

马尾松毛虫害等级的 Fisher 判别分析^{*}

许章华¹ 李聪慧² 刘健^{3,4} 余坤勇^{3,4} 龚从宏^{3,4} 唐梦雅^{3,4}

(1. 福州大学环境与资源学院, 福州 350108; 2. 北京林业大学自然保护区学院, 北京 100083;

3. 福建农林大学 3S 技术应用研究所, 福州 350002; 4. 福建农林大学林学院, 福州 350002)

摘要: 马尾松毛虫是我国森林最具危害性的害虫之一,开展该虫害的遥感预测研究有助于全面、及时地发现可能发生虫害的区域,为有效采取森防检疫措施,降低虫害造成的直接、间接损失提供技术支持。以福建省为研究区,于 2012 年 2—5 月份依次开展了三明市、将乐县、沙县、南平市、华安县、云霄县、南安市、安溪县、莆田市、长汀县、建阳市、宁德市及福清市等 13 个监测站马尾松毛虫越冬代调查,调查指标包括 GPS 坐标、松林冠层光谱、气候、地形、林分、虫源、人文环境等。对上述因子进行分析与提取,经相关分析得到 13 个可有效预测马尾松毛虫害等级的子因子;利用 Fisher 判别分析,建立无危害、轻度危害、中度危害、重度危害等 4 个虫害等级的判别函数,建模组样本自检结果显示 4 个等级的判别精度分别为 71.0%、84.9%、87.2%、100%,总判别精度为 80.8%,验证组样本检验结果显示判别精度为 73.00%,预测准确率为 75.99%,表明利用 Fisher 判别分析预测马尾松毛虫害等级具有较好的可行性。

关键词: 马尾松毛虫 虫害等级 Fisher 判别 相关分析

中图分类号: TP79; P237 **文献标识码:** A **文章编号:** 1000-1298(2014)06-0275-09

引言

马尾松毛虫 (*Dendrolimus punctatus* Walker) 是我国破坏最严重的森林害虫,具有危害面积大、暴发周期短、成灾最严重的特点。据福建省近 50 年的观测数据,马尾松毛虫害平均 3.69 年出现一次发生高峰;近 20 年来,年均发生面积 8 万 hm^2 。武红智等^[1]利用 GIS 空间分析技术,探讨了马尾松毛虫害的发生、发展与道路、居民地的相关关系,结果表明,马尾松毛虫的早期危害点或虫源地往往都在远离道路与居民地的位置,并呈现由远及近的发展趋势;亦即松毛虫的成灾地区可能位于人烟稀少、交通不便的偏僻山区,仅依靠传统的人力踏查工作量十分巨大,且不可能完全发现灾害区,因此虫害实际发生面积可能远高于传统监测结果。遥感技术的不断发展及在森林资源中的日益成熟应用为森林虫害的遥感监测提供了新的解决方案,我国 HJ、CBERS、ZY、天绘、高分等系列卫星的发射更为森林健康的实时监测与动态预测创造了新的平台和契机。

马尾松毛虫害的卫星遥感监测研究已有一段历史^[2-8]。相较于松毛虫害监测,其预测的实际应用价值更为巨大。然而,当前对马尾松毛虫害的预测

工作还较多地限制于单纯的数学模型测算,尚未与遥感技术有机结合。在南方山地丘陵区域,开展马尾松毛虫害的遥感监测、预测工作具有较大难度,本文经大量实验研究,获取与松毛虫害密切相关的诸因子,在此基础上利用 Fisher 判别分析对无危害、轻度危害、中度危害及重度危害等 4 个虫害等级进行预测。

1 研究区概况

研究区设置在海峡西岸经济区的核心区——福建省。福建省拥有极为丰富的森林资源,是全国南方重点集体林区;商品林产业较为发达,杉木、马尾松、桉树是其主要用材树种。

2 研究方法

2.1 数据收集

马尾松毛虫在福建省一般一年发生 2~3 代,闽西北以 2 代为主,偶尔发生 3 代,沿海地区则一般发生 3 代。一般情况下,松毛虫幼虫第 1 代出现在 5 月份,第 2 代在 7 月份,而第 3 代则在 9 月份。11 月份左右,松毛虫进入越冬代,直至次年 5 月份左右开始活动。于 2012 年 2—5 月份依次开展了三明市、

收稿日期: 2013-08-29 修回日期: 2013-09-16
^{*} 国家自然科学基金资助项目(30871965)和福州大学科研启动基金资助项目(510049)
作者简介: 许章华,讲师,主要从事资源环境遥感、地理学与城乡规划研究,E-mail: fafuxzh@163.com
通讯作者: 刘健,教授,博士生导师,主要从事森林经营管理与 3S 技术应用研究,E-mail: fjliujian@126.com

将乐县、沙县、南平市、华安县、云霄县、南安市、安溪
县、莆田市、长汀县、建阳市、宁德市及福清市等 13
个监测站的马尾松毛虫越冬代调查。原则上,每县
(市)选择 12 个小班,每个小班调查 4 个测点,测定
其地形、林分等因子;同时携带两部麦哲伦手持
GPS,当两部 GPS 定位结果基本一致时作记录;松林
冠层光谱采用中国合肥仪思特光电技术有限公司生
产的 ISI921VF-256 型野外地物光谱辐射计测定;
同时收集气候、林分、虫源及人文环境等相关资料。
本文随机选择其中 392 组数据开展实验研究。

马尾松毛虫是一类周期性森林害虫,越冬代虫
口密度、虫株率、虫情级数等因子的调查可以较为准
确地预测当年的虫害发生等级与危害级别。松毛虫
“上树”后,其危害性开始直观地表现出来。虫害的
识别与监测,应立足于为害期;而对于虫害预测、预
警而言,必须在越冬代或之前就要较为精确地作出
判断,这样才可有效地指导虫害预防方案的制定与
预防措施的落实。本研究经中国资源卫星应用中心
网站,选择了对应于松毛虫 2010 年为害期、2010—
2011 年越冬代的数景 HJ-1 卫星数据(2 级),即
HJ-1A 与 HJ-1B 获取的 CCD 多光谱数据;对影像数
据进行大气校正、几何校正、镶嵌、裁剪及云去除等
预处理工作。

2.2 松林信息提取

松林是马尾松毛虫的寄主,有效提取松林信息
是实现松毛虫害天基空间监测与预警的基础性工
作。传统的森林资源调查靠一类、二类完成,人力、
财力、物力消耗极大,且无法实现全区域范围的调
查。1987 年以来,卫星影像逐渐在森林资源调查与
研究领域得以应用,并成为大区域尺度资源调查的
重要数据,精确到“种”的信息提取技术成为林业行
业普遍关心的热点问题^[9-10]。马尾松毛虫的主要
寄主植物有马尾松、湿地松、黑松、火炬松等,松毛虫
对上述各种植物的为害机制相似,寄主对虫害的响
应也基本上一致;加之福建省湿地松、黑松与火炬松
的分布范围极为有限,马尾松分布面积占绝对优势,故
将寄主松属植物归于一类,将其林分统称为“松林”。

利用遥感技术实现信息自动提取基本上可分为
3 个环节:分类特征波段的选择;训练样区的选取;
训练分类器和分类精度的评价。其中,分类特征波
段的选择最为重要,也是决定信息挖掘深度、广度的
关键步骤。综观已有文献,该环节可分为 4 个层次:
①不经增强变换,仅利用原始影像各波段进行分析,
选择最佳波段组合。②利用图像增强技术,对原始
影像进行增强变换,以增强后的信息作为信息提取
的特征波段。③引入空间纹理特征,将其与光谱信

息结合,作为信息提取的特征波段。④引入其他来
源的遥感影像或海拔高度、坡度、坡向、坡位等非遥
感数据,作为信息提取的联合因子^[11-14]。本文研究
区福建省是典型的山地丘陵区,地形复杂、植被类型
丰富,同物异谱、异物同谱现象较为普遍,单纯依靠
光谱特征实现松林信息提取难度极大;参考本团队
近年来开展南方山区竹林及其他森林类型信息提取
的经验^[15-16],提出“基于光谱片层-面向纹理类”的
松林信息提取算法。“片层”定义为边缘围绕的连
成一片的统计上均匀的区域,而分片所形成的不同
区域可以说是空间上连接在一起的像元,即将光谱
值或像元灰度相似的像元分割出来,形成单独的层。
该算法的基本原理是:在运用光谱分析技术实现
“片层”分割的基础上,对该“片层”进行纹理特征分
析,利用纹理信息识别光谱特征无法有效识别的同
谱异物的地物信息。依此实现 2010 年为害期、
2010—2011 年越冬代两期的松林信息提取,采用精
度评估法对提取精度进行评价,结果显示:两期提取
精度分别为 83.36%、84.36%,总 Kappa 系数分别
为 0.609 6、0.602 6。

2.3 相关因子获取与分析

2.3.1 气候因子分析与获取

气候主要通过影响马尾松毛虫及马尾松等宿主
的生物学特性而对虫害分布产生宏观性作用。由于
气候条件不同,我国南北带马尾松毛虫害的发生与
分布规律存在差异,南带每年可发生 3~4 代,而北
带可能仅发生 2 代^[17]。同时,气候直接影响松林的
分布范围^[18-20],总体上看,马尾松的分布极广,北可
至河南及山东南部,南可至两广,这与松毛虫的分布
区域是一致的。作为喜光喜温的阳性树种,马尾松
适生于年均温 13~22℃、年降水量 800~1 800 mm、
绝对最低温度大于 -10℃ 的区域。归结起来,影响
马尾松毛虫害空间分布的气候因子涉及均温、积温、
湿度、降雨量、降雨天数、大风天数、霜期等。经上
述分析,结合数据的可获取性,本文将气候因子的选
择分为两大类:定量因子与定性因子。定量因子包
括目标年上年平均气温(℃)、平均气温距平(℃)、降
雨量(mm)、降雨量距平(mm)、日照时数(h)及日照
时数距平(h)等;定性因子的选择参考陈绘画
等^[21]、朱建华等^[22]的研究成果,包括目标年上年
4 月至当年 3 月旬平均气温、旬降雨量及旬日照时
数的距平状态:异常偏低(少)、偏低(少)、正常、偏
高(多)、异常偏高(多)分别赋值为 1、2、3、4、5。

气候因子由气候台站测定得到,台站分布位置
与野外调查的位置不同,因此,首先应当使气候要素
与松林分布区相匹配;台站在全省范围内分布较为

均匀,故可采用空间插值的方法对气温、降水、日照及距平点状数据予以空间化。空间插值常用于将离散点的测量数据转换为连续的数据曲面,是空间模拟的最常用方法。本文采用样条函数法中的规则样条函数法。在 ArcGIS 平台下,采用默认值(weight 为 0.1, Number of points 为 12),得到 2010 年福建省平均气温、平均气温距平、降雨量、降雨量距平、日照时数及日照时数距平分布图。同理,获取 2010 年 4 月至 2011 年 3 月旬平均气温、旬降雨量及旬日照时数的距平状态分布图。

2.3.2 地形因子分析与获取

影响马尾松毛虫害空间分布的地形因子主要包括海拔高度、坡度和坡向。从海拔高度上看,多数研究认为,100~300 m 是马尾松毛虫害最易发生的区域,而当海拔高度在 800 m 以上时,则进入了一个“安全范围”。从坡度上看,不同区域有所差异,但 15°~35°范围分布概率更大;而随着灾害的扩散,将逐步向其他坡度级蔓延^[1]。从坡向上看,阳坡是马尾松毛虫害最易发生的区域,半阳坡次之,而阴坡、半阴坡、平地等位置虫害分布则少得多。魏初奖等^[23]根据区划理论,将福建省马尾松毛虫灾区分成 3 个区和 3 个危险程度等级,分别是沿海重灾区、闽西北常灾区、闽中山地微灾区。由此可见,沿海、闽西北、闽中在地形方面的分异对马尾松毛虫害的空间分布产生重要影响,在越适于马尾松毛虫生长发育的地形下,虫害等级总体越高;反之亦然。综上,选择海拔高度、坡度、坡向、坡位作为马尾松毛虫害预警的地形因子。

利用 ArcGIS 3D Analyst 工具,基于 DEM 数据派生出海拔高度、坡度、坡向等地形因子。在提取地形因子基础上,对这些因子进行分级,分级标准如表 1 所示。

表 1 福建省地形因子分级

Tab. 1 Topographical factors grades of Fujian Province			
地形因子	级别	分级区间	编码
海拔高度	低	< 500 m	1
	中	[500 m,1 000 m)	2
	高	≥ 1 000 m	3
坡度	平坡	< 5°	1
	缓坡	[5°,15°)	2
	斜坡	[15°,25°)	3
	陡坡	≥ 25°	4
坡向	阳坡	[135°,225°)	1
	半阳坡	[90°,135°), [225°,270°)	2
	半阴坡	[45°,90°), [270°,315°)	3
	阴坡	[315°,360°], [0,45°)	4
	无坡向	- 1	5

坡位因子则由坡位指数(Topographic position index, TPI) 来替代,该指数由 Weiss 于 2001 年在 ESRI 国际大会上提出,是用来描述地形部位的一个参数,其基本思想是:用某点高程与周围一定范围内平均高程的差,结合该点的坡度,来确定其在坡面上所处的部位。TPI 对于地貌分类具有十分重要的意义^[24~25],其表达式为

$$R_{TPI} = (z - \bar{z}, G)$$

式中 R_{TPI} ——坡位指数
 z ——地表某点的高程
 $\bar{z}$ ——该点周围一定范围内的平均高程
 G ——该点的坡度

坡位指数与坡位的对应关系如表 2 所示。

表 2 坡位指数与坡位的对应关系
Tab. 2 Corresponding relationship between TPI and topographic position

分类代码	坡位	R_{TPI}
1	山脊	大于 S_D
2	上坡	$(0.5S_D, S_D]$
3	中坡	$(-0.5S_D, 0.5S_D)$, 坡度大于 5°
4	平坡	$(-0.5S_D, 0.5S_D)$, 坡度小于 5°
5	下坡	$(-S_D, 0.5S_D]$
6	山谷	小于 $-S_D$

注: S_D 是 TPI 与领域像元高程的标准偏差。

2.3.3 林分因子分析与获取

森林内部的结构同样也影响着马尾松毛虫害的空间分布,这些林分因子主要有树种结构、林龄、郁闭度等。松木纯林树种单一、地被物稀少、林分结构单调、食物链简单,导致整个生态系统对害虫的自然调控能力差;混交林,尤其是多层的松木与阔叶树混交林,林分郁闭度高,林内光照强度小而湿度大,不利于马尾松毛虫的生长发育。早在 1960 年,人们就已经通过调查发现以阔叶树为主的混交林分,很少发生或不发生松毛虫害^[26]。许多研究表明,松毛虫首先在纯林暴发,之后向周围林分发展。对比纯林与混交林的马尾松毛虫种群,可发现前者幼虫生长发育速率、化蛹率、羽化率、产卵量、食物利用率、转化率及消耗率等方面均显著高于后者,死亡率亦然。老叶、嫩叶对于马尾松毛虫幼虫的生长发育意义重大。松毛虫小幼虫喜食老叶,由于老针叶水分少,酸的含量亦较低,且纤维素、可溶性碳水化合物及蛋白质含量较高,营养物质较丰富,有助于松毛虫的生长发育。但由于嫩叶多汁,幼虫食后死亡率高,这对第 1 代的幼龄幼虫尤其具有现实意义。由于越冬代幼虫可能将老叶基本食尽,萌发出的新叶势必无法满足第 1 代小幼虫的食料要求,故虫口密度将会迅速

下降。此外,森林下部的植被状况也影响着马尾松毛虫的生长发育。前人总结发现,马尾松毛虫害多发于 10 年生左右密度较小的马尾松纯林。本文选择树种结构、林龄(龄组)、郁闭度、植被覆盖度等 4 个林分因子作为马尾松毛虫害预警技术体系构建的指标。上述各指标的获取方式简述如下:

(1)树种结构。纯林与混交林是两种典型的树种结构,也是本文所要区分的两个级别。纯林定义为一个树种(组)蓄积量(未达到起测径级的按标数计算)占总蓄积量(株数)的 65% 以上的乔木林地;混交林则指任何一种树种(组)蓄积量(未达到起测径级时按数计算)占总蓄积量不足 65% 的乔木林地。一个 HJ-1 CCD 影像像元所对应的地面面积为 900 m²,往往都在同一个小班范围内,生长状况接近,故近似地可以用一个树种(组)面积占总面积的百分比来区分纯林与混交林;面积占像元大小 65% 以上即为纯林,反之为混交林。由此,树种组成的获取问题转化为松林占像元的丰度问题,亦即混合像元分解问题。采用线性波谱分离法计算各像元的松林丰度图;当优势树种所占比例达 0.65 以上时,则定义为纯林;反之,则为混交林。依此标准,对松林丰度图进行分级,混交林编码为 1,松林纯林编码为 2。

(2)林龄(龄组)。分析幼龄林、中龄林、近熟林、成熟林、过熟林等 5 个龄组松林的光谱特征可知,在 HJ-1 CCD 多光谱影像 4 个可见光波段中,各龄组松林的 B₄(近红外波段)值普遍最大,包含的信息量最为丰富,且各龄组间的均值亦有较大差异,由大到小表现为幼龄林、中龄林、过熟林、成熟林、近熟林的特征,其他波段的排序特征与其类似,但差异较小;各龄组松林的 NDVI 特征则与可见光有较大不同,由大到小表现为成熟林、近熟林、过熟林、幼龄林、中龄林的特征,其中近熟林与成熟林的差异较小。依据分龄组松林可见光波段及 NDVI 的排序特征,构建一个新的植被指数 AGVI (Age group vegetation index)

$$I_{AGVI} = \frac{B_4}{I_{NDVI}} = \frac{B_4(B_4 + B_3)}{B_4 - B_3}$$

(2)

式中 I_{AGVI} ——龄组植被指数
 I_{NDVI} ——归一化差值植被指数
 B_4 、 B_3 ——影像第 4、3 波段灰度

在 0 ~ 255 灰度范围内,该指数即可以增大不同龄组间的绝对差异,又能较好地扩大相对差异,降低混淆概率。在此基础上,采用 K-均值法对样本进行聚类,经实验,聚类数设为 10,得到 AGVI 的最终聚类中心及最终聚类中心间距;由此判定龄组与聚类结

果的对应关系。幼龄林、中龄林、近熟林及成过熟林的编码依次为:1、2、3、4(将成熟林与过熟林两个龄组合并)。

(3)郁闭度和植被覆盖度。以实测数据与遥感影像图为基础,建立郁闭度与各特征波段、NDVI 的关系式,求得松林的郁闭度分布图。

$$C_D = a_1B_1 + a_2B_2 + a_3B_3 + a_4B_4 + a_5I_{NDVI} + a_0$$

(3)

式中 C_D ——郁闭度
 a_0 、 a_1 、 a_2 、 a_3 、 a_4 、 a_5 ——待定系数
 B_1 、 B_2 ——影像第 1、2 波段灰度

依据植被覆盖度与 I_{NDVI} 的物理关系,求得各像元的植被覆盖度。植被覆盖度与 NDVI 的物理关系表达式为

$$V_F = \frac{I_{NDVI} - I_{NDVI_s}}{I_{NDVI_g} - I_{NDVI_s}}$$

(4)

式中 V_F ——植被覆盖度
 I_{NDVI_g} ——纯植被覆盖像元的 NDVI 值
 I_{NDVI_s} ——纯土壤覆盖像元的 NDVI 值

2.3.4 虫源和天敌因子分析与获取

上代马尾松毛虫危害程度及防治效果对下代的发生与发展具有直接的影响。马尾松毛虫的大暴发是其种群增殖的结果,当种群数量增加并达到一定临界值时,便可暴发成灾。上代的种群数量为下代提供了基数,代与代之间存在着明显的继承关系,即上代虫口密度大、虫情级数高,在未得到有效防治的前提下,下代在获得较高的增殖基数后,危害范围与等级必将蔓延与提高。完整的自然生态系统中,必然存在着物物相克,以维持整个生态系统的平衡。马尾松毛虫的天敌多达百种,包括鸟类、寄生蜂、寄生蝇及捕食性昆虫等,这些天敌对松毛虫的生长发育具有重要的抵制作用。甚至有人认为,天敌是影响马尾松毛虫发生消长的主导因子。其理由是,自然界生物群落中的各个种群间,常存在相互制约的作用。天敌数量将随着松毛虫的数量增加而增加,且增加速度通常较之于松毛虫快得多,从而迅速地将松毛虫抑制下去;而此后天敌数量则锐减,因而不久以后,松毛虫又再度上升。

虫源亦是开展马尾松毛虫害天基空间预警研究的重要因子,故本文试图实现福建省马尾松毛虫害等级的 Fisher 判别,须将 2010 年的虫害分布情况考虑其中。通过咨询福建省林业厅森防检疫总站、福建农林大学,实地调查 13 个县(市)的有关专家及基层工作人员,均认为对于马尾松毛虫害而言,其天敌较多,在福建省区域范围内,基本上不存在空间差异性,鉴于此,本文舍弃天敌因子。选择可反映马尾松毛虫为害现状的若干因子,利用 QUEST 决策树算

法实现 2010 年无危害、轻度危害、中度危害、重度危害等 4 个虫害等级信息的提取,依次编码为 1、2、3、4。

2.3.5 人文环境因子分析与获取

人文环境因子主要包含两层内容,一是人们生活所依赖的交通与居民地,二是人类对森林的直接或间接作用。森林砍伐、经营管理等都会影响松毛虫害的发生;而人类活动所产生的大气污染、水土流失、土壤结构破坏等后果势必也影响着马尾松及松毛虫的空间分布,并可能造成它们生物学特性的突变。近年来,随着森林可持续经营理念的贯彻落实,总体上看,人为活动对虫害的加剧作用有所减缓。

通过收集福建省各县(市)2010 年林业总产值及地区 GDP 数据,以二者比值命名为“林业产值指数(FOVI)”,将该指数作为人为干扰强度的判定指标:比值越大,表示人为干扰越强,反之亦然。收集的数据显示:福建省各县(市)林业总产值与 GDP 的最大比值在 40% 以上,故假设当林业总产值与地区 GDP 的比值达到 50% 时,林业产值指数 FOVI 为 1,则可将 FOVI 的大小限定在 0~1 范围内。其表达式为

$$I_{FOVI} = \frac{R_i}{50} \tag{5}$$

式中 I_{FOVI} ——林业产值指数
 R_i ——林业总产值与地区 GDP 比值

经实验,反距离权重插值(Inverse distance weighting interpolation, IDW)所得到的 2010 年林业总产值与地区 GDP 比值与收集的数据吻合性较强,故采用该法得到福建省 2010 年林业总产值与地区 GDP 比值的空间分布图。依据式(5),计算得到 2010 年林业产值指数分布图。

2.3.6 相关分析

通过对气候、地形、林分、虫源和天敌、人文环境等因子进行分析,计算得到 148 个子因子;采用相关分析法,筛选得到海拔高度(X_1)、龄组(X_2)、2010 年虫害等级(X_3)、2010 年 8 月下旬平均气温(X_4)、2011 年 1 月上旬平均气温(X_5)、2010 年 8 月中旬降雨量(X_6)、2011 年 3 月上旬降雨量(X_7)、2010 年 7 月中旬日照时数(X_8)、2010 年 12 月中旬日照时数(X_9)、2011 年 1 月上旬极低气温(X_{10})、郁闭度(X_{11})、2010 年平均气温(X_{12})及 2010 年日照时数(X_{13})共 13 个子因子。

2.4 Fisher 判别分析

当收集到一个新的样本数据,要确定该样本数据应当归于已知类别中的哪一类时,该问题即属于判别分析的范畴。判别分析(Discriminant

analysis),又称为分辨法,是在分类确定的条件下,依据研究对象的各种特征值判别其类型归属的一种多变量统计分析方法^[27-28]。判别分析的基本原理是:按照一定的判别准则,建立一个或者多个判别函数,以研究对象的大量样本资料为基础,确定判别函数中的待定系数,并计算判别指标;依此确定新样本数据的归属^[29-30]。Fisher 判别分析法(Fisher discriminant analysis, FDA)由 Fisher 于 1936 年提出,是最为基本的一类判别分析方法,亦称为线性判别分析法,其基本思想是将 k 组 m 维样本数据投影到某一个方向,使投影后各组之间尽可能地分开,依据组内方差尽量小、组间方差尽量大的一元方差分析原则确定判别函数,基于一定的判别准则,确定新样本的归属^[31-35]。

3 结果与分析

3.1 马尾松毛虫害等级的判别函数建立

将 392 组数据随机分为建模组与验证组,两组样本量分别为 292、100。在 SPSS 软件下导入 292 组建模组样本数据,运行 Fisher 判别分析,得到 3 个判别函数

$$Y_1 = 1.679X_1 + 0.432X_2 + 0.251X_3 + 3.425X_4 - 0.811X_5 + 2.401X_6 - 6.844X_9 + 0.770X_{11} + 1.545X_{12} + 0.005X_{13} - 38.711 \tag{6}$$

$$Y_2 = 3.173X_1 - 0.020X_2 - 0.146X_3 - 0.507X_4 - 1.456X_5 + 1.814X_6 - 0.099X_9 + 0.833X_{11} - 0.263X_{12} - 0.001X_{13} + 6.780 \tag{7}$$

$$Y_3 = 0.458X_1 + 0.807X_2 + 0.301X_3 + 0.201X_4 - 0.473X_5 - 0.421X_6 + 1.816X_9 - 1.206X_{11} - 0.839X_{12} - 0.002X_{13} + 12.859 \tag{8}$$

在 Fisher 判别分析中,判别方程的解释量可以用其方差所占的比例来衡量。

表 3 显示,第 1 个判别函数方差所占比例为 60.4%,说明其可解释样本 60.4% 的信息;第 2 个判别函数方差所占比例 37.2%,联合 1、2 判别函数可以解释样本 97.6% 的信息,利用 1、2 两个函数可以实现绝大多数样本虫害等级的判别;当联合运用 3 个判别函数时,所有的样本均可以得到解释。当使用第 1 个判别模型无法实现样本类别的明确判断时,可结合第 2 个判别式进行判断,依次类推。

表 4 所示为各典型判别函数在各类别(即 4 个虫害等级)的中心值。以判别函数式(6)为例,其在虫害等级 1(无危害)的中心得分是 0.075,在虫害等级 2(轻度危害)的中心得分是 0.955,在虫害等级 3(中度危害)的中心得分是 -1.710,在虫害等级 4(重度危害)的中心得分是 -3.245。在此基础上,

可以通过比较样本函数值与 4 个虫害等级中心值的距离来判别其归属,离某个中心越近,即将该样本判别为该类别。

表 3 典型判别式函数

Tab.3 Typical discriminant functions

函数	特征值	方差所占 比例/%	累积 比例/%	正则 相关性
式(6)	1.496	60.4	60.4	0.774
式(7)	0.922	37.2	97.6	0.693
式(8)	0.059	2.4	100	0.235

表 4 典型判别函数在各类别的中心值

Tab.4 Central values of typical discriminant functions
in different categories

虫害等级	函数		
	式(6)	式(7)	式(8)
1	0.075	1.248	-0.013
2	0.955	-0.754	0.062
3	-1.710	-0.742	-0.473
4	-3.245	-0.504	0.632

图 1 为运用式(6)和式(7)判别函数对样本进行分组的示意,据此可以清晰地分辨各个组别。总体上看,4 个等级样本在空间上具有较好的可分性,同一等级样本聚集性较强,有助于样本虫害等级的判别;但亦存在误判的可能性。

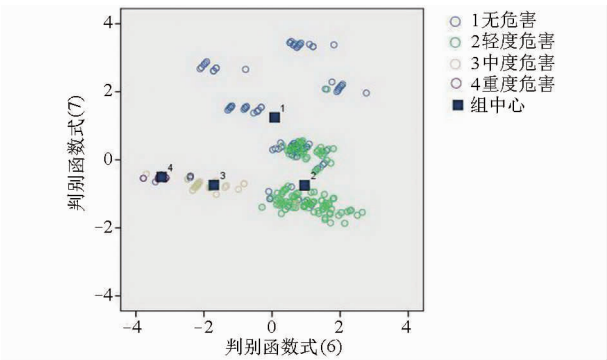


图 1 典型判别函数组中心图

Fig.1 Group centers of typical discriminant functions

3.2 松毛虫害 Fisher 判别效果检验

3.2.1 判别效果检验概述

本文对松毛虫害 Fisher 判别效果的检验分为两部分:其一,利用建模数据进行回代估计;其二,采用验证组数据对 Fisher 判别函数的效果进行检验,分析判别精度与准确率。准确率的计算方法为:若划分的虫害等级与实际等级完全一致,计 1;无危害与有危害之间互相混淆,计 0;在轻度、中度、重度 3 个有危害等级间,若相差 1 级以内为基本符合,计 0.75;若相差 2 级范围,计 0.5;以其平均值作为虫害等级提取的准确率。

3.2.2 建模组样本自检结果分析

以上述 3 个判别函数为依据,判定 292 个样本的归属类别(表 5)。

表 5、表 6 显示,292 个建模数据中,4 个等级的判别精度分别为 71.0%、84.9%、87.2%、100%,总判别精度为 80.8%。

表 5 Fisher 判别分析自检结果(部分)

Tab.5 Self-test results based on Fisher
discriminant analysis (part)

序号	实际虫害 等级	预测虫害 等级	式(6)	式(7)	式(8)
1	2	2	1.560	-1.708	1.002
2	2	2	1.579	-1.271	0.279
3	2	2	1.502	-1.354	0.400
4	1	1	0.608	0.348	-0.149
5	1	1	0.531	0.265	-0.028
6	1	1	0.647	0.390	-0.209
7	2	2	0.502	0.174	0.065
8	2	2	1.405	0.177	1.619
9	2	2	1.482	0.260	1.499
10	1	1	0.185	0.319	-0.923
11	3	3	-0.823	-0.702	1.114
12	3	3	-2.235	-0.768	-1.127
13	3	3	-1.726	-0.704	-0.440
14	3	3	-1.726	-0.704	-0.440
15	1	3	-1.764	-0.746	-0.380
16	1	4	-3.781	-0.544	-0.268
17	1	4	-3.311	-0.522	0.479
18	4	4	-3.349	-0.563	0.539
19	4	4	-3.272	-0.480	0.419
20	4	4	-3.272	-0.480	0.419

表 6 Fisher 判别分析自检精度分析

Tab.6 Self-test accuracy analysis based on Fisher
discriminant analysis

预测虫害 等级	参数	虫害等级				合计
		1	2	3	4	
1	个数	76	24	2	5	107
	精度/%	71.0	22.4	1.9	4.7	100
2	个数	16	107	1	2	126
	精度/%	12.7	84.9	0.8	1.6	100
3	个数	1	2	34	2	39
	精度/%	2.6	5.1	87.2	5.1	100
4	个数	0	0	0	19	19
	精度/%	0	0	0	100	100

3.2.3 验证组样本检验分析

将 100 组验证组样本分别代入判别函数式(6)~(8),计算其函数值;利用最小距离法判定样本虫害等级,并与实际值进行比较(表 7)。

表 7 显示,100 个验证组样本中,有 73 个样本

的虫害等级被正确判别,而其余的 27 个样本的虫害等级判别错误,判别精度 73.00%,预测准确率为

75.99%,表明利用 Fisher 判别分析预测马尾松毛虫害等级具有较好的可行性。

表 7 验证组样本 Fisher 判别结果

Tab.7 Fisher discriminant results of validation group samples

样本号	得分	判别等级	实际等级	样本号	得分	判别等级	实际等级	样本号	得分	判别等级	实际等级	样本号	得分	判别等级	实际等级
1	-0.293	1	1	26	0.388	1	1	51	-2.080	3*	1	75	-2.749	4	4
2	-0.409	1	1	27	0.349	1	1	52	-2.003	3*	1	77	-2.788	4	4
3	-1.273	3	3	28	-0.814	1	1	53	-4.546	4*	1	78	-2.788	4	4
4	-1.080	3	3	29	-1.717	3	3	54	-4.469	4*	1	79	-0.755	1	1
5	-0.945	3	3	30	-0.814	1	1	55	-4.546	4*	1	80	-0.253	1	1
6	-0.293	1	1	31	-0.853	3*	1	56	-1.395	3*	1	81	-1.555	3	3
7	-0.332	1	1	32	-0.776	1	1	57	-1.723	3*	1	82	-1.304	3	3
8	-1.196	3	3	33	-0.776	1*	2	58	-1.144	3*	2	83	-1.381	3	3
9	-1.196	3	3	34	-1.509	3	3	59	-1.608	3*	1	84	-1.014	3	3
10	1.014	2	2	35	-1.903	3	3	60	-0.210	1*	2	85	-1.401	3	3
11	1.053	2	2	36	-1.903	3	3	61	-1.931	3*	2	86	-1.227	3	3
12	0.112	1*	2	37	-1.509	3	3	62	-1.893	3*	2	87	-1.265	3	3
13	1.516	2	2	38	-1.956	3	3	63	-0.558	1*	3	88	-1.111	3	3
14	1.175	2	2	39	-1.563	3	3	64	-0.997	3*	2	89	-1.555	3	3
15	1.213	2	2	40	-2.046	3	3	65	-2.412	3	3	90	-1.516	3	3
16	0.073	1*	2	41	-0.737	1	1	66	-1.942	3*	1	91	-1.478	3	3
17	0.860	2	2	42	-0.904	3	3	67	-2.057	3*	1	92	-1.987	3	3
18	0.324	1*	2	43	-1.768	3	3	68	-2.019	3*	1	93	-1.265	3	3
19	0.899	2	2	44	-1.768	3	3	69	-2.303	3	3	94	-1.014	3	3
20	1.021	2	2	45	-1.717	3	3	70	-0.937	3	3	95	-1.304	3	3
21	1.021	2	2	46	-1.717	3	3	71	-0.937	3	3	96	-1.304	3	3
22	0.157	1	1	47	-0.699	1	1	72	-2.167	3	3	97	-2.455	3*	1
23	0.157	1	1	48	-1.956	3*	1	73	-2.529	4	4	98	-2.493	4*	1
24	0.157	1	1	49	-2.072	3*	1	74	-2.058	3	3	99	-2.339	3	3
25	0.157	1	1	50	-1.918	3*	1	76	-2.788	4	4	100	-2.378	3	3

注：* 表示误判。

4 结束语

针对当前马尾松毛虫害等级遥感预测成果不足的现状,依托野外实测数据、HJ-1 CCD 多光谱数据及气候、林分等部门数据,建立马尾松毛虫害等级的 Fisher 判别函数。建模组自检结果显示,无危害、轻度危害、中度危害及重度危害等 4 个虫害等级的判

别精度分别为 71.0%、84.9%、87.2%、100%,总判别精度为 80.8%;验证组样本的检验结果表明,总精度 73.00%,预测准确率为 75.99%。上述结果证实 Fisher 判别具备马尾松毛虫害 4 个等级的判别能力,也证明国产 HJ-1 CCD 数据对森林虫害有显著的响应。

参 考 文 献

1 武红智,陈改英. 基于 GIS 的马尾松毛虫灾害空间扩散规律分析[J]. 遥感学报, 2004, 8(5): 475-480.
Wu Hongzhi, Chen Gaiying. Research on development rules of forest hazard caused by Masion pine caterpillars based on GIS[J]. Journal of Remote Sensing, 2004, 8(5): 475-480. (in Chinese)

2 戴昌达,雷莉萍,胡德永,等. 卫星遥感监测松毛虫灾[J]. 遥感信息, 1991(3): 32-34.

3 侯陶谦. 中国松毛虫[M]. 北京: 科学出版社, 1987.

4 武红敢,崔恒建,乔彦友,等. 密郁闭林分针叶林失叶量遥感监测模型初探[J]. 国土资源遥感, 1995(2): 29-35.
Wu Honggan, Cui Hengjian, Qiao Yanyou, et al. Modeling defoliation conditions in moderate canopy density and closed coniferous forest using Landsat TM data[J]. Remote Sensing for Land & Resources, 1995(2): 29-35. (in Chinese)

5 武红敢,黄建文,乔彦友,等. 利用航天遥感资料监测森林失叶量的可行性研究[J]. 林业科技通讯, 1994(7): 25.

6 武红敢,黄建文,乔彦友,等. 松毛虫早期灾害点遥感监测研究初报[J]. 林业科学, 1995, 31(4): 379-384.

Wu Honggan, Huang Jianwen, Qiao Yanyou, et al. A preliminary study on remote sensing detection of damage by pine caterpillar [J]. *Scientia Silvae Sinicae*, 1995, 31(4): 379–384. (in Chinese)

7 杨俊泉, 陈尚文, 沈建中, 等. 马尾松毛虫危害区植被指数时序变化特征研究[J]. *国土资源遥感*, 1997(4): 7–13.
Yang Junquan, Chen Shangwen, Shen Jianzhong, et al. Study for the character of NDVI time-series variation of pine caterpillar moth injury region by NOAA satellite image[J]. *Remote Sensing for Land & Resources*, 1997(4): 7–13. (in Chinese)

8 亓兴兰, 刘健, 陈国荣, 等. 应用 MODIS 遥感数据监测马尾松毛虫害研究[J]. *西南林学院学报*, 2010, 30(1): 42–46, 54.
Qi Xinglan, Liu Jian, Chen Guorong, et al. Studies on monitoring of *Dendrolimus punctatus* