

DOI: 10.13758/j.cnki.tr.2022.03.025

蒙莉娜, 丁建丽, 张振华. 耦合植被物候的新疆绿洲土壤盐分信息挖掘. 土壤, 2022, 54(3): 629–636.

## 耦合植被物候的新疆绿洲土壤盐分信息挖掘<sup>①</sup>

蒙莉娜<sup>1,2</sup>, 丁建丽<sup>1,2\*</sup>, 张振华<sup>1,2</sup>

(1 新疆大学资源与环境科学学院, 乌鲁木齐 830046; 2 新疆大学绿洲生态教育部重点实验室, 乌鲁木齐 830046)

**摘要:** 随着土壤环境问题涉及的尺度日趋增大, 小区域斑块化盐渍化信息的提取难以了解土壤环境总体的变化趋势。本文以野外监测的南北疆典型绿洲区域——渭库绿洲和艾比湖流域为分析靶区, 通过实测数据建立土壤-环境关系, 并通过 MODIS EVI 数据反演得到植被物候特征, 耦合植被物候、植被指数、盐度指数、地表温度和地形参数作为随机森林(random forest, RF)模型的输入因子, 预测新疆绿洲区域土壤盐分含量信息并绘制土壤盐分空间分布图。结果表明: 通过深入挖掘植被物候信息, 物候参数在预测土壤盐分方面具有较高的相对重要性, 代表生物积累量的 LSI 和 SSI 参数表征土壤盐渍化的能力较强, 优于其他几个物候参数。耦合物候参数后土壤盐分信息预测精度明显提高, 决定系数  $R^2$  从 0.53 提升到 0.61。经模型反复迭代进一步筛选出适合研究区的 23 个环境参数, 大幅提升了预测精度( $R^2 = 0.73$ , RMSE = 5.19, MAE = 3.59)。从得到的盐渍化空间分布特征来看, 新疆绿洲大部分区域分布的是非盐渍化土和轻盐渍化土, 且普遍分布在绿洲内部, 中度及以上盐渍化土多分布在绿洲外围, 总体盐渍化水平依次为: 伊犁平原 < 北疆绿洲 < 南疆绿洲 < 东疆绿洲。

**关键词:** 土壤盐渍化; 植被物候; 机器学习; 数字土壤制图

**中图分类号:** S153      **文献标志码:** A

### Soil Salinization Mining in Xinjiang Based on Vegetation Phenology

MENG Lina<sup>1,2</sup>, DING Jianli<sup>1,2\*</sup>, ZHANG Zhenhua<sup>1,2</sup>

(1 College of Resources and Environmental Sciences, Xinjiang University, Urumqi 830046, China; 2 Laboratory of Oasis Ecology Under Ministry of Education, Xinjiang University, Urumqi 830046, China)

**Abstract:** With the increasing scale of soil environmental problems, it is difficult to understand the overall changing trends of soil environment by extracting the information on patchy salinization in small areas. In this study, the Weigan-Kuqa River delta oasis and Ebinur River basin were taken as the research area, topsoil samples (0–10 cm) were collected from 209 of typical sites based on the representative grade sampling method, the electric conductivity (EC) and salt contents of the samples were determined, and then the relationship between soil EC and various environment variables were established. The characteristics of vegetation phenology were inverted by MODIS EVI data and were coupled with vegetation index, salinity index, surface temperature and topographic parameters as input factors of the random forest (RF) model, and then the information of soil salt content in the oasis regions in Xinjiang was deduced and the spatial distribution maps of soil salt content was drawn. The results show that digging deeper into the information of vegetation phenology promote significantly the importance of the phenological parameters in predicting soil salinity, the large seasonal integral (LSI) and small seasonal integral (SSI) representing bioaccumulation are better in characterizing soil salinization than other phenological parameters. The accuracy of saltness prediction is significantly improved after coupling with the phenological parameters with the coefficient of determination  $R^2$  increased from 0.53 to 0.61. 23 environmental parameters suitable for the study area are screened out after the iterative selection, which significantly improve the prediction accuracy ( $R^2 = 0.73$ , RMSE = 5.19, MAE = 3.59). According to the spatial distribution of salinization, non-salinized and lightly salinized soils are distributed in most areas of the Xinjiang oases, and they are generally distributed in the interior of the oases, while the moderate and above salinized soils are mostly distributed in the periphery of the oases. The total salinization level is in order of Yili plain < Northern

①基金项目: 国家自然科学基金项目(41961059、41261090、41771470)资助。

\* 通讯作者(watarid@xju.edu.cn)

作者简介: 蒙莉娜(1995—), 女, 新疆石河子人, 硕士, 主要从事干旱区环境演变遥感研究。E-mail: menglina\_xj@163.com

xinjiang oasis < Southern Xinjiang oasis < Eastern Xinjiang oasis.

