器采用Adam,学习率设为0.001,批大小为16,训练轮次为50。各参数的设置综合考虑了模型复杂度与训练稳定性,确保在验证集上误差收敛且无明显过拟合。

2.2 单/双时间节点建模

研究采用单/双时间节点建模。单时间节点建模-全时期反演是指选取月初数据建模,利用此模型对全时期的逐日数据进行反演。双时间节点建模-分段反演是指以14 d为一个反演周期,选取月初和月中两日数据分别建立模型,月

初模型用于反演月初至月中逐日土壤水分,月中模型用于反演月中至月末逐日土壤水分。节点选择综合考虑了土壤水分变化特征、反演精度、模型稳定性及业务操作效率,14 d反演周期能够有效平衡代表性和操作可行性,相比3 d或7 d更适合作为分段反演的周期。

2.3 评价指标

本研究采用决定系数(Determination Coefficient, R^2)、平均绝对误差(Mean Absolute Error, MAE)、均方根误差(Root Mean Square Error, RMSE)评价模型精度及模型准确性。

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

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

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (3)$$

其中: y_i 为土壤体积含水量观测值; $\hat{y}_i$ 为土壤体积含水量模型反演值; $\bar{y}$ 为土壤体积含水量观测值的均值; n 为样本数。

3 结果与分析

3.1 土壤水分关键特征因子筛选

通过皮尔逊相关性和特征重要性分析筛选各时期的土壤水分关键特征因子,其中相关性分析选取显著性 $P < 0.05$ 的因子,特征重要性选取大于0.05的因子作为特征因子。由图3和图4可知,裸土-播种期关键因子为FMC、WM、LST、TCI、NDWI、BVI、VHI;出苗-拔节期关键因子为FMC、WM、DEM、LST、TCI、WI、NDWI、LSWI、NDIIB7、NDVI、VCI;抽雄-灌浆期关键因子为BD、FMC、WM、LST、TCI、NDWI、LSWI、NMDI、NDIIB7、NDVI、BVI、TVI、VHI;成熟期关键因子为BD、FMC、WM、LST、LSWI、NMDI、BVI。

各时期输入因子中FMC、WM与LST均为关键特征因子,表明土壤参数与地表温度对土壤水分具有持续性的影响。其中前三个时期TCI与NDWI均为关键特征因子,表明冠层温度与土壤含水量相关指数对土壤水分具有一定的影响。此外,随着植被覆盖度增加,出苗-拔节期与抽雄-灌浆期的冠层含水量和作物形态及绿度相关指数的影响增大。

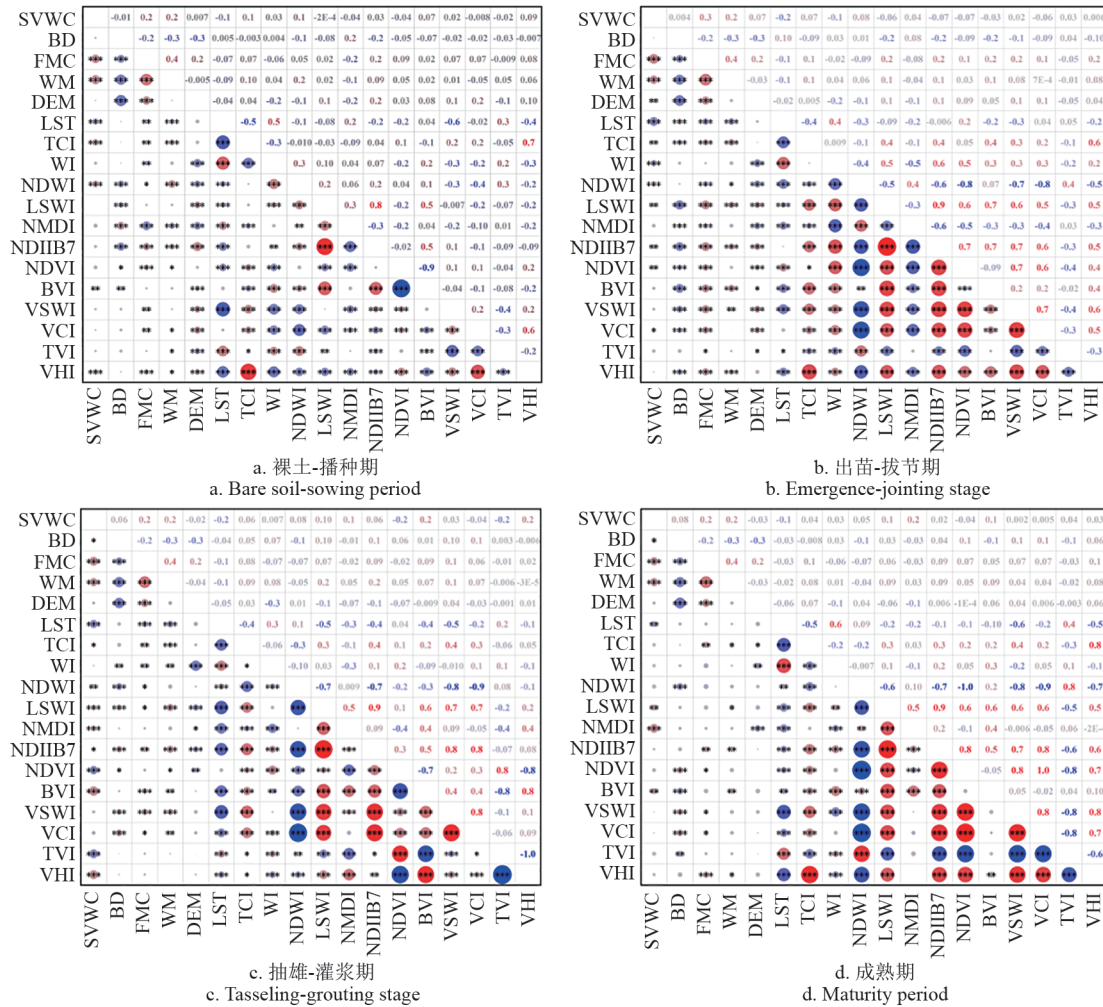


图3 分时期土壤体积含水量与筛选数据集相关性分析

Fig. 3 Correlation analysis between soil volumetric water content and selected datasets for different periods

