基于 Landsat 8 数据的荒漠土壤水分遥感反演

高 琪[1]，彭 杰[1]，冯春晖[1]，宋 奇[1]，马自强[2]，王玉珍[1]

(1.塔里木大学 植物科学学院，新疆 阿拉尔 843300；
2.北京大学 地球与空间科学学院，遥感与地理信息系统研究所，北京 100000)

摘 要：[目的] 分析荒漠土壤水分变化特征，为南疆干旱区荒漠土壤水分遥感监测提供理论依据和方法支持。[方法] 以 Landsat 8 数据构建干旱地区荒漠土壤水分建模指示因子，通过优选的 26 个光谱指数、地表温度(T_s)和地形数据(DEM)为建模因子，分别以偏最小二乘(PLSR)、支持向量机(SVM)和随机森林(RF)3 种方法构建土壤水分反演模型，对模型进行验证和对比，选取最优模型反演空台里克土壤水分空间分布。[结果] ①TVDI，NR，GLI 等 26 个优选的光谱指数中，T_s 和 DEM 与土壤水分均达极显著相关，可作为南疆干旱区荒漠土壤水分遥感建模的指示因子；②对比 3 种模型，RF 模型建模集和预测集的 R^2 分别为0.93，0.91，预测集 RPD 为 3.90，各评价指标均为最高，PLSR 模型精度次之，SVM 模型精度最低；③以 RF 模型反演研究区表层土壤水分，在不同土地利用分类中土壤水分分布特征存在明显差异，特别在盐结皮区域的差异尤为突出。[结论] 利用光谱指数、环境因子和地形数据构建的多因子、多指数综合的模型能较高精度地反演干旱区荒漠表层土壤水分，对研究该地区土地荒漠化和生态环境治理具有一定参考价值。

关键词：Landsat 8；荒漠土壤水分；光谱指数；PLSR；SVM；RF

文献标识码：B **文章编号**：1000-288X(2021)01-0125-07 **中图分类号**：S152.7

文献参数：高琪，彭杰，冯春晖，等. 基于 Landsat 8 数据的荒漠土壤水分遥感反演[J]. 水土保持通报，2021，41(1)：125-131. DOI：10.13961/j.cnki.stbctb.2021.01.018；Gao Qi，Peng Jie，Feng Chunhui，et al. A study on inversion for remote sensing of desert soil moisture based on Landsat 8 data [J]. Bulletin of Soil and Water Conservation，2021，41(1)：125-131.

A Study on Inversion for Remote Sensing of Desert Soil Moisture Based on Landsat 8 Data

Gao Qi[1]，Peng Jie[1]，Feng Chunhui[1]，Song Qi[1]，Ma Ziqiang[2]，Wang Yuzhen[1]

(1. *College of Plant Science，Tarim University，Alar，Xinjiang* 843300，*China*；2. *School of Earth and Space Sciences，Peking University，Institute of Remote Sensing and Geographic Information System，Beijing* 100000，*China*)

Abstract：[Objective] The variation characteristics of desert soil moisture were analyzed to provide a theoretical basis and methodology for remote-sensing monitoring of soil water content in the arid desert of Southern Xinjiang Uygur Autonomous Region. [Methods] The desert soil moisture modeling indicators were constructed based on Landsat 8 data. An optimal 26 spectral index，land surface temperature (T_s)，and digital elevation model data (DEM) were selected as modeling factors，and the soil water inversion model was constructed using the partial least squares regression (PLSR)，support vector machine (SVM)，and random forest (RF) algorithms. After model validation and comparison，the spatial distribution of soil moisture in Kongtailike was retrieved using the optimal model. [Results] ① The temperature vegetation dryness index，NR，GLI，and other 26 preferred spectral indices，as well as T_S and DEM，were significantly correlated with soil moisture. They could be used as indicators for remote-sensing modeling of desert soil moisture in the arid area of Southern Xinjiang. ② Among the three models，the R^2 of calibration and validation based on the RF model were 0.93 and 0.91，respectively，and the RPD of validation was 3.90，which was the highest. The

收稿日期：2020-10-07 **修回日期**：2020-11-12

资助项目：国家重点研发计划项目“土壤综合观测与智能服务平台研发与应用”(2018YFE0107000)资助。

第一作者：高琪(1996—)，男(汉族)，新疆维吾尔自治区昌吉市人，硕士研究生，主要研究方向为干旱区生态环境遥感。Email：gaoqizky@163.com。

通讯作者：彭杰(1977—)，男(汉族)，湖南省邵阳市人，教授，主要从事干旱区资源环境遥感监测机理与模型研究。Email：pjzky@163.com。

PLSR model accuracy was the second best, and the SVM model accuracy was the lowest. ③ The surface soil moisture in the study area was retrieved by the RF model, and the characteristics of soil moisture distribution in different ground classifications were noticeably different, especially in the salt crust region. [Conclusion] The comprehensive use of spectral index, environmental factors, and terrain factors could result in the inversion of the soil water content in arid areas with a higher accuracy, providing scientific value for the desertification and ecological environment control in this area.