**Key words:** Soil salinization; Vegetation phenology; Machine learning; Digital soil mapping

土壤盐渍化是严重威胁农业生产的全球性问题,我国受盐渍化影响的土地面积( $3.6 \times 10^7 \text{ hm}^2$ )占全国可利用土地资源总量的 4.88%<sup>[1]</sup>。绿洲区域作为新疆重要的粮食储备区,受盐渍化的影响,制约了农业生产的可持续发展。随着现代数学发展以及遥感数据类型日益广泛和易于获取,一种新兴、高效的技术方法——数字化制图(digital soil mapping, DSM)在土壤空间分布制图研究中得以应用,并得到了飞速发展<sup>[2]</sup>。目前主流的数字土壤制图方法主要有两种,第一种是基于要素相关性,包括统计学方法、机器学习方法和基于专家知识的方法,第二种是基于空间自相关,包括全局自相关和局域自相关分析,其中克里格插值法是应用最为广泛的一种方法。

构建反演模型量化的区域土壤盐分动态监测,有助于盐渍化区域制定改良措施和优化政策,已有大量研究尝试不同方法对土壤盐渍化进行反演。Vermeulen 和 van Niekerk<sup>[3]</sup>运用 k 邻近、支持向量机和随机森林等方法,结合 DEM 预测土壤盐害影响范围,证明高程数据在监测灌区盐分积累方面具有很大潜力。Zhang 等<sup>[4]</sup>运用 MODIS 时间序列植被数据,通过内插法模拟年际植被指数变化,结合季度植被指数对黄河三角洲的土壤盐分进行数字制图,结果表明结合长时间序列植被信息较大地提升了预测精度。Wang 等<sup>[5]</sup>结合环境变量预测新疆渭库绿洲土壤盐度信息比较了 13 种算法及其性能,证明随机森林方法具有较好效果。

以往大多数研究证明植被指数在盐分预测方面表现良好<sup>[6-7]</sup>,但多数研究仅针对特定时期进行,MODIS 长时间序列数据能满足长期观测的条件,而植被物候作为连续的环境参量在之前的盐分制图工作中被忽略或低估<sup>[8]</sup>。此外对于土壤盐渍化监测,大都集中在田间尺度和流域尺度<sup>[9-11]</sup>。随着土壤环境问题涉及尺度愈发扩大,当前土地规划、生态文明及粮食安全需要决策者从宏观角度进行把握<sup>[12]</sup>。因此,本研究以新疆绿洲区域作为研究靶区,借助随机森林(random forest, RF)模型,结合研究区域的典型采样点,耦合物候参数和地表参数反演土壤盐分信息,通过绘制新疆绿洲地区土壤盐渍化空间分布图,深入理解绿洲区域土壤盐渍化的形成及分布,以期物候信息在大尺度的盐渍化特征监测应用中提供参考。

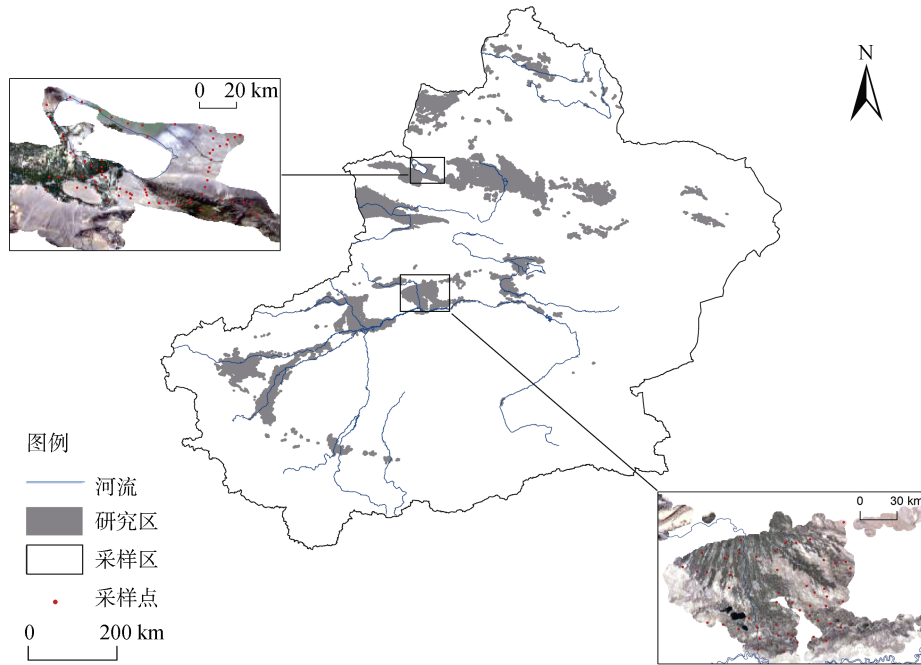
## 1 材料与方法

### 1.1 研究区概况

新疆维吾尔自治区( $73^\circ 40' \sim 96^\circ 18' \text{ E}$ ,  $34^\circ 25' \sim 48^\circ 10' \text{ N}$ ),深居欧亚大陆腹地,远离海洋,四周为高山环抱,海洋气流通达性不强,形成明显的温带大陆性气候。年均降水量 50 ~ 200 mm,蒸发量高达 1 500 ~ 3 000 mm,蒸发强烈造成地表土壤无机盐含量升高,形成大面积盐碱土。除气候因素外,人为因素加速土壤次生盐渍化,导致约 32% 的耕地出现次生盐渍化。依据第二次新疆地区土壤普查结果,新疆土壤可划分为 7 个土纲及 32 个土类,其中风沙土、棕漠土、棕钙土、寒冻土及石质土占比较大。新疆主要植被类型为旱生灌木、小半灌木和乔木等,其稀疏根系发达且较耐干旱,农田植被主要由玉米、小麦和棉花等构成,大多分布在绿洲内部水资源丰富的区域。新疆绿洲面积约占总面积的 5% ~ 8%,集中了新疆 90% 的耕地、人口和国内生产总值。由于水资源的不合理利用及不科学的耕作方式,致使绿洲生态环境恶化,盐渍化现象依然普遍,严重影响了绿洲的可持续发展。

