

基于参数优化 SVM 方法识别盐生植被钠离子光谱特征

邓来飞^{1,2}, 张 飞^{1,2,3*}, 齐亚霄^{1,2}, 袁 婕^{1,2}

1. 新疆大学资源与环境科学学院智慧城市与环境建模自治区普通高校重点实验室, 新疆 乌鲁木齐 830046
2. 新疆大学绿洲生态教育部重点实验室, 新疆 乌鲁木齐 830046
3. 中亚地理信息开发利用国家测绘地理信息局工程技术研究中心, 新疆 乌鲁木齐 830002

摘 要 新疆盐渍土地分布广、面积大,在这些盐渍土地上生长着多种类型的盐生植物,它们对改良盐渍土地、维护生态稳定、促进生态平衡具有重要的现实意义。有关研究表明,许多盐土植物大量吸收钠,钠与钾都能增加细胞渗透压,以适应高盐环境,产生膨压而促进细胞的伸长,因而对其生长是有益的,能部分代替钾的功能。因此掌握盐生植物的钠特征,有助于了解盐生植物对生态环境的长期适应和响应,使用高光谱技术实现有效诊断叶片钠特征。首先,对实测冠层高光谱数据,采用离散小波变换(DWT)和 db5 小波对原始光谱进行 9 层小波分解,求取最佳分解层数为 5 层。其次,对光谱数据进行 5 层 db5 小波分解,对分解后的高频分量和低频分量建立了小波植被指数,筛选出可敏感表征钠离子含量的小波植被指数。最后,利用 SVR, LS-SVR, PSO-SVR 和 PSO-LS-SVR 模型建立盐生植被钠离子含量的估算模型,并与由原始光谱构建的光谱植被指数建立的估算模型进行比较。此外,引入偏最小二乘回归模型 PLSR 作为对比,评价参数优化的支持向量机方法在高光谱技术估测盐生植被叶片钠离子含量的优势。结果表明:(1)5 种模型预测结果表明,PSO 能有效优化 SVR 和 LS-SVR 模型参数(c, g),提高模型精度和预测能力,优化后的模型具有预测精度高、泛化能力强以及稳健性能好等特点。(2)综合小波指数构建的模型是综合多尺度、多分辨率数据的反演模型,其能从不同侧面反映植被的信息,因而综合小波指数构建的 4 种模型优于单一小波指数构建的模型。(3)对比两种类型的植被指数反演结果,单一小波植被指数构建 Na^+ 含量的预测模型可取得较好的预测效果,单一光谱指数估测 Na^+ 含量效果不佳,这是因为小波变换可以减少原始光谱的噪声,凸显光谱的细节信息,增强其反演 Na^+ 含量的精度;综合小波植被指数构建的模型精度和预测效果优于综合光谱指数构建的模型,原始光谱经小波变换后,可凸显更多的细节信息,提高高光谱反演叶片 Na^+ 含量的能力。

关键词 盐生植被; 钠离子; 高光谱; 小波变换; 支持向量机

中图分类号: O433.4 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2020)01-0247-08

引 言

新疆盐渍土地分布广、面积大,在这些盐渍土地上生长着多种类型的盐生植物,它们对改良盐渍土地、维护生态稳定、促进生态平衡具有重要的现实意义。因此,有关盐生植物的相关研究受到了众多学者的重视。有关研究发现,许多盐土植物大量吸收钠,钠与钾都能增加细胞渗透压,以适应高盐环境,产生膨压而促进细胞的伸长,因而对其生长是有益的,能部分代替钾的功能。因此,掌握盐生植物的钠特征,

有助于了解盐生植物对生态环境的长期适应和响应。

随着高光谱遥感技术的发展,众多学者利用高光谱遥感估算植被营养元素、监测植被含水量^[1]、识别植被信息与提取特征参数、监测植被生长状况与评估生物量^[1-3]等。植被生理参数估算采用的研究方法主要包括基于统计的经验或半经验模型和植被辐射传输模型。现阶段,由于统计方法建模简单、建模结果可以有效适用,因而目前在估算植被生物化学成分中使用较为广泛。在统计方法中,利用高光谱技术估算植被生物化学成分的常见模型有传统线性模型和非线性回归模型,如偏最小二乘回归^[4]、最小二乘支持向量机回归、

收稿日期: 2018-11-09, 修订日期: 2019-03-12

基金项目: 国家自然科学基金项目(新疆联合本地优秀青年人才培养专项)(U1503302), 新疆维吾尔自治区天山英才项目(400070010209), 自治区人民政府地方公派出国留学成组配套项目(L06)资助

作者简介: 邓来飞,女,1991年生,新疆大学资源与环境科学学院硕士研究生 e-mail: serenity_dlf@163.com

* 通讯联系人 e-mail: zhangfei3s@163.com

BP 神经网络等。Zhang 等^[5]采用统计分析的方法筛选了盐渍土环境下棉花叶片 Na^+ , K^+ , Ca^{2+} 和 Mg^{2+} 等离子含量和水分含量与光谱反射率的敏感波段, 构建了基于叶片离子含量和相对水分含量与光谱特征指数的简单线性和指数函数模型。Mirzaie 等^[6]使用偏最小二乘回归和主成分回归分析对植被含水量反演精度的影响。田旷达等筛选出表征小麦叶绿素含量的指数 REP, 使用最小二乘支持向量回归模型反演叶绿素含量, 并使用该模型进行 OMIS 影像的填图。李媛媛^[7]等使用主成分回归、偏最小二乘回归和 BP 神经网络模型建立了乳熟期的玉米叶片 SPAD 值和光谱特征参数的估算模型。