Keywords: Landsat 8; desert soil moisture; spectral index; partial least squares regression (PLSR); support vector machine (SVM); random forest (RF)

水资源是地球生态系统的重要组分,并在全球水汽循环中发挥着积极作用,土壤水分作为水资源的重要分支,是连接植被、土壤和大气的关键因子,随着全球气候变暖,三者间物质和能量循环加剧,从而土壤水分蒸散严重,干旱现象频繁发生,如何合理的监测土壤水分运移,改善旱情是当今全球的热点问题[1-4]。传统土壤水分监测是基于地面有限点的测定,不但耗损大量经济成本和人力资源,而且空间连续性缺乏,时空信息获取困难,导致传统方法难以满足大范围土壤水分监测的需求[5]。遥感技术因具有宏观性、高效性、时效性、经济成本低和地面限制因素少等独特优势,已逐渐发展为现今土壤水分研究的关键方法[6]。

20 世纪 60 年代至今,以遥感监测土壤水分的技术日渐成熟,并逐渐形成了微波遥感法、热惯量法、温度—植被干旱指数法和光谱指数法等[7-8]。其中,微波遥感在监测过程中易受地形、植被覆盖度和地表温度的影响[9-10];热惯量法存在卫星数据获取时效性差、遥感图像解译效果低等劣势[11];温度-植被干旱指数法不适用于高植被覆盖区域土壤水分的监测;光谱指数法在监测过程中虽对遥感数据要求较多,但在植被和土地类型多样且覆盖程度复杂区域有较好的监测效果,因此,以光谱指数法监测地貌和植被类型复杂的空台里克地区土壤水分变化,其适应度更高,监测效果更为显著。鉴于光谱指数监测土壤水分的特性,国内外诸多学者以光谱指数法对土壤水分监测进行了大量研究,如 Ghulam 等[12]通过构建 Nir-Red 特征空间,描述植被覆盖信息,发现改进垂直干旱指数(modified perpendicular drought index, MPDI)比传统指数能更好的反演旱情;张学艺等[13]结合 MODIS-EVI 得到改进型干旱指数(modified temperature vegetation dry ness index, MTVDI),通过构建各指数与土壤含水量的相关函数,表明经改进后 MTVDI 的监测效果显著;姚云军等[14]发现并提出了短波红外土壤湿度指数(shortwave infrared soil moisture index, SIMI),以 SIMI 反映土壤水分信息有较好的效果。诸多研究结果表明基于单一光谱指数研究均能反映土壤水分表征旱情,但面对植被覆盖复杂、地理因素多样,同时外延到研究区以外大范围监测研究时,其适用性尚未验证。为此,有学者在光谱指数法的基础上构建综合干旱监测模型,如孙灏等[15]比较 13 种典型光谱指数的适用特点,构建了农业干旱遥感监测的分类体系;葛少青等[16]在综合 NDPI,TVDI 和 PDI 共 3 种指数监测干旱区土壤水分取得了良好效果;Du 等[17]考虑植被、土壤和降水亏缺等干旱因素,以主成分变换构建综合干旱监测指数(spatial drought index,SDI),表明综合光谱指数监测土壤水分的精度和适用性均优于单一指数,同时以光谱指数监测土壤水分的发展逐步呈现为单指标向多指标、单因素向多因素发展的整体趋势。

迄今为止,以多种光谱指数、多因素为模型因子对土地利用类型和地表景观复杂多样区域土壤水分监测研究的相关报道甚少,其监测效果有待进一步验证。本文以南疆干旱区阿克苏地区的空台里克冲积扇为研究区,该研究区以沙漠—绿洲过渡带为主要土地利用类型,同时地表景观类型多样且交错分布,其中,包括高覆盖区域植被—中覆盖区域植被—低覆盖区域植被、裸地和农田;区域内土壤质地空间分布变异性强,从而地表粗糙度差异较大,特别是在盐结皮区域和非盐结皮区域之间差异较为显著。为更好的探究综合光谱指数法对地表景观复杂多样的空台里克区域荒漠土壤水分监测的效果及适用性,因此,本研究以 Landsat 8 数据拟构建一种基于多个光谱指数,同时考虑地形(digital elevation model, DEM)和环境数据等因子的土壤水分综合监测模型,为干旱区旱情监测和荒漠土壤水分空间分布研究提供一定理论依据和方法支持。

1 材料与方法

1.1 研究区概况

研究区位于南疆阿克苏地区温宿县境内的空台里克冲积扇(40°40′10″—41°32′36″N,80°45′09″—81°12′45″E),北靠天山山脉,西临塔里木盆地边缘,地势北高南低,西高东低[18],平均海拔 1 200 m,面积

约 2 176 km^2。研究区属于典型沙漠—绿洲过渡带,光热资源充足,气候干燥,年平均积温 4 200 ℃,无霜期 212 d,年平均降水 55.1 mm,而年平均蒸发却高达 2 428 mm;区域交通单一,以南北方向长 110 km 和东西方向长 80 km 的省级公路为主且贯穿整个研究区,沿省级公路两侧由北向南土地利用类型依次为沙漠、沙漠—绿洲过渡带、河流和农田;土壤质地以沙土、沙壤土为主;地表植被类型分布多样,以柽柳、骆驼刺,盐惠木等荒漠植被为主且呈高(植被全覆盖)—中(植被半覆盖)—低(裸地)的趋势集中分布于大面积沙漠—绿洲过渡带,小面积农田以棉花种植为主。