### 1.2 土壤盐分数据采集

根据土壤盐渍化状况的不同,本研究在新疆地区选择两个不同的研究靶区采集土样,分别是位于天山以南的渭干河-库车河绿洲(简称渭库绿洲)和天山以北的艾比湖流域。基于相似性的土壤制图方法无需要求采样点的数量或分布,它根据每个地点与采样点的环境相似性来预测土壤盐度值。在这两个地区选择代表性采样区,保证其包括不同地理环境和典型气候区,融合两个地区的实测数据,使之具备与新疆其他区域近似的地理和气候环境。从研究区中心到外围按均匀分层采样,采用“S”型布点避开城镇用地,选择大量的典型覆被类型,例如绿洲荒漠交错带、盐渍地、沙地、荒漠进行采样,尽量保证样点均匀地分布在样区内部,如图 1 所示。采用网格法布设样方,在地势较为平坦之处选取土壤样方( $30 \text{ m} \times 30 \text{ m}$ ),保持土壤性质相对一致,环境因素相似。采用五点法进行表层土壤(0 ~ 10 cm)样品的采集,将测试的数据求平均值作为本样点的实际观测值。采集的土样自然风干清除石块和杂草后,研磨过 2 mm 孔筛,制备水土质量比为 5 : 1 的浸提液,采用 Cond7310 土壤测试仪测



(该图基于国家测绘地理信息局标准地图服务网站下载的审图号为GS(2019)3266号的标准地图制作,底图无修改;下同)

图1 采样点分布图

Fig.1 Design drawing of sampling sites

定土壤上清液电导率,采用残渣烘干法测定土壤全盐量。采样时间为2019年7月,剔除无效样点,选取209个采样点(渭库绿洲113个,艾比湖96个)进行分析。

### 1.3 地表覆盖数据获取

本研究将新疆绿洲区域作为研究靶区,采用王智等<sup>[13]</sup>的分区方法,首先按照新疆行政区边界将全疆划分为北疆、南疆、东疆和伊犁4个大区。其次,结合DEM数据在四大分区基础上对北疆和伊犁地区以海拔小于1 300 m为基础划分出平原区域,而南疆和东疆地区则以海拔小于2 000 m为阈值同样划分出平原地区。最后,运用2009年MODIS 500分辨率的土地利用产品数据MCD12Q1,提取出平原地区的耕地和城镇居民用地,以深入了解绿洲边缘与部分交错带区域的土壤盐分分布,在消除破碎斑块后,对提取出的两类土地利用数据向外做3 km缓冲区划为绿洲区域。

### 1.4 地表参数获取

土壤盐渍化在不同尺度下受土壤、气候、地形和生物等因素的综合影响,对于不同研究区,影响土壤形成发展最主要的环境因子可能有所不同<sup>[12,14]</sup>。本文选择红波段(R)、蓝波段(B)和近红外波段(NIR)作为反射率参数,并以蓝、红、近红外波段组合出23种地表遥感参数,见表1。

### 1.5 物候参数计算

在高植被生物量水平下NDVI容易受到土壤背景变化、噪声、气溶胶消光系数和饱和度的影响<sup>[23]</sup>,因此选取EVI进行植被物候参数的提取<sup>[15]</sup>。本文所选用的MODIS EVI植被指数产品MOD13A2数据虽然已通过16 d的MVC预处理,但仍不可避免大气和噪声等的影响,直接提取对物候参数影响较大。通过TIMESAT 3.2的Data plotting模块对比Asymmetric Gaussian、Double Logistic和Savitzky-Golay三种平滑拟合方法,发现Savitzky-Golay滤波法对时间序列包络数据保留更加完整且对年生长季原始植被指数拟合更精准,因此,选用Savitzky-Golay滤波法重建物候曲线。

新疆地区横跨不同气候带且植被类型多样,不同土地覆盖类别其物候曲线也会有所差异。采用阈值法通过对植被指数设定阈值条件来确定植被物候的始末期指标EOS和SOS,该方法可以忽略地表的覆盖类型和人为活动的影响,根据研究区植被不同的生理特点,形象地描述植被的生长发育过程<sup>[24-25]</sup>。根据相关研究<sup>[26]</sup>,本文将提取生长季始末日期的系数设为30%,其计算公式如下:

$$EVI(SOS) = (EVI_{max} - EVI_{min1}) \times 30\% \quad (1)$$

$$EVI(EOS) = (EVI_{max} - EVI_{min2}) \times 30\% \quad (2)$$

式中:EVI(SOS)和EVI(EOS)分别表示最后确定的阈

值,  $EVI_{max}$  为植被指数最大值,  $EVI_{min1}$  为植被指数曲线上阶段最小值,  $EVI_{min2}$  为曲线下阶段最小值。

