

# 微分光谱连续小波系数估测雅氏落叶松尺蠖危害下的落叶松失叶率

黄晓君<sup>1,2,3,4</sup>, 颜耀文<sup>1\*</sup>, 包玉海<sup>2,3</sup>, 包刚<sup>2,3</sup>, 青松<sup>2,3</sup>, 包玉龙<sup>2,3,4</sup>

1. 兰州大学资源环境学院, 甘肃 兰州 730000
2. 内蒙古师范大学地理科学学院, 内蒙古 呼和浩特 010022
3. 内蒙古自治区遥感与地理信息系统重点实验室, 内蒙古 呼和浩特 010022
4. 内蒙古自治区蒙古高原灾害与生态安全重点实验室, 内蒙古 呼和浩特 010022

**摘要** 害虫引起的林木失叶会严重威胁森林健康。森林虫害遥感监测与评价中快速、准确获取失叶信息十分重要。基于此, 针对雅氏落叶松尺蠖引起的落叶松失叶灾象, 在蒙古国开展受害林木光谱测量和失叶率估测试验。首先通过光谱实测数据的处理, 得到微分光谱反射率(DSR, 对光谱反射率求一阶导数)和微分光谱连续小波系数(DSR-CWC, 利用 Biorthogonal, Coiflets, Daubechies 和 Symlets 等 4 种小波系的 36 个母小波基函数对 DSR 进行连续小波变换), 分析 DSR 和 DSR-CWC 对失叶率的敏感性, 进而借助 MATLAB 的 Findpeaks(Fp)函数自动寻找 DSR 和 DSR-CWC 的敏感波段并确定其对应的敏感特征, 然后利用连续投影算法(SPA)对敏感特征进行降维处理, 最后利用敏感特征建立偏最小二乘回归(PLSR)和支持向量机回归(SVMR)失叶率估测模型, 并与逐步多元线性回归(SMLR)模型进行比较。研究结果表明: ①DSR-CWC 与 DSR 相比, 对失叶率变化的敏感性更显著且敏感波段亦较多, 其敏感波段主要分布于三个吸收谷(440~515, 630~760 和 1 420~1 470 nm)和三个反射峰(516~620, 761~1 000 和 1 548~1 610 nm)范围内。说明 DSR-CWC 能够增强光谱反射和吸收特征。②Fp 与 SPA 结合模式(Fp-SPA)不仅能够快速、客观选择敏感特征, 而且对特征有效降维, 是一种光谱敏感特征选择的有效方法。③4 种小波系的最优母小波基分别为 bior2.4, coif2, db1 和 sym6, 其中 db1 的失叶率估测性能最稳定, 精度最高。④对 DSR 进行连续小波变换能够提高失叶率估测精度, 在 DSR-CWC 中 db1-PLSR 模型( $R_M^2=0.9340$ ,  $RMSE_M=0.0890$ )提高的最为显著, 比 DSR-PLSR 的  $R_M^2$  提高了 0.0475 并且比 DSR-PLSR 的  $RMSE_M$  降低了 0.0249。⑤利用 DSR-CWC 建立的 PLSR 和 SVMR 模型估测精度类似, 其精度优于 SMLR 模型。可见, DSR-CWC 比 DSR 失叶率估测更有潜力, 可为森林虫害遥感监测中提供重要参考。

**关键词** 雅氏落叶松尺蠖; 落叶松失叶率; 微分光谱连续小波系数; Findpeaks 函数; 连续投影算法

**中图分类号:** O433.4 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2019)09-2732-07

## 引言

森林虫害往往使林木失叶, 威胁森林健康安全<sup>[1]</sup>。失叶率(leaf loss rate, LLR 是指单位面积上整体树冠叶片损失量占全部叶片量的百分比)是表征森林失叶的参量, 是森林虫害监测重要指标。因而快速、准确获取林木失叶率信息在森林虫害监测和森林健康评价等方面有重要意义。

蒙古高原森林资源丰富, 但森林虫害频发, 其中蒙古国雅氏落叶松尺蠖(*Eramis Jacobsoni* Djak)是常见的严重危害

落叶松的害虫, 是蒙古高原北部独有一年一代森林害虫, 其幼虫在 6 月—7 月暴食针叶, 使落叶松严重失叶, 导致成片落叶松林死亡。多年来蒙古国大部落叶松林发生了雅氏落叶松尺蠖灾害, 使森林生态系统受到严重破坏。更值得注意的是该虫逐年往东传播, 已经对我国大兴安岭森林生态系统安全造成威胁<sup>[2]</sup>。在此背景下, 快速监测雅氏落叶松尺蠖灾害, 趁早掌握该虫发生发展规律, 为害虫防控提供重要依据是势在必行。目前遥感是植物虫害监测的重要手段, 与传统人工调查相比有监测范围广、效率高和成本低等优势<sup>[3]</sup>。植物虫害光谱分析及建模是后续遥感大面积监测植物虫害的实

收稿日期: 2019-03-28, 修订日期: 2019-06-06

基金项目: 国家自然科学基金项目(41861056, 61631011), 内蒙古自然科学基金项目(2018MS04008), 内蒙古科技计划项目(201702116)资助