1.2　遥感数据获取及预处理

通过美国地质调查局 USGS 官网(https://earthexplorer.usgs.gov/)下载过境时间为 2019 年 6 月 15 日,轨道号为 146/31,146/32,云量小于 10%,空间分辨率为 30 m 的 2 景 Landsat 8 遥感影像。通过 ENVI 5.3 软件完成影像预处理,具体包括:辐射定标、FLAASH 大气校正、影像镶嵌、裁剪和波段运算,经预处理后对研究区遥感影像进行支持向量机分类(SVM Classical);以遥感影像提取比辐射率和黑体辐射亮度等数据信息用于地表温度计算。通过 ArcGIS 10.2 软件获取各光谱指数信息。

通过地理空间数据云(http://www.gscloud.cn/)下载空间分辨率为 30 m 的 DEM 数据,以 ArcGIS 10.2 软件完成洼地填充、影像裁剪、几何精度校正和高程数据提取。

1.3　土壤样品采集与测定

本研究于 2019 年 6 月 15 日 Landsat 8 卫星过境当日完成土壤样品采集,共计 160 个,为使野外采样充分反映研究区土壤质地及植被覆盖,同时考虑到研究区交通较为单一的情况,沿东西走向长 110 km 的省级公路两侧以近似等间距布设 20 个采样样方,各样方内分布 8 个采样点且遵循如下采样原则:①远离省级公路 300 m 以上;②避开河流和农业用水渠等;③依据土地利用分类结果,在不同覆盖类型内均布设样方;④以五点法采集各样方表层 0—20 cm 的土样并均匀混合。为防止土样水分蒸发而导致含水量数据误差,将采集的土样快速装入自封袋中密封保存,带回实验室称取 50 g 左右,采用质量烘干法测定土壤含水率,每个样品设置 3 次重复,各重复间相对误差控制在 5%以内,取 3 个重复样的平均值为该样品的土壤水分实际测定值。剩余土壤样品放置于室内风干,经充分研磨后统一过筛(2 mm),使用稀释溶液法测定 pH 值和电导率(EC)。

1.4　建模方法与模型精度评价

本研究选取经典线性模型偏最小二乘法(partial least squares regression, PLSR)、二分类模型支持向量机(support vector machine, SVM)和机器学习模型随机森林(random forest, RF)3 种方法,分别通过 The Unscrambler10.5 和 *R* 语言软件构建模型。将 160 个样品实测数据在 Excel 软件中依次由小到大排列,以 2∶1比例分成 107 个建模集样品,53 个预测集样品,通过决定系数(R^2)、均方根误差(RMSE)、平均绝对误差(MAE)、相对分析误差(RPD)对模型精度验证;当 R^2 值越大、RMSE 值越小,说明模型精度越高,预测效果越好;MAE 表示模型误差范围,其值越低,模型精度越高;相对 RPD 值而言,当 RPD<1.5 时,表明模型没有预测能力;当 1.5≤RPD<2.0,RPD≥3.0,表明模型预测能力较差,只能进行粗略估算;当 2.0≤RPD<2.5 时,表明模型预测能力较好;当 2.5≤RPD<3.0 时,表明模型预测能力很好;当 RPD≥3 时,表明模型预测能力极出色,能进行高精度估算[19]。其中,各精度评价指标通过如下(1)—(5) 公式计算:

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

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

$$\mathrm{MAE}=\frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i-y_i| \tag{3}$$

$$\mathrm{RPD}=\frac{\mathrm{SD}}{\mathrm{RMSE}} \tag{4}$$

$$\mathrm{SD}=\sqrt{\frac{\sum_{i=1}^{n}(y_i-\bar{y})^2}{n-1}} \tag{5}$$

1.5　研究方法

1.5.1　*光谱指数*　不同光谱指数描述的地表干旱特征有所差异,进而反映土壤水分表征旱情的结果也不一致[15],由此造成单一光谱指数监测土壤水分精度较低,本研究在此基础上针对空台里克地区植被覆盖和土地利用等复杂情况,选取 TVDI,NR 和 NDVI 等 31 个光谱指数并充分考虑了不同指数对土壤水分反映的基本特征,以多个光谱指数为建模因子构建反演模型,探究研究区土壤水分变化规律。

1.5.2　*地表温度(T_s)*　地表温度是区域尺度地表物理参量的一个关键因子,作为地球表面能量平衡的重要指示指标,地表温度能间接的反映土壤水分状态表达旱情特征。其计算方法见公式(6)—(8):

热红外辐射亮度值的表达式如下:

$$L_\lambda=[\varepsilon B(T_s)+(1-\varepsilon)L\downarrow]\varepsilon+L\uparrow \tag{6}$$

式中 ε 为地面比辐射率数据；T_s 为地面实际温度；$B(T_s)$为黑体在地表实际温度的热辐射亮度；τ 为大气；$L\uparrow$为大气向上的辐射亮度；$L\downarrow$为大气向下由地表反射的辐射亮度。

黑体辐射亮度 $B(T_s)$的计算公式如下：

$$B(T_s)=\frac{L_\lambda-L\uparrow-\tau(1-\varepsilon)L\downarrow}{\tau\varepsilon} \tag{7}$$

式中：L_λ 表示大气校正后的热红外波段辐射定标值。

得到黑体辐射亮度数据后通过普朗克函数计算地表温度(T_s)。

$$T_s=\frac{K_2}{\ln\left(\frac{K_1}{B(T_s)}+1\right)} \tag{8}$$

式中：T_s 为地表实际温度；$B(T_s)$为黑体在 T_s 的热辐射亮度；K_1，K_2 为卫星预设常数；Landsat 8 预设 $K_1=774.89$ W/(m^2 · sr · μm)，$K_2=1\,321.08$ K。