注: 图中红色为正相关, 蓝色为负相关, 颜色越深, 圆越大相关性越强。* 为显著 $P < 0.05$, ** 为较显著 $P < 0.01$, *** 为极显著 $P < 0.001$ 。

Note: In the figure, red represents a positive correlation and blue represents a negative correlation. The darker the color and the larger the circle, the stronger the correlation. * is significant $P < 0.05$, ** is more significant $P < 0.01$, *** is highly significant $P < 0.001$.

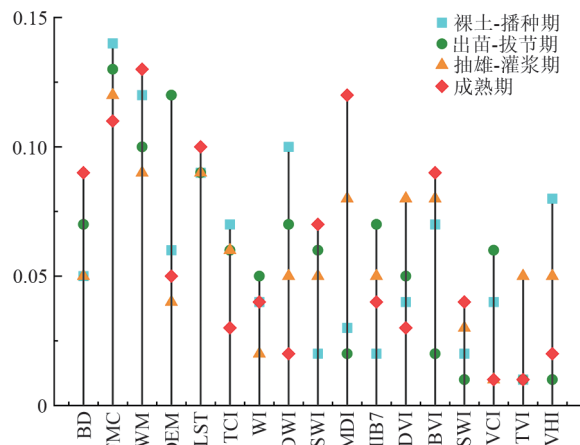


图4 分时期特征重要性分布图

Fig. 4 Feature importance distribution for different periods

3.2 建模时间节点的选取对反演精度的影响

根据关键因子筛选结果, 以随机森林模型为例, 采用单/多时间节点建模, 对比分析两种建模时间节点的选取对反演精度的影响。基于单时间节点建模的模型反演结果 R^2 多在 0.6 以下, 证明模型能够捕捉一定程度的土壤水分变化, 但整体精度有限(图5)。其中出苗-拔节期反演精度最差, 可能是此期间土壤墒情、植被指数等都快速变化, 单时间节点建模与中后期条件存在较大差异, 从而影响模型反演结果。运用双时间节点建模有效降低了上述影响因素对模型反演结果的影响, 大大提高了模型反演值的精度。

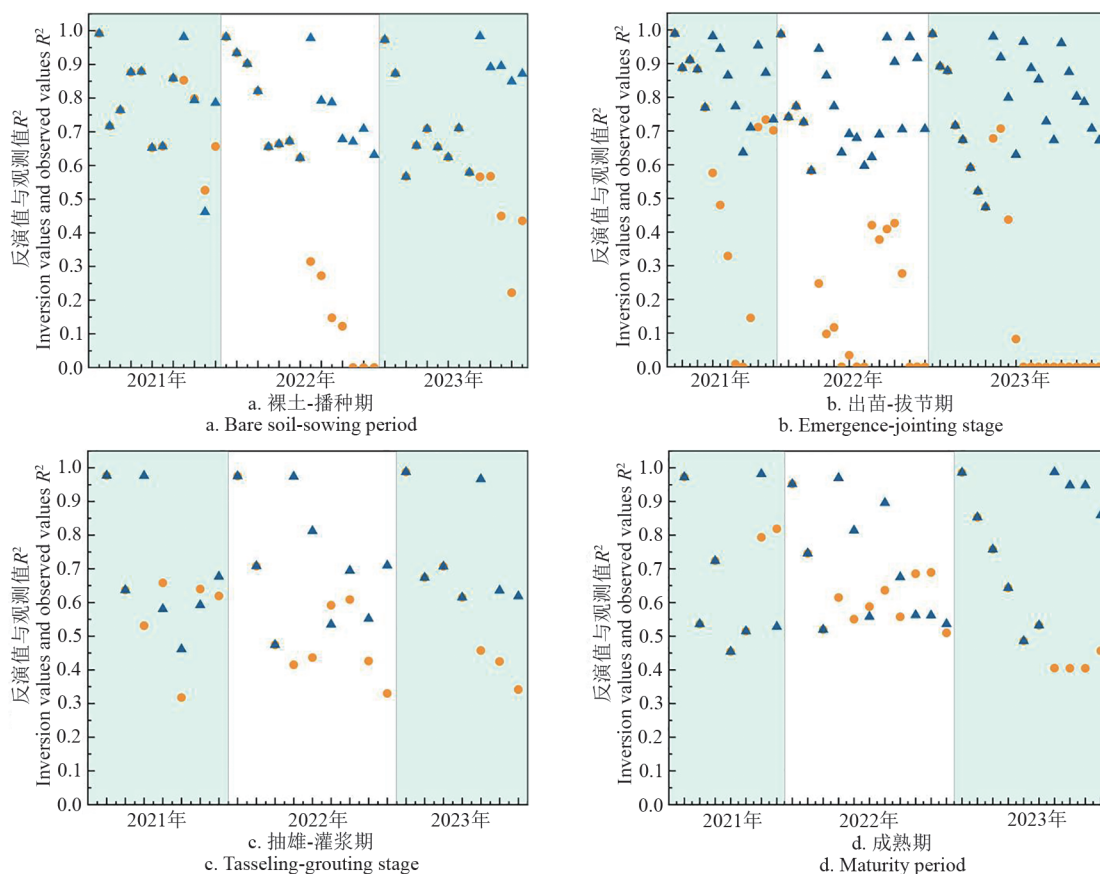


图5 分时期单/双时间节点模型反演值与观测值 R^2 对比图

Fig. 5 Comparison of R^2 between modeled and observed soil moisture for single-/dual-time-node models for different periods

注:图中●为单时间节点,▲为双时间节点。

Note: In the figure, ● represents a single phase and ▲ represents a dual phase.

对比单/双时间节点下的反演精度,单时间节点模型反演精度 R^2 主要分布于 $[0, 0.7)$ 之间(表3),表明单时间节点建模多数反演精度不高。双时间节点模型反演精度 R^2 主要分布于 $[0.6, 1]$ 之间,表明多数样本具有较高的反演精度。说明双时间节点建模策略能够更准确地捕捉土壤水分随时间变化的动态特征,提高模型整体的反演稳定性与精度。

3.3 模型训练精度

基于双时间节点构建随机森林、XGBoost和ConvLSTM模型在4个时期的反演值与实测值的决定系数 R^2 均高于0.875,总体拟合结果较好。其中,随机森林模型总体验证集的 R^2 在0.965~0.993之间,均方根误差RMSE在0.005~0.010(cm^3/cm^3)之间,平均绝对误差MAE在0.002~0.004(cm^3/cm^3)之间;XGBoost模型总体验证集的反演值与观测值的决定系数 R^2 在

0.951~0.984之间,均方根误差RMSE在0.007~0.011(cm^3/cm^3)之间,平均绝对误差MAE在0.005~0.007(cm^3/cm^3)之间;ConvLSTM模型总体验证集的反演值与观测值的决定系数 R^2 在0.875~0.982之间,均方根误差RMSE在0.010~0.017(cm^3/cm^3)之间,平均绝对误差MAE在0.006~0.012(cm^3/cm^3)之间(表4)。因此,随机森林模型的反演值与观测值的拟合结果最好。ConvLSTM模型的反演值与观测值的拟合结果较多数据出现低估情况,拟合结果相对较差。故3种模型在验证集的拟合效果为随机森林模型>XGBoost模型>ConvLSTM模型(图6)。

3.4 模型精度验证及对比分析

为了评估模型在未知样本的泛化能力,利用以上3种模型反演了2021年至2023年辽宁省不同时期的土壤水分情况,同时为了更真实地评估模型的泛化能力,本研究在进行验证时排除了建

表3 单/双时间节点下反演值与观测值 R^2 范围占比

R^2 范围 R^2 range	裸土-播种期		出苗-拔节期		抽雄-灌浆期		成熟期	
	Bare soil-sowing period		Emergence-jointing stage		Tasseling-grouting stage		Maturity period	
	单 Single	双 Dual	单 Single	双 Dual	单 Single	双 Dual	单 Single	双 Dual
	占比/% Proportion	占比/% Proportion	占比/% Proportion	占比/% Proportion	占比/% Proportion	占比/% Proportion	占比/% Proportion	占比/% Proportion
[0.9, 1]	12.20	19.51	7.02	28.07	13.04	26.09	11.11	25.93
[0.8, 0.9)	14.63	21.95	7.02	19.30	0.00	4.35	7.41	14.81
[0.7, 0.8)	12.20	21.95	15.79	22.81	8.70	13.04	14.81	11.11
[0.6, 0.7)	24.39	29.27	3.51	21.05	30.43	30.43	18.52	7.41
[0.5, 0.6)	12.20	4.88	7.02	7.02	8.70	17.39	29.63	33.33
[0, 0.5)	24.38	2.44	59.64	1.75	39.13	8.70	18.52	7.41
[0, 1]	100	100	100	100	100	100	100	100

注:表中单为单时间节点,双为双时间节点。

Note: Single represents the single time-node model, and Dual represents the dual time-node model.

表4 模型测试集评价指标对比表

Table 4 Comparison of model test set evaluation indicators

模型	评价指标	裸土-播种期	出苗-拔节期	抽雄-灌浆期	成熟期
Model	Evaluation index	Bare soil-sowing period	Emergence-jointing stage	Tasseling-grouting stage	Maturity period
随机森林 Random forest	R^2	0.973~0.993	0.965~0.992	0.973~0.991	0.973~0.991
	RMSE(cm^3/cm^3)	0.005~0.007	0.005~0.008	0.006~0.007	0.005~0.010
	MAE(cm^3/cm^3)	0.002~0.002	0.002~0.003	0.003~0.003	0.002~0.004
XGBoost	R^2	0.954~0.984	0.954~0.983	0.958~0.976	0.951~0.978
	RMSE(cm^3/cm^3)	0.007~0.010	0.008~0.001	0.009~0.010	0.010~0.011
	MAE(cm^3/cm^3)	0.005~0.007	0.005~0.006	0.006~0.008	0.005~0.007
ConvLSTM	R^2	0.951~0.974	0.875~0.982	0.946~0.979	0.891~0.961
	RMSE(cm^3/cm^3)	0.011~0.012	0.010~0.017	0.011~0.013	0.012~0.017
	MAE(cm^3/cm^3)	0.007~0.008	0.006~0.010	0.008~0.009	0.008~0.012

