doi:10.6041/j.issn.1000-1298.2019.04.022

基于高光谱特征的松材线虫岭回归估测模型研究

张素兰[1,2]　黄金龙[1]　秦　林[3]　李宏群[4]

(1.长江师范学院大数据与智能工程学院,重庆 408100;2.昆士兰大学园艺科学中心,布里斯班 4072;
3.长江师范学院电子信息工程学院,重庆 408100;4.长江师范学院三峡库区环境监测与灾害防治工程研究中心,重庆 408100)

摘要:以 2017 年 6—8 月获取的重庆永胜林场马尾松光谱反射率为数据源,对绿光区(490 ~ 560 nm)、黄光区(560 ~ 590 nm)、红光区(620 ~ 680 nm)、红边(680 ~ 780 nm)、近红外区(780 ~ 1 100 nm)最大反射率和反射率总和、绿峰(500 ~ 670 nm)反射高度、红谷(560 ~ 760 nm)吸收深度等 14 个高光谱特征参数进行岭迹分析,筛选出非共线性特征参数,构建松材线虫岭回归估测模型。结果表明:红边和近红外区反射率最大值、红边和近红外区反射率总和、红谷吸收深度岭迹曲线变化稳定且不趋于零,可用于岭回归建模。当岭迹参数 $k=0.2$ 时,上述 5 个高光谱特征参数岭迹趋于稳定,根据 k 值计算岭回归系数,构建松材线虫岭回归估测模型。模型决定系数 R^2 为 0.868 6,均方根误差 RMSE 为 0.273 5,平均估测精度为 87.15%,可为松材线虫病害早期监测和防治研究提供技术支持。

关键词:松材线虫;马尾松;高光谱特征;岭回归;估测模型

中图分类号:S763;TP79　　**文献标识码**:A　　**文章编号**:1000-1298(2019)04-0196-07

Ridge Regression Model for Estimating Pine Wilt Disease Based on Hyperspectral Characteristics

ZHANG Sulan[1,2]　HUANG Jinlong[1]　QIN Lin[3]　LI Hongqun[4]

(1. *Institute of Big Data and Intelligent Engineering*, *Yangtze Normal University*, *Chongqing* 408100, *China*
2. *Centre for Horticultural Science*, *The University of Queensland*, *Brisbane* 4072, *Australia*
3. *College of Electronic Information Engineering*, *Yangtze Normal University*, *Chongqing* 408100, *China*
4. *Hyperspectral Remote Sensing Monitoring Center for Ecological Environment of the Three Gorges Reservoir Area*, *Yangtze Normal University*, *Chongqing* 408100, *China*)

Abstract: Pine wilt disease (PWD) caused by the pine wood nematode, *Bursaphelenchus xylophilus*, is considered as the most destructive forest-invasive alien species and may cause serious economic losses. A ridge regression model was proposed based on the hyperspectral characteristics to estimate the degrees of pine wilt disease for *Pinus massoniana* in Yongsheng forest of Chongqing, Southwest China. The spectral reflectance and quantitated pet levels for *Pinus massoniana* were measured from June to August 2017. And then the ridge trace analysis was operated on 14 spectral characteristics, which covered maximum and sum of reflectance ranging in green region (490 ~ 560 nm), yellow region (560 ~ 590 nm), red region (620 ~ 680 nm), red edge (680 ~ 780 nm), near-infrared region (780 ~ 1 100 nm), as well as the reflectance height of green peak (500 ~ 670 nm) and absorption depth of red valley (560 ~ 760 nm). Furthermore, the hyperspectral characteristic parameters with less collinearity were selected to construct the estimation model of PWD with ridge regression. The results demonstrated that ridge trace curves for the maximum of reflectance in red edge, near-infrared region, the sum of reflectance in the red edge, near-infrared region, as well as absorption depth of red valley were stable, which were not close to zero. Therefore, those five spectral characteristics could be considered in ridge regression modeling; when the ridge trace parameter k was 0.2, the ridge traces of the above five hyperspectral characteristic parameters became stable, and then the ridge regression coefficients were calculated. Finally, a regression estimation

收稿日期:2018-10-25　修回日期:2019-02-09

基金项目:国家自然科学基金项目(61601060)、国家留学基金委项目(201709955001)、重庆市科委基础与前沿研究计划项目(cstc2016jcyjA0437)、重庆市教委高校优秀成果转化项目(KJZH17132)和重庆市教委科学技术研究项目(KJ1501201)

作者简介:张素兰(1984—),女,副教授,博士,昆士兰大学访问学者,主要从事农业遥感技术研究,E-mail:slzhang@cqu.edu.cn

通信作者:黄金龙(1989—),男,讲师,博士,主要从事大数据处理技术研究,E-mail:h.jinlong@yznu.edu.cn

model of PWD was built with determination coefficient R^2 of 0.868 6, root-mean-square error (RMSE) of 0.273 5, and average estimation accuracy of 87.15%. The research provided both scientific support and application reference for monitoring forest pet disease with remote sensing technology.