1.5.3 *地形数据*(DEM) 研究区地势北高南低，自西北向东南倾斜，考虑到地势高低起伏的特征是影响土壤水分、植被分布和气候特征等地域分异的关键因子[20]，因此，本研究以 DEM 数据作为建模因子，分析因地形变化引起土壤水分分异的规律，更进一步了解空台里克地区土壤水分分布的总体特征。

2 结果与分析

2.1 土壤属性统计性描述

表 1 对实地采集的 160 个表层土壤样品基础属性进行统计描述。由表 1 可知，表层土壤水分最小值为 4.30%，而最大值为 44.29%，极差高达 39.99%，平均值为 21.58%，说明表层土壤水分(0—20 cm)分布差异较大，引起这种差异的可能原因为不同土壤质地和植被覆盖，从而造成研究区表层各部分土壤的持水能力差异。EC，pH 值的极差分别为 78.53 dS/m，1.8，平均值分别为 24.01 dS/m，8.16，变异系数分别为 44.57%和 5.12%，同时表层土壤水分变异系数为 30.74%。而变异系数一般分为强变异(>50%)、中等变异(20%～50%)和弱变异(0～20%)[21]，引起土壤水分和 EC 值中等变异一是植被覆盖差异较大，高植被覆盖区域植被对土壤水分和养分元素需求较高，低覆盖区域直接受太阳高温辐射作用较强；二是研究区人为活动现象明显，新开垦农耕田、撂荒地不断增加以及人为设立保护区存在，导致土壤结构人为破坏严重，使土壤水分和盐分较强变异。土壤属性的偏度和峰度值分别描述数据偏斜程度和各特征值的高低分布情况，表中所有数据偏度值均大于 0，峰度均与 3 有一定差距，表明数据总体呈现偏态分布且均不符合正态分布规律。

表 1 土壤样品基础属性

土壤属性	最小值	最大值	均值	变异系数/%	峰值	偏度
土壤水分/%	4.30	44.29	21.58	30.74	1.07	0.49
EC/(dS · m^{-1})	1.07	79.60	24.01	44.57	5.16	1.60
pH 值	7.37	9.17	8.16	5.12	−0.79	0.13

2.2 建模因子与土壤水分的相关性分析

为探究各建模因子与土壤水分相关性，将地表温度(T_s)和地形数据(DEM)以及经波段运算的 31 个光谱指数与实测土壤水分在 Excel 2019 软件中进行相关分析并附各因子计算公式(具体见表 2)。通过表 2 可知，26 个光谱指数与土壤水分为极显著相关，其中，温度植被干旱指数(TVDI)与土壤水分相关性最高，为 0.53，表明 26 个光谱指数均能在不同程度内指示土壤水分，但 TVDI 指示效果优于其余光谱指数；绿度差值植被指数(green difference vegetation index，GDVI)、土壤改良植被指数 2(modified soil adjustment vegetation，MSAVI2)和非线性植被指数(nonlinear vegetation index，NLI)的相关系数分别为 0.06，0.08，0.15，与土壤水分均未达到显著或极显著相关，表明 3 个光谱被指数均不能有效的指示土壤水分信息，不适宜作为建模因子，地表温度和地形与土壤水分相关系数分别为−0.41 和−0.43，均为极显著负相关，表明两者均能影响土壤水分分布且相关性越高，土壤含水量则越低。因此，本研究在相关系数的基础上筛选出 26 个光谱指数，地表温度和地形为反演模型建模因子。

2.3 不同建模方法对比分析

表 3 为 PLSR，SVM，RF3 种模型方法构建土壤水分建模集和预测集的精度统计表。通过表 3 可知，建模集中 3 种模型方法的精度差异较大，RF 模型的 R^2 最高，为 0.93，而 RMSE 和 MAE 均为最小值，分别为 1.49，1.09；SVM 模型的 R^2 最低，为 0.43，RMSE 和 MAE 分别为 3.66，2.82，均为建模集中最大值；PLSR 模型性能介于二者间。综上所述，建模集中，RF 模型性能明显优于 PLSR 和 SVM 方法，反演效果优劣排序依次为 RF>PLSR>SVM。在预测集中，3 种模型的各指标相较于建模集均无明显下降，表明模型比较稳定，但三者相比较而言，仍然是 RF 模型的各项指标明显优于 PLSR 和 SVM，而 PLSR 和 SVM 二者间相比，PLSR 模型的 R^2 明显高于 SVM 模型，而 RMSE 和 MAE 要低于 SVM 模型。就 RPD 指标而言，PLSR 模型和 SVM 模型的 1.5≤RPD<2.0，表明这两种模型只能对土壤水分进行粗略估计。RF 模型的 RPD 值为 3.90(3.90>3.0)，表明 RF 模型有

出色预测能力,能精准反映研究区土壤水分。综合考察各模型建模集与预测集的评价指标,3 种模型的预测性能和稳定性从高到低排序依次为 RF,PLSR 和 SVM,表明神经网络模型比线性模型和二分类模型监测反演土壤水分更理想,模型稳定性更强,精度更高。

