

基于毛竹叶片理化参数的刚竹毒蛾危害检测研究

黄旭影¹, 许章华^{1,2,3,4,5*}, 林璐¹, 石文春¹, 余坤勇⁴, 刘健⁴, 陈崇成², 周华康⁶

1. 福州大学环境与资源学院, 福建 福州 350116
2. 空间数据挖掘与信息共享教育部重点实验室, 福建 福州 350116
3. 福建省水土流失遥感监测评估与灾害防治重点实验室, 福建 福州 350116
4. 福建省资源环境监测与可持续经营利用重点实验室, 福建 三明 365004
5. 福州大学信息与通信工程博士后科研流动站, 福建 福州 350116
6. 福建省南平市延平区林业局, 福建 南平 353000

摘要 虫害检测算法研究是开展虫害快速、准确监测, 制定精准森防检疫措施的重要基础。以毛竹叶片为研究尺度, 基于刚竹毒蛾危害下的寄主外部形态与内部生理现象总结, 选择并实测叶损量 LL、相对叶绿素含量 RCC、相对含水量 RWC、原始光谱的 733.66~898.56 nm 值($\rho_{733.66\sim 898.56}$)、一阶微分光谱的 562.95~585.25 nm 值($\rho'_{562.95\sim 585.25}$)与 706.18~725.41 nm 值($\rho'_{706.18\sim 725.41}$)等理化参数, 随机划分实验组(63组)和验证组(37组)并设计5次重复实验; 分别运用 Fisher 判别分析、BP 神经网络、随机森林等三种方法建立刚竹毒蛾危害等级的检测模型, 从检测精度、Kappa 系数及 R^2 等指标对模型的检测效果予以分析和比较。结果显示, Fisher 判别分析、BP 神经网络、随机森林的检测精度分别为 69.19%, 65.41%, 83.78%, Kappa 系数分别为 0.576 9, 0.532 4 和 0.778 8, R^2 分别为 0.722 2, 0.582 6 和 0.870 9, 总体而言, 三种方法均具备刚竹毒蛾危害的检测能力, 随机森林的检测效果最优, Fisher 判别分析次之, 再次为 BP 神经网络; 从分等级来看, 随机森林的检测精度亦优于 Fisher 判别分析与 BP 神经网络, 但 3 种方法对中度危害等级的检测精度均有所不足。该成果可为刚竹毒蛾危害及其他病虫害检测算法的选择提供参考, 并为进一步建立冠层、遥感影像像元等尺度的虫害检测模型奠定基础。

关键词 刚竹毒蛾; 毛竹叶片; Fisher 判别分析; BP 神经网络; 随机森林

中图分类号: O433 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2019)03-0857-08

引言

毛竹具有用途广、经济与生态价值高等特点, 是我国竹产业经营与发展的最主要竹种。随着我国人工林面积的增加和纯林化程度的上升, 竹林所面临的病虫害威胁形势依然严峻。据统计, 当前已记载的竹林害虫达 630 余种, 其中 60 多种已先后在全国各竹区周期性或暴发性发生, 给当地政府及林农造成巨大的经济损失, 成为竹林健康与生态安全的主要威胁。刚竹毒蛾 (*Pantana phyllostachysae* Chao) 属鳞翅目毒蛾科竹毒蛾属, 是竹最主要的食叶性害虫, 其暴发成灾时短时间内即可将整片竹林噬食殆尽, 如同火烧一般。当前, 刚

竹毒蛾的发生情况多采用人工调查的方法获取, 但由于虫害的突发性和迁飞性, 所得信息往往不够全面, 且数据滞后性较大, 难以满足林业现代数字化管理及森林可持续发展战略的实施。

近年来, 遥感技术的日趋成熟为森林虫害的大面积、多时相信息获取及预警、监测体系的构建奠定了良好基础^[1-2]。获取影像中对虫害响应敏感的特征信息是解决虫害遥感快速、准确识别的方向, 故地面微观机理总结对遥感宏观监测研究有着重要的理论指导意义^[3-4]。选择有效的检测算法是微观机理研究的重要内容, 综观国内外研究成果, 已有诸多算法被应用于病虫害识别研究领域^[5-8]。Fisher 判别分析是多元统计分析中的常用算法, 于 1936 年被首次提出, 已有

收稿日期: 2017-09-29, 修订日期: 2018-01-15

基金项目: 国家自然科学基金项目(41501361, 41401385), 中国博士后科学基金面上项目(2018M630728), 福建省资源环境监测与可持续经营利用重点实验室开放基金项目(ZD1403), 福州大学人才基金项目(XRC-1345)资助

作者简介: 黄旭影, 1993 年生, 福州大学环境与资源学院硕士研究生 e-mail: 332153458@qq.com

* 通讯联系人 e-mail: fafuxzh@163.com

虫害监测的部分成果,如许章华等^[9]建立了不同马尾松毛虫等级的 Fisher 判别函数,证实该方法预测马尾松毛虫等级的可行性;但该方法在虫害检测研究中的经验尚待积累^[10]。1986 年, Rumelhart 和 McClelland 等提出 BP 神经网络,现已在虫害检测研究中得以应用,例如, Bai 等^[11]提出了一种基于粗糙集理论和 BP 神经网络的虫害识别模型,为胡杨林虫害防治工作提供了有效技术支持。随机森林则是 Breiman 和 Cutler 于 2001 年提出的一种新型算法,自提出以来即被广泛应用于生态学、医学、经济学等领域中^[12-14],但多集中于“是”和“否”的二分模型,且尚未见运用于虫害检测研究的报道。