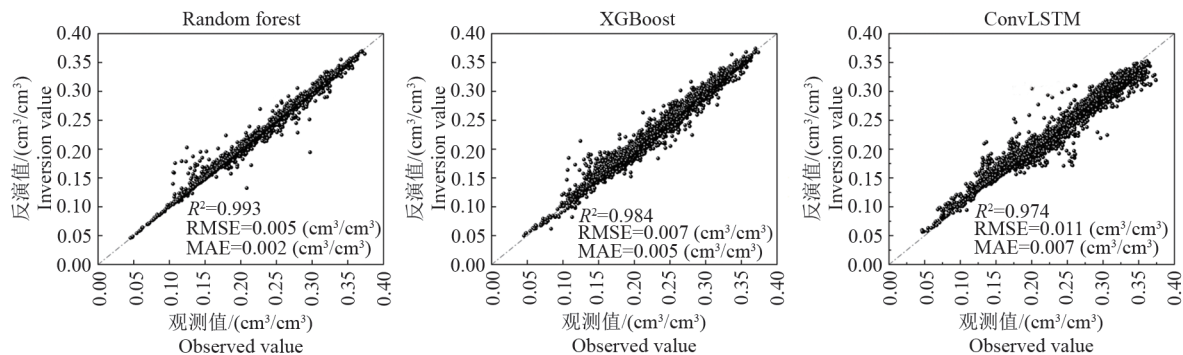
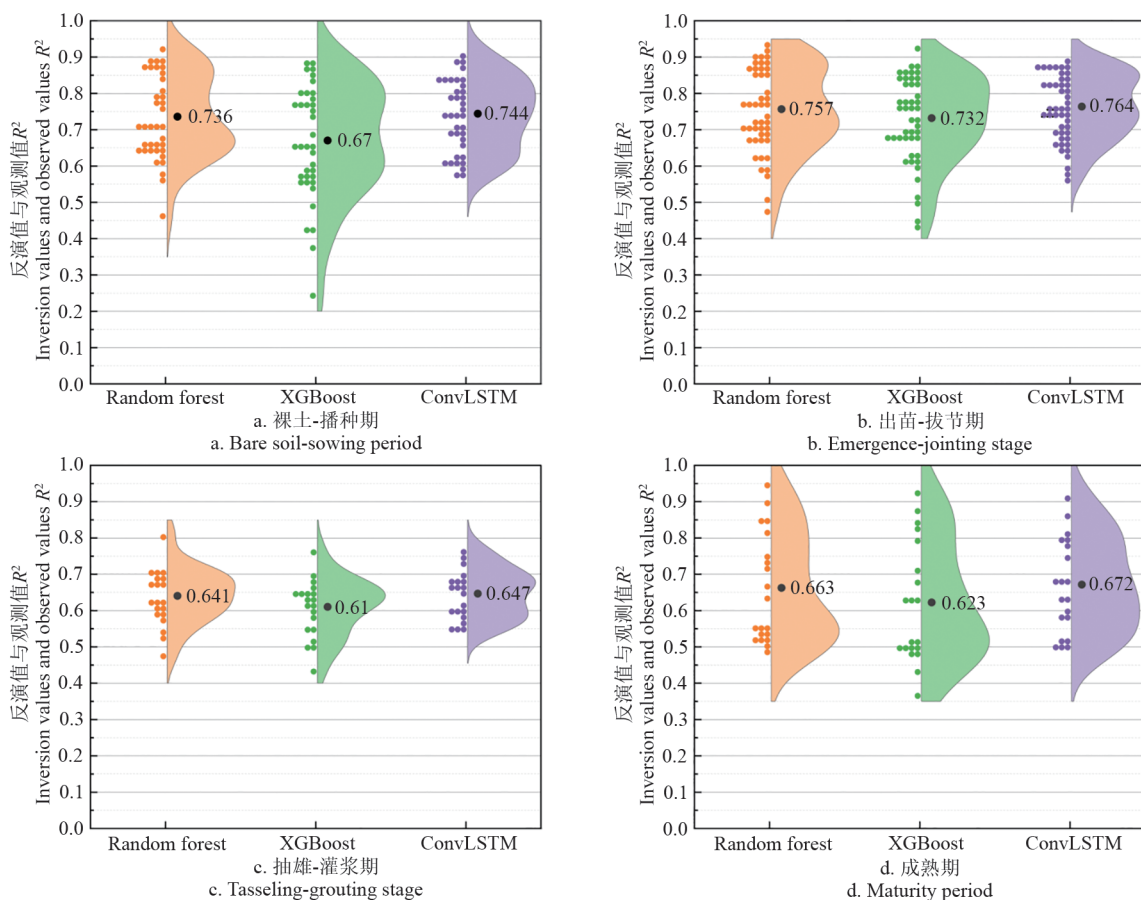


图6 不同模型的反演值与观测值拟合图($n=82\ 790$)

Fig. 6 Fitting between retrieved and observed values of different models ($n=82\ 790$)

模日期对应的样本数据,避免因模型对该时间节点的“记忆”而产生的过拟合现象。结果表明,反演精度表现为ConvLSTM模型>随机森林模型>XGBoost模型,该结果与模型在验证集的结果不一致(图7)。主要原因在于,随机森林和XGBoost作为传统机器学习,虽在验证集上实现

高度拟合,但面对新数据易因局部数据异常,导致模型反演结果分布较为分散,波动较大;相比之下,ConvLSTM模型倾向于输出平滑的预测结果,减少了局部极值的误判,导致反演结果中高精度与低精度的较少,反演结果分布更为集中,表现出更高的稳定性。

图7 分时期模型反演值与观测值 R^2 分布图Fig. 7 Distribution of R^2 between retrieved and observed soil moisture for different periods