根据平滑和参数设置, 用 TIMESAT 软件计算出 SOS、EOS 和 LEN 等 10 个物候参量, 其代表的生态学意义见表 2。

表 1 地表参量定义与计算  
Table 1 Definition and calculation of surface parameters

参数类别	参选变量	公式	参考文献
植被指数	差值植被指数(DVI)	$NIR-R$	[15]
	增强型植被指数(EVI)	$2.5[(NIR-R)/(NIR+6\times R-7.5\times R+1)]$	[16]
	广义差值植被指数(GDVI)	$(NIR^2-R^2)/(NIR^2+R^2)$	[17]
	土壤调节植被指数(SAVI)	$[(NIR-R)/(NIR+R+L)]\times(1+L)$	[18]
	改进型土壤调节植被指数(MSAVI)	$[2NIR+1-((2NIR+1)^2-8(NIR-R))^{0.5}]/2$	[15]
	归一化植被指数(NDVI)	$(NIR-R)/(NIR+R)$	[15]
盐分指数	归一化盐分指数(NDSI)	$(R-NIR)/(R+NIR)$	[19]
	盐分指数(S1)	$B/R$	[19]
	盐分指数(S2)	$(B-R)/(B+R)$	[19]
	盐分指数(SI)	$(B\times R)^{0.5}$	[19]
	盐分指数(SIT)	$(R/NIR)\times 100$	[19]
	亮度指数(BI)	$(R^2+NIR^2)^{0.5}$	[20]
特征空间	温度植被干旱指数(TVDI)	$(T_s-T_{min})/(T_{max}-T_{min})$	[21]
地表温度 波段	地表温度(LST)		[22]
	红波段(R)		
	蓝波段(B)		
地形指数	近红外波段(NIR)		
	高程(ELE)		[15]
	河道基准线(CNBL)		[16]
	相对坡度位置(RSP)		[17]
	谷底平整度(MRRTF)		[18]
	总集水面积(TCA)		[15]
	谷深(VD)		[15]

注: B: 蓝波段; R: 红波段; NIR: 近红外波段; L 是气溶胶和大气相关参数;  $T_{max}=a+b\times NDVI$ ;  $T_{min}=c+d\times NDVI$ (a、b、c、d 分别为干湿边的拟合系数)。

表 2 物候参量定义及生态学释义  
Table 2 Definition and ecological interpretation of phenological parameters

物候参量	定义	生态学定义
生长季开始时间(SOS)	生长季开始日期	植被开始返青绿化时间
生长季结束时间(EOS)	生长季结束日期	植被开始泛黄凋零时间
生长季长度(LEN)	生长季伊始与截止所经时长	植被生长季所需时长
基准值(BV)	曲线左右两侧最小值的均值	生物量基准值
峰值(PV)	生长期植被指数的最大值	光合作用最高水平和最大值到达的时间值
生长季中值时间(MOS)	曲线左缘增加与右缘回溯之间的均值日期点	植被生长期中段时间
左导数(LD)	从左缘最低点到最高值直线的斜率	植被生长速率
右导数(RD)	从右缘最高值回溯到最低点直线的斜率	植被衰败速率
生长季大生物累积量(LSI)	从生长季开始到结束的整季拟合函数曲线的积分	生长季光合作用累计值
生长季小生物累积量(SSI)	从生长季节开始到结束整季生长拟合函数曲线与基准值的积分	

## 1.6 建模方法及模型评价指标

随机森林(RF)方法是一种利用多棵决策树对样

本进行训练并预测的机器学习算法, 该算法的优势在于具备非线性挖掘能力, 数据的分布不需要符合任何

假设<sup>[27]</sup>。RF 模型需要用户定义 3 个参数：终端节点树(ntree)、作为每棵树的预测特征的特征个数(mtry), 以及每个终端节点的最小值<sup>[28]</sup>。ntree 的默认值为 500, 但一般认为它不足以产生可靠的结果<sup>[29]</sup>, 本研究中通过遍历确定 ntree 的最优值为 1 000。在使用 mtry 确定单一树与模型中其他树之间的相关性时, 随着 mtry 值的增加每棵树和树之间的相关强度逐渐增加<sup>[30]</sup>。本研究以预测误差最小化为目标函数, 通过遍历比较确定 mtry 的最优值为 9, 终端节点大小为 5。同时采用较常用的交叉验证法将样点数据随机分为 5 个子集, 其中 4/5 作为训练集, 1/5 用于模型验证。选择决定系数( $R^2$ )、均方根误差(RMSE)和平均绝对误差(MAE)来评估预测模型性能, 经过反复迭代获得最终的精度。 $R^2$  范围在 0 ~ 1 之间, 越接近 1, 模型的稳定性越高, RMSE 和 MAE 越小, 模型的预测能力越强、精度也越高。

$$R^2 = 1 - \frac{\sum_{i=1}^n (V_{pi} - V_{oi})^2}{\sum_{i=1}^n (\bar{V}_{oi} - V_{oi})^2} \quad (3)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (V_{oi} - V_{pi})^2} \quad (4)$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |V_{oi} - V_{pi}| \quad (5)$$