表 2　建模因子与土壤水分相关系数

建模因子	相关系数	计算公式
归一化植被指数(NDVI)	0.45**	(NIR−R)/(NIR+R)
综合光谱响应指数(COSRI)	0.27**	(B+G)/(R+NIR)×(NIR−R)/(NIR+R)
增强植被指数(EVI)	0.48**	2.5×(NIR−R)/(NIR+6×R−7.5×B+1)
绿色叶绿素指数(GCI)	0.20*	(NIR/G)−1
差异植被指数(DVI)	0.32**	NIR−R
绿地调整植被指数(GSAVI)	0.21**	1.5×〔(NIR−G)/(NIR+G+0.5)〕
绿化率植被指数(GRVI)	0.20*	NIR/G
绿色优化土壤调节植被指数(GOSAVI)	0.22**	(NIR−G)/(NIR+G+0.16)
绿色归一化差异植被指数(GNDVI)	0.23**	(NIR−G)/(NIR+G)
绿叶指数(GLI)	0.51**	〔(G−R)+(G−B)〕/(2×G)+R+B
绿度差值植被指数(GDVI)	0.06	NIR−G
叶面积指数(LAI)	0.48**	3.618×〔2.5×(NIR−R)/(NIR+6×R−7.5×B+1)〕−0.118
红外植被指数(IPVI)	0.45**	NIR/(NIR+R)
全球植被水分指数(GVMI)	0.46**	(NIR+0.1)−(SWIR+0.02)〕/〔(NIR+0.1)+(SWIR+0.02)
绿色植被指数(GVI)	−0.36**	0.284 8×B−0.243 5×G+0.543 6×R+0.724 3×NIR+0.084×SWIR1−0.18×SWIR2
土壤改良植被指数 2(MSAVI2)	0.08	0.5×{2×(NIR+1)−SQRT〔(2×NIR+1)²−8×(NIR−R)〕}
土壤改良植被指数(MSAVI)	0.44**	〔2R+1−$\sqrt{(2NIR+1)^2-8(NIR-R)}$〕/2
修正简单比(MSR)	0.43**	(NIR/R)−1/〔SQRT(NIR/R)〕+1
植被供水指数(VSWI)	−0.42**	NDVI/LST
近红外归一化指数(NNIR)	0.36**	NIR/(NIR+R+G)
非线性指数(NLI)	0.15	(NIR²−R)/(NIR²+R)
归一化差异水分指数(NDWI)	0.44**	(NIR−SWIR)/(NIR+SWIR)
比值植被指数(RVI)	0.41**	NIR/R
重正化差分植被指数(RDVI)	0.39**	SQRT〔(NIR−R/NIR+R)×(NIR−R)〕
垂直干旱指数(PDI)	−0.42**	〔R+M(MR+1)〕/SQRT(M²+1)
优化土壤调节植被指数(OSAVI)	0.44**	(NIR−R)/(NIR+R+0.16)
红光归一化指数 NR	−0.51**	R/(NRI+R+G)
可见大气阻力指数(VARI)	0.38**	(G−R)/(G+R−B)
温度植被干旱指数(TVDI)	0.53**	(LST−LST_{min})/(LST_{max}−LST_{min})
短波红外垂直失水指数(SPSI)	−0.36**	1/$\sqrt{M^2+1}$(SWIR+MNIR)
土壤调节植被指数(SAVI)	0.44**	(1+L)×〔(NIR−G)/(NIR+G+L)〕
地表温度(T_s)	−0.41**	$T_s=K_2/\ln[K_1/B(T_s)+1]$
地形(DEM)	−0.43**	

注:① * 为显著相关,* * 为极显著相关;②样品数(n)为 160;③B,G,R,NIR,SWIR 均为 Landsat 8 波段反射率,B 为蓝波段;G 为绿波段;R 为红波段;NIR 为近红外波段;SWIR 为短波红外;④M 为光谱特征空间的基线斜率;L 为土壤调节因子,其值介于 0～1 之间;LST_{min}为地表温度最小值;LST_{max}为地表温度最大值;K_1,K_2 为 Landsat 8 卫星预设常数;$B(T_s)$为热辐射亮度。

表 3　不同模型下土壤水分反演精度

模型	建模集			预测集			
	R^2	RMSE	MAE	R^2	RMSE	MAE	RPD
PLSR	0.72	3.12	2.37	0.63	3.63	2.76	1.79
SVM	0.43	3.66	2.82	0.32	3.81	2.91	1.69
RF	0.93	1.49	1.09	0.91	1.91	1.44	3.90

2.4　不同建模方法可信度检验

分析土壤水分在 PLSR,SVM 和 RF 模型预测结果下与实测土壤水分的相关性,进一步确定预测土壤水分数据的精准性(见图 1)。通过图 1 可以看出,RF 模型(c)实测水分与预测水分 R^2 最高,为 0.91,表明

基于机器学习方法的 RF 模型反演土壤水分最优，PLSR(a)和 SVM(b)的 R^2 为 0.63，0.32，相比较 RF 模型，对土壤水分反演效果较差，但都具有一定的可信度。