作者简介: 黄晓君, 1984 年生, 内蒙古师范大学地理科学学院讲师 e-mail: hxj3s@qq.com

\* 通讯联系人 e-mail: xieyw@lzu.edu.cn

验理论基础<sup>[4-5]</sup>。因此,通过光谱分析技术,研究雅氏落叶松尺蠖危害下落叶松失叶率估算显得尤为重要。

害虫引起的森林失叶遥感监测是很多学者关注的研究热点<sup>[3]</sup>。目前在失叶遥感监测中以分类为主,而失叶率估测很少见且精度有待提高<sup>[6]</sup>。如 Rahimzadeh-Bajgiran 等通过 Landsat 遥感数据计算归一化差异水分指数、增强型植被指数和归一化植被指数,对云杉失叶进行了4级分类,总体分类精度为70%左右<sup>[7]</sup>。Hawrylo 等利用 Sentinel-2 遥感数据计算几种植被指数,采用K近邻(KNN)、随机森林(RF)和支持向量机(SVM)等机器学习算法建立了失叶分类模型,其总体准确率分别为78%,75%和78%<sup>[6]</sup>。Spruce 等采用 MODIS NDVI 时间序列数据集,对舞毒蛾引起的森林失叶进行了两级分类,其总体分类准确率为79%<sup>[8]</sup>。Philip A Townsend 等借助 Landsat 遥感数据计算归一化差异红外指数和水分胁迫指数,建立了失叶率估测模型,其 RMS 误差估计值为14.9%<sup>[9]</sup>。从参考文献发现森林失叶监测多集中在多光谱遥感应用上,而高光谱遥感的应用鲜有报道。高光谱与多光谱相比,对植物结构和生理的细微变化有更强感知能力。在植物虫害监测中有更大潜力,利用高光谱估测受害林木失叶率,可为森林失叶高光谱遥感监测提供有力实验依据,并对森林虫害遥感监测的发展具有重要意义。

尽管现有学者利用地面高光谱数据来区分不同程度的害虫引起的林木失叶状况<sup>[10]</sup>,但失叶率高光谱估测尚未见报道。一般来讲,植物虫害高光谱检测模型的实现过程由光谱敏感性分析、敏感光谱特征提取和高光谱检测模型构建等三部分组成<sup>[2]</sup>。首先选择一种敏感性分析方法从高光谱高维数据中确定有意义的敏感波段。敏感性分析方法有很多,如:相关性分析、判别分析、主成分分析、方差分析、敏感性分析、遗传算法、后向特征消去法和变量重要性分析等<sup>[2, 10-11]</sup>。然后从敏感波段中提取敏感光谱特征。光谱特征形式有多种,如:光谱反射率、微分光谱特征、去除包络线光谱特征、光谱小波特征和光谱植被指数等<sup>[12]</sup>。敏感光谱特征确定之后可利用多元线性回归、支持向量机、偏最小二乘法和决策树等数理统计和数据挖掘算法实现植物虫害检测模型<sup>[11, 13-14]</sup>。本文针对雅氏落叶松尺蠖危害下落叶松失叶灾象,通过微分光谱连续小波系数,利用偏最小二乘回归和支持向量机回归算法,构建失叶率估测模型,揭示高光谱对失叶率估测的潜力,旨在为利用高光谱遥感技术进行雅氏落叶松尺蠖灾害监测与预警提供参考。

## 1 实验部分

### 1.1 数据获取

在蒙古国杭爱省的 Ikhtamir, Tsenkher 和 Battsengel 等雅氏落叶松尺蠖灾区进行了光谱测量和失叶率估测的试验研究,其中 Ikhtamir 和 Tsenkher 试验区作为建模训练区,而 Battsengel 试验区作为模型验证区。

#### 1.1.1 高光谱数据

采用 ASD 地物光谱仪对不同程度失叶的落叶松进行了光谱测量。该光谱仪光谱分辨率为 3 nm@350~1 000 nm 和

10 nm@1 000~2 500 nm; 采样间隔为 1.4 nm@350~1 000 nm 和 2 nm@1 000~2 500 nm; 数据间隔为 1 nm。在整个测定过程中,探头使用视场角为 25°且垂直向下,高度要求离测定对象 0.2 m 左右,所有光谱在晴朗无云的天气于北京时间 10:30—14:30 期间测定(图1)。每个样本树垂直方向上分为上、中、下3个层次,然后各层次上选择典型一条树枝进行观测20次,每样本树观测前后均使用参考白板对光谱进行校正,将 20×3 次的光谱反射率的平均值来代表一个样本树的冠层光谱反射率值。从 Ikhtamir 和 Tsenkher 训练区分别采集 26 棵和 20 棵样本树的光谱曲线,用于失叶率估测模型的建立;从 Battsengel 验证区采集 20 棵样本树的光谱曲线,用于模型精度的验证。

#### 1.1.2 失叶率数据

现场失叶率估算是与光谱测量同时进行的,即对光谱测量的3条树枝计算失叶率。失叶率计算方法是首先在光谱测量范围(以光谱测量探头在树枝上的投影点为中心,以0.09 m 为半径的圆,见图1)内记录损失和健康针叶数量;然后通过式(1)分别计算3条树枝的失叶率,再取平均值作为当前样本树的失叶率。

$$LLR = \frac{N_l}{N_h + N_l} \times 100\% \quad (1)$$

式中 LLR 为树枝失叶率,  $N_h$  和  $N_l$  分别为健康和损失的针叶计数。

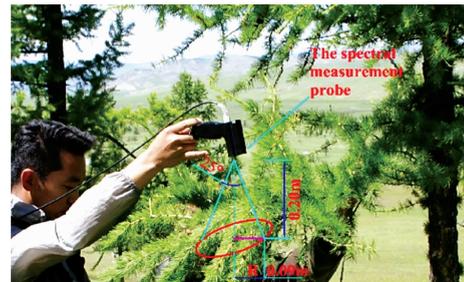


图1 落叶松树枝光谱测量

Fig. 1 Spectral measurement of larch branches

### 1.2 光谱数据预处理

高光谱数据在获取的过程中难免受到背景和仪器噪声的干扰,不可避免地产生异常光谱。利用 ViewSpecPro 软件剔除异常光谱,同时进行了平滑处理得到不同失叶率落叶松冠层预处理的光谱曲线。然后采用 ENVI 软件对预处理的光谱反射率求一阶导数,获得微分光谱反射率(记为 DSR),其目的为降低背景干扰,净化光谱数据及突出光谱特征有用信息<sup>[2]</sup>。