式中:  $V_{oi}$  为样本点  $i$  处的实测值;  $V_{pi}$  为点  $i$  处的预测值;  $n$  为总样点个数。

## 2 结果

### 2.1 物候参数提取结果

对绿洲区域的 10 种物候参数进行提取, 根据其对于盐渍化预测的相对重要性, 仅对生长季大生物累积量(LSI)和生长季小生物累积量(SSI)2 个物候参数进行展示。从图 2 可以看出, 两类物候参量分布情况高度一致, 均呈现由绿洲外围向绿洲内部递增状态, 从侧面可以反映出土壤盐渍化信息, 即在绿洲外围盐渍化情况一般较重, 植被长势不佳, 生物量较低, 值域在 1.32 ~ 3.14。对比发现 LSI 和 SSI 在北疆区域分布几乎一致, 整个伊犁地区生物累积量最大集中在 4.58 ~ 11.04, 说明其植被长势良好, 而在南疆绿洲区域存在 LSI 明显高于 SSI 的情况, 说明南疆绿洲区域大生物累积量高而小生物累积量较低, 这可能与该地区本身的气候干燥有关。纵观 LSI 可以发现, 其在北疆绿洲区域和南疆绿洲区域的分布情况基本一致, 大部分区域集中于 4.61 ~ 9.07。一般正常生长的植被生物量较高, 而受盐渍化或病虫害影响的植被长势受限, 积累生物量会呈现较低水平, 因此, LSI 和 SSI 参数在监测土壤盐渍化水平中具有极高的应用潜力。

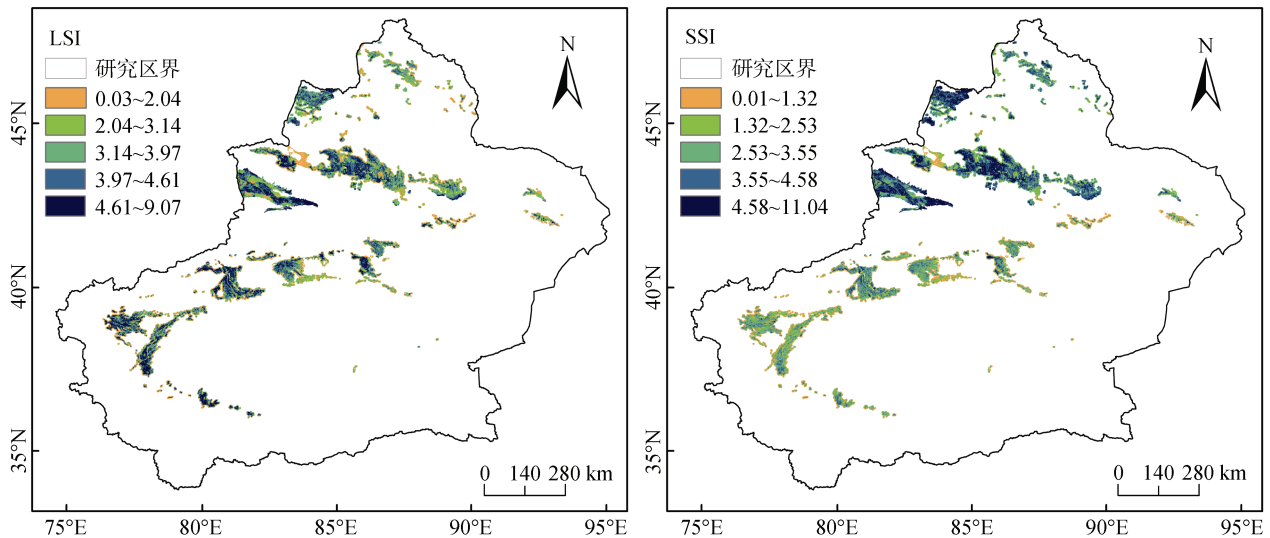


图 2 生物累积量分布图  
Fig. 2 Biomass accumulation patterns

### 2.2 模型精度验证及变量遴选

由表 3 可得, 耦合物候参数的预测模型精度( $R^2 = 0.61$ ,  $RMSE = 9.14$ ,  $MAE = 7.76$ )高于仅用陆表参数的预测模型( $R^2 = 0.53$ ,  $RMSE = 10.75$ ,  $MAE = 9.44$ ),

说明在融合土壤湿度、植被和地形等因素的同时, 加入受不同程度盐渍化影响的植被长势情况所对应的物候参数, 能最大化地提升模型预测精度。将物候信息这类长时间序列反映年际变化的数据源加入到环

表 3 模型精度验证结果  
Table 3 Accuracy verification results of RF model

模型参数	R <sup>2</sup>	RMSE	MAE
耦合物候参数	0.61	9.14	7.76
无物候参数	0.53	10.75	9.44
筛选结果参数	0.73	5.19	3.59

境因子中,能更好地阐明复杂景观异质性条件下土壤盐渍化与环境之间的响应关系。但其均方根误差 RMSE 相差不大,可能是由于这 33 个参数之间存在多重共线性。Keskin 等<sup>[31]</sup>认为,当变量增加到一定程度时其精度将不再继续提升,继续增加变量会与仅有少量变量一样使模型运作变差甚至出现精度“负增长”的情况。因此,有必要对参数进行筛选来确定更合理的预测变量,本研究结合变量相对重要性,对比反演精度,经模型反复筛选最终选出 23 个取得较好预测结果的参数。