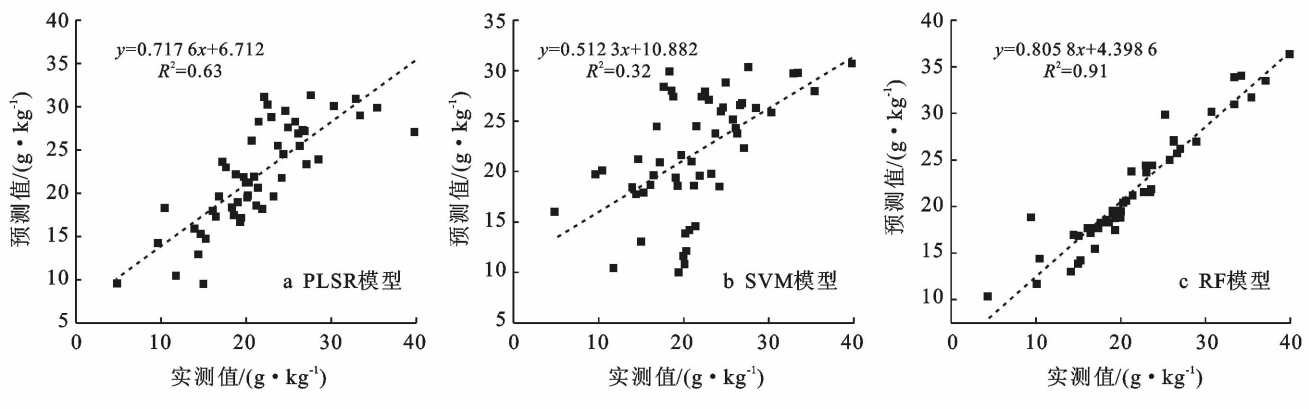


图 1 不同模型反演数据与实测数据相关性

2.5 对比不同模型的土壤水分空间分布

图 2 为最优模型(RF)反演表层(0—20 cm)荒漠土壤水分的分布情况，通过图 2 可知，土壤含水量以 0～5%和 5%～12%集中分布于研究区东北部；15%～20%或>20%大面积分布于北部、中部和南部区域。

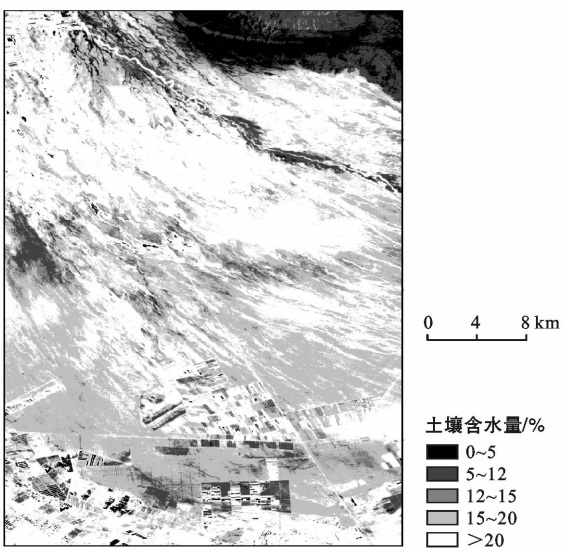


图 2 最优模型反演土壤水分分布结果

表 4 对土壤水分反演结果的分布面积进行统计，从表 4 可知，0～5%土壤含水量面积最小，为 104 km^2；>20%的分布面积最大，为 1 041.18 km^2；0～15%之间的分布面积总占比为 18.21%；而 15%～20%和>20%为 81.79%，表明研究区土壤含水量低值较少，总体呈偏高的分布特征。为更进一步探究研究区土壤水分变化特征，结合土地利用分类结果和最优模型土壤水分分布结果(图 2)可知，东北部为沙漠区，土壤水分含量集中为 0～12%，南部多为高作物覆盖农田区，土壤含水量多为 15%～20%和>20%，同时农田内存在无作物覆盖的新开垦田和撂荒地，土壤含水量为 0～15%，中部、北部为典型沙漠—绿洲过渡带区，土壤含水量多集中为 15%～20%或者>20%，表明在不同土地利用类型下，土壤水分含量存在较大差异。

表 4 土壤水分分布面积及比例

分 级	像素数	分布面积/km^2	面积比例/%
0～5%	115 553	104.00	4.78
5%～12%	171 182	154.06	7.08
12%～15%	153 564	138.21	6.35
15%～20%	820 590	738.53	33.94
>20%	1 156 871	1 041.18	47.85

3 讨 论

3.1 多因素建模可行性

本研究以 26 个光谱指数、环境数据因子和地形因子(DEM)构建 PLSR，SVM 和 RF3 种模型反演土壤水分。研究区地势北高南低，由 1 400 m 向 980 m 依次降低，而土壤含水量与地形相关系数为－0.44，达极显著相关，表明 DEM 变化对土壤水分分布有较大影响，而本研究引入 DEM 数据作为模型因子，校正因地形引起土壤水分反演精度的偏差，该结果与蔡亮红等[20]在 TVDI 基础上引入 DEM 数据校正，校正后 TVDI 能更好的反映土壤水分状况的研究结论较为一致。地表温度在土壤湿度和地、气相互循环过程中有重要作用，同时能间接反映土壤水分状况表征旱情分布，因而对地表温度数据的研究是不可或缺的内容[21-22]，本研究地表温度与土壤水分为极显著负相关，表明土壤水分随地表温度的增大而迅速减少，以

地表温度为环境数据因子,构建土壤水分综合模型,对土壤水分高精准反演有重要价值,与蔡国印等[23]以热惯量法研究土壤水分的结果较为相似。结果表明在模型中引入地形和地表温度数据对土壤水分反演模型的精度有较大提升。