从评价指标来看(表5),裸土-播种期反演值 R^2 均值达0.744,总体 R^2 在0.574以上, RMSE低于 $0.035(\text{cm}^3/\text{cm}^3)$, MAE低于 $0.025(\text{cm}^3/\text{cm}^3)$ 。该阶段植被覆盖较少,干扰的因素较少。此时主要受土壤参数及地表温度影响,气象条件稳定,建模所需输入因子较少,减少了复杂因素的干扰,模型反演精度较高。出苗-拔节期反演值 R^2 均值达0.764,总体 R^2 在0.560以上, RMSE在 $0.033(\text{cm}^3/\text{cm}^3)$ 以下, MAE在 $0.027(\text{cm}^3/\text{cm}^3)$ 以下。此阶段作物生长迅速,地表覆盖度增加,植被对遥感信号的影响增强,但气象条件仍较稳定。模型能够有效捕捉遥感指数与土壤水分动态变化的耦合关系,从而实现较高的反演精度。抽雄-灌浆期反演值 R^2 均值为0.647,总体 R^2 在0.548以上, RMSE在 $0.036(\text{cm}^3/\text{cm}^3)$ 以下, MAE在 $0.027(\text{cm}^3/\text{cm}^3)$ 以下。该阶段进入主汛期,降雨增多,蒸散发强烈,气象条件复杂,土壤水分空间分布差异增大,模型较其他时期反演精度略有

下降。成熟期反演值 R^2 均值为0.672,总体 R^2 在0.499以上, RMSE在 $0.046(\text{cm}^3/\text{cm}^3)$ 以下, MAE在 $0.032(\text{cm}^3/\text{cm}^3)$ 以下。该阶段降雨逐渐减少,蒸散发增强,植被进入衰老期,植被指数影响下降,反演精度较上一时期有所提升。

3.5 辽宁省干旱监测业务应用

考虑作物发育期,选取每个时期中期一日数据作为典型代表,并与CLDAS数据进行对比。以2021年4月8日、2021年6月1日、2021年8月7日、2023年9月15日的反演结果为例(图8),干旱等级划分依据见表2。

从时期来看,辽宁省干旱呈季节性变化,裸土-播种期辽宁省西部朝阳建平及中部城市存在小范围中、轻旱,而大部分地区处于无旱状态,这一时期土壤水分相对充足,干旱发生率较低;进入出苗-拔节期干旱范围明显扩大,中部及西部地区干旱最为严重,这一时期作物需水量增加,而降雨较少,干旱风险显著上升;在抽雄-灌浆期

表 5 模型反演结果评价指标对比表
Table 5 Comparison of evaluation metrics for model inversion results

模型	评价指标	裸土-播种期	出苗-拔节期	抽雄-灌浆期	成熟期
Model	Evaluation index	Bare soil-sowing period	Emergence-jointing stage	Tasseling-grouting stage	Maturity period
随机森林 Random forest	R^2	0.462~0.934	0.474~0.944	0.474~0.812	0.486~0.948
	RMSE(cm^3/cm^3)	0.012~0.042	0.012~0.031	0.017~0.036	0.011~0.037
	MAE(cm^3/cm^3)	0.008~0.026	0.008~0.027	0.012~0.026	0.007~0.031
XGBoost	R^2	0.243~0.885	0.431~0.924	0.433~0.763	0.365~0.926
	RMSE(cm^3/cm^3)	0.014~0.037	0.016~0.032	0.020~0.036	0.014~0.040
	MAE(cm^3/cm^3)	0.013~0.030	0.011~0.025	0.014~0.027	0.009~0.029
ConvLSTM	R^2	0.574~0.910	0.560~0.883	0.548~0.767	0.499~0.911
	RMSE(cm^3/cm^3)	0.014~0.035	0.015~0.033	0.019~0.032	0.017~0.046
	MAE(cm^3/cm^3)	0.010~0.025	0.010~0.027	0.014~0.025	0.017~0.032

全省大部分地区保持无旱或轻旱状态,这一时期辽宁省进入汛期,在一定程度上缓解了干旱;成熟期西北部、中部及凤城地区干旱再次出现,这一时期汛期结束并持续高温再次引发干旱。结合农业生产实际,出苗-拔节期干旱对作物生长和产量影响最大,应重点监测并采取适当灌溉措施。

从精度来看,将反演结果与观测值及 CLDAS-V2.0 产品值进行了对比分析。从各生育期来看,反演结果能够较准确地捕捉干旱区域的位置与强度,存在部分低值区高估、部分高值区低估情况,但总体与观测值较为一致,能够有效进行干旱预警。相比之下 CLDAS-V2.0 产品存在普遍高估现象,尤其在出苗-拔节期和成熟期这两个干旱高发时期的低值区结果不够精确,无法做到干旱的及时预警。这与王卓颖^[29]等人的研究结论,CLDAS-V2.0 土壤水分产品多数时间存在高估现象一致。从空间分辨率来看,CLDAS-V2.0 空间分辨率为 $0.062\ 5^\circ \times 0.062\ 5^\circ$,本研究为 $0.002\ 5^\circ \times 0.002\ 5^\circ$ 。对比结果表明,本研究结果在反映干旱空间格局和干旱强度方面具有较高的可靠性和稳定性,可为区域干旱监测和农业水分管理提供准确的数据支撑,并明显优于 CLDAS 产品,能够更好地满足精细化农业应用的需求。