### 1.3 模型建立与验证

#### 1.3.1 连续小波变换

连续小波变换是一种能同时将频域和时域局部化的线性变换方法,利用母小波基将高光谱反射率数据在不同尺度和位置上变换成一系列连续小波能量系数。这种系数可直接与全波段相对应,能够提供植物冠层光谱吸收特征与反射特征的形状及位置的可解释信息<sup>[12]</sup>。在光谱预处理的基础上,利用连续小波对不同程度失叶的林木冠层微分光谱反射率进行

变换, 获得微分光谱连续小波系数(记为 DSR-CWC)。在连续小波变换中采用了 Biorthogonal, Coiflets, Daubechies 和 Symlets 小波系的 36 种母小波基, 如: ① Biorthogonal: bior1.3, bior1.5, bior2.2, bior2.4, bior2.6, bior2.8, bior3.1, bior3.3, bior3.5, bior3.7, bior3.9, bior4.4, bior5.5 和 bior6.8 等 14 个母小波基; ② Coiflets: coif1, coif2, coif3, coif4 和 coif5 等 5 个母小波基; ③ Daubechies: db1, db2, db3, db4, db5, db6, db7, db8, db9 和 db10 等 10 个母小波基; ④ Symlets: sym2, sym3, sym4, sym5, sym6, sym7 和 sym8 等 7 个母小波基。为避免太多的分解尺度, 笔者仅在  $2^1, 2^2, 2^3, 2^4, 2^5, 2^6, 2^7, 2^8, 2^9$  和  $2^{10}$  等 10 个尺度上<sup>[12]</sup>, 采用 MATLAB2016b 软件完成了连续小波变换处理。

### 1.3.2 敏感连续小波特征提取

失叶率与 DSR-CWC 的决定系数  $R^2$  值大于 0.64 ( $p < 0.0001$ ) 对应的波段设为敏感波段<sup>[12]</sup>。利用 MATLAB 软件的 Findpeaks(Fp) 函数[见式(2)]自动寻找  $R^2$  峰值并确定敏感波段, 然后通过敏感波段位置, 分别提取 10 个尺度上的小波系数。

$$[P, L] = \text{Findpeaks}(x, \text{'minpeakheight'}, h, \text{'minpeakdistance'}, d) \quad (2)$$

式中,  $x$  为决定系数  $R^2$  和波段组成的矩阵,  $P$  为  $R^2$  峰值,  $L$  为  $R^2$  峰值相应的波段位置,  $h$  为  $R^2$  峰值的最小高度(设为 0.64),  $d$  为相邻  $R^2$  峰值间的最小波段间隔数(设为 10 nm)。

在此基础上, 利用连续投影算法(SPA)对每个尺度的小波系数进行降维处理, 并确定敏感连续小波特征。SPA 是一种将可变共线最小化的前向循环选择方法, 从一个波长开始, 每次迭代时使用向量空间中的投影算法选择一个新波长, 直到预定波长数为止。因而 SPA 以克服敏感波段间的共线性来选择重要波长, 能够建立稳定的模型。SPA 的实现步骤如下:

建模样本树数量和敏感波段波长数分别为  $m$  和  $n$ , 则两者组成的微分光谱连续小波系数矩阵为  $\mathbf{X}_{m \times n}$ , 初始迭代向量为  $\mathbf{x}_{k(0)}$ , 需要提取的连续小波特征个数(也就是连续小波系数个数)为  $t$ 。

①首次迭代:  $d=1$ , 从微分光谱连续小波系数矩阵中任选 1 列向量  $\mathbf{x}_j$ , 记为  $\mathbf{x}_{k(0)}$ , 非选入列向量的集合记为  $Q$

$$Q = \{j, 1 \leq j \leq n, j \notin \{k(0), \dots, k(d-1)\}\}$$

②计算  $\mathbf{x}_j$  对  $Q$  中向量的投影向量

$$P\mathbf{x}_j = \mathbf{x}_j - (\mathbf{x}_j^T - \mathbf{x}_{k(d-1)}^T)\mathbf{x}_{k(d-1)}(\mathbf{x}_{k(d-1)}^T - \mathbf{x}_{k(d-1)}^T)^{-1}$$

③记录最大投影的序号

$$k(d) = \arg(\max \|P\mathbf{x}_j\|), j \in Q$$

④最大投影作为下轮投影向量

$$\mathbf{x}_j = P\mathbf{x}_j, j \in Q$$

⑤ $d=d+1$ , 如果  $d < t$ , 回到②继续投影。

最终降维后的敏感小波系数为  $\{\mathbf{x}_{k(d)} = 0, \dots, t-1\}$ 。

### 1.3.3 失叶率估测模型

利用偏最小二乘回归(PLSR)和支持向量机回归(SVMR)算法建立失叶率估测模型, 并与逐步多元线性回归(SMLR)模型进行了比较。PLSR 是一种多元线性数据统计分析法, 旨在解决存在许多(可能相关的)预测变量和相对较

少样本的情况。在建模中它从自变量空间和因变量空间里寻求某些相关程度最大的成分线性组合, 以便更好地解释因变量的变异信息, 相当于普通线性回归、主成分分析、典型相关分析的组。SVMR 是一种基于统计学习理论的模式识别方法。利用统计学习理论的 VC-维数和最小结构风险原理, 建立支持向量机。它在保证数据逼近精度的同时能降低逼近函数的复杂度, 对有限样本和非线性问题方面特别有优势。该模型以径向基函数(RBF)为核函数, 它能反映数据点和成本损失函数(正则化参数)之间的相似性。另外, 考虑到正则项参数  $C$  和核参数  $g$  对模型性能的影响, 我们采用交叉验证确定最优参数  $C$  与  $g$ , 确保 SVMR 模型性能。