Key words: pine wilt disease; *Pinus massoniana*; hyperspectral characteristic; ridge regression; evaluation model

0　引言

松材线虫病是由松材线虫寄生在松树上引起的毁灭性病害,1982 年首次传入中国,并被确定为我国主要的森林病虫害[1-2]。松树感染初期,感病枝梢针叶由绿变黄再到红褐色,并由局部发展到整树针叶出现萎蔫,直到全株枯萎死亡。松树一旦感染松材线虫病,整株松树最快 40 d 左右即可枯死,成片松林从最初少数死树到林相被毁只需 5 年左右[3]。因此,及时准确地掌握松材线虫病发生的地理位置、面积、危害等级等,对于松材线虫病的防治尤为重要[4]。目前监测技术手段仍是现场观察,费时费力,且获得的信息具有滞后性,使得松材线虫病很难彻底控制和根除,严重破坏了森林生态系统[5]。

近年来,随着信息技术的发展,高光谱遥感技术作为一种新的监测手段被广泛运用于植物生长监测中[6-8]。当植物受到病虫危害时,绿叶中细胞活性、含水率和叶绿素等生理指标都会变化,引起反射光谱特性上的差异,特别是红色区和近红外区的光谱特征的差异[9]。因此,基于高光谱遥感数据,分析病虫害危害程度与原始光谱、光谱特征、植被指数等之间的关系,确定不同种类病虫害监测的敏感波段,建立反演模型,可对森林病虫害进行有效监测[10]。NASI 等[11]利用光谱指数和 K 近邻分类器估测云杉受皮甲虫危害程度,对于判别健康和死亡两种类型的总体准确率可达 90%。CALDERON 等[12]根据生化参数和病害指数,利用支持向量机(SVM)和线性判别分析(LDA)分类方法对橄榄树黄萎病进行分类,其中 LDA 在初始和低严重程度上分类效果较优,分类准确率可达到 71.4% 和 75.0%。

在松材线虫病害高光谱遥感监测方面,国内外学者主要研究了敏感波段、植被指数及光谱特征检测松材线虫病害[13-16]。KIM 等[13]基于现有 10 个高光谱植被指数,引入了绿边与红边面积指数,判断是否感染松材线虫病。JU 等[14]研究了健康马尾松感染松材线虫病的最佳光谱波段或波段组合,发现 759 nm 处的波谱的一阶导数为区别马尾松是否染病的最有效波段。黄明祥等[15]对受松材线虫感染的马尾松树的时间序列光谱变化进行了分析,得出近红外最大一阶导数(FD)值、红边 FD 值和蓝色边缘 FD 值之和的比值可作为松材线虫检测的有效因素。张衡等[16]分析了波长 593 nm 处光谱反射率的一阶导数光谱特征,结合叶绿素质量分数,判断马尾松是否感染了松萎蔫病,对肉眼可见感病特征前的早期阶段具有良好的监测效果。

研究表明,松树光谱特征与其健康程度具有较大的相关性,但利用其相关性估测松林感染松材线虫病等级的研究鲜有报道。本文以重庆永胜林场马尾松为研究对象,利用 70 个采样区的野外高光谱数据,分析松材线虫危害下马尾松绿针的光谱特征,构建最大反射率、反射率总和、绿峰反射率高度、红谷吸引深度及其比率和归一化值等 14 个特征参数。利用岭迹分析筛选最优特征参数,进行岭回归建模反演感染松材线虫病害等级,并对模型估算精度进行验证,以期为实现松材线虫害早期监测和防治提供支持。

1　数据获取

1.1　数据采集

重庆松材线虫病从 6 月初开始发病, 7 月发生的病树最多,8 月逐渐下降,本研究采样时间为 2017 年 6—8 月,共设置 70 个采样区。从各个采样区内选取受害程度不一的典型松树针叶进行测量。数据采集过程中使用 FieldSepc4 型野外光谱分析仪,作用波段为 400 ~ 1 100 nm,光谱分辨率为 1 nm。根据实际需求,测量过程中设定光谱平均次数为 5 次,暗电流平均次数为 10 次,白板平均次数为 5 次,共采集了 70 个马尾松植株的有效光谱数据。使用 FieldSepc4 型野外光谱分析仪配套软件 RS3,优化 FieldSpec4 型仪器采集的原始计数、辐射亮度/辐射照度等相关数据,计算其光谱反射率。