4 讨论

本研究中,我们提出了 ConvLSTM 模型用于土壤水分反演,并取得了良好的反演效果。目前多数研究只使用常用的随机森林^[38]、XGBoost 等简单的机器学习模型,对深度学习模型的探索较少^[39,40]。本研究主要利用 FY-3D 数据计算的遥

感指数数据和土壤数据作为输入数据,利用随机森林、XGBoost、ConvLSTM 模型构建土壤体积含水量反演模型,结果表明 ConvLSTM 模型反演结果更为良好。在模型结构上,ConvLSTM 模型可以捕捉遥感指数与土壤参数之间的复杂关系。但在模型构建、调优方面比传统的机器学习模型更为复杂,在运行时间上更久,同时在训练数据的输入上要求更高,构建模型所需的数据量更大。其次,深度学习模型反演的结果很难解释变量之间的因果关系^[41,42]。

本研究方法虽在空间分辨率与反演精度方面均优于 CLDAS-V2.0 产品,显著提高了监测的精细化水平。但 CLDAS-V2.0 为全国产品,适用范围广,局地预报精度略低。本研究方法虽然在辽宁省表现良好,但在其他省份或不同气候区的适用性仍需进一步验证。表 6 列举了已有土壤水分反演研究,结果表明在同样时间分辨率下,本文方法在空间分辨率及反演精度上均表现出良好的结果,高分辨率反演结果能够更好地揭示区域土壤水分空间差异,尤其在农田尺度的精准灌溉、干旱监测与作物水分胁迫评估中具备更高的应用价值,但在数据选取上也存在一定的局限性。如受云层的影响,本研究中去掉云层部分,导致云层部分对应的土壤水分无法得到有效的监测。本研究在影响因子的选取上考虑了植被、地表温度、土壤参数、DEM 等,构建土壤水分反演模型。土壤水分的影响因素较为复杂,例如地表反照率、蒸发、散发、降水等因素也会影响土壤体积含水量^[43]。这些因素将在未来的研究中进行考虑,以探索更适用于土壤水分反演的模型,以便更为准确、快速地监测土壤水分。

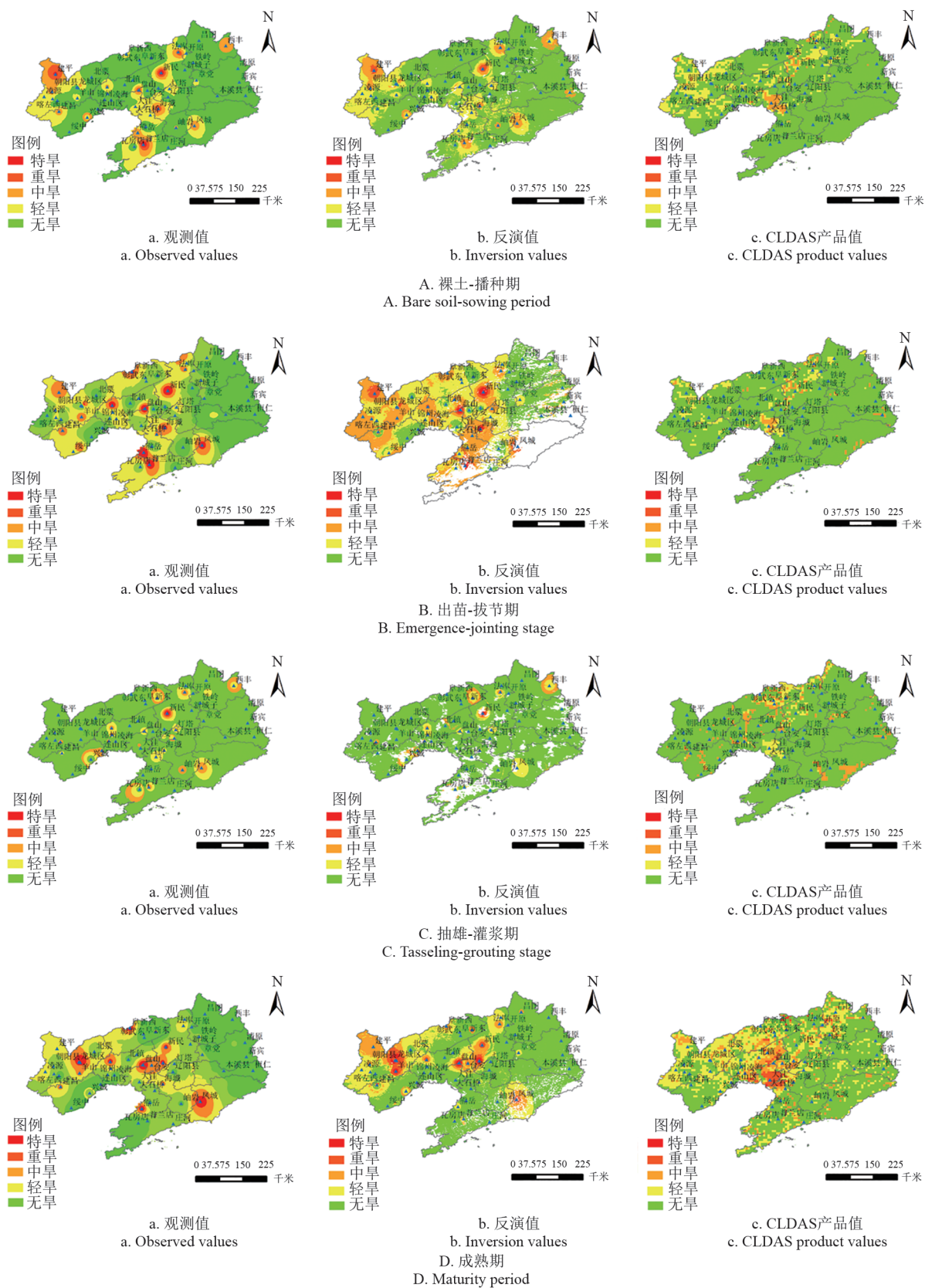


图8 分时期干旱等级图

Fig. 8 Drought-grade maps for different periods

表 6 已有表层土壤水分结果对比
Table 6 Comparison of published surface soil moisture results

方法 Method	卫星数据 Satellite data	时间分辨率 Temporal resolution	空间分辨率 Spatial resolution	精度 Accuracy	来源 Reference
RF/GBM	MODIS	Daily	1 km	R^2 : 0.48~0.62	[6]
ConvLSTM	FY-3D	Daily	250 m	R^2 : 0.50~0.91	本文方法