### 1.3.4 模型精度验证

为验证模型精度的可靠性, 训练数据与验证数据的采集选择在不同地区。利用决定系数( $R^2$ )和均方根误差 RMSE 作为训练模型(记为  $R_c^2$  和  $\text{RMSE}_c$ )和验证模型(记为  $R_v^2$  和  $\text{RMSE}_v$ )精度评定的基础指标。为验证模型精度更为客观, 利用训练与验证的平均决定系数 [ $R_M^2 = 0.5 \times (R_c^2 + R_v^2)$ ] 和平均均方根误差 [ $\text{RMSE}_M = 0.5 \times (\text{RMSE}_c + \text{RMSE}_v)$ ] 作为模型精度评价指标。 $R_M^2$  的最大值为 1, 其值越接近于 1, 说明模型拟合程度越好, 估测精度越高;  $\text{RMSE}_M$  的最小值为 0, 其值越接近于 0, 说明模型预测误差越小, 估测精度越高。

## 2 结果与讨论

### 2.1 DSR 与 DSR-CWC 的失叶率敏感性对比

图 2 显示了在 350~1 800 nm 范围内不同失叶率林木冠层 DSR 和 DSR-CWC 的曲线, 其中 DSR-CWC 是采用 bior2.4( $2^9$  尺度)母小波基函数变换的。从图 2 可知, 雅氏落叶松尺蠖引起的落叶松失叶, 导致 DSR 和 DSR-CWC 的变化。对 DSR 而言, 在 505~535 和 680~760 nm 波段内出现峰, 而在 550~590 和 600~620 nm 波段内出现谷。随失叶率增加, 峰的高度和谷的深度逐渐降低趋势[图 2(a)]。对 DSR-CWC 来讲, 在 450~495, 545~588, 648~688, 738~784 和 1 130~1 160 nm 波段内出现谷, 而在 500~545 和 688~738 nm 波段内出现峰。当失叶率上升时, 峰的高度和谷的深度呈明显降低趋势[图 2(b)]。与 DSR 相比, 随失叶率减小 DSR-CWC 峰高度更高, 波谷深度更深, 不同失叶率对应的 DSR-CWC 差异更大, 也就是说 DSR-CWC 对不同失叶率的响应特征更为显著。因此, 利用 DSR-CWC 与失叶率的关系估测失叶率是可行的。

为进一步分析 DSR-CWC 对不同失叶率落叶松冠层的敏感性, 将 DSR 及 DSR-CWC 逐波段进行 Pearson 相关性分析, 得到各波长上的决定系数  $R^2$  分布图(图 3), 其中图 3(b)表示了基于  $2^1 \sim 2^{10}$  尺度 Biorthogonal, Coiflets, Daubechies 和 Symlets 小波系 36 个母小波基的  $R^2$  在 350~1 800 nm 波段上的分布(相邻母小波基之间小刻度表示尺度, 刻度间隔为 2)。由图可知, DSR 及 DSR-CWC 与失叶率均存在显著相关。对 DSR 来说, 在 505~535, 550~590, 600~620, 680~760, 1 126~1 188, 1 276~1 340, 1 530~1 576, 1 730~1 768 nm 波段内其  $R^2$  均大于 0.64, 其中 560~580 和 713~

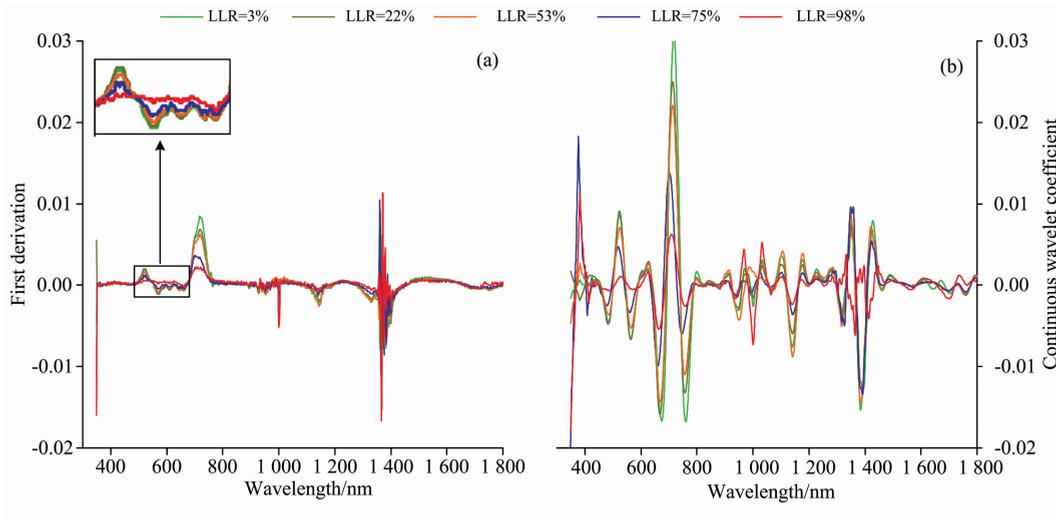


图 2 不同失叶率落叶松冠层的 DSR 曲线 (a) 和 DSR-CWC 曲线 (b)