通过 RF 模型迭代 100 次获得环境协变量的相对

重要性(relative importance, RI),并绘制重要性玫瑰图。如图 3 所示,在全部参量(图 3A)中,S1、S2、TCA、MOS 和 SOS 5 个参数重要性较高,说明盐分指数、典型物候期和地形特征是监测土壤盐渍化的重要指标。DVI、MSAVI 和 SAVI 等植被指数在相对重要性的排名中都接近中间和偏后的位置,说明单纯的植被覆盖指数并不能很好地作为反映土壤盐渍化的指标,因为在盐渍化较重区域,盐生植被一般长势较好。在仅地表参量(图 3B)中,植被指数相对重要性有了一定提升,遥感反射率参数也有一定重要性,进一步说明地表可见光的反射率能有效地反映土壤盐渍化的具体分布情况。在筛选参数(图 3C)中,作为物候参数代表生物光合作用的 SSI 和 LSI 起重要作用,因为不同类型植被对土壤盐分含量响应程度不同,最终会导致生物积累量的不同,所以 LSI 和 SSI 对土壤盐渍化的表征能力总体优于其他几种物候参数。

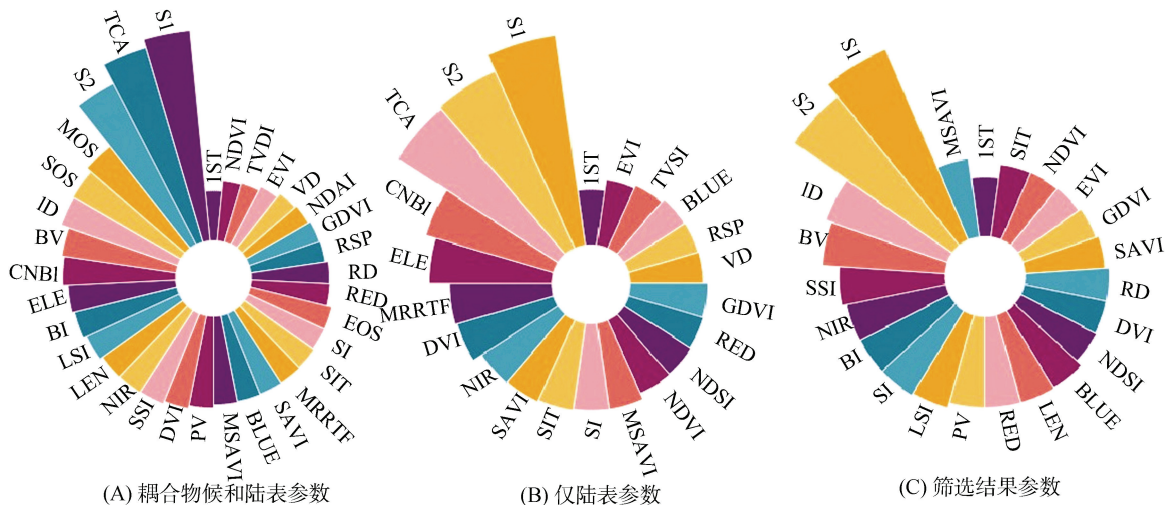


图 3 环境协变量相对重要性(%)  
Fig. 3 Relative importance of environmental covariables

### 2.3 盐渍化空间分布特征

将筛选出的 23 个参数作为预测变量输入 RF 模型,根据乔木等<sup>[32]</sup>拟定的新疆盐渍土含盐量划分标准(非盐渍化 < 8 g/kg; 8 g/kg < 轻盐渍化 < 10 g/kg; 10 g/kg < 中盐渍化 < 15 g/kg; 15 g/kg < 重盐渍化 < 20 g/kg; 极重度盐渍化 > 20 g/kg),得到新疆绿洲区域 2019 年土壤盐渍化空间分布如图 4 所示。总体上看,新疆绿洲大部分区域分布的是非盐渍化土和轻盐渍化土,普遍分布在绿洲内部,含盐量大多在 8 g/kg 以下,中度及以上盐渍化土多分布在绿洲外围,尤其是东疆绿洲和南疆绿洲表现较为明显,即含盐量都在 10 g/kg 以上。位于南疆,在塔里木河流

经的区域周边普遍是盐渍化土,特别是渭库南面区域为重度盐渍化土,主要是由于该区地下水位较高,加之蒸发作用强烈,盐分易累积于地表形成盐渍化土,这与野外实地调研结果一致。而在伊犁地区沿伊犁河的绝大部分区域都是非盐渍化土,据相关研究<sup>[33]</sup>,伊犁河沿岸有大量地表水,在自然的洗盐作用下降低了土壤含盐量。