5 结论

土壤水分反演是一个复杂的过程,涉及多种数据源和方法。本研究基于 FY-3D/MERSI 卫星遥感数据,结合遥感数据、土壤参数等,构建了随机森林、XGBoost 及 ConvLSTM 模型,并对辽宁省 2021—2023 年土壤体积含水量进行反演与验证。主要结论如下:

(1)不同时期的影响因子存在显著差异,反映出土壤水分受土壤参数、地表热量及植被生长状态的综合影响。FMC、WM 与 LST 在各时期均为关键特征因子,植被覆盖度高的时期,冠层含水量和作物形态及绿度相关指数的影响增大。

(2)双时间节点建模较单时间节点 R^2 平均提升 0.45 以上,表明多时间节点建模有助于增强模型的预测稳定性。

(3)ConvLSTM 模型表现最优且具有泛化能力。与传统的机器学习相比,ConvLSTM 模型更具备实际应用潜力,模型 R^2 达到 0.875 以上, RMSE 低于 $0.017 \text{ cm}^3/\text{cm}^3$, MAE 低于 $0.012 \text{ cm}^3/\text{cm}^3$ 。

(4)辽宁省干旱呈明显的季节性变化,反演结果具有良好的时空一致性,与 CLDAS-V2.0 产品相比,在反映干旱空间格局和强度方面具有更高的可靠性和稳定性,能够为精细化农业水分管理提供精准的数据支撑。

综合来看,基于 ConvLSTM 的土壤水分反演方法可实现高精度、逐日尺度的土壤水分监测,优于 CLDAS-V2.0 等常规产品,在空间分辨率与应用精度上均具有优势,为辽宁省土壤水分监测提供技术支持,可为干旱预警与响应等方面提供更加精准与精细化的数据支撑。

参考文献

[1] 陈国茜,杨耀晨,李素云,等.基于随机森林模型的青海高原土壤水分反演技术研究[J].遥感技术与应用,2024,39(06):1417-1428.
[2] 郝萍萍,韩金朝,刘晓,等.土壤水分含量测定方法的技术现状与发展趋势[J].江苏农业科学,2024,52

(17):11-20.
[3] 江笑薇,白建军,刘宪峰.基于多源信息的综合干旱监测研究进展与展望.地球科学进展[J],2019,34(03):275-287.
[4] 张浩杰,薛华柱,赵天杰,等.联合极轨和倾斜轨道被动微波传感器的青藏高原土壤水分变化监测[J].遥感技术与应用,2024,39(06):1404-1416.
[5] 赵天杰.被动微波反演土壤水分的 L 波段新发展及未来展望[J].地理科学进展,2018,37(02):198-213.
[6] 付平凡,杨晓静,姜波,等.融合多源数据的高分辨率土壤水分模拟模型构建及应用[J].农业工程学报,2025,41(05):96-106.
[7] 李万涛,杨明龙,唐秀娟,等.基于 Sentinel 多源遥感数据的农田地表土壤水分反演[J].南方农业学报,2025,56(01):87-96.
[8] 蔡斌,陆文杰,郑新江.气象卫星条件植被指数监测土壤状况[J].国土资源遥感,1995(04):45-50+66.
[9] 张涛,张立新,蒋玲梅,等.裸露土壤参数空间分布对被动微波遥感反演土壤水分的影响[J].中国科学:地球科学,2012,42(09):1403-1413.
[10] 张添,黄春林,沈焕锋.土壤水分对土壤参数的敏感性及其参数优化方法研究[J].地球科学进展,2012,27(06):678-685.
[11] 刘昫昊,李雪冬,费龙,等.基于特征选择与遗传神经网络的土壤水分反演[J].中国农业气象,2024,45(10):1095-1108.
[12] 刘娣,孙佳倩,余钟波.基于机器学习模型的多层土壤湿度反演[J].河海大学学报(自然科学版),2024,52(03):7-14.
[13] Hanade Houmma I, El Mansouri L, Gadal S, et al. Prospective analysis of spatial heterogeneity influence on the concordance of remote sensing drought indices: a case of semi-arid agrosystems in Morocco (Moulouya and Tensift watersheds) [J]. Geocarto International,2022,37(27):14899-14924.
[14] Chen J, Jin Q,Chao J. Design of deep belief networks for short-term prediction of drought index using data in the Huaihe River basin[J]. Mathematical Problems in Engineering2012, 2012(1): 235929.
[15] Wu T, Xu L, Lv Y, et al. Integrating causal inference with ConvLSTM networks for spatiotemporal forecasting of root zone soil moisture[J]. Journal of Hydrology, 2025, 659: 133246.
[16] ElSaadani M, Habib E, Abdelhameed A M, et al.