当前刚竹毒蛾危害下的寄主特征多集中于外部形态与内部生理现象的初步阐述与总结^[15-16],缺乏响应机理及检测算法的研究。刚竹毒蛾以毛竹叶片为食,初期取食竹叶表皮,后期将竹叶食成缺刻;失叶状态下的寄主光合作用能力下降,水分缺失,进而使未被取食的竹叶及新生叶片出现发黄、病斑等症状^[17-18]。根据前人总结及实地踏查结果,本工作选择并实测毛竹叶片各理化参数,分别构建刚竹毒蛾危害等级的 Fisher 判别分析、BP 神经网络与随机森林检测模型,并分析比较这三种方法的检测效果,为该虫害的遥感快速、准确监测提供参考。

1 试验区概况

试验区为福建省南平市延平区,该区地处福建省中部偏北,经纬度范围为 117°50′—118°40′E, 26°15′~26°52′N,土地总面积约 2 659.7 km²。延平区处闽中大谷地的最低处,境内峰峦岑嶂,高、低丘陵蜿蜒,构成主要的地貌形态。该地区属中亚热带海洋季风气候,夏季炎热,冬无严寒,春早秋迟,夏长冬短,雨季明显,潮湿多雾,多静风;年均气温 19.3℃,年降水量 1 669 mm,无霜期 268 d。延平区气候温和,森林资源丰富,是中国南方三大杉木产区之一。截至 2015 年底,全区共有林业用地 2 123.33 km²,其中有林地面积 1 943.33 km²,生态公益林面积 424 km²,森林覆盖率达 73.3%,森林蓄积量约 1 787.04 万 m³,其中,茫荡山自然保护区内乔木品种超过欧洲大陆总和(除杉、松外),区内有竹林 392 km²,其中毛竹林 372.67 km²,中小径竹 19.6 km²,故延平区有“森林之窗,本甲全闽”之称。近 10 年来,延平区各类森林病虫害发生面积近 4.08 万 hm²,主要包括刚竹毒蛾、马尾松毛虫、松材线虫病等,并呈逐年上升的趋势。

2 实验部分

2.1 刚竹毒蛾危害等级划分

不同于林分尺度依靠虫株率、虫口数量/密度的方法,本研究叶片尺度危害等级的确定采用综合判定法:(1)根据刚竹毒蛾的危害机制以及国家林业局发布的《林业有害生物发生及成灾标准》,将单株失叶率(无危害:0%、轻度危害:0%~25%、中度危害:25%~50%、重度危害:>50%)及虫口数量(无危害:<10 条、轻度危害:10~30 条、中度危

害:31~80、重度危害:>80 条)列入虫害等级划分的参考因子;(2)以植物保护、森林保护等学科背景的高校学者及长期从事森防检疫工作的林业从业人员为对象,利用专家咨询法对虫害等级进行最终判定。

2.2 刚竹毒蛾危害下毛竹叶片理化参数的测定与处理

于 2017 年 2 月底刚竹毒蛾越冬代幼虫初期赴延平区开展调研工作。刚竹毒蛾为典型的食叶性害虫,受害后毛竹叶片呈现斑点、缺刻等叶损特征,同时伴随叶片水分散失,叶绿素降低等状态(图 1)。

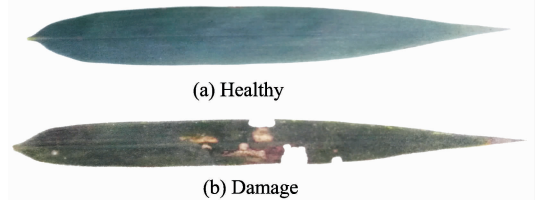


图 1 健康与受害毛竹叶片照片

Fig. 1 Photos of moso bamboo leaves in healthy state (a) and damaged by *P. Chao* (b)

据此,选择可反映失叶、叶绿素、叶水分等维度的理化参数:叶损量、相对叶绿素含量、相对含水量,同时运用单因素方差分析法(One-way ANOVA)对各参数的虫害响应能力予以检测,当无危害叶片与受害叶片间的差异显著时($p < 0.05$),则将其确定为刚竹毒蛾危害检测模型的自变量因子;此外,作为虫害检测中最为基本的信息,光谱往往可以反映叶片的细微变化,故其亦是必需考量的指标。上述各参数的测定与处理过程为:

(1) 叶损量(Leaf loss, LL)

基于参照板拍摄不同虫害等级叶片照片后将其导入电脑,利用 Photoshop 等软件调节照片亮度及对比度,以确保病斑及缺刻的可辨性;借助 CAD、Excel 等计算病斑、缺刻等叶损面积,并将其与叶片总面积相除而得叶损量;随后利用单因素方差分析法检测其虫害响应能力,结果显示健康叶片与受害叶片间的 LL 差异达极显著水平($p < 0.01$),故将其列入刚竹毒蛾危害检测模型的自变量因子。LL 的计算公式如式(1)

$$LL = A_{LL}/A \times 100\% \quad (1)$$

式中:LL 为叶损量; A_{LL} 为叶损面积; A 为叶片总面积。

(2) 相对叶绿素含量(Relative chlorophyll content, RCC)

叶绿素含量常用的测定方法有丙酮研磨法、浸提法、便携式叶绿素仪测定法等,为保证数据的时效性及测定精度,采用 TYS-4N 植物营养测定仪对毛竹叶片进行无损测定,以 SPAD 值表示 RCC。该仪器的工作原理在于测定叶片在叶绿素的两个吸收光波长范围内的透光系数,以此确定叶绿素的相对含量。同样运用单因素方差分析法检测其虫害响应能力,结果显示健康叶片和受害叶片间的 SPAD 值亦存在极显著差异($p < 0.01$),据此将其列入自变量因子。

(3) 相对含水量(Relative water content, RWC)

相对含水量由叶片鲜重与干重计算而得。将叶片装入密封袋之后,去除叶片根部后利用电子天平逐片速测其鲜重;

随后经 105 °C 杀青后在 80 °C 下烘至恒重, 分别称取各叶片干重, 利用式(1)计算其单叶含水量。单因素方差分析结果显示, 健康叶片与受害叶片间具有极显著差异 ($p < 0.01$), 由此将其列入自变量因子。

$$RWC = (FW - DW) / FW \times 100\% \quad (2)$$

式中: RWC 为相对含水量, FW 为叶片鲜重, DW 为叶片干重。

(4) 特征波长光谱

采用合肥仪思特光电技术有限公司生产的 ISI921VF-256 野外地物光谱辐射计于野外测定毛竹叶片光谱数据。该设备波长范围为 380~1 050 nm, 计 256 个波段, 光谱分辨率为 4 nm, 视角角为 3°。为保证光谱数据的准确性, 每当测量位置发生变化时进行一次标准白板校正; 每片竹叶分别测定近叶尖处、叶中、近叶基处三个部位, 每个部位取连续测定三次的平均值, 将三个部位的平均光谱数据作为该叶片光谱值, 获取各虫害等级毛竹叶片的光谱数据(图 2); 运用光谱微分法对所得光谱数据进行处理, 得到各虫害等级毛竹叶片的一阶微分光谱数据(图 3)。

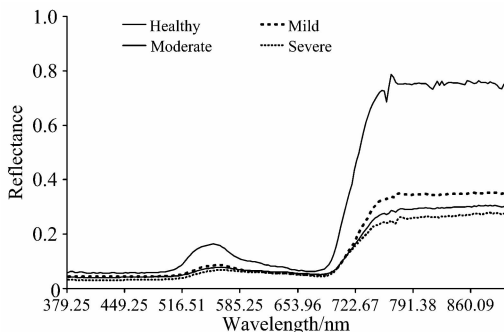


图 2 各虫害导致毛竹叶片原始光谱反射率曲线

Fig. 2 Reflectance curves of moso bamboo leaves in different pest levels

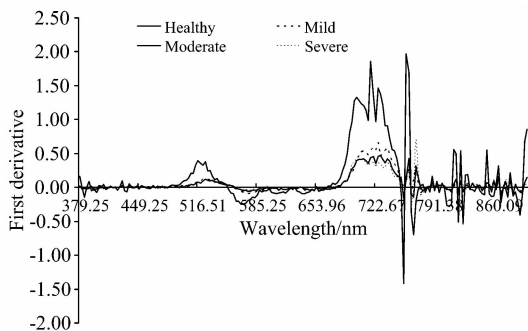


图 3 各虫害导致毛竹叶片一阶微分光谱反射率曲线

Fig. 3 Reflectance first derivative curves of moso bamboo leaves in different pest levels

为保证各虫害等级间的可辨性, 采用两两分组的方式进行分析(即无危害-轻度危害、无危害-中度危害、无危害-重度危害、轻度危害-中度危害、轻度危害-重度危害、中度危害-重度危害), 利用单因素方差分析获取各虫害等级叶片间具有极显著差异 ($p < 0.01$) 的波长, 当四组(或以上)差异同

时达极显著水平 ($p < 0.01$) 时, 则记录该波长。利用欧式距离、相关系数与光谱角匹配等三种判别方法分析波长的虫害判别能力, 若通过两种或两种以上方法检验, 则将其确定为特征波长。由于实测高光谱波长数多, 信息量大, 考虑到与遥感影像的对接问题, 剔除小于 10 nm 的波长。经此步骤, 入选的特征波长光谱为: $\rho'_{733.66 \sim 898.56}$, $\rho'_{562.95 \sim 585.25}$ 和 $\rho'_{706.18 \sim 725.41}$, 即原始光谱的 733.66~898.56 nm、一阶微分光谱的 562.95~585.25 与 706.18~725.41 nm 处的值。

2.3 虫害检测模型

(1) Fisher 判别分析 (Fisher discriminant analysis, FDA): 其基本思想是投影, 即基于方差分析的思想(组内方差尽量小, 组间方差尽量大), 将 n 组 p 维数据投影至某一个方向, 使其成为一维数据, 再依据相应的判别准则确定其类别。其判别函数为式(3)

$$y(X) = \hat{C}_1^T X \quad (3)$$

式中: \hat{C}_1 为最大特征值对应的特征向量; X 为样本自变量矩阵。

(2) BP 神经网络 (BP neural networks, BPNN): 是一种多层前向型网络, 其核心思想是运用梯度下降法求解以网络误差平方为目标函数的最小值。BP 神经网络的优点在于其具备较强的非线性映射能力以及强大的自学及自适应能力, 且能给出完整的数理推导。当前主要应用于函数逼近、模式识别、数据分类及压缩等方面。3 层神经网络是最常见的 BP 神经网络结构, 其示意图见图 4。

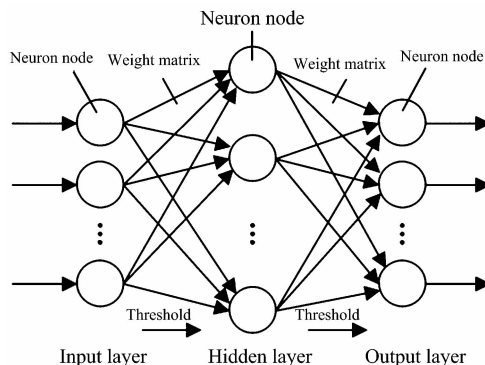


图 4 BP 神经网络三层结构示意图

Fig. 4 Structure of 3-layer BP neural networks

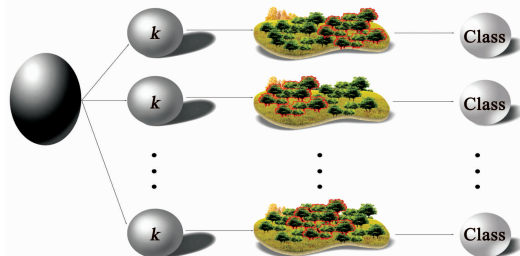


图 5 随机森林结构示意图

Fig. 5 Modeling schematic diagram of random forest

(3) 随机森林 (Random forest, RF): 实质是对决策树算法的一种改进, 其基本思想为利用自助抽样法 (bootstrap) 从

原始训练集中有放回的抽取 k 个样本作为新的训练集；分别对其进行决策树建模后生成 k 个分类树，生成随机森林；最后根据各个分类树的结果(Class)决定新样本的归属(图 5)。该算法的优势在于能够有效处理高维数据，且无需降维，对于缺省值及非平衡的数据亦能获得较好的分类精度。

2.4 分组设计及模型评价

为避免实验误差，将所采集的样本数据随机划分为实验组(样本数 63)和验证组(样本数 37)并作 5 次重复，依次采用 Fisher 判别分析、BP 神经网络及随机森林法进行计算，设为试验 1—5。

精度、Kappa 系数和判定系数 R^2 是常用的分类评价指标，本研究采用此三个指标评价各模型的虫害检测效果。检测精度、Kappa 系数、 R^2 越高，表明该模型的检测效果越

好；反之亦然。

3 结果与讨论

3.1 基于 Fisher 判别分析的刚竹毒蛾危害检测

基于实验组数据，以叶损量 LL、相对叶绿素含量 RCC、相对含水量 RWC、原始光谱的 733.66 ~ 898.56 nm 值($\rho'_{733.66 \sim 898.56}$)、一阶微分光谱的 562.95 ~ 585.25 nm 值($\rho'_{562.95 \sim 585.25}$)与 706.18 ~ 725.41 nm 值($\rho'_{706.18 \sim 725.41}$)等理化参数为自变量，建立无危害、轻度危害、中度危害及重度危害的 Fisher 线性判别函数(表 1)；将验证组样本分别代入以上 4 个判别函数中，取函数值达到最大的判别函数，将样本归入其所对应的虫害等级。

表 1 Fisher 线性判别函数系数

Table 1 Linear discriminant functions of Fisher discriminant analysis

Test	Pest level	Characteristic coefficient						
		LL	RCC	RWC	$\rho'_{733.66 \sim 898.56}$	$\rho'_{562.95 \sim 585.25}$	$\rho'_{706.18 \sim 725.41}$	c
1	Non-damage	26.256 9	1.576 8	92.977 9	-0.492 1	-95.230 2	21.412 9	-60.961 3
	Mild damage	29.677 1	1.606 3	84.786 3	-0.500 8	-85.712 8	8.158 5	-44.389 0
	Moderate damage	36.103 4	1.444 1	74.478 3	-0.214 9	-62.841 4	-5.221 4	-37.973 5
	Severe damage	58.855 1	1.421 5	77.908 7	-0.195 5	-49.143 2	-5.218 6	-48.793 4
2	Non-damage	22.963 5	1.650 8	73.715 6	-0.108 5	-116.263 2	-1.530 5	-58.200 0
	Mild damage	27.076 6	1.792 1	65.777 7	-0.136 4	-108.657 1	-15.690 9	-43.934 2
	Moderate damage	35.892 6	1.752 2	60.138 0	0.116 5	-94.945 4	-28.786 1	-43.033 2
	Severe damage	50.934 8	1.492 2	61.768 9	0.143 4	-80.136 2	-28.046 0	-43.298 9
3	Non-damage	24.528 4	1.667 5	78.181 5	-0.522 1	-96.986 8	19.694 2	-55.821 4
	Mild damage	25.122 7	1.894 0	74.993 5	-0.583 0	-118.582 2	2.492 6	-45.439 7
	Moderate damage	35.521 0	1.596 7	67.054 5	-0.164 4	-92.136 5	-16.241 1	-37.954 5
	Severe damage	52.590 7	1.581 4	68.059 5	-0.189 4	-84.119 4	-12.867 7	-45.176 3
4	Non-damage	22.513 2	1.408 5	81.208 7	-0.605 0	-92.934 0	30.210 1	-55.255 1
	Mild damage	27.086 4	1.519 7	78.174 2	-0.615 7	-100.527 8	12.195 9	-42.112 8
	Moderate damage	31.077 3	1.374 0	63.632 2	-0.167 3	-76.688 6	-9.236 1	-35.243 6
	Severe damage	49.143 4	1.234 3	71.944 8	-0.266 0	-66.308 6	-1.836 0	-42.058 9
5	Non-damage	28.549 8	1.645 9	74.091 8	-0.493 7	-179.041 0	13.292 8	-60.889 5
	Mild damage	28.264 5	1.589 0	70.044 4	-0.494 5	-114.451 4	6.186 0	-41.481 0
	Moderate damage	34.227 1	1.479 4	59.386 3	-0.147 1	-88.429 8	-10.148 7	-36.962 4
	Severe damage	50.770 3	1.331 3	63.624 1	-0.161 6	-73.889 3	-6.896 8	-41.463 5