3.2　土壤水分空间分布特征

通过最优模型 RF 反演空台里克地区土壤水分,结合图 2 和表 4 可知,东北部沙漠土壤含水量最低,集中为 0～12%,中部沙漠—绿洲过渡带土壤含水量最高,集中在 15～20%和>20%且分布面积为 81.79%,由于沙漠和沙漠—绿洲过渡带土壤质地相差较大,土壤结构疏密度和土壤持水能力不同,因而在不同土地类型下土壤含水量差异较大,这与姜红等[24]研究结论较为一致。相比较,沙漠—绿洲过渡带土壤含水量普遍较高,由于卫星遥感在监测中具有不确定性,空台里克是典型的高盐渍化区域,盐分表聚现象严重,地表上层形成大量盐结皮,而下层土壤水分受盐结皮封阻蒸发困难,导致下湿上干的现象,与 Peng 等[25]对空台里克地区盐渍化程度研究结论相符;南部农耕田土壤含水量呈 0～12%,15～20%和>20%的空间分布特征,而农耕田内存在大量新开垦农田和撂荒地,在不同作物覆盖区域,土壤水分的运移速率有所差异,导致变异性较高(44.57%),而含水量低值区(新开垦田、撂荒地)受人为因素干扰,土壤表层结构破坏严重,下层土壤水分受 6 月高温蒸发上移现象明显,与王思楠等[26]研究结论较为相似。以光谱指数法研究土壤水分,受植被覆盖、大气、土壤背景粗糙度等自然因素、遥感传感器等自身因素影响较为严重,后续研究需进一步探索不同影响因素对土壤水分的贡献率,筛选出一套适合研究区甚至新疆南部类似干旱区域的高精度土壤水分反演方法。

4　结 论

本研究有效结合卫星遥感和地面实测数据,通过 PLSR,SVM 和 RF 3 种方法构建多指数、多因素的综合监测模型反演土壤水分,选取最优模型对南疆阿克苏地区空台里克表层荒漠土壤(0—20 cm)水分预测分析。得到以下基本结论:

(1) 各因子与土壤水分相关性分析表明 TVDI,NR,NDVI,EVI,GOSAVI 等 26 个光谱指数,环境因子地表温度(T_s)、地形(DEM)均与土壤水分极显著相关,适合作为南疆荒漠土壤水分遥感监测的指示因子。

(2) 对比 PLSR,SVM 和 RF3 种模型各评价指标精度,RF 模型的 R^2 为 0.91,RMSE 为 1.49,同时 RPD 值高达 3.90,表明 RF 为土壤水分反演最优模型。

(3) 以最优模型 RF 反演研究区表层土壤水分,呈现沙漠 0～12%集中分布,沙漠—绿洲过渡带 15%～20%和>20%广泛分布,农田区域交错分布的土壤水分总体分布特征,其结果与土地利用分类特征和表 1 描述性统计结果较为一致。

[参 考 文 献]

[1] 包青岭,丁建丽,王敬哲,等. 基于 VIC 模型模拟的干旱区土壤水分及其时空变化特征[J]. 生态学报,2020,40(9):3048-3059.

[2] Anagnostopoulos V, Petropoulos G P, Ireland G, et al. A modernized version of a 1 D soil vegetation atmosphere transfer model for improving its future use in land surface interactions studies [J]. Environmental Modelling & Software, 2017,90:147-156.

[3] 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.

[4] Leng Pei, Song Xiaoning, Duan Sibo, et al. A practical algorithm for estimating surface soil moisture using combined optical and thermal infrared data [J]. International Journal of Applied Earth Observation and Geoinformation, 2016,52:338-348.

[5] 蔡亮红,丁建丽. 基于改进植被指数土壤水分遥感反演[J]. 干旱区地理,2017,40(6):1248-1255.

[6] 李华朋,张树清,高自强,等. MODIS 植被指数监测农业干旱的适宜性评价[J]. 光谱学与光谱分析,2013,33(3):756-761.

[7] Njoku E G, Jackson T J, Lakshmi V, et al. Soil moisture retrieval from AMSR-E [J]. Ieee Transactions on Geoscience and Remote Sensing, 2003,41(2):215-229.

[8] 李柏贞,周广胜. 干旱指标研究进展[J]. 生态学报,2014,34(5):1043-1052.

[9] De Jeu R A M, Holmes T R H, Parinussa R M, et al. A spatially coherent global soil moisture product with improved temporal resolution [J]. Journal of Hydrology, 2014,516:284-296.

[10] Pan Ming, Sahoo A K, Wood E F. Improving soil moisture retrievals from a physically-based radiative transfer model [J]. Remote Sensing of Environment, 2014,140:130-140.

[11] Price J C. On the analysis of thermal infrared imagery: The limited utility of apparent thermal inertia [J]. Remote Sensing of Environment, 1985,18(1):59-73.

[12] Ghulam A, Qin Qiming, Zhan Zhiming. Designing of the perpendicular drought index [J]. Environmental Geology, 2007,52(6):1045-1052.