Fig. 2 (a) DSR curve and (b) DSR-CWC curve of different LLR for larch canopy

718 nm 波段内其  $R^2$  达到 0.8 以上 [图 3(a)]。这是因为 560~580 nm 波段对绿色有强反射, 当失叶率增加时, 落叶松针叶数量减少, 冠层绿色明显减弱进而使光谱的绿色反射逐渐衰减。而 713~718 nm 为红边波段, 当失叶率上升时, 落叶松叶面积减少, 叶绿素含量下降, 红边波段色素吸收减弱导致其光谱反射增强。对 DSR-CWC 而言, 在综合顾及 10 个尺度的前提下, 除了 bior3.1 其余母小波基函数的小波系数基本在可见光区(420~760 nm)和红外区的近红外反射肩(761~1 350 nm)、近红外吸收谷(1 416~1 500 nm)、中红外反射峰(1 501~1 795 nm)等波段内与失叶率均存在极显著相关, 其中 3 个吸收谷(440~515, 630~760 和 1 420~1 470 nm)和 3 个反射峰(516~620, 760~1 000 和 1 548~1 610 nm)波段内其  $R^2$  达到 0.8 以上。这是因为林木冠层反射率受叶色素、叶片内部结构和水分含量等因子的控制, 如: 440~515 和 630~760 nm 波段主要吸收叶绿素 a、b, 1 420~1 470 nm 波段主要吸收针叶水分, 这些波段通常反射率较低, 但是当失叶率增加时, 林木叶绿素和水分含量随之减少, 致使这些波段的林木冠层反射率上升。516~620 nm 对绿色具有强反射特性, 当落叶松被雅氏落叶松尺蠖侵染时, 林木失叶率增加, 冠层颜色由绿色变红色, 导致该波段反射减弱。761~1 000 nm 波段光谱除了很小部分的吸收, 大部分为反射, 使得其反射率较高。当落叶松受害程度加剧时, 针叶细胞结构被破坏, 失叶量增加, 叶面积减少, 致使该波段反射率降低。1 548~1 610 nm 波段反射率受 1 450 和 1 940 nm 的强吸收叶片水分特征的影响, 它是这两个吸收光谱位置中间形成的反射峰, 当林木失叶率增加时, 叶片水分下降, 使得 1 450 和 1 940 nm 的水分吸收减弱, 进而导致 1 548~1 610 nm 波段反射率相对上升。连续小波变换能够增强上述三个反射峰和三个吸收谷光谱特征。因此, 这些波段光谱反射率对应的 DSR-CWC 与失叶率存在了极显著相关性。可见, DSR-CWC 与 DSR 相比对失叶率的敏感性更明显, 可作为估测失叶率的重要指标。

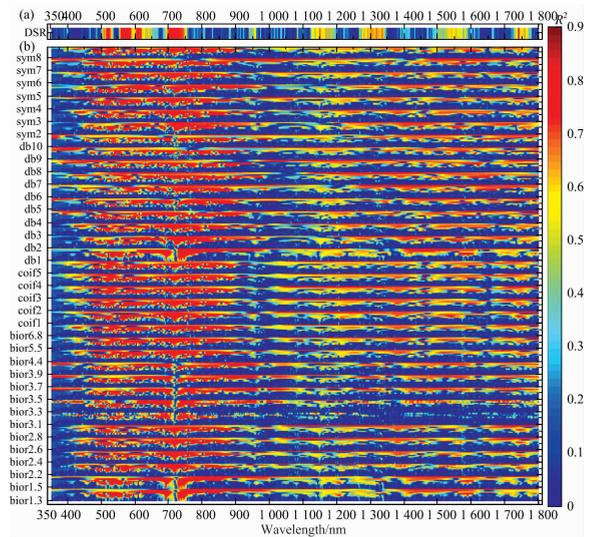


图 3 DSR (a) 及 DSR-CWC (b) 与失叶率的相关关系图

Fig. 3 Correlation diagram between (a) DSR and LLR, and (b) DSR-CWC and LLR

### 2.2 基于 Fp-SPA 的敏感 DSR-CWC 提取

利用 Fp 和 Fp-SPA 提取的 DSR 和 DSR-CWC 的敏感波段如图 4 所示。由图可知, 采用 Fp 自动获取的敏感波段数较多, 如 DSR: 16 个、Biorthogonal 的 14 个母小波基: 55~220 个、Coiflets 的 5 个母小波基: 53~104 个、Daubechies 的 10 个母小波基: 54~146 个和 Symlets 的 7 个母小波基: 57~96 个。显然与 DSR 相比, DSR-CWC 的敏感波段更多。为压缩模型输入变量冗余和减少输入变量多重共线性, 在基于 Fp 的敏感波段提取的基础上, 利用 SPA 算法对其进一步处理, 获得最终敏感波段。结果显示, DSR 和 DSR-CWC 的最终敏感波段数为 DSR: 4 个、Biorthogonal: 2~13 个、Coiflets: 2~17 个、Daubechies: 2~13 个和 Symlets: 1~13 个。降维处理后, DSR 的敏感波段为 714, 941, 1 296 和 1 759

nm, 而 DSR-CWC 的敏感波段更好捕捉到了光谱的吸收和反射特征, 如 Biorthogonal 的 bior2.4 的敏感波段为 440, 454, 490, 518, 543, 599, 685, 717, 1 059, 1 151, 1 539, 1 552 和 1 748 nm, Coiflets 的 coif2 的敏感波段为 490, 526, 568, 642, 708, 718, 739, 780, 829, 971, 1 222, 1 433, 1 599, 1 749 和 1 757 nm, Daubechies 的 db2 的敏感波段为 470, 601, 688, 720, 756, 830, 920, 1 301, 1 548, 1 762 和 1 767 nm, Symlets 的 sym4 的敏感波段为 470, 518, 538, 567, 718, 738, 1 196, 1 221, 1 586 和 1 760 nm。可见, 上述波段较好捕获了蓝光波段(440, 454 和 470 nm)和红光波段(642, 685, 688, 708, 717, 718, 720, 738, 739 和 756 nm)对叶绿素的强吸收特征、绿光波段(518, 526, 538, 543, 567 和 568 nm)对叶绿素的反射特征、近红外波段(780, 829, 830, 920 和 971 nm)对针叶细胞组织的强反射特征、短波红外波段(1 433 nm)对针叶水分含量的吸收特征以及中红外波段(1 548, 1 552, 1 586 和 1 599 nm)的反射特征。这与敏感性分析结果基本相吻合。显然, Fp-SPA 模式能够有效捕获可见光区、近红外和中红外区的 DSR-CWC 的敏感波段。

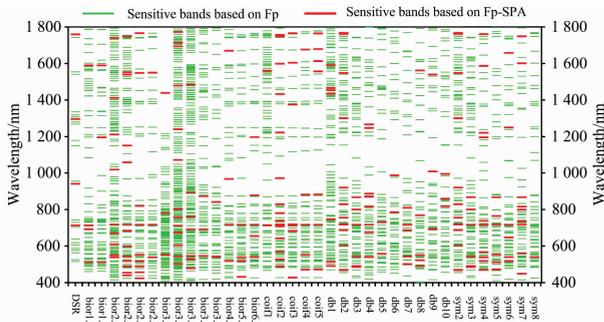


图 4 DSR 和 DSR-CWC 的敏感波段分布

Fig. 4 Sensitive band distribution of DSR and DSR-CWC

### 2.3 失叶率估测模型及精度评价

从 66 棵样本树的光谱和失叶率数据中, 选取训练区的 46 棵样本树数据作为建模数据, 验证区的 20 棵样本树数据作为精度验证数据。通过 MATLAB 软件平台, 利用 DSR 和 4 种小波系 36 个母小波基的 DSR-CWC, 分别构建了 PLSR 和 SVMR 失叶率估测模型, 并与 SMLR 模型进行了估测精度比较(见表 1)。结果表明:

在基于 Biorthogonal 小波系 14 个母小波基 DSR-CWC 的不同模型中利用 bior2.4 的 PLSR(bior2.4-PLSR)模型估测精度最优( $R_M^2=0.9332$ ,  $RMSE_M=0.0885$ ), 与 SMLR 相比其  $R_M^2$  提高了 0.0773,  $RMSE_M$  降低了 0.0391。在基于 Coiflets 小波系 5 个母小波基 DSR-CWC 的不同模型中采用 coif2 的 SVMR(coif2-SVMR)模型估测精度最优( $R_M^2=0.9274$ ,  $RMSE_M=0.0958$ ), 与 SMLR 相比其  $R_M^2$  提高了 0.0655,  $RMSE_M$  降低了 0.0226。在基于 Daubechies 小波系 10 个母小波基 DSR-CWC 的不同模型中运用 db1 的 PLSR(db1-PLSR)模型估测精度最优( $R_M^2=0.9340$ ,  $RMSE_M=0.0890$ ), 与 SMLR 相比其  $R_M^2$  提高了 0.0092,  $RMSE_M$  降低了 0.0156。在基于 Symlets 小波系 7 个母小波基 DSR-

CWC 的不同模型中利用 sym6 的 PLSR(sym6-PLSR)模型估测精度最优( $R_M^2=0.9225$ ,  $RMSE_M=0.0958$ ), 与 SMLR 相比其  $R_M^2$  提高了 0.0618,  $RMSE_M$  降低了 0.0275。

在 PLSR 和 SVMR 模型中, 利用 DSR-CWC 比 DSR 估测精度有明显提高, 具体表现为: bior2.4-PLSR, coif2-PLSR, db1-PLSR 和 sym6-PLSR 的  $R_M^2$  比 DSR-PLSR 分别提高了 0.0467, 0.0339, 0.0475 和 0.0360,  $RMSE_M$  分别降低了 0.0254, 0.0170, 0.0249 和 0.0181; bior2.4-SVMR, coif2-SVMR, db1-SVMR 和 sym6-SVMR 的  $R_M^2$  比 DSR-SVMR 分别提高了 0.0196, 0.0195, 0.0251 和 0.0096,  $RMSE_M$  分别降低了 0.0200, 0.0153, 0.0188 和 0.0126。

基于 bior2.4, coif2, db1 和 sym6 的 PLSR 和 SVMR 模型估测精度相对类似, 其  $R_M^2$  均大于 0.91, 其中 db1-PLSR 模型精度相比之下达到最高。在 36 个母小波基中 db1 对 PLSR, SVMR 和 SMLR 3 种模型均表现出最好的估测效果, 说明 db1 为精度和稳定性兼备的最有估测潜力的母小波基。

表 1 不同模型估测精度比较

Table 1 Comparison of estimation accuracies among different models

Accuracy index	$R_M^2$			RMSE <sub>M</sub>		
	PLSR	SVMR	SMLR	PLSR	SVMR	SMLR
DSR	0.8865	0.9079	0.8862	0.1139	0.1111	0.1169
bior2.4	0.9332	0.9275	0.8559	0.0885	0.0911	0.1276
coif2	0.9204	0.9274	0.8549	0.0969	0.0958	0.1195
db1	0.9340	0.9330	0.9248	0.0890	0.0923	0.1046
sym6	0.9225	0.9175	0.8607	0.0958	0.0985	0.1233

为进一步揭示模型精度和稳定性, 利用 DSR 和 DSR-CWC 构建的不同模型建模集和验证集的估测值与实测值之间分别进行 1:1 直线拟合分析, 如图 5 所示。由图 5 得知, 基于 bior2.4, coif2, db1 和 sym6 的 DSR-CWC 与 DSR 相比, 其 PLSR 和 SVMR 的建模集和验证集的大部分数据点分布在 1:0.5 和 1:1.5 直线范围内且较均匀分布在 1:1 直线两侧, 说明利用 DSR-CWC 建立的两种模型拟合效果均较好, 估测精度较高。对 DSR-CWC 而言, PLSR 和 SVMR 与 SMLR 相比, 分别在 bior2.4, coif2, db1 和 sym6 中建模集和验证集数据点均表现较好拟合效果, 而 SMLR 的验证集数据点离散程度较明显, 表明利用 DSR-CWC 构建的 PLSR 和 SVMR 模型比 SMLR 模型稳定性好并精度高。