SVM 模型具有严谨的数学基础, 通过统计学习中的 VC 维理论和寻求结构风险最小化原理来提高泛化能力, 且能在保证数据精度的同时降低逼近函数的复杂度, 特别是对小样本、非线性问题的求解具有诸多优势。在 SVM 模型中, 决定模型预测精度的主要参数是惩罚因子 c 和核函数参数 σ , 一般采用网格划分的方式、遗传算法、粒子群算法等对 c 和 σ 进行寻优。本文以实测的植被高光谱数据和叶片 Na^+ 含量为数据源, 构建对 Na^+ 含量有较好预测能力的小波植被指数, 采用支持向量机 (support vector regression, SVM)、最小二乘支持向量机 (least squares-support vector regression, LS-SVM)、粒子群优化的支持向量机 (particle swarm optimization for SVM, PSO-SVM)、和粒子群优化的最小二乘支持向量机 (particle swarm optimization for LSSVM, PSO-LSSVM) 4 种方法构建适用于该区域的盐生植被叶片 Na^+ 含量的高光谱估算模型, 为使用高光谱技术监测西北干旱区盐生植被生长提供技术参考。

1 实验部分

1.1 研究区概况

艾比湖湿地国家自然保护区位于新疆博尔塔拉蒙古自治州 (博州) 境内东北角 (E 82°36'—83°50', N 44°30'—45°10'), 总面积 2 670.85 km², 其中水域面积为 500~600 km²。在行政区划上地跨博州精河县、博乐市和阿拉山口口岸。研究区属典型的温带干旱大陆气候, 干旱少雨、气温变化剧烈, 年平均气温 6~8 °C, 年降水量约 160 mm, 蒸发量 1 600 mm 以上, 日照数约 2 800 h。近年来, 由于入湖水量减少, 湖面萎缩, 水位下降, 加之艾比湖腹地位于阿拉山口主风通道区, 保护区土壤盐渍化严重。主要植物类型有梭梭、胡杨、碱蓬、怪柳、盐穗木、盐节木、花花柴、芦苇等。

1.2 数据收集

于 2017 年 5 月对研究区进行综合调查, 根据研究区的可达性和植被分布情况, 环湖均匀布设 37 个采样点, 如图 1。每个样点选择数种典型盐生植被。采用 ASD Field Spec3 型地物光谱仪 (光谱范围 350~2 500 nm, 光谱分辨率 3 nm, 采样间隔 1 nm), 测量时间为 10:00—14:00, 在风力小、天气晴朗条件下进行植株冠层高光谱测量。测量前用白板定标, 为减少环境对仪器的影响, 每隔一段时间用白板参考。仪器探头置于植株上方 10 cm 处, 与地面保持垂直, 每一样

本重复测量 10 次, 对 10 条光谱曲线取均值作为原始光谱反射率, 采用移动窗口平滑对光谱数据进行平滑处理。将采集光谱的植株叶片取下, 放于液氮罐中, 返回室内测植株 Na^+ 含量, 方法参见文献^[8]。

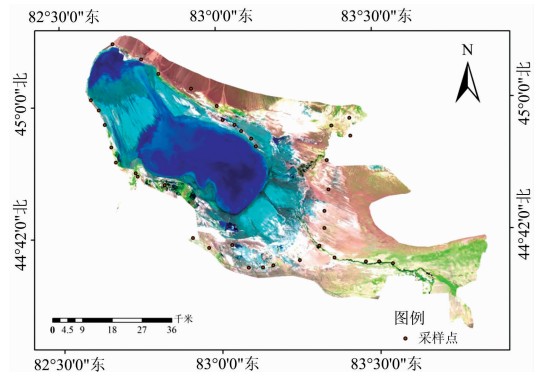


图 1 研究区示意图

Fig. 1 Schematic diagram of the study area

1.3 光谱小波变换

小波分析的最主要特点就是通过小波函数的伸缩和平移表征信号的局部频谱特征, 而植被的各种理化成分的吸收或反射特征具有明显的局部特征, 因此, 对植被光谱数据使用小波分析进行提取, 可充分表征光谱信息。离散小波变换可减少冗余信息, 可有效地从光谱信号中提取信息, Liu 等^[9]研究表明 db5 小波能精准的探测作物光谱的奇异性, 可有效提取植被光谱信息。本文采用离散小波变换 (DWT) 和 db5 母小波对原始光谱进行 9 层小波分解, 借鉴李军等^[10]提出的相关系数法确定最佳小波分解层次。

1.4 植被指数构建

选择归一化型 (NDSI)、比值型 (RSI) 和差值型指数 (DSI) 构建原始光谱植被指数; 并对小波分解后重构的低频和高频分量构建上述 3 种小波植被指数。公式如表 1。

表 1 构建的各植被指数公式

Table 1 The indices formulas constructed in this study

光谱植被指数	小波植被指数
$\text{NDSI}_{(\lambda_1, \lambda_2)} = \frac{R_{\lambda_1} - R_{\lambda_2}}{R_{\lambda_1} + R_{\lambda_2}}$	$\omega\text{NDSI}_{(\lambda_1, \lambda_2)} = \frac{\omega R_{\lambda_1} - \omega R_{\lambda_2}}{\omega R_{\lambda_1} + \omega R_{\lambda_2}}$
$\text{RSI}_{(\lambda_1, \lambda_2)} = \frac{R_{\lambda_1}}{R_{\lambda_2}}$	$\omega\text{RSI}_{(\lambda_1, \lambda_2)} = \frac{\omega R_{\lambda_1}}{\omega R_{\lambda_2}}$
$\text{DSI}_{(\lambda_1, \lambda_2)} = R_{\lambda_1} - R_{\lambda_2}$	$\omega\text{DSI}_{(\lambda_1, \lambda_2)} = \omega R_{\lambda_1} - \omega R_{\lambda_2}$