- Assessment of a spatiotemporal deep learning approach for soil moisture prediction and filling the gaps in between soil moisture observations[J]. *Frontiers in Artificial Intelligence*, 2021, 4: 636234.
- [17] Yinglan A, Wang G, Hu P, et al. Root-zone soil moisture estimation based on remote sensing data and deep learning[J]. *Environmental Research*, 2022, 212: 113278.
- [18] 赵天杰,施建成,徐红新,等. 闪电河流域水循环和能量平衡遥感综合试验[J]. *遥感学报*, 2021, 25(04): 871-887.
- [19] 王岩,王敬宜,冯锐,等. 基于FY-3D/MERSI数据的东北地区干旱监测方法研究[J]. *干旱地区农业研究*, 2023, 41(04): 289-297+317.
- [20] 李伟,宋睿,刘明江,等. 光学遥感在区域灌溉监测中应用的研究进展[J]. *排灌机械工程学报*, 2024, 42(11): 1157-1165.
- [21] Kogan F N. Application of vegetation index and brightness temperature for drought detection [J]. *Advances in Space Research*, 1995, 15(11): 91-100.
- [22] Bajgain R, Xiao X M, Wagle P, et al. Sensitivity analysis of vegetation indices to drought over two tall grass prairie sites. *ISPRS [J]. ISPRS Journal of Photogrammetry and Remote Sensing*, 2015, 108: 151-160.
- [23] 黄友昕,胡茂胜,沈永林,等. MODIS 干旱指数结合RBFNN反演冬小麦返青期土壤湿度[J]. *农业工程学报*, 2019, 35(12): 81-88.
- [24] Carlson T N, Gillies R R, Perry E M. A method to make use of thermal infrared temperature and NDVI measurements to infer surface soil water content and fractional vegetation cover[J]. *Remote Sensing Reviews*, 1994, 9(1/2): 161-173.
- [25] Zhang Z, Xu W, Shi Z, et al. Establishment of a comprehensive drought monitoring index based on multisource remote sensing data and agricultural drought monitoring[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 2113-2126.
- [26] Xing N, Huang W, Xie Q, et al. A transformed triangular vegetation index for estimating winter wheat leaf area index[J]. *Remote Sensing*, 2019, 12(1): 16.
- [27] Kogan F. World droughts in the new millennium from AVHRR-based vegetation health indices[J]. *Eos, Transactions American Geophysical Union*, 2002, 83(48): 557-563.
- [28] Gao B C. NDWI: a normalized difference water index for remote sensing of vegetation liquid water from space[J]. *Remote Sensing of Environment*, 1996, 58(3): 257-266.
- [29] 王卓颖,刘杨晓月,杨骥,等. 风云卫星系列及CLDAS土壤水分产品多尺度精度验证与评价:以青藏高原那曲地区为例[J]. *中国农业气象*, 2021, 42(09): 788-804.
- [30] 全国农业气象标准化技术委员会(SAC/TC 539). 农业干旱等级: GB/T 32136-2015[S]. 北京: 中国标准出版社, 2015.
- [31] 王晓,刘海新,孙振宇,等. 融合多源数据与机器学习算法的综合干旱监测模型构建——以河北省为例[J]. *水资源与水工程学报*, 2024, 35(06): 207-219.
- [32] Adab H, Morbidelli R, Saltalippi C, et al. Machine learning to estimate surface soil moisture from remote sensing data[J]. *Water*, 2020, 12(11): 3223.
- [33] Luo L, Li Y, Guo F, et al. Research on robust inversion model of soil moisture content based on GF-1 satellite remote sensing[J]. *Computers and electronics in agriculture*, 2023, 213: 108272.
- [34] 王思超,蔡玉林,朱子正,等. XGBoost和DSCGAN模型协同的SMAP土壤水分降尺度方法[J]. *地球信息科学学报*, 2024, 26(09): 2213-2225.
- [35] Zhang Y, Xie D, Tian W, et al. Construction of an integrated drought monitoring model based on deep learning algorithms[J]. *Remote Sensing*, 2023, 15(3): 667.
- [36] 王然,赵建辉,杨会巾,等. 基于RIME-CNN-SVR模型的麦田土壤水分反演[J]. *农业工程学报*, 2024, 40(15): 94-102.
- [37] Kasim A A, Leng P, Li Y X, et al. Remote sensing of root zone soil moisture: A review of methods and products[J]. *Journal of Hydrology*, 2025: 133002.
- [38] 张菊英,覃光华,郭立. 土壤水分监测站点优选与预测模型选择的研究[J]. *中国农村水利水电*, 2025(03): 37-43+54.
- [39] Wu Z, Cui N, Zhang W, et al. Estimating soil moisture content in citrus orchards using multi-temporal sentinel-1A data-based LSTM and PSO-LSTM models[J]. *Journal of Hydrology*, 2024, 637: 131336.
- [40] Kumar K C A, Reddy G P O, Masilamani P, et al. Integrated drought monitoring index: A tool to monitor agricultural drought by using time-series datasets of space-based earth observation satellites[J]. *Advances in Space Research*, 2021, 67(1): 298-315.
- [41] Zhao Q, Ma X, Yao W, et al. A drought monitoring method based on precipitable water vapor and precipitation[J]. *Journal of Climate*, 2020, 33(24): 10727-10741.
- [42] Gavahi K, Abbaszadeh P, Moradkhani H. How does precipitation data influence the land surface data assimilation for drought monitoring? [J]. *Science of the Total Environment*, 2022, 831: 154916.
- [43] 王霞迎,折育霖,张双成,等. 基于Sentinel-1和Landsat 8遥感数据的高原地区土壤水分反演[J]. *测绘通报*, 2024(07): 140-146.