1.2　数据处理

对 70 个光谱数据进行归一化处理,计算式为

$$\rho_i = \frac{R_i - R_{\min}}{R_{\max} - R_{\min}} \tag{1}$$

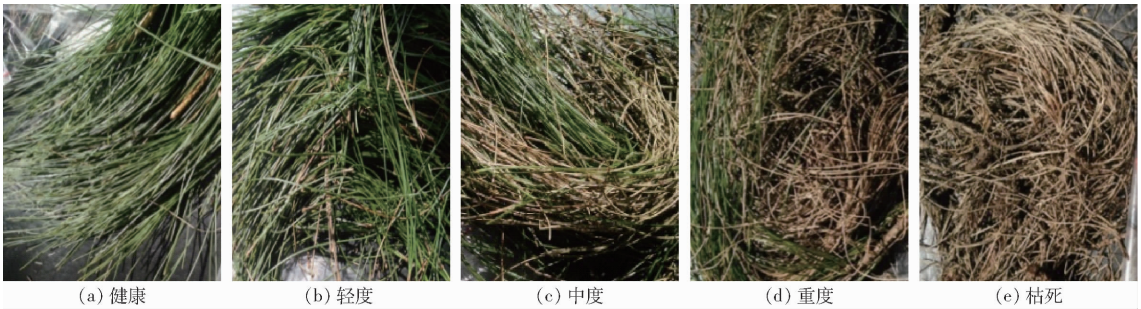
式中　ρ_i——样本 i 的归一化光谱反射率

R_i——样本 i 光谱反射率

$R_{\max}$、$R_{\min}$——样本最大、最小光谱反射率

经采集样本专业人员鉴定,按照松树针叶受害

程度,依次划分为健康、轻度、中度、重度和枯死 5 个等级,如图 1 所示[17],并量化为虫害等级 1,2,…,5,从而得到样本光谱数据及虫害等级数据,如表 1 所示。



(a) 健康 (b) 轻度 (c) 中度 (d) 重度 (e) 枯死

图 1 不同受害程度的松树针叶

Fig. 1 Needle leaves suffering from different levels of disease

表 1 70 个采样区数据

Tab. 1 Data for 70 sampling areas

样本号	坐标		海拔/m	光谱反射率			虫害等级
	经度/(°)	纬度/(°)		ρ_{400}	…	ρ_{1100}	
1	29.806 7	107.232 4	386	0.035 9	…	0.701 2	2
2	29.806 9	107.233 8	427	0.029 3	…	0.657 8	3
3	29.807 2	107.234 6	433	0.013 4	…	0.576 8	3
4	29.807 4	107.235 2	442	0.036 4	…	0.858 8	2
⋮	⋮	⋮	⋮	⋮		⋮	⋮
70	29.806 4	107.234 4	413	0.022 7	…	0.956 1	2

2 研究方法

2.1 光谱特征参数构建

光谱特征是物质在电磁波的作用下,在特定波长位置形成的反映物质成分和结构信息的光谱吸收和反射特征[18]。由于受松材线虫害侵蚀,马尾松针叶色素、含水率以及细胞构造发生剧烈变化,使得其光谱反射率曲线产生明显的变化特征[19]。图 2 为健康、轻度、中度、重度和枯死样本所对应光谱反射率曲线[17]。由图 2 可知,健康样本在可见光 550 nm 有明显的反射峰,在 680 nm 处有明显的吸收谷,在 680 ~ 780 nm 波段(红边)反射率急剧上升,在 780 ~ 1 100 nm 波段(近红外),反射率相对平稳且有最大值。随着虫害程度的加深,样本反射峰、吸收谷、红边反射率陡峭程度、近红外最大反射率降低。特别当样本处于枯死状态时,反射峰和吸收谷明显消失,红边及近红外区反射率被逐渐拉平。这种光谱上的差异使得应用高光谱遥感技术估测病害程度成为可能。

根据文献[13,15,18,20]对波段的划分,选择绿光区(490 ~ 560 nm)、黄光区(560 ~ 590 nm)、红光区(620 ~ 680 nm)、红边(680 ~ 780 nm)、近红外(780 ~ 1 100 nm)波段。计算上述波段范围内光谱反射率最大值和反射率总和,500 ~ 670 nm 波段范围内绿峰反射高度和 560 ~ 760 nm 波段范围内红谷

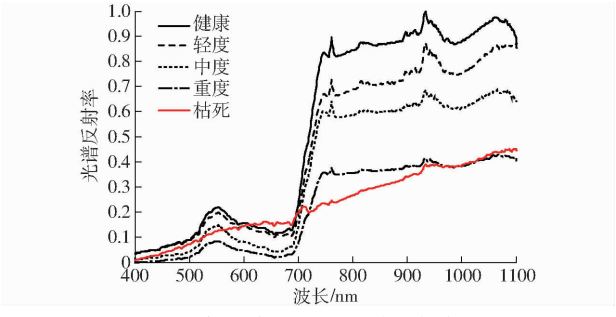


图 2 健康与感病植株光谱反射率曲线

Fig. 2 Spectral reflectance curves for healthy and infected pines

吸收深度、其比值和归一化值,共 14 个特征参数,分别表示为 M_g、M_y、M_r、M_{re}、M_{inr}、S_g、S_y、S_r、S_{re}、S_{inr}、G_H、R_D、D_H、$D_{_H}$,如表 2 所示。由于从光谱特征上可区分枯死样本,因此后续研究只针对健康、轻度、中度、重度 4 种类型。

2.2 岭回归估计模型

在估测模型构建过程中,特征参数筛选尤为关键。上述 14 个特征参数中,有些特征参数相关性较强,不适合全部用于建模。为了提高模型估计准确度,需删除一些相关性较强的特征参数。常用的变量优选方法主要有:基于自适应重加权[21]、多元逐步线性回归[22]和遗传算法[23]等。岭估计是一种改进的最小二乘法,其本质是一种放弃最小二乘法的无偏性与部分精确度,寻求效果稍差但更符合实际情况的回归过程,其在特征提取方面效果较优,并可

表 2　光谱特征参数定义

Tab. 2　Definition of spectral characteristic parameters

变量	参数	公式	定义
X_1	M_g	$M_g = \max(\rho_{490\sim560})$	绿光区(490 ~ 560 nm)反射率最大值
X_2	M_y	$M_y = \max(\rho_{560\sim590})$	黄光区(560 ~ 590 nm)反射率最大值
X_3	M_r	$M_r = \max(\rho_{620\sim680})$	红光区(620 ~ 680 nm)反射率最大值
X_4	M_{re}	$M_{re} = \max(\rho_{680\sim780})$	红边(680 ~ 780 nm)反射率最大值
X_5	M_{inr}	$M_{inr} = \max(\rho_{780\sim1\,100})$	近红外区(780 ~ 1 100 nm)反射率最大值
X_6	S_g	$S_g = \sum \rho_{490\sim560}$	绿光区(490 ~ 560 nm)反射率总和
X_7	S_y	$S_y = \sum \rho_{560\sim590}$	黄光区(560 ~ 590 nm)反射率总和
X_8	S_r	$S_r = \sum \rho_{620\sim680}$	红光区(620 ~ 680 nm)反射率总和
X_9	S_{re}	$S_{re} = \sum \rho_{680\sim780}$	红边(680 ~ 780 nm)反射率总和
X_{10}	S_{inr}	$S_{inr} = \sum \rho_{780\sim1\,100}$	近红外区(780 ~ 1 100 nm)反射率总和
X_{11}	G_H	$G_H = 1 - \dfrac{R_S + \dfrac{R_E - R_S}{\lambda_E - \lambda_S}(\lambda_C - \lambda_S)}{R_C}$	绿峰反射高度，R_S、R_C、R_E分别为反射特征起点、中心点和结束点处的反射率，λ_S、λ_C、λ_E 分别为反射特征起点、中心点和结束点处的波长
X_{12}	R_D	$R_D = 1 - \dfrac{R_C}{R_S + \dfrac{R_E - R_S}{\lambda_E - \lambda_S}(\lambda_C - \lambda_S)}$	红谷吸收深度
X_{13}	D_H	$D_H = R_D / G_H$	红谷吸收深度与绿峰反射高度比值
X_{14}	$D_{_H}$	$D_{_H} = \dfrac{R_D - G_H}{R_D + G_H}$	红谷吸收深度与绿峰反射高度归一化值

简化模型和提高模型的鲁棒性[24]。因此本研究采用岭回归分析优选特征参数，并用于估测模型构建。

多元线性回归模型可表示为

$$\boldsymbol{Y} = \boldsymbol{X}\beta + \boldsymbol{\varepsilon} \tag{2}$$

式中　$\boldsymbol{X}$——特征参数组成的观测阵

$\boldsymbol{Y}$——等级量化值组成的向量阵

β——回归系数　　$\boldsymbol{\varepsilon}$——随机误差

其最小二乘估计系数 $\hat{\beta}$ 为

$$\hat{\beta} = (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y} \tag{3}$$

式中　$\boldsymbol{X}^{\mathrm{T}}$——观测阵 $\boldsymbol{X}$ 的转置矩阵

如果特征参数之间存在较强的共线性，会导致 $|\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}| \approx 0$，使得 $(\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X})^{-1}$ 对角线上的值很大，从而导致 $\hat{\beta}$ 产生较大的变化，产生不准确的估计系数。为了提高估计系数的准确度，借助岭估计剔除多重共线性特征参数。

待定参数的岭估计岭回归系数计算公式为

$$\beta(k) = (\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X} + k\boldsymbol{I}_p)^{-1}\boldsymbol{X}^{\mathrm{T}}\boldsymbol{Y} \tag{4}$$

式中　k——岭迹参数

$\boldsymbol{I}_p$——p 阶单位矩阵

p——建模样本数

为了改善方阵 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ 特征根趋于 0 的情况，式(4)中引入岭迹参数 k，使所求估计系数尽可能达到最优值。将估计系数 β 作为 k 的函数，当 k 在[0, 1]范围内变化时，以 k 为横坐标，$\beta(k)$ 为纵坐标，其在平面直角坐标中所对应的图形即为岭迹。岭迹可直观地反映各特征参数对虫害等级估测的作用和相互关系，借助岭迹分析，可有效地选取影响虫害等级的主要特征参数。

基于岭迹分析筛选特征参数的原则有[24]：①随着岭迹参数 k 值增加，岭回归系数很不稳定的特征参数删除。②随着岭迹参数 k 值增加，岭回归系数迅速趋于零的特征参数删除。③其岭回归系数虽然稳定但其绝对值比较小的特征参数删除。根据岭迹分析，当剩余特征参数的岭迹趋于基本稳定时，取岭迹参数 k 的值，根据岭迹参数 k 计算得特征参数的待定系数，进而可得虫害等级估测模型(式(2))。

3　结果与分析

对于采样的 70 个光谱数据，根据光谱曲线去除枯死样本，剩余 56 个光谱数据。为了使建模数据集和测试数据集可以充分反映研究区松材线虫害程度，将 56 个样本数据按光谱反射率平均值从大到小进行排序，等间隔抽取 35 个作为建模数据集，21 个作为测试数据集。对于用于建模的 35 个样本提取其特征参数与虫害等级，对特征参数组成的自变量 $\boldsymbol{X}$，计算其方阵 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$，然后对其进行主成分分析，分析其是否存在线性关系，计算 $\boldsymbol{X}^{\mathrm{T}}\boldsymbol{X}$ 的特征值及特征向量，结果如表 3 所示。其中 9 个特征参数特征值较小，表明他们之间线性相关性较大，因此需剔除。但需要剔除哪些特征参数，还需借助岭迹图来进行判断。将特征参数组成的自变量 $\boldsymbol{X}$ 进行中心标准化，利用 Matlab 进行岭回归分析，得到各特征参数岭迹如图 3 所示。

表 3 方阵 X^TX 特征值

Tab. 3 Eigenvalue of square matrix X^TX

变量	参数	特征值	变量	参数	特征值
X_1	M_g	0	X_8	S_r	0.024 5
X_2	M_y	0	X_9	S_{re}	0.418 9
X_3	M_r	0.000 1	X_{10}	S_{inr}	3.413 0
X_4	M_{re}	0.000 1	X_{11}	G_H	7.416 1
X_5	M_{inr}	0.001 4	X_{12}	R_D	70.972 1
X_6	S_g	0.003 9	X_{13}	D_H	255.478 6
X_7	S_y	0.011 8	X_{14}	$D_{_H}$	1 574 829.252 1

基于岭迹分析筛选特征参数原理，最符合条件的特征参数有 4 个，分别为 X_4（M_{re}）、X_5（M_{inr}）、X_9（S_{re}）、X_{10}（S_{inr}），其岭迹变化平稳且不趋于 0，其次是 X_{12}（R_D）。分别计算健康、轻度、中度、重度 4 种样本（图 2）特征参数 X_4、X_5、X_9、X_{10}、X_{12}，如表 4 所示。由表 4 可知，随着虫害程度的加深，特征参数 X_4、X_5、X_9、X_{10}逐渐减小，与虫害程度呈负相关；特征参数 X_{12}总体呈上升趋势，与虫害程度呈正相关，与岭迹分析结果一致。由此可见，红谷、红边和近红外区域是反映虫害程度的敏感区域，因此利用高光谱特征估算虫害程度是可行的。

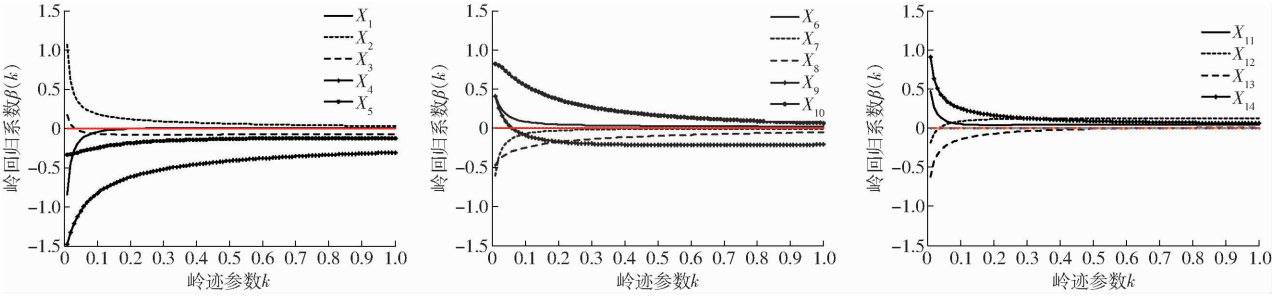


图 3 特征参数岭迹图

Fig. 3 Ridge traces for characteristic parameters

表 4 不同虫害程度样本特征参数

Tab. 4 Values of characteristic parameters for samples in different healthy levels

特征参数	健康	轻度	中度	重度
X_4（M_{re}）	0.897 0	0.728 1	0.640 4	0.377 2
X_5（M_{inr}）	1.000 0	0.874 1	0.685 9	0.429 7
X_9（S_{re}）	57.250 1	46.298 7	39.993 1	23.460 5
X_{10}（S_{inr}）	297.757 0	254.346 1	207.522 7	128.006 6
X_{12}（R_D）	0.790 7	0.772 2	0.890 1	0.907 8

选取特征参数 X_4、X_5、X_9、X_{10}、X_{12}，由图 3 可知，当 $k=0.2$ 时，上述 5 个特征参数趋于稳定。因此取 $k=0.2$ 进行岭回归建模，得到基于 5 个特征参数的虫害等级估测模型为

$$Y=4.554\,4-0.261\,2X_4-0.011\,8X_5-0.069\,3X_9-0.002\,7X_{10}+1.791\,6X_{12} \tag{5}$$