其中, R_{λ_1} 和 R_{λ_2} 为原始光谱反射率, ωR_{λ_1} 和 ωR_{λ_2} 分别为小波分解后重构得到的低频或高频分量。

1.5 模型构建与精度评价

LS-SVR 是在传统 SVR 的基础上采用不同的损失函数, 同时将 SVR 中的不等式约束条件转化为等式约束条件, 在此过程中求解的是线性方程组, 从而简化问题, 提高运行效率。粒子群 (PSO) 是一种基于种群和适应度的全局优化算法, 其易于理解、便于实现, 多用于参数寻优、模型分类、神经网络训练预测等领域^[11]。PSO 可以有效优化 SVR 和 LS-

SVR 模型参数(c, g), 广泛应用于非线性回归模型参数估计中, 具体介绍参见文献[12-13]。

本文依据 Kennard-Stone 算法选择建模样本与预测样本, 数量比为 2 : 1, 模型精度评估选取决定系数(coefficient of determination, R²), 均方根误差(root mean squared error, RMSE)和平均相对分析误差(residual predictive deviation, RPD)。

2 结果与讨论

2.1 样品的 Na⁺ 统计特征

表 2 为 Na⁺ 含量样本统计特征, Na⁺ 变异系数为

61.97%, 属于中等变异, Kolmogorov-Smirnov 检验所得 p 值为 0.797, 大于 0.1, 样本符合正态分布。

2.2 小波植被指数

2.2.1 光谱小波变换及最佳分解层数

由于光谱仪器和环境噪声影响, 水分吸收谷附近存在巨大噪声, 水分吸收谷波段附近本文不作讨论, 研究的光谱波段为: 350~1 350, 1 440~1 800 和 2 020~2 400 nm。由 db5 小波母函数对盐生植被高光谱曲线进行 9 层分解后, 对各层高频系数进行单支重构, 并将其与叶片 Na⁺ 含量进行相关性分析。如图 2, 4~5 层的相关系数高于 1~3 层, 并且从第 5 层开始相关系数的波动幅度减小, 渐渐趋于稳定。由此, 本研究的最佳分解层数确定为 5 层。

表 2 Na⁺ 含量样本统计特征
Table 2 Statistical characteristics of the content of Na⁺

样本数	最小值/ (mg · kg ⁻¹)	最大值/ (mg · kg ⁻¹)	均值/ (mg · kg ⁻¹)	标准差/ (mg · kg ⁻¹)	变异 系数/%	Kolmogorov-Smirnov 检验 渐进显著性(双侧)
60	0.034 986	0.735 06	0.327 11	0.202 71	61.97	0.797