(下转第 151 页)

curity based on GIS [J]. Science of the Total Environment, 2020,704:135427.
[13] 刘秀丽,郭丕斌,张勃,等. 采煤与脆弱生态复合区生态安全评价:以山西为例[J]. 干旱区研究,2018,35(3):677-685.
[14] 何珍珍,王宏卫,杨胜天,等. 塔里木盆地中北部绿洲生态安全评价[J]. 干旱区研究,2018,35(4):963-970.
[15] 李京京,吕哲敏,石小平,等. 基于地形梯度的汾河流域土地利用时空变化分析[J]. 农业工程学报,2016,32(7):230-236.
[16] 郭利刚,冯珍珍,刘庚,等. 基于物元模型的汾河流域土地生态安全评价[J]. 生态学杂志,2020,39(6):2061-2069.
[17] 山西省统计局. 山西省统计年鉴[M]. 北京:中国统计出版社,2006-2018.
[18] 中华人民共和国统计局. 中国县(市)社会经济统计年鉴[M]. 北京:中国统计出版社,2006-2018.
[19] 张凤太,王腊春,苏维词. 基于物元分析-DPSIR 概念模型的重庆土地生态安全评价[J]. 中国环境科学,2016,36(10):3126-3134.
[20] 麦丽开·艾麦提,满苏尔·沙比提,张雪琪,等. 基于 PSR-EEES 模型的叶尔羌河平原绿洲生态安全预警测度[J]. 中国农业大学学报,2020,25(2):130-141.
[21] 魏小芳,赵宇鸾,李秀彬,等. 基于"三生功能"的长江上游城市群国土空间特征及其优化[J]. 长江流域资源与环境,2019,28(5):1070-1079.
[22] 王富喜,毛爱华,李赫龙,等. 基于熵值法的山东省城镇化质量测度及空间差异分析[J]. 地理科学,2013,33(11):1323-1329.
[23] 吕广斌,廖铁军,姚秋昇,等. 基于 DPSIR-EES-TOPSIS 模型的重庆市土地生态安全评价及其时空分异[J]. 水土保持研究,2019,26(6):249-258,266.
[24] 李俊翰,高明秀. 滨州市生态系统服务价值与生态风险时空演变及其关联性[J]. 生态学报,2019,39(21):7815-7828.
[25] 邓楚雄,钟小龙,谢炳庚,等. 洞庭湖区土地生态系统的服务价值时空变化[J]. 地理研究,2019,38(4):844-855.
[26] 王劲峰,徐成东. 地理探测器:原理与展望[J]. 地理学报,2017,72(1):116-134.
[27] 孙泽乾,杨晓霞,曾于珈. 基于地理探测器的重庆市旅游发展水平空间分异影响因子研究[J]. 西南师范大学学报(自然科学版),2019,44(4):57-63.
[28] 马勇,童昀,任洁,等. 公众参与型环境规制的时空格局及驱动因子研究:以长江经济带为例[J]. 地理科学,2018,38(11):1799-1808.

(上接第 131 页)

[13] 张学艺,李剑萍,秦其明,等. 几种干旱监测模型在宁夏的对比应用[J]. 农业工程学报,2009,25(8):18-23.
[14] 姚云军,秦其明,赵少华,等. 基于 MODIS 短波红外光谱特征的土壤含水量反演[J]. 红外与毫米波学报,2011,30(1):9-14.
[15] 孙灏,陈云浩,孙洪泉. 典型农业干旱遥感监测指数的比较及分类体系[J]. 农业工程学报,2012,28(14):147-154.
[16] 葛少青,张剑,孙文,等. 3 种干旱指数在干旱区沼泽湿地土壤水分遥感反演中的应用[J]. 生态学报,2018,38(7):2299-2307.
[17] Du Lingtong, Tian Qingjiu, Yu Tao, et al. A comprehensive drought monitoring method integrating MODIS and TRMM data [J]. International Journal of Applied Earth Observation and Geoinformation, 2013,23(1):245-253.
[18] 彭杰. 荒漠土壤盐渍化遥感监测与开垦方案分析[D]. 浙江 杭州:浙江大学,2019.
[19] 刘新路,彭杰,冯春晖,等. 基于电磁感应仪数据的南疆棉田土壤电导率反演模型研究[J]. 土壤学报,2020,57(3):646-655.
[20] 蔡亮红,丁建丽,魏阳. 基于多源数据的土壤水分反演及空间分异格局研究[J]. 土壤学报,2017,54(5):1057-1067.
[21] 刘宁. 不同土地利用方式下黄河三角洲土壤特性空间变异研究[D]. 山东 泰安:山东农业大学,2007.
[22] 董雪,田静,吴骅,等. 考虑土壤水分影响的比辐射率方法在地表温度反演中的应用[J]. 资源科学,2017,39(8):1592-1604.
[23] 蔡国印. 基于 MODIS 数据的地表温度、热惯量反演研究及其在土壤水分、地气间热交换方面的应用[D]. 北京:中国科学院研究生院(遥感应用研究所),2006.
[24] 姜红,玉素甫江·如素力,拜合提尼沙·阿不都克日木,等. 基于支持向量机回归算法的土壤水分光学与微波遥感协同反演[J]. 地理与地理信息科学,2017,33(6):30-36.
[25] Peng Jie, Biswas A, Jiang Qingsong, et al. Estimating soil salinity from remote sensing and terrain data in southern Xinjiang Province, China [J]. Geoderma, 2019,337:1309-1319.
[26] 王思楠,李瑞平,李夏子. 基于综合干旱指数的毛乌素沙地腹部土壤水分反演及分布[J]. 农业工程学报,2019,35(13):113-121.