### 3 讨论

不同盐度的土壤具备不同光谱特征,土壤盐分随土壤样品光谱反射率增加,这使得利用多光谱波段及其衍生的光谱指数进行表土盐分监测和制图成为可

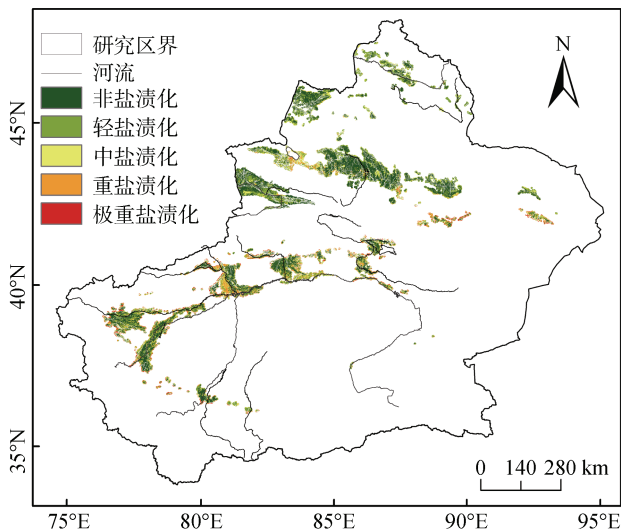


图4 土壤盐分含量(0~10 cm)空间分布预测  
Fig. 4 Spatial distribution prediction of soil salt content

能。盐度指数与植被指数在盐分预测方面表现良好,然而,每种植被类型的生境和耐受性有很大差异,为了监测土壤盐渍化,必须考虑盐生植被的光谱反射率,特别是在严重盐渍化的干旱和半干旱地区<sup>[34]</sup>。本研究除了加入常用的植被指数,还挖掘了物候参数对土壤盐分的预测能力,发现代表生物光合作用的SSI和LSI在土壤盐分预测中起重要作用。此外,新疆地区气候干旱,蒸发强烈,土壤中的可溶性盐随水分运输积聚表层,土壤水分与盐分关系密切,而温度影响水热分布从而影响着土壤盐分水平梯度变化,反演地表温度LST和温度植被干旱指数TVDI对盐分的预测具有较大潜力。Masoud等<sup>[35]</sup>利用Landsat影像反演地表温度,预测土壤盐渍化取得较好效果并推理土壤盐分未来可能的变化。

研究区地质环境复杂,气候环境跨度较大,导致土壤盐碱类型多样。研究中通过采样区域的物候与非物候特征得出的建模关系,运用到全疆其他区域时,通过预测发现其含盐量的总体分布特征符合实际,说明全疆的绿洲区域在总体的环境变量上存在某些内在与外在的相似性,使得通过小区域来推广到大尺度区域的研究具有一定的可行性和现实意义。考虑到具体环境的差异性,今后可以从以下两方面进行深入研究:

1) 新疆地区的物候验证数据杂乱且难以获取,对植被物候节点始末的预测也略显不足。因此,在今后的野外采样过程中应该对典型样区的植被物候数据予以阶段性记录并观测,以寻找最佳物候始末期,提高盐渍化推理制图精度。

2) 若能研究区划分成具有不同等级的区域系统,针对具体环境筛选合适的变量,将进一步提高大尺度土壤盐分信息的反演精度。探讨不同分区的土壤盐分信息,也有助于决策者因地制宜地进行生态安全的保护和农业生产的建设。

## 4 结论

本文基于机器学习算法,利用少量典型样点,结合关键环境协变量预测新疆绿洲盐渍化土壤的空间分布情况,得出全疆绿洲区域盐渍化水平是伊犁平原<北疆绿洲<南疆绿洲<东疆绿洲。充分挖掘不同物候参数对土壤盐渍化程度的响应,发现代表生物积累量的LSI和SSI能较好地反映土壤盐渍化信息。在耦合物候参数后模型对土壤盐渍化的预测精度明显提升,决定系数从0.53提升到0.61,经模型反复迭代,最终遴选出适合该研究区的23个环境协变量,模型决定系数提升至0.73。

## 参考文献:

- [1] Zhang J F. Coastal saline soil rehabilitation and utilization based on forestry approaches in China[M]. Berlin, Heidelberg: Springer Berlin Heidelberg, 2014.
- [2] 朱阿兴, 杨琳, 樊乃卿, 等. 数字土壤制图研究综述与展望[J]. 地理科学进展, 2018, 37(1): 66-78.
- [3] Vermeulen D, van Niekerk A. Machine learning performance for predicting soil salinity using different combinations of geomorphometric covariates[J]. Geoderma, 2017, 299: 1-12.
- [4] Zhang T T, Qi J G, Gao Y, et al. Detecting soil salinity with MODIS time series VI data[J]. Ecological Indicators, 2015, 52: 480-489.
- [5] Wang F, Shi Z, Biswas A, et al. Multi-algorithm comparison for predicting soil salinity[J]. Geoderma, 2020, 365: 114211.
- [6] Brunner P, Li H T, Kinzelbach W, et al. Generating soil electrical conductivity maps at regional level by integrating measurements on the ground and remote sensing data[J]. International Journal of Remote Sensing, 2007, 28(15): 3341-3361.
- [7] Zhang C, Lu D S, Chen X, et al. The spatiotemporal patterns of vegetation coverage and biomass of the temperate deserts in Central Asia and their relationships with climate controls[J]. Remote Sensing of Environment, 2016, 175: 271-281.
- [8] 何宝忠, 丁建丽, 刘博华, 等. 渭库绿洲土壤盐渍化时空变化特征[J]. 林业科学, 2019, 55(9): 185-196.
- [9] Peng J, Biswas A, Jiang Q S, et al. Estimating soil salinity from remote sensing and terrain data in southern Xinjiang Province, China[J]. Geoderma, 2019, 337: 1309-1319.