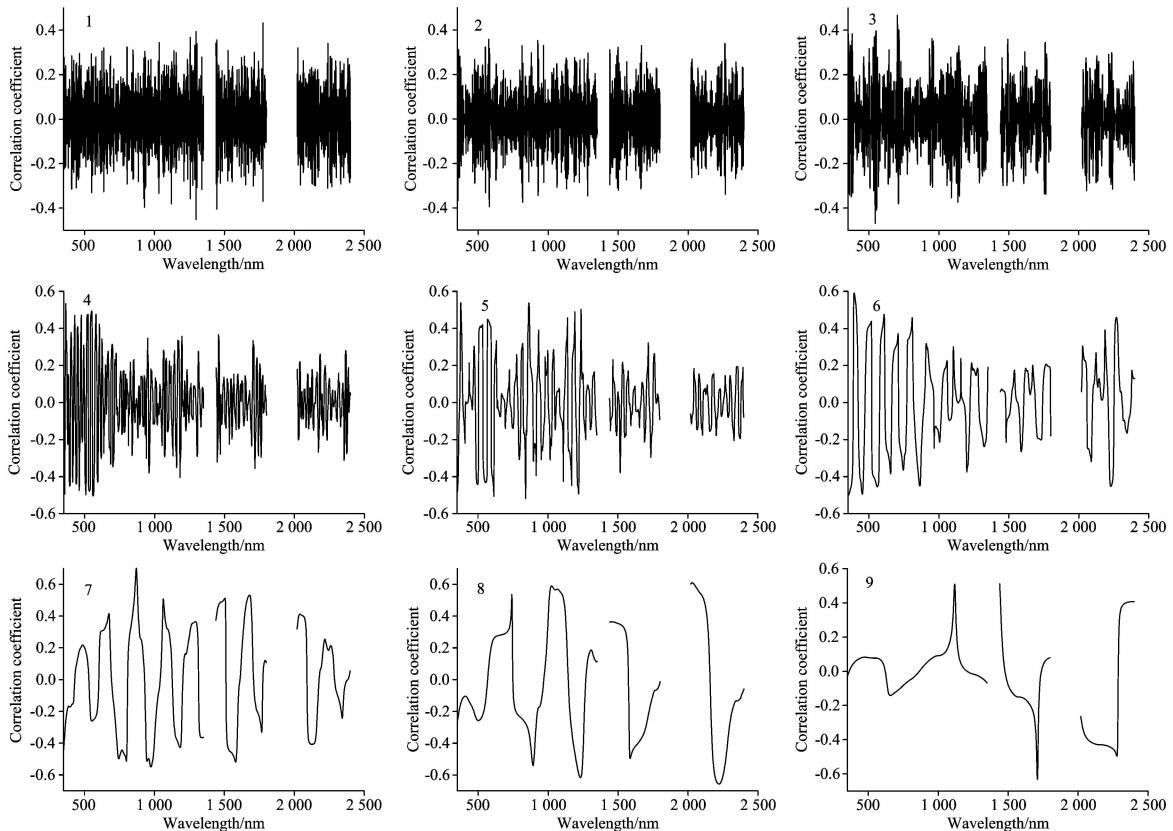


图 2 9 层小波分解高频系数与叶片 Na⁺ 含量相关系数图

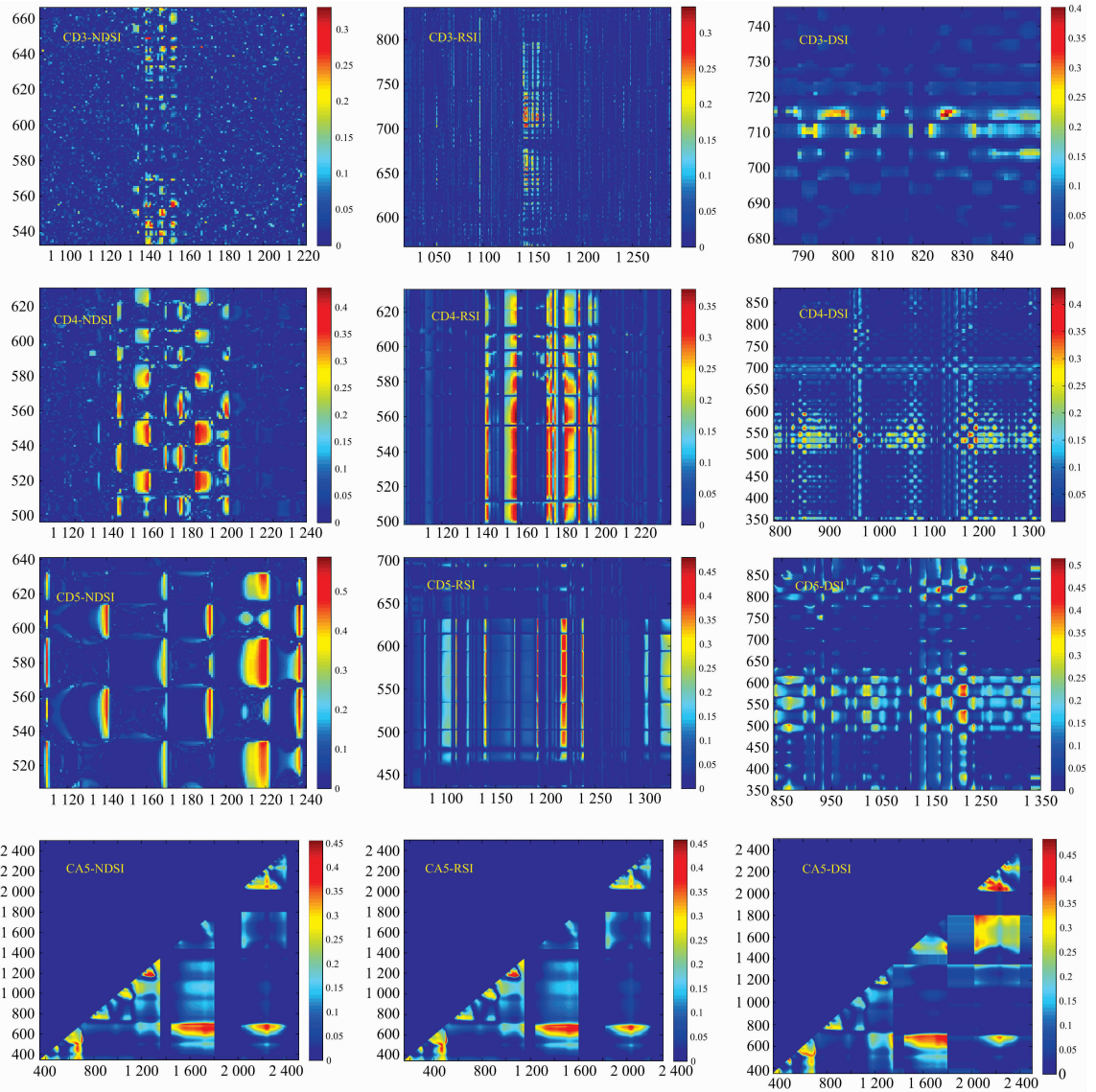
Fig. 2 Correlation coefficient between high frequency coefficient of 9-layer wavelet decomposition and Na⁺ content of leaves

2.2.2 小波植被指数与叶片 Na⁺ 的关系

使用 db5 小波母函数对光谱数据进行 5 层分解, 得到 1 个低频分量 CA5, 5 个高频分量 CD1, CD2, CD3, CD4 和 CD5, 对低频和高频分量构建 ωNDSI, ωRSI, ωDSI 指数, 并建立各小波植被指数与 Na⁺ 含量的决定系数等势线图, 如

图 3。

由图 3 可知, 由低频分量 CA5 构建的 NDSI 指数 R² 较高的波段集中在可见光、近红外区域, 范围为: X: 1 440~1 800 nm, Y: 600~720 nm 和 X: 1 200~1 300 nm, Y: 1 160~1 230 nm, 决定系数最高在 (R_{1 277}, R_{1 171}) 处, R² 为

图 3 小波植被指数与 Na^+ 含量决定系数图Fig. 3 Determination coefficient map of wavelet vegetation index and Na^+ content