Note: c is a constant

表 2 Fisher 判别分析结果

Table 2 Detection effects of Fisher discriminant analysis

Evaluation indicator	Detection accuracy/%	Kappa coefficient	R^2
Test 1	70.27	0.590 1	0.728 1
Test 2	70.27	0.595 8	0.738 4
Test 3	67.57	0.554 2	0.706 1
Test 4	64.86	0.516 1	0.685 8
Test 5	72.97	0.628 5	0.752 5

根据上述结果，分别计算试验 1—5 的检测精度及 Kappa 系数，同时对检测结果和实测结果进行回归分析，求取其 R^2 (表 2)。结果显示，耦合以上指标的 Fisher 判别函数对刚

竹毒蛾危害具备一定的检测能力，其检测精度在 64% 以上，Kappa 系数在 0.51 以上，而 R^2 均高于 0.68，其中试验 5 的检测精度、Kappa 系数及 R^2 最高，其次为试验 2，再次为试验 1，试验 3 从次，最后为试验 4。

分别计算各虫害等级的检测精度并求其平均值(图 6)，可以看出，不同危害等级的检测精度有较大差异，轻度危害的精度最高，达 93.33%，无危害次之，为 75.00%，再次为重度危害，为 68.57%，中度危害的精度则最低，仅为 36.00%。

3.2 基于 BP 神经网络的刚竹毒蛾危害检测

通过 Matlab 平台建立“6—10—4”的 3 层网络结构的误差反向传播模型，输入层到隐含层传递函数设为 tansig、隐

含层到输出层传递函数设为 logsig、训练方式选择 Levenberg-Marquardt 算法、迭代次数为 1 000，并导入验证组数据予以仿真。

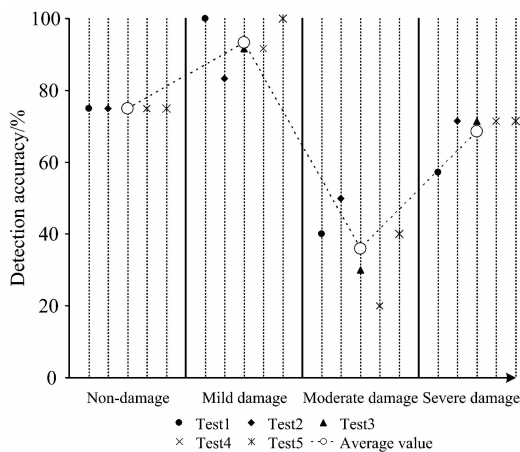


图 6 各虫害导致 Fisher 判别精度

Fig. 6 Detection accuracy of Fisher discriminant analysis in different pest levels

同理分别计算试验 1—5 的检测精度、Kappa 系数及 R^2 (表 3)，结果显示，试验 1—3 的检测精度最高，均为 67.57%，其次为试验 5，试验 4 的检测精度最低，为 59.46%；从 Kappa 系数来看，试验 3 最高，为 0.561 7，试验 2 次之，再次为试验 1，试验 5 从次，试验 4 最低，为 0.462 7；从 R^2 来看，试验 1 与实际结果的吻合度最高，为 0.706 0，试验 5 次之，再次为试验 3，试验 2 从次，试验 4 最低，为 0.298 0。由此可见 BP 神经网络亦具备一定的刚竹毒蛾危害检测能力，但总体上稍弱于 Fisher 判别分析。

表 3 BP 神经网络分析结果

Table 3 Detection effects of BP neural networks

Evaluation indicator	Detection accuracy / %	Kappa coefficient	R^2
Test 1	67.57	0.558 6	0.706 0
Test 2	67.57	0.560 0	0.602 0
Test 3	67.57	0.561 7	0.603 0
Test 4	59.46	0.462 7	0.298 0
Test 5	64.86	0.519 0	0.704 0

分等级来看(图 7)，无危害的检测精度最高，达 82.50%，轻度危害次之，为 66.67%，再次为重度危害，为 57.14%，中度危害的精度依然最低，为 52.00%，不同危害等级的检测精度仍有所差异。

3.3 基于随机森林的刚竹毒蛾危害检测

将上述六个指标设为自变量，因变量设为无危害、轻度危害、中度危害及重度危害；决策树数量 ntree 设为 5 000，节点分割变量 mtry 设为 5，据此建立随机森林模型。

结果显示(表 4)，随机森林对刚竹毒蛾危害的检测效果较好，其检测精度在 81% 以上，Kappa 系数至少达到了 0.74 以上， R^2 则均高于 0.84，其中试验 2 的检测精度最高，其次

为试验 1、试验 3 及试验 4，再次为试验 5；从 Kappa 系数来看，试验 2 最高，其次为试验 3，再次为试验 4，试验 1 从次，试验 5 最低；从 R^2 来看，试验 2 依然最高，其次为试验 3，再次为试验 1，试验 4 从次，试验 5 最低。

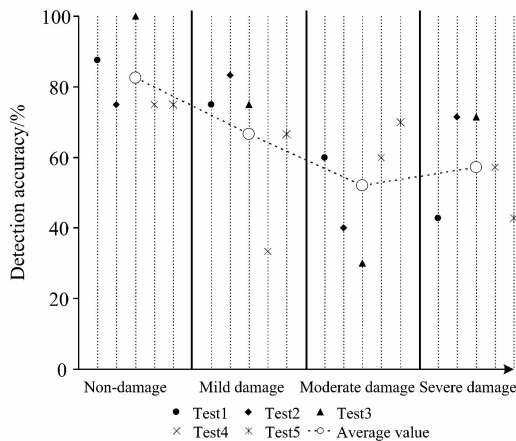


图 7 各虫害等级 BP 神经网络判别精度

Fig. 7 Detection accuracy of BP neural networks in different pest levels

表 4 随机森林分析结果

Table 4 Detection effects of random forest

Evaluation indicator	Detection accuracy / %	Kappa coefficient	R^2
Test 1	83.78	0.777 3	0.872 9
Test 2	86.49	0.815 4	0.890 8
Test 3	83.78	0.778 0	0.873 2
Test 4	83.78	0.777 8	0.871 6
Test 5	81.08	0.745 3	0.845 8