- [10] 蒙莉娜, 丁建丽, 王敬哲, 等. 基于环境变量的渭干河-库车河绿洲土壤盐分空间分布[J]. 农业工程学报, 2020, 36(1): 175–181.
- [11] 张振华, 丁建丽, 王敬哲, 等. 集成土壤-环境关系与机器学习的干旱区土壤属性数字制图[J]. 中国农业科学, 2020, 53(3): 563–573.
- [12] 丁建丽, 王飞. 干旱区大尺度土壤盐度信息环境建模——以新疆天山南北中低海拔冲积平原为例[J]. 地理学报, 2017, 72(1): 64–78.
- [13] 王智, 师庆三, 王涛, 等. 1982—2006 年新疆山地-绿洲-荒漠系统植被覆盖变化时空特征[J]. 自然资源学报, 2011, 26(4): 609–618.
- [14] 杨琳, 朱阿兴, 李宝林, 等. 应用模糊 c 均值聚类获取土壤制图所需土壤-环境关系知识的方法研究[J]. 土壤学报, 2007, 44(5): 784–791.
- [15] Mahmoudabadi E, Karimi A, Haghnia G H, et al. Digital soil mapping using remote sensing indices, terrain attributes, and vegetation features in the rangelands of northeastern Iran[J]. Environmental Monitoring and Assessment, 2017, 189(10): 500.
- [16] Lobell D B, Lesch S M, Corwin D L, et al. Regional-scale assessment of soil salinity in the red river valley using multi-year MODIS EVI and NDVI[J]. Journal of Environmental Quality, 2010, 39(1): 35–41.
- [17] Wu W C, Mhaimeed A S, Al-Shafie W M, et al. Mapping soil salinity changes using remote sensing in Central Iraq[J]. Geoderma Regional, 2014, 2/3: 21–31.
- [18] Huete A, Didan K, Miura T, et al. Overview of the radiometric and biophysical performance of the MODIS vegetation indices[J]. Remote Sensing of Environment, 2002, 83(1/2): 195–213.
- [19] Allbed A, Kumar L, Aldakheel Y Y. Assessing soil salinity using soil salinity and vegetation indices derived from IKONOS high-spatial resolution imageries: Applications in a date palm dominated region[J]. Geoderma, 2014, 230/231: 1–8.
- [20] Taghizadeh-Mehrjardi R, Minasny B, Sarmadian F, et al. Digital mapping of soil salinity in Ardakan region, central Iran[J]. Geoderma, 2014, 213: 15–28.
- [21] 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.
- [22] Qin Z, Karnieli A, Berliner P. A mono-window algorithm for retrieving land surface temperature from Landsat TM data and its application to the Israel-Egypt border region[J]. International Journal of Remote Sensing, 2001, 22(18): 3719–3746.
- [23] Lu L L, Kuenzer C, Wang C Z, et al. Evaluation of three MODIS-derived vegetation index time series for dryland vegetation dynamics monitoring[J]. Remote Sensing, 2015, 7(6): 7597–7614.
- [24] 胡召玲, 戴慧, 侯飞, 等. 中国东北城乡植被物候时空变化及其对地表温度的响应[J]. 生态学报, 2020, 40(12): 4137–4145.
- [25] 马新萍, 白红英, 贺映娜, 等. 基于 NDVI 的秦岭山地植被遥感物候及其与气温的响应关系——以陕西境内为例[J]. 地理科学, 2015, 35(12): 1616–1621.
- [26] 何宝忠, 丁建丽, 王飞, 等. 基于物候特征的盐渍化信息数据挖掘研究[J]. 生态学报, 2017, 37(9): 3133–3148.
- [27] 王飞, 杨胜天, 丁建丽, 等. 环境敏感变量优选及机器学习算法预测绿洲土壤盐分[J]. 农业工程学报, 2018, 34(22): 102–110.
- [28] Friedman J H, Meulman J J. Multiple additive regression trees with application in epidemiology[J]. Statistics in Medicine, 2003, 22(9): 1365–1381.
- [29] Yang R M, Zhang G L, Liu F, et al. Comparison of boosted regression tree and random forest models for mapping topsoil organic carbon concentration in an alpine ecosystem[J]. Ecological Indicators, 2016, 60: 870–878.
- [30] Prasad A M, Iverson L R, Liaw A. Newer classification and regression tree techniques: Bagging and random forests for ecological prediction[J]. Ecosystems, 2006, 9(2): 181–199.
- [31] Keskin H, Grunwald S, Harris W G. Digital mapping of soil carbon fractions with machine learning[J]. Geoderma, 2019, 339: 40–58.
- [32] 乔木, 周生斌, 卢磊, 等. 新疆渭干河流域土壤盐渍化时空变化及成因分析[J]. 地理科学进展, 2012, 31(7): 904–910.
- [33] 黄志方. 伊犁河流域农田盐渍化特征及驱动因子分析[D]. 乌鲁木齐: 新疆大学, 2010.
- [34] Scudiero E, Skaggs T H, Corwin D L. Regional-scale soil salinity assessment using Landsat ETM + canopy reflectance[J]. Remote Sensing of Environment, 2015, 169: 335–343.
- [35] Masoud A A, Koike K. Arid land salinization detected by remotely-sensed landcover changes: A case study in the Siwa region, NW Egypt[J]. Journal of Arid Environments, 2006, 66(1): 151–167.