0.452。由低频分量 CA5 构建的 RSI 指数 R^2 较高的波段集中在可见光、近红外区域, 范围为: $X: 1\ 440\sim 1\ 800\ \text{nm}$, $Y: 590\sim 710\ \text{nm}$ 和 $X: 1\ 200\sim 1\ 300\ \text{nm}$, $Y: 1\ 160\sim 1\ 230\ \text{nm}$ 和 $X: 630\sim 700\ \text{nm}$, $Y: 400\sim 600\ \text{nm}$, 决定系数最高在 $(R_{1\ 281}, R_{1\ 172})$ 处, R^2 为 0.457。由低频分量 CA5 构建的 DSI 指数 R^2 较高的波段集中在可见光、近红外区域, 范围为: $X: 1\ 430\sim 1\ 800\ \text{nm}$, $Y: 590\sim 720\ \text{nm}$, $X: 2\ 020\sim 2\ 400\ \text{nm}$, $Y: 1\ 460\sim 1\ 800\ \text{nm}$ 和 $Y: 2\ 020\sim 2\ 270\ \text{nm}$, 决定系数最高在 $(R_{2\ 232}, R_{2\ 046})$ 处, R^2 为 0.483。由高频分量 CD3 构建的 NDSI 指数 R^2 较高的波段位于可见光和近红外区域, 范围为: $X: 1\ 135\sim 1\ 160\ \text{nm}$, $Y: 530\sim 670\ \text{nm}$, 决定系数最高在 $(R_{1\ 141}, R_{554})$ 处, R^2 为 0.32。由高频分量 CD3 构建 RSI 指数 R^2 较高的波段位于近红外区域, 范围为: $X: 1\ 140\sim 1\ 160\ \text{nm}$, $Y: 700\sim 745\ \text{nm}$, 决定系数最高在 $(R_{1\ 143}, R_{729})$ 处, R^2 为 0.336。由高频分量 CD3 构建 DSI 指数 R^2 较高的

波段位于可见光和近红外区域, 范围为: $X: 785\sim 830\ \text{nm}$, $Y: 705\sim 720\ \text{nm}$, 决定系数最高在 (R_{826}, R_{715}) 处, R^2 为 0.403。由高频分量 CD4 构建 NDSI 指数 R^2 较高的波段位于可见光和近红外区域, 范围为: $X: 1\ 140\sim 1\ 200\ \text{nm}$, $Y: 500\sim 585\ \text{nm}$, 决定系数最高在 $(R_{1\ 182}, R_{550})$ 处, R^2 为 0.436。由高频分量 CD4 构建 RSI 指数 R^2 较高的波段位于可见光和近红外区域, 范围为: $X: 1\ 140\sim 1\ 200\ \text{nm}$, $Y: 500\sim 630\ \text{nm}$, 决定系数最高在 $(R_{1\ 157}, R_{559})$ 处, R^2 为 0.378。由高频分量 CD4 构建 DSI 指数 R^2 较高的波段位于可见光和近红外区域, 范围为: $X: 820\sim 1\ 320\ \text{nm}$, $Y: 500\sim 600\ \text{nm}$ 决定系数最高在 (R_{960}, R_{548}) 处, R^2 为 0.425。由高频分量 CD5 构建 NDSI 指数 R^2 较高的波段位于可见光和近红外区域, 范围为: $X: 1\ 190\sim 1\ 240\ \text{nm}$, $Y: 500\sim 635\ \text{nm}$, 决定系数最高在 $(R_{1\ 237}, R_{612})$ 处, R^2 为 0.586。由高频分量 CD5 构建 RSI 指数 R^2 较高的波段位于可见光和近红外区域, 范围为:

X: 1 215~1 225 nm, Y: 505~635 nm, 决定系数最高在 ($R_{1\ 218}$, $R_{5\ 564}$)处, R^2 为 0.438。由高频分量 CD5 构建 DSI 指数 R^2 较高的波段位于可见光、近红外区域, 范围为: X:

1 130~1 220 nm, Y: 490~830 nm 决定系数最高在 ($R_{1\ 237}$, $R_{6\ 13}$)处, R^2 为 0.515。详见表 3。(高频分量 CD1, CD2 构建的指数效果不好, 予以省略)

表 3 基于小波变换构建的植被指数

Table 3 Vegetation index based on wavelet transform

小波系数类型	指数类型					
	NDSI		RSI		DSI	
	波段组合	决定系数	波段组合	决定系数	波段组合	决定系数
CA5	($R_{1\ 277}$, $R_{1\ 171}$)	0.452	($R_{1\ 281}$, $R_{1\ 172}$)	0.457	($R_{2\ 232}$, $R_{2\ 046}$)	0.483
CD3	($R_{1\ 141}$, $R_{5\ 54}$)	0.314	($R_{1\ 143}$, $R_{7\ 29}$)	0.337	($R_{8\ 26}$, $R_{7\ 15}$)	0.403
CD4	($R_{1\ 182}$, $R_{5\ 50}$)	0.436	($R_{1\ 157}$, $R_{5\ 59}$)	0.378	($R_{9\ 60}$, $R_{5\ 48}$)	0.425
CD5	($R_{1\ 237}$, $R_{6\ 12}$)	0.586	($R_{1\ 218}$, $R_{5\ 64}$)	0.438	($R_{1\ 237}$, $R_{6\ 13}$)	0.515
光谱植被指数	($R_{1\ 275}$, $R_{1\ 172}$)	0.464	($R_{1\ 275}$, $R_{1\ 172}$)	0.466 7	($R_{2\ 231}$, $R_{2\ 040}$)	0.533

表 3 是从等势线图中挑选出较好的小波植被指数用于估算盐生植被叶片 Na^+ 含量。本文选取 R^2 大于或等于 0.4 的指数作为估算叶片 Na^+ 含量的敏感波段组合, 分别是: NDSI(CA_5) ($R_{1\ 277}$, $R_{1\ 171}$), RSI(CA_5) ($R_{1\ 281}$, $R_{1\ 172}$), DSI(CA_5) ($R_{2\ 232}$, $R_{2\ 046}$), DSI(CD_3) ($R_{8\ 26}$, $R_{7\ 15}$), NDSI(CD_4) ($R_{1\ 182}$, $R_{5\ 50}$), DSI(CD_4) ($R_{9\ 60}$, $R_{5\ 48}$), NDSI(CD_5) ($R_{1\ 237}$, $R_{6\ 12}$), RSI(CD_5) ($R_{1\ 218}$, $R_{5\ 64}$), DSI(CD_5) ($R_{1\ 237}$, $R_{6\ 13}$)。此外, 挑选由原始光谱构建的 NDSI($R_{1\ 275}$, $R_{1\ 172}$), RSI($R_{1\ 275}$, $R_{1\ 172}$)和 DSI($R_{2\ 231}$, $R_{2\ 040}$)指数, 用于估算叶片 Na^+ 含量。

2.3 叶片 Na^+ 含量估算模型

2.3.1 基于小波植被指数的叶片 Na^+ 含量反演

以挑选的 9 个小波植被指数作为自变量, 构建 Na^+ 的估算模型, 模型结果表明以低频分量 CA5 构建的 DSI 指数和 高频分量 CD5 构建的 DSI 和 NDSI 指数的光谱特征值构建的 Na^+ 估测模型效果较好, 并以这 3 个小波植被指数作为输入变量, 构建 Na^+ 含量综合反演模型, 并引入 PLSR 作为对比, 结果见表 4。

总体上, 小波植被指数 DSICA5($R_{2\ 232}$, $R_{2\ 046}$) 构建的各估算模型中, 对比各模型建模集和预测集的 R^2 , RMSE, RPD 可知, PSO-SVR 模型预测效果最好, 模型建模集 R^2 为 0.494, RMSE 为 0.145、预测集 R^2 为 0.656, RMSE 为 0.124、RPD 为 1.533; PSO-LS-SVR 模型预测效果次之, 建模集 R^2 为 0.483, RMSE 为 0.147、预测集 R^2 为 0.634, RMSE 为 0.129、RPD 为 1.474; LS-SVR 模型效果次于 PSO-LS-SVR 模型, 其 R^2 和 RPD 稍低于 PSO-LS-SVR, RMSE 稍高于 PSO-LS-SVR; SVR 模型预测效果最差; PLSR 模型预测结果出现负值。小波植被指数 DSICD5($R_{1\ 237}$, $R_{6\ 13}$) 构建的各估算模型中, PSO-LS-SVR 模型预测效果最好, 建模集 R^2 为 0.51, RMSE 为 0.144, 预测集 R^2 为 0.561, RMSE 为 0.107, RPD 为 1.645; PSO-SVR 模型次之, SVR 模型预测效果最差; PLSR 模型预测结果出现负值。小波植被指数 NDSICD5($R_{1\ 237}$, $R_{6\ 12}$) 构建的各估算模型中, PSO-LS-SVR 模型建模集 R^2 为 0.528, RMSE 为 0.144, 预测集 R^2 为 0.758, RMSE 为 0.092, RPD 为 2.035; PSO-SVR 和 LS-SVR 模型精度次之; SVR 和 PLSR 模型精度最

低。光谱小波变换后的低频分量(CA5)实际为分辨率降低后的植被光谱, 其表征光谱峰谷特征的细节信息又被分解到高频部分, 这是高频分量(CD5)构建的 NDSI 指数估算 Na^+ 含量预测模型精度优于低频分量(CA5)构建的 DSI 指数估算 Na^+ 含量预测模型精度的原因。

表 4 小波植被指数的模型反演结果

Table 4 Model inversion results of wavelet vegetation index

光谱指数	模型	建模集		验证集		
		R^2	RMSE	R^2	RMSE	RPD
DSICA5($R_{2\ 233}$, $R_{2\ 046}$)	PLSR	—	—	—	—	—
	SVR	0.466	0.252	0.610	0.202	0.941
	LS-SVR	0.481	0.252	0.631	0.196	0.970
	PSO-SVR	0.494	0.145	0.656	0.124	1.533
	PSO-LS-SVR	0.483	0.147	0.634	0.129	1.474
DSICD5($R_{1\ 237}$, $R_{6\ 13}$)	PLSR	—	—	—	—	—
	SVR	0.492	0.231	0.503	0.167	1.057
	LS-SVR	0.509	0.142	0.511	0.136	1.292
	PSO-SVR	0.503	0.146	0.556	0.119	1.479
	PSO-LS-SVR	0.510	0.144	0.561	0.107	1.645
NDSICD5($R_{1\ 237}$, $R_{6\ 12}$)	PLSR	0.446	0.183	0.478	0.197	0.949
	SVR	0.509	0.251	0.664	0.210	0.889
	LS-SVR	0.510	0.144	0.697	0.094	1.984
	PSO-SVR	0.528	0.147	0.758	0.105	1.782
	PSO-LS-SVR	0.528	0.144	0.758	0.092	2.035
综合小波指数	PLSR	0.552	0.176	0.665	0.165	1.211
	SVR	0.523	0.230	0.748	0.188	1.063
	LS-SVR	0.554	0.241	0.785	0.186	1.075
	PSO-SVR	0.549	0.138	0.770	0.102	1.960
	PSO-LS-SVR	0.642	0.136	0.778	0.094	2.126

以 DSICA5($R_{2\ 232}$, $R_{2\ 046}$), DSICD5($R_{1\ 237}$, $R_{6\ 13}$) 和 NDSICD5($R_{1\ 237}$, $R_{6\ 12}$) 这 3 个指数作为输入变量, 构建的 Na^+ 含量综合反演模型中, PSO-LS-SVR 模型预测效果最佳, 建模集 R^2 为 0.642, RMSE 为 0.136、预测集 R^2 为 0.778, RMSE 为 0.094、RPD 为 2.126; PSO-SVR 模型预测效果次之, 建模集 R^2 为 0.549, RMSE 为 0.138、预测集 R^2 为 0.77,