## 3 结论

以受雅氏落叶松尺蠖危害的落叶松为研究对象, 利用 66 棵样本落叶松冠层光谱数据和失叶率数据, 通过 Fp-SPA 模式捕捉 DSR 和 DSR-CWC 的敏感特征建立 PLSR 和 SVMR 失叶率估测模型, 得出了以下结论:

(1) DSR-CWC 与 DSR 相比对受害落叶松失叶率变化的敏感性更显著并且敏感波段更多。除了 bior3.1 其余 35 个母

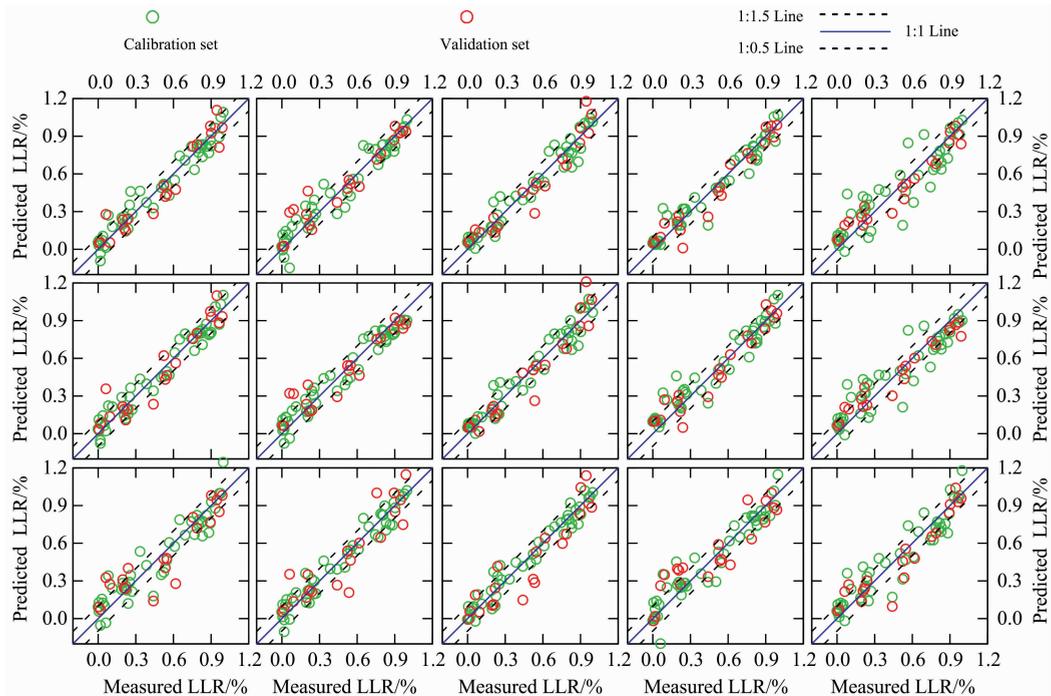


图 5 不同模型的 1 : 1 直线拟合效果比较

Fig. 5 Comparison of 1 : 1 Line fitting effects among different models

小波基的 DSR-CWC 的敏感波段主要分布于三个吸收谷 (440 ~ 515, 630 ~ 760 和 1 420 ~ 1 470 nm) 和三个反射峰 (516 ~ 620, 761 ~ 1 000 和 1 548 ~ 1 610 nm) 波段范围内。可见, 对雅氏落叶松尺蠖危害下的落叶松失叶, DSR-CWC 能够较好挖掘叶色素、叶片内部结构和水分含量变化引起的光谱响应特征。

(2) Fp-SPA 模式是一种选择敏感光谱小波特征的有效方法。它不仅能够快速、客观选择敏感特征, 而且对特征有效降维, 即由上百个特征降维到几个关键特征, 故可减少模型计算量还能够提高模型稳定性。

(3) 在基于 DSR-CWC 的估测失叶率中, Biorthogonal, Coiflets, Daubechies 和 Symlets 等 4 种小波系的最优母小波基分别为 bior2.4, coif2, db1 和 sym6, 其中 db1 比其余母小波基对 PLSR, SVMR 和 SMLR 模型均表现出较高估测精度, 且在 1 : 1 直线上其模型的建模集和验证集数据点拟合效果较好, 说明所有母小波基中 db1 的估测失叶率性能最为

稳定。

(4) 与 DSR 相比, DSR-CWC 提高了失叶率估测精度。在 PLSR、SVMR 和 SMLR 三种模型中 db1-PLSR 对估测精度的提高最为明显, 在 1 : 1 直线上比 DSR-PLSR 表现出更好的拟合效果, db1-PLSR 的  $R_M^2$  和  $RMSE_M$  分别为 0.934 0 和 0.089 0, 比 DSR-PLSR 的  $R_M^2$  提高了 0.047 5, 比 DSR-PLSR 的  $RMSE_M$  降低了 0.024 9。显然利用 DSR-CWC 估测失叶率更有潜力。

(5) 基于 DSR-CWC 的 PLSR 和 SVMR 模型失叶率估测精度相对类似, 且优于 SMLR 模型, 尤其 bior2.4-PLSR 和 coif2-SVMR 较为明显。bior2.4-PLSR 与 bior2.4-SMLR 相比其  $R_M^2$  提高了 0.077 3,  $RMSE_M$  降低了 0.039 1, coif2-SVMR 与 coif2-SMLR 相比其  $R_M^2$  提高了 0.065 5,  $RMSE_M$  降低了 0.022 6。说明在雅氏落叶松尺蠖危害下, 利用 DSR-CWC 的 PLSR 和 SVMR 模型的失叶率估测精度具有较高可靠性, 对森林虫害监测具有参考价值。