使用决定系数 R^2 及均方根误差 RMSE 评价估测模型预测能力。R^2 用于判定模型的稳定程度，其值越接近 1 则估测模型的稳定性越好；RMSE 用于表征模型的准确度，其值越小则估测模型精度越高。

为了比较基于特征参数 X_4、X_5、X_9、X_{10}和 X_{12}的岭回归估测模型稳定性和准确度，实验中分别对 14 个特征参数进行二进制编码 $G=[X_1\quad X_2\quad \cdots\quad X_{14}]$。当特征参数 X_i用于建模时，将其编码为 1，否则为 0。当每个特征参数分别编码为 0、1 时，共构建 $2^{14}-1$ 个岭回归估测模型，其均方根误差（RMSE）统计如图 4 所示。由图 4 可知，基于 X_4、X_5、X_9、X_{10}、X_{12} 构建的估测模型 $R^2=0.868\,6$，RMSE 为 0.273 5；基于 14 个特征参数的多元岭回归估测模型 $R^2=0.859\,5$，RMSE 为 0.288 0，且其余估测模型的 RMSE 均大于 0.273 5，表明岭迹分析能够优选特征参数。

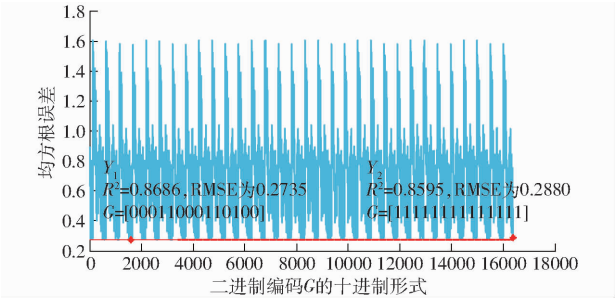


图 4 不同特征参数对应的岭回归估测模型均方根误差（RMSE）

Fig. 4 Root-mean-square error (RMSE) of ridge regression models with different characteristic parameters

将剩余 21 个样本数据使用式（5）进行估测，所得到的结果进行验证，验证计算式为

$$E_i=\left(1-\left|\frac{Y_{ei}-Y_{ri}}{Y_{ri}}\right|\right)\times 100\% \tag{6}$$

式中 E_i——样本 i 的估测精度

Y_{ri}——样本 i 的实测数据

Y_{ei}——样本 i 的估测数据

其结果如表 5 及图 5 所示。由表 5 可知，模型 Y_1决定系数 R^2 为 0.829 4，均方根误差 RMSE 为 0.328 0，平均精度为 87.15%，与模型 Y_2相比其决定

系数与精度均较大，而均方根误差较小。从图 5 可知，模型 Y_1 得到的估测数据与实测数据紧密分布在 $y=x$ 周围，估测值与实际值相差不大。因此估测模型岭回归估测模型 Y_1 具有较好的估测能力。

表 5　估测模型验证结果

Tab. 5　Validation results of estimation model

模型	R^2	均方根误差	平均精度/%
$Y_1 = 4.554\,4 - 0.261\,2X_4 - 0.011\,8X_5 - 0.069\,3X_9 - 0.002\,7X_{10} + 1.791\,6X_{12}$	0.829 4	0.328 0	87.15
$Y_2 = 0.014\,2 + 0.046\,2X_1 + 0.048\,8X_2 + 0.014\,9X_3 - 0.261\,2X_4 - 0.011\,8X_5 - 0.313\,1X_6 + 0.785\,7X_7 - 0.110\,8X_8 - 0.069\,3X_9 - 0.002\,7X_{10} + 1.731\,5X_{11} + 1.791\,6X_{12} + 2.409\,6X_{13} + 0.061\,9X_{14}$	0.801 0	0.353 5	86.54

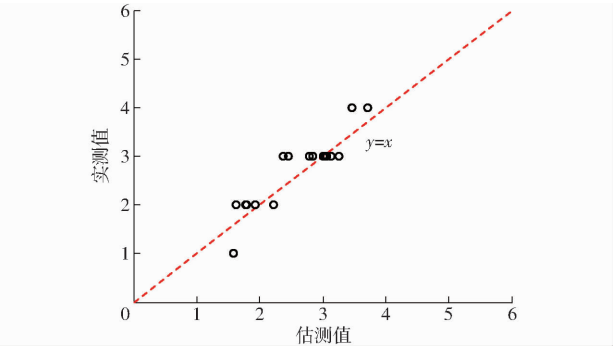


图 5　岭回归估测模型 Y_1 验证

Fig. 5　Validation of ridge regression estimation model Y_1

4　结论

（1）健康样本在绿光区（490 ~ 560 nm）和红光区（620 ~ 680 nm）有明显的反射峰和吸收谷，在红边内（680 ~ 760 nm）反射率急剧上升，在近红外区域内（780 ~ 1 100 nm）反射率相对平稳且有最大值。随着虫害程度的加深，样本反射峰、吸收谷、红边反射率陡峭程度、近红外最大反射率降低。