RMSE 为 0.102、RPD 为 1.96；LS-SVR，SVR 和 PLSR 模型精度最低。相比于单一小波指数，以综合小波指数构建的各模型预测精度均有所提高，它是综合多尺度、多分辨率数据的反演模型，其更能从不同侧面反映植被的信息。综上，4 种模型预测结果表明，PSO 能有效优化 SVR 和 LS-SVR 模型参数(c , g)，提高模型精度和预测能力。由于篇幅限制，仅列出由综合小波指数构建的 PSO-LS-SVR 模型的最终优化参数结果，如图 4。

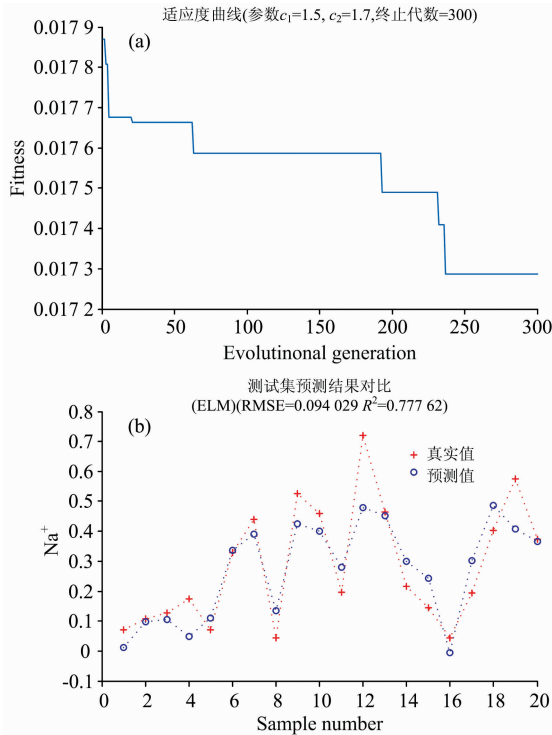


图 4 PSO-LS-SVR 模型参数优化和预测效果

Fig. 4 Parameter optimization and prediction based on PSO-LS-SVR model

表 5 光谱植被指数的各模型反演结果

Table 5 Inversion results of spectral vegetation index

光谱指数	模型	建模集		验证集		
		R^2	RMSE	R^2	RMSE	RPD
综合光谱指数	PLSR	0.505	0.228	0.536	0.140	1.428
	SVR	0.541	0.233	0.559	0.191	1.046
	LS-SVR	0.520	0.237	0.666	0.196	1.020
	PSO-SVR	0.542	0.137	0.665	0.103	1.940
	PSO-LS-SVR	0.550	0.137	0.781	0.094	2.126

2.3.2 基于光谱植被指数的叶片 Na^+ 含量反演

分别以挑选的 3 个光谱植被指数作为输入变量，构建 Na^+ 含量的预测模型，并以这 3 个光谱植被指数为输入变量，构建 Na^+ 含量综合反演模型，并引入 PLSR 作为对比，如表 5。

3 个光谱植被指数分别构建的各模型预测效果均不太理想(省略)。以这 3 个光谱植被指数作为输入变量，构建的各综合反演模型结果，如表 5，PSO-LS-SVR 模型最优，建模集 R^2 为 0.55，RMSE 为 0.137、预测集 R^2 为 0.781，RMSE 为 0.094、RPD 为 2.126；PSO-SVR 次之，建模集 R^2 为 0.542，RMSE 为 0.137、预测集 R^2 为 0.665，RMSE 为 0.103、RPD 为 1.940；LS-SVR，SVR 和 PLSR 模型精度最低。4 种模型预测结果表明，PSO 能有效优化 SVR 和 LS-SVR 模型参数(c , g)，提高模型精度和预测能力。相比于单一光谱指数，由于加入了多波段的信息，以综合光谱指数构建的 SVR 模型预测精度均有所提高。

对比两种类型的植被指数反演结果，单一小波植被指数构建 Na^+ 含量的预测模型可取得较好的预测效果，单一光谱指数估测 Na^+ 含量效果不佳，这是因为小波变换可以减少原始光谱的噪声，凸显光谱的细节信息，增强其反演 Na^+ 含量的精度；综合小波植被指数构建的模型是综合多尺度、多分辨率数据的反演模型，其模型精度和预测效果优于综合光谱指数构建的模型。光谱经过小波分解后，低频部分保留了植被原始光谱特征，但相对于原始光谱，其分辨率有所降低，体现原始光谱峰谷特征信息被分解到高频部分，其可凸显光谱的细节信息，利用该种小波信号建立的指数可提高叶片 Na^+ 含量反演精度。

3 结 论

(1) 5 种模型预测结果表明，PSO 能有效优化 SVR 和 LS-SVR 模型参数(c , g)，提高模型精度和预测能力。(2) 基于综合小波指数构建的模型是综合多尺度、多分辨率数据的反演模型，其更能从不同侧面反映植被的信息，因而综合小波指数构建的 4 种模型优于单一小波指数构建的模型。(3) 对比两种类型的植被指数反演结果，单一小波植被指数构建 Na^+ 含量的预测模型可取得较好的预测效果，单一光谱指数估测 Na^+ 含量效果不佳，这是因为小波变换可以减少原始光谱的噪声，凸显光谱的细节信息，增强其反演 Na^+ 含量的精度；综合小波植被指数构建的模型精度和预测效果优于综合光谱指数构建的模型，原始光谱经小波变换后，可凸显更多的细节信息，提高高光谱反演叶片 Na^+ 含量的能力。