从分等级来看(图 8)，无危害的精度最高，达 100.00%，轻度危害次之，为 98.33%，再次为重度危害为 91.43%，中度危害的精度则最低，为 50.00%，可见随机森林对中度危害的识别能力仍稍弱，对不同危害等级的检测能力亦有所不同。

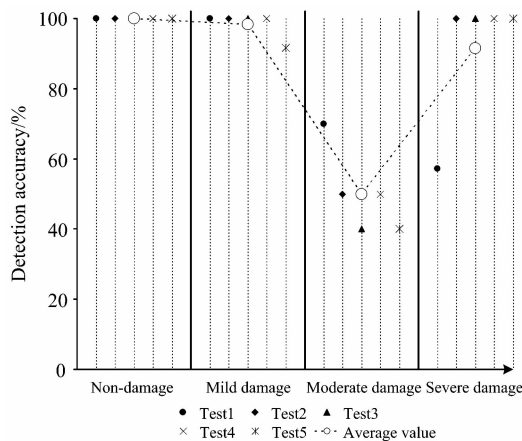


图 8 各虫害等级随机森林判别精度

Fig. 8 Detection accuracy of random forest in different pest levels

3.4 虫害检测效果比较

进一步汇总试验 1—5 的检测精度、Kappa 系数及 R^2 ，并将其绘成散点图(图 9)，可以看出，比较 Fisher 判别分析和 BP 神经网络，随机森林的检测精度分别高出 14.59% 和 18.38%；Kappa 系数分别提高 0.201 8 和 0.246 4；根据 R^2 结果显示，随机森林所给出的分类结果与实测虫害等级的拟合效果较好，较 Fisher 判别分析提高 0.148 7，较 BP 神经网络则提高 0.283 3。总体上，随机森林对刚竹毒蛾危害的检测效果最优，其次为 Fisher 判别分析，再次为 BP 神经网络。

从分等级比较(表 5)，三种模型对不同危害等级的检测能力均表现出差异性，中度危害等级的检测效果均相对较差，其中，BP 神经网络的检测精度最高，为 52.00%，其次为随机森林，为 50.00%，Fisher 判别分析的精度最低，为 36.00%；而对于无危害、轻度危害和重度危害等级，随机森林的检测精度最高，Fisher 判别分析对无危害等级的检测能力稍弱，但轻度危害和重度危害的检测精度则高于 BP 神经

网络。

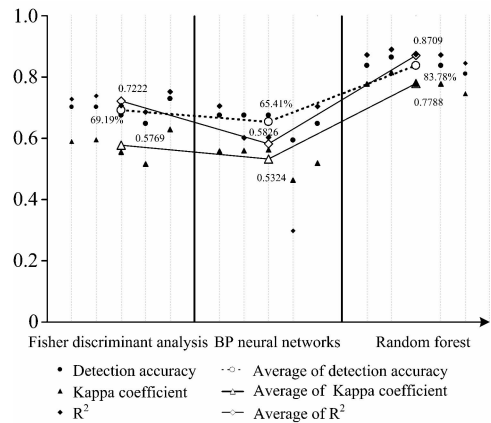


图 9 检测结果对比

Fig. 9 Detection effects comparison of three algorithms

表 5 检测算法比较

Table 5 Analysis results of different pest levels

Evaluation indicator	Detection algorithm	Pest level			
		Non-damage	Mild damage	Moderate damage	Severe damage
Detection accuracy/%	Fisher discriminant analysis	75.00	93.33	36.00	68.57
	BP neural networks	82.50	66.67	52.00	57.14
	Random forest	100.00	98.33	50.00	91.43

3.5 讨论

刚竹毒蛾是竹最主要的食叶害虫，亦是我国林业部门的重点防治对象，建立有效的检测算法对刚竹毒蛾的防治工作具有重大意义。遥感是实现虫害大面积监测的重要手段，但其需要一定的理论与方法支撑。综观国内外研究现状可知，当前对于刚竹毒蛾危害方面的研究甚少，其寄主的表像变化及内部理化参数指标尚未总结完整，采用何种算法能够有效实现虫害检测亦需探讨。模型法是当前较为常用的虫害检测方法^[19-20]，其可通过耦合多个自变量，进而减少虫害检测过程中的干扰信息，提升检测精度，故本研究将实测刚竹毒蛾危害下不同虫害等级毛竹叶片的叶损量 LL、相对叶绿素含量 RCC、相对含水量 RWC 及 $\rho_{733.66\sim 898.56}$ ， $\rho'_{562.95\sim 585.25}$ 和 $\rho'_{706.18\sim 725.41}$ 等三个特征波长的光谱值作为自变量，分别建立 Fisher 判别分析、BP 神经网络及随机森林模型，检测上述三种方法的虫害检测能力。

根据实地踏查结果及所测数据可以看出，相同虫害等级下的毛竹叶片其形状、叶绿素含量、含水量、叶片光谱等不尽相同，甚至存在相同虫害等级叶片间某项或某几项参数差异大于不同虫害等级的现象，这是由于植物的各项理化参数是随母体生长状况而变化的，而其外部的生长环境，诸如土壤肥力、坡度坡向、气候等亦会影响植物的生长，进而使其理化性质发生变化^[21-22]，同时，在判定危害等级时，对叶片状态采用了综合考量，更符合实际与应用推广需求。因此，欲实现不同虫害等级叶片间的有效区分，需选择一种能够有效处理离散型变量且容错率较大的检测算法。Fisher 判别分

析的基本思路是将高维空间中的数据投影至低维空间，再依据相应的准则对其进行分类，但需要注意的是，其面向的对象需为线性可分，而对于线性不可分的样本，该方法是无法判定其分类的。BP 神经网络对训练集数据的要求较高，从本研究所采集的数据来看，即便是相同虫害等级下的叶片其各指标数据亦存在差异，故所训练出来的模型可能陷入局部最优状态，导致模型无法表现出训练集数据应有的规律，进而影响其泛化能力。随机森林的自主抽样方式使其对样本数据噪声及异常值有着较高的容忍度，且在创建随机森林过程中，对于泛化误差采用的是无偏估计，有效提升了模型的泛化能力，表现出优于其他算法的检测精度^[23]。由此可见，从方法原理看，随机森林较之于其他两种方法具有一定优势。

本研究尝试将 Fisher 判别分析、BP 神经网络与随机森林三种算法应用于刚竹毒蛾危害的检测研究中，总体而言，三种算法均可应用于刚竹毒蛾危害检测，但相对来说，随机森林更能够有效处理离散型变量，可作为刚竹毒蛾危害检测的优选算法，Fisher 判别分析则可作为次选算法，再次为 BP 神经网络。从分等级检测结果可知，三种算法对中度危害等级的检测能力均较弱，这是由于相同刚竹毒蛾危害等级下的寄主其各项理化参数在不同时点并不相同，即使是被重度蚕食的叶片，短期内其叶绿素含量、相对含水量等甚至有可能接近于健康叶片；此外，考虑到经济、生态效益，在确定重度危害时并未选择几近或完全枯黄的叶片作为样本，有可能使中度危害和重度危害间存在一定的混淆性而导致误分率上升。

4 结 论

经单因素方差分析,将叶损量 LL、相对叶绿素含量 RCC、相对含水量 RWC、原始光谱的 733.66~898.56 nm 值 ($\rho_{733.66\sim 898.56}$)、一阶微分光谱的 562.95~585.25 nm 值 ($\rho'_{562.95\sim 585.25}$)与 706.18~725.41 nm 值 ($\rho'_{706.18\sim 725.41}$)等理化参数确定为刚竹毒蛾危害检测模型的自变量因子,利用 Fisher 判别分析、BP 神经网络及随机森林对其进行虫害等级检测,通过检测精度、Kappa 系数及 R^2 等指标对三种方法的虫害检测效果予以综合评价。

(1)三种方法均具备刚竹毒蛾危害的检测能力,其中随机森林的检测效果最佳,其检测精度、Kappa 系数、 R^2 分别为 83.78%, 0.778 8 和 0.870 9;其次为 Fisher 判别分析,其

检测精度、Kappa 系数、 R^2 依次为 69.19%, 0.576 9 和 0.722 2;再次为 BP 神经网络,其检测精度、Kappa 系数及 R^2 分别为 65.41%, 0.532 4 和 0.582 6。

(2)各方法对不同危害等级的检测能力均表现为差异性,随机森林对不同危害等级的检测效果均优于 Fisher 判别分析和 BP 神经网络,而 Fisher 判别分析与 BP 神经网络对四个虫害等级的检测精度则互有高低;三种方法对中度危害的检测能力有所不足,精度均低于 60%。

(3)本研究综合考虑了刚竹毒蛾危害下的毛竹叶片各理化参数,证实上述参数对该虫害响应的敏感性及其虫害检测能力;将 Fisher 判别分析、BP 神经网络与随机森林引入刚竹毒蛾危害检测研究,并予以综合分析、比较,为森林病虫害检测及遥感监测提供重要支持。

References

- [1] Park J, Sim W, Lee J. *Journal of Forest and Environmental Science*, 2016, 32(4): 384.
- [2] Mária H, Ján F, Rastislav J. *Forest Ecology and Management*, 2017, 391(1): 349.
- [3] Croft H, Chen J M, Zhang Y, et al. *Remote Sensing of Environment*, 2013, 133(12): 128.
- [4] Zarco-Tejada P J, Miller J R, Harron J, et al. *Remote Sensing of Environment*, 2004, 89(2): 189.
- [5] Peixoto M D S, Barros L C D, Bassanezi R C. *Applied Mathematics*, 2014, 5(5): 1133.
- [6] Singh C B, Jayas D S, Paliwal J, et al. *Journal of Stored Products Research*, 2009, 45(3): 151.
- [7] ZHANG Jian-hua, JI Rong-hua, YUAN Xue, et al(张建华, 冀荣华, 袁雪, 等). *Transactions of the Chinese Society of Agricultural Machinery(农业机械学报)*, 2011, 42(8): 178.
- [8] LIN Qi-nan, HUANG Guo-hua, CHEN Ling, et al(林起楠, 黄华国, 陈玲, 等). *Spectroscopy and Spectral Analysis(光谱学与光谱分析)*, 2016, 36(8): 2538.
- [9] XU Zhang-hua, LI Cong-hui, LIU Jian, et al(许章华, 李聪慧, 刘健, 等). *Transactions of the Chinese Society of Agricultural Machinery(农业机械学报)*, 2014, 45(6): 275.
- [10] XU Zhang-hua, LIU Jian, YU Kun-yong(许章华, 刘健, 余坤勇). *Acta Agriculturae Universitatis Jiangxiensis(江西农业大学学报)*, 2015, (6): 1016.
- [11] Bai Tiecheng, Meng Hongbing, Yao Jianghe. *Neural Computing and Applications*, 2014, 25(7-8): 1699.
- [12] Louisr I, Anantham P, Stephenn M, et al. *Forest Ecology & Management*, 2008, 254(3): 390.
- [13] YANG Fan, LIN Chen, ZHOU Qi-feng, et al(杨帆, 林琛, 周绮凤, 等). *System Engineering Theory and Practice(系统工程理论与实践)*, 2012, 32(4): 815.
- [14] PENG Ling, XU Su-ning, PENG Jun-huan(彭令, 徐苏宁, 彭军还). *Journal of Jilin University(吉林大学学报)*, 2016, 46(1): 175.
- [15] LIANG Guang-hong(梁光红). *Journal of Fujian Forestry Science and Technology(福建林业科技)*, 2003, 30(1): 34.
- [16] CHEN De-liang, ZHAO Ren-you, QU Qiao-wen(陈德良, 赵仁友, 瞿巧文). *Forest Pest and Disease(中国森林病虫)*, 2006, 25(6): 14.
- [17] LI Yuan-yuan, ZHANG Kai, LI Shuang-wen, et al(李媛媛, 张凯, 李霜雯, 等). *Scientia Silvae Sinicae(林业科学)*, 2015, 51(3): 93.
- [18] ZHOU Xin-hua, LI Yue-qiao, HUANG Zheng, et al(周新华, 厉月桥, 黄拯, 等). *Journal of Central South University of Forestry & Technology(中南林业科技大学学报)*, 2017, 37(3): 39.
- [19] Anees A, Aryal J, O'Reilly M M, et al. *Isprs Journal of Photogrammetry & Remote Sensing*, 2016, 122: 167.
- [20] XU Zhang-hua, LIU Jian, GONG Cong-hong, et al(许章华, 刘健, 龚从宏, 等). *Journal of Central South University of Forestry & Technology(中南林业科技大学学报)*, 2012, 32(10): 72.
- [21] GONG Zhao-ning, ZHAO Ya-li, ZHAO Wen-ji, et al(宫兆宁, 赵雅莉, 赵文吉, 等). *Acta Ecologica Sinica(生态学报)*, 2014, 34(20): 5736.
- [22] JI Rong-hua, ZHENG Li-hua, DENG Xiao-lei, et al(冀荣华, 郑立华, 邓小蕾, 等). *Transactions of the Chinese Society of Agricultural Machinery(农业机械学报)*, 2014, 45(8): 269.
- [23] Rodriguez-Galiano V, Mendes M P, Garcia-Soldado M J, et al. *Science of the Total Environment*, 2014, 476-477(4): 189.

Pantana Phyllostachysae Chao Damage Detection Based on Physical and Chemical Parameters of Moso Bamboo Leaves

HUANG Xu-ying¹, XU Zhang-hua^{1,2,3,4,5*}, LIN Lu¹, SHI Wen-chun¹, YU Kun-yong¹, LIU Jian¹, CHEN Chong-cheng², ZHOU Hua-kang⁶

1. College of Environment and Resources, Fuzhou University, Fuzhou 350116, China

2. Key Lab of Spatial Data Mining & Information Sharing, Ministry of Education, Fuzhou 350116, China

3. Fujian Provincial Key Laboratory of Remote Sensing of Soil Erosion and Disaster Protection, Fuzhou 350116, China

4. Fujian Provincial Key Laboratory of Resources and Environment Monitoring & Sustainable Management and Utilization, Sanming 365004, China

5. Postdoctoral Research Station of Information and Communication Engineering, Fuzhou University, Fuzhou 350116, China

6. Yanping District Forestry Bureau of Nanping, Nanping 353000, China

Abstract Pest detection algorithm research is an important guarantee to precisely and rapidly monitor the forest pest and forest protection and quarantine. Based on the external morphology of the host and its internal physiological phenomena, taking the leaf loss (LL), relative chlorophyll content (RCC), relative water content (RWC), and the three spectral values of the characteristic wavelengths ($\rho_{733.66\sim 898.56}$, $\rho'_{562.95\sim 585.25}$, $\rho'_{706.18\sim 725.41}$) as the experimental data which were randomly divided into experimental group (63) and verification group (37) with 5 repeated tests, then the models of Fisher discriminant analysis, random forest and BP neural networks for pest levels were constructed. The detection accuracy, Kappa coefficient and R^2 were used to comprehensively compare the detection effects of these three algorithms. The results showed that the detection accuracy of Fisher discriminant analysis, BP neural networks and random forest were 69.19%, 65.41% and 83.78%, and Kappa coefficient were 0.576 9, 0.532 4 and 0.778 8, and R^2 were 0.722 2, 0.582 6 and 0.870 9. Overall, all of these algorithms have the capability of pest detection, among which, the detection effect of the random forest is the best, and Fisher discriminant analysis is secondly, and BP neural networks is thirdly. Besides, the accuracy of random forest detection is superior to that of Fisher discriminant analysis and BP neural networks in non-damage, mild damage and severe damage, but these three methods have insufficient detection accuracy for moderate damage level. The results could be a reference to the selection of detection algorithm in *P. chao* and other types of diseases and insect pests, building a strong foundation for further study.

Keywords *Pantana phyllostachysae* Chao; Moso bamboo leaves; Fisher discriminant analysis; BP neural networks; Random forest

(Received Sep. 29, 2017; accepted Jan. 15, 2018)

* Corresponding author