（2）根据岭迹图可知，红边反射率最大值 X_4、近红外区反射率最大值 X_5，红边反射率总和 X_9、近红外区反射率总和 X_{10} 和红谷吸收深度 X_{12} 岭迹曲线变化稳定且不趋于零，符合岭迹分析筛选标准，因此选取这个 5 个特征参数用于岭回归建模。

（3）当 $k = 0.2$ 时，特征参数 X_4、X_5、X_9、X_{10} 和 X_{12} 岭迹趋于稳定。计算岭回归系数 $\beta(k = 0.2)$，进行岭回归建模，得松材线虫岭回归估测模型，其决定系数为 0.868 6，均方根误差为 0.273 5。

（4）使用测试数据集对岭回归估测模型精度进行验证，并将结果与基于 14 个光谱特征的岭回归估测模型结果进行对比，可得其决定系数为 0.829 4，均方根误差为 0.328 0，平均精度为 87.15%，均优于基于 14 个特征参数的岭回归估测模型。

参　考　文　献

[1] GAO R, WANG Z, SHI J, et al. Effect of *Bursaphelenchus xylophilus* infection on leaf photosynthetic characteristics and resource-use efficiency of *Pinus massoniana*[J]. Ecology & Evolution, 2017, 7(10): 3455 - 3463.

[2] GAO R, SHI J, HUANG R, et al. Effects of pine wilt disease invasion on soil properties and Masson pine forest communities in the Three Gorges reservoir region, China[J]. Ecology & Evolution, 2015, 5(8): 1702 - 1716.

[3] 范军祥，黄焕华，钱明惠，等. 松材线虫病的诊断方法探讨[J]. 广东林业科技，2008，24(5)：52 - 55.
FAN Junxiang, HUANG Huanhua, QIAN Minghui, et al. Study on the diagnotic methods of pinewood nematode [J]. Guangdong Forestry Science and Technology, 2008, 24(5): 52 - 55. (in Chinese)

[4] CARNEGIE A J, VENN T, LAWSON S, et al. An analysis of pest risk and potential economic impact of pine wilt disease to *Pinus plantations* in Australia[J]. Australian Forestry, 2018, 81(1): 24 - 36.

[5] 徐华潮，骆有庆，张廷廷，等. 松材线虫自然侵染后松树不同感病阶段针叶光谱特征变化[J]. 光谱学与光谱分析，2011，31(5)：1352 - 1356.
XU Huachao, LUO Youqing, ZHANG Tingting, et al. Changes of reflectance spectra of pine needles in different stage after being infected by pine wood nematode[J]. Spectroscopy and Spectral Analysis, 2011, 31(5): 1352 - 1356. (in Chinese)

[6] STONE C, MOHAMMED C. Application of remote sensing technologies for assessing planted forests damaged by insect pests and fungal pathogens: a review[J]. Current Forestry Reports, 2017, 3(2): 75 - 92.

[7] 韩文霆，李广，苑梦婵，等. 基于无人机遥感技术的玉米种植信息提取方法研究[J/OL]. 农业机械学报，2017，48(1)：139 - 147.
HAN Wenting, LI Guang, YUAN Mengchan, et al. Extraction method of maize planting information based on UAV remote sensing techonology[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2017, 48(1): 139 - 147. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag = 1&file_no = 20170118&journal_id = jcsam. DOI: 10.6041/j.issn.1000-1298.2017.01.018. (in Chinese)

[8] 吕雅慧，张超，郧文聚，等. 高分辨率遥感影像农田林网自动识别[J/OL]. 农业机械学报，2018，49(1)：157 - 163.

LÜ Yahui, ZHANG Chao, YUN Wenju, et al. Automatic recognition of farmland shelterbelts in high spatial resolution remote sensing data[J/OL]. Transactions of the Chinese Society for Agricultural Machinery, 2018, 49(1):157 - 163. http://www.j-csam.org/jcsam/ch/reader/view_abstract.aspx? flag = 1&file_no = 20180120&journal_id = jcsam. DOI:10.6041/j.issn.1000-1298.2018.01.020. (in Chinese)

[9] 乔红波, 程登发, 孙京瑞, 等. 麦蚜对小麦冠层光谱特性的影响研究[J]. 植物保护, 2005, 31(2):21 - 26.
QIAO Hongbo, CHENG Dengfa, SUN Jingrui, et al. Effects of wheat aphid on spectrum reflectance of the wheat canopy[J]. Plant Protection, 2005, 31(2):21 - 26. (in Chinese)

[10] 汪红,舒清态,吴秋菊,等. 成像光谱遥感在森林病虫害监测中的应用研究[J]. 中国植保导刊, 2018, 38(1):24 - 28.
WANG Hong, SHU Qingtai, WU Qiuju, et al. Application of imaging spectral remote sensing techniques in monitoring of forestry disease and insect pests[J]. China Plant Protection, 2018, 38(1):24 - 28. (in Chinese)

[11] NASI R, HONKAVAARA E, LYYTIKAINEN-SAARENMAA P, et al. Using UAV-based photogrammetry and hyperspectral imaging for mapping bark beetle damage at tree-level[J]. Remote Sensing, 2015, 7(11): 15467 - 15493.

[12] CALDERON R, NAVAS CORTES J A, ZAREO-TEJADA P J. Early detection and quantification of verticillium wilt in olive using hyperspectral and thermal imagery over large areas[J]. Remote Sensing, 2015, 7(5): 5584 - 5610.

[13] KIM S R, LEE W K, LIM C H, et al. Hyperspectral analysis of pine wilt disease to determine an optimal detection index[J]. Forests, 2018, 9(3): 115.

[14] JU Y, PAN J, ZHANG H, et al. Detection of *Bursaphelenchus xylophilus* infection in *Pinus massoniana* from hyperspectral data[J]. Nematology, 2014, 16(10):1197 - 1207.

[15] 黄明祥, 龚建华, 李顺, 等. 松材线虫病害高光谱时序与敏感特征研究[J]. 遥感技术与应用, 2012, 27(6):954 - 960.
HUANG Mingxiang, GONG Jianhua, LI Shun, et al. Study on pine wilt disease hyper-spectral time series and sensitive features[J]. Remote Sensing Technology & Application, 2012, 27(6):954 - 960. (in Chinese)

[16] 张衡, 潘洁, 巨云为, 等. 基于高光谱数据的马尾松松萎蔫病早期监测[J]. 东北林业大学学报, 2014, 42(11):115 - 119.
ZHANG Heng, PAN Jie, JU Yunwei, et al. Early detection of pine wilt disease in *Pinus massioniana* with hyperspectral data[J]. Journal of Northeast Forestry University, 2014, 42(11):115 - 119. (in Chinese)

[17] 张素兰, 覃菊, 唐晓东, 等. 松材线虫危害下马尾松光谱特征与估测模型研究[J]. 光谱学与光谱分析, 2019, 39(3): 865 - 872.
ZHANG, Sulan, QIN Ju, TANG Xiaodong, et al. Spectral characteristics and evaluation model of Pinus Massoniana suffering from *Bursaphelenchus xylophilus* disease[J]. Spectroscopy and Spectral Analysis,2019, 39(3): 865 - 872. (in Chinese)

[18] 尼加提·卡斯木, 师庆东, 王敬哲,等. 基于高光谱特征和偏最小二乘法的春小麦叶绿素含量估算[J]. 农业工程学报, 2017, 33(22):208 - 216.
NIJAT Kasim, SHI Qingdong, WANG Jingzhe, et al. Estimation of spring wheat chlorophyll content based on hyperspectral features and PLSR model[J]. Transactions of the CSAE, 2017, 33(22):208 - 216. (in Chinese)

[19] 王震, 张晓丽, 安树杰. 松材线虫病危害的马尾松林木光谱特征分析[J]. 遥感技术与应用, 2007,22(3):367 - 370.
WANG Zhen, ZHANG Xiaoli, AN Shujie. Spectral characteristics analysis of pinus massoniana suffered by *Bursaphelenchus xylophilus*[J]. Remote Sensing Technology and Application, 2007, 22(3):367 - 370. (in Chinese)

[20] GITELSON A A, MERZLYAK M N, LICHTENTHALER H K. Detection of red edge position and chlorophyll content by reflectance measurements near 700 nm[J]. Journal of Plant Physiology, 1996, 148(3 - 4): 501 - 508.

[21] 于雷, 朱亚星, 洪永胜, 等. 高光谱技术结合 CARS 算法预测土壤水分含量[J].农业工程学报, 2016, 32(22):138 - 145.
YU Lei, ZHU Yaxing, HONG Yongsheng, et al. Determination of soil moisture content by hyperspectral technology with CARS algorithms[J]. Transactions of the CSAE, 2016, 32(22):138 - 145. (in Chinese)

[22] 张德虎, 田海清, 武士钥, 等. 河套蜜瓜糖度可见近红外光谱特征波长提取方法研究[J]. 光谱学与光谱分析, 2015, 35(9): 2505 - 2509.
ZHNAG Dehu, TIAN Haiqing, WU Shiyue, et al. Study on extraction methods of characteristic wavelength of visible near infrared spectroscopy used for sugar content of Hetao muskmelon[J]. Spectroscopy and Spectral Analysis, 2015, 35(9): 2505 - 2509. (in Chinese)

[23] 孙俊, 路心资, 张晓东, 等. 基于高光谱图像的红豆品种 GA - PNN 神经网络鉴别[J/OL]. 农业机械学报, 2016, 47(6): 215 - 221.
SUN Jun, LU Xinzi, ZHANG Xiaodong, et al. Identification of red bean variety with probabilistic GA - PNN based on hyperspectral imaging[J/OL]. Transactions of the Chineses Society for Agricultural Machinery, 2016, 47(6):215 - 221. http://www.j-scam.org/jcsam/ch/read/view_abstract.aspx? flag = 1&file_no = 20160628&journal_id = jcsam. DOI:10.6041/j.issn.1000-1298.2016.06.028. (in Chinese)

[24] 何晓群, 刘文卿. 应用回归分析[M]. 4 版. 北京: 中国人民大学出版社, 2015.