References

- [1] Yi Q X, Bao A M, Wang Q, et al. *Computers & Electronics in Agriculture*, 2013, 90(1): 144.
- [2] Jin X, Yang G, Li Z, et al. *Precision Agriculture*, 2018, 19(1): 1.
- [3] Qi Y, Dennison P E, Jolly W M, et al. *Remote Sensing of Environment*, 2014, 150(7): 198.
- [4] Zhang C, Wang Q, Liu F, et al. *Measurement*, 2017, 97: 149.
- [5] Zhang L, Zhou Z, Zhang G, et al. *European Journal of Remote Sensing*, 2014, 47: 593.
- [6] Mirzaie M, Darvishzadeh R, Shakiba A, et al. *International Journal of Applied Earth Observation & Geoinformation*, 2014, 26: 1.
- [7] LI Yuan-yuan, CHANG Qing-rui, LIU Xiu-ying, et al(李媛媛, 常庆瑞, 刘秀英, 等). *Transactions of the Chinese Society of Agricultural Engineering(农业工程学报)*, 2016, 32(16): 135.
- [8] LI Zhe, ZHANG Fei, FENG Hai-kuan, et al(李哲, 张飞, 冯海宽, 等). *Acta Optica Sinica(光学学报)*, 2017, 22(11): 1128002.
- [9] Liu M, Liu X, Wu L, et al. *Computers & Geosciences*, 2011, 37(9): 1254.
- [10] LI Jun, LI Pei-jun, GUO Jian-cong, et al(李军, 李培军, 郭建聪). *Progress in Natural Science(自然科学进展)*, 2007, 17(11): 1500.
- [11] XIAO Yan, JIANG Qi-gang, WANG Bin, et al(肖艳, 姜琦刚, 王斌, 等). *Transactions of the Chinese Society of Agricultural Engineering(农业工程学报)*, 2016, 32(4): 211.
- [12] ZHANG Jian-wei, JIANG Qi, LIU Xuan-ran, et al(张建伟, 江琦, 刘轩然, 等). *Transactions of the Chinese Society of Agricultural Engineering(农业工程学报)*, 2017, 33(11): 75.
- [13] YANG Liu, SUN Jin-hua, FENG Zhong-ke, et al(杨柳, 孙金华, 冯仲科, 等). *Transactions of the Chinese Society of Agricultural Machinery(农业机械学报)*, 2016, 47(8): 273.

Identification of Sodium Ion Spectral Characteristics of Halophytes Based on Parameter Optimized SVM Method

DENG Lai-fei^{1, 2}, ZHANG Fei^{1, 2, 3*}, QI Ya-xiao^{1, 2}, YUAN Jie^{1, 2}

1. College of Resources & Environmental Science, Xinjiang University, Urumqi 830046, China

2. Key Laboratory of Oasis Ecology, Xinjiang University, Urumqi 830046, China

3. Engineering Research Center of Central Asia Geoinformation Development and Utilization, National Administration of Surveying, Mapping and Geoinformation, Urumqi 830002, China

Abstract There are a wide range of saline soils in Xinjiang. It covers a large area. Various types of halophytes which have prominent significance for improving the saline lands, maintaining ecological stability and promoting ecological balance grow on these saline soils. Studies have shown that many halophytes absorb a large amount of sodium. Both sodium and potassium can increase the cell osmotic pressure to adapt to the high-salt condition, producing turgor pressure and promoting cell elongation. So it is beneficial to its growth and can partially replace the function of potassium. Thus, mastering the sodium characteristics of halophytes is helpful to understand the long-term adaptation and response of halophytes to the ecological environment. In this paper, HyperSpectral technique was used to effectively explore the characteristics of leaf sodium. Firstly, the discrete wavelet transform (DWT) and db5 wavelet were used to decompose the original spectral in 9 layers, and the optimal decomposition layer is 5 layers. Secondly, the original spectral were decomposed by db5 wavelet in 5 layers, and the wavelet vegetation indices were created by the decomposed high-frequency components and low-frequency components. We selected the wavelet vegetation indices which could sensitively characterize sodium ion content of halophytes. Finally, the SVR, LS-SVR, PSO-SVR and PSO-LS-SVR models were used to estimate the sodium ion content of halophytes vegetation. The results were compared to the models created by the spectral vegetation indices of original spectral. In addition, we used the partial least squares regression model as a comparison to evaluate the advantages of the parameter-optimized support vector machine method in estimating the sodium ion content of the leaves of the halophytes vegetation using hyperspectral techniques. The results showed that: (1) The prediction results of the five models showed that PSO can effectively optimize the parameters (c , g) of SVR and LS-SVR models, and improve the accuracy and prediction ability of the models. The optimized models had the advantages of high prediction accuracy, strong generalization ability and good robustness performance. (2) The model constructed by the multiple wavelet index was an inversion model of integrated multi-scale and multi-resolution data, which can reflect the vegetation information from different aspects. Therefore, the four models constructed by the multiple wavelet index were superior to the models constructed by the

single wavelet index. (3) Contrasting the inversion results of two types index, the Na^+ content prediction model constructed by a single wavelet vegetation index can achieve a better prediction results. The single spectral index is not effective in estimating Na^+ content, which is because the wavelet transform can reduce the noise of the original spectral and highlight the detailed feature of the spectral, improving the prediction results. The model accuracy and prediction effect of the integrated wavelet vegetation index were better than those of the integrated spectral index. More spectral subtle feature can be highlighted by wavelet transform, thus improving the ability of retrieving Na^+ content in leaves by HyperSpectral method.

Keywords Halophyte; Sodium ion; HyperSpectral; Wavelet analysis; Support vector machine

(Received Nov. 9, 2018; accepted Mar. 12, 2019)

* Corresponding author

敬告读者——《光谱学与光谱分析》已全文上网

从 2008 年第 7 期开始在《光谱学与光谱分析》网站(www.gpxygpx.com)“在线期刊”栏内发布《光谱学与光谱分析》期刊全文,读者可方便地免费下载摘要和 PDF 全文,欢迎浏览、检索本刊当期的全部内容;并陆续刊出自 2004 年以后出版的各期摘要和 PDF 全文内容。2009 年起《光谱学与光谱分析》每期出版日期改为每月 1 日。

《光谱学与光谱分析》期刊社