## References

- [1] White J, Wulder M, Brooks D, et al. Remote Sensing of Environment, 2005, 96(3-4): 340.
- [2] HUANG Xiao-jun, XIE Yao-wen, BAO Yu-hai(黄晓君, 颀耀文, 包玉海). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2018, 38(3): 905.
- [3] Rullansilva C D, Olthoff A E, Delgado d l M J A, et al. Forest Systems, 2013, 22(3): 377.
- [4] Wulder M A, Dymond C C, White J C, et al. Forest Ecology & Management, 2006, 221(1-3): 27.
- [5] Niemann K O, Quinn G, Stephen R, et al. Canadian Journal of Remote Sensing, 2015, 41 (3): 191.
- [6] Hawryło P, Bednarz B, Weżyk P, et al. European Journal of Remote Sensing, 2018, 51(1): 194.
- [7] Rahimzadeh-Bajgiran P, Weiskittel A, Kneeshaw D, et al. Forests, 2018, 9 (6): 357.
- [8] Spruce J P, Sader S, Ryan R E, et al. Remote Sensing of Environment, 2011, 115(2): 427.

- [ 9 ] Townsend P A, Singh A, Foster J R, et al. *Remote Sensing of Environment*, 2012, 119(8): 255.
- [10] Adelabu S, Mutanga O, Adam E, et al. *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, 2013, 7(1): 177.
- [11] Fassnacht F E, Latifi H, Ghosh A, et al. *Remote Sensing of Environment*, 2014, 140(1): 533.
- [12] Cheng T, Rivard B, Sánchezzofeifa G A, et al. *Remote Sensing of Environment*, 2010, 114(4): 899.
- [13] Lausch A, Heurich M, Gordalla D, et al. *Forest Ecology & Management*, 2013, 308(4): 76.
- [14] Shi Y, Huang W J, González-Moreno P, et al. *Remote Sensing*, 2018, 10(4): 525.

## Estimation of Leaf Loss Rate in Larch Infested with *Erannis Jacobsoni* Djak Based on Differential Spectral Continuous Wavelet Coefficient

HUANG Xiao-jun<sup>1, 2, 3, 4</sup>, XIE Yao-wen<sup>1\*</sup>, BAO Yu-hai<sup>2, 3</sup>, BAO Gang<sup>2, 3</sup>, QING Song<sup>2, 3</sup>, BAO Yu-long<sup>2, 3, 4</sup>

1. College of Earth and Environmental Sciences, Lanzhou University, Lanzhou 730000, China

2. College of Geographical Science, Inner Mongolia Normal University, Huhhot 010022, China

3. Inner Mongolia Key Laboratory of Remote Sensing & Geography Information System, Huhhot 010022, China

4. Inner Mongolia Key Laboratory of Disaster and Ecological Security on the Mongolia Plateau, Huhhot 010022, China

**Abstract** Defoliation caused by insect pests severely threatens the health and safety of forests; the rapid and accurate acquisition of information regarding leaf loss is of considerable significance to the remote sensing monitoring and estimation of forest pests. Based on this, we conducted spectral measurements of infested trees and tested leaf loss rate estimation owing to larch defoliation caused by *Erannis jacobsoni* Djak in Mongolia. Differential spectral reflectance (DSR, first derivative of spectral reflectance) and continuous wavelet coefficient of differential spectral reflectance (DSR-CWC, continuous wavelet transform of DSR carried out using 36 mother wavelet basis functions of four wavelet families: biorthogonal, coiflets, daubechies and symlets) were obtained based on the processing of spectral measurement data. The sensitivity of DSR and DSR-CWC with respect to the estimation of leaf loss rate was analyzed, following which the sensitive bands of DSR and DSR-CWC were automatically identified using the Findpeaks (Fp) function of MATLAB and the sensitive features identified. Dimension reduction of the sensitive features was processed using a successive projections algorithm (SPA). Partial least squares regression (PLSR) and support vector machine regression (SVMR) models for estimating leaf loss rate were established based on these sensitive features and their effectiveness was compared with that of stepwise multiple linear regression (SMLR) models. The results showed that: ①DSR-CWC was determined to be more sensitive than DSR to changes in leaf loss rate in infested larch, with more sensitive bands, mainly distributed in three absorption valleys (440~515, 630~760 and 1 420~1 470 nm) and three reflection peaks (516~620, 761~1 000 and 1 548~1 610 nm). This finding reflects the fact that DSR-CWC can enhance spectral reflection and absorption characteristics. ②The use of the combination pattern of Fp and SPA (Fp-SPA) was an effective method for the selection of sensitive spectral features that could not only select these features quickly and objectively but also effectively reduce dimensions. ③The optimal mother wavelet bases for the four wavelet families respectively were bior2. 4, coif2, db1, and sym6; db1 had the most stable performance and accuracy for leaf loss rate estimation. ④The continuous wavelet transform of DSR could improve the accuracy of leaf loss estimation; db1-PLSR ( $R_M^2=0.934\ 0$ ,  $RMSE_M=0.089\ 0$ ) exhibited the most obvious improvement, achieving an  $R_M^2$  that was 0.047 5 higher than that of DSR-PLSR and an  $RMSE_M$  that was 0.024 9 lower than that of DSR-PLSR. ⑤The estimation accuracy of the PLSR and SVMR models established based on DSR-CWC was either similar to or better than that of the SMLR models. DSR-CWC thus estimated leaf loss rate more effectively than DSR did. It can be seen that DSR-CWC has more potential than DSR in estimating leaf loss rate, and it can provide important reference for remote sensing monitoring of forest pests.

**Keywords** *Erannis jacobsoni* Djak; Leaf loss rate of larch; Differential spectral continuous wavelet coefficient; Findpeaks function; Continuous projection algorithm

\* Corresponding author

(Received Mar. 28, 2019; accepted Jun. 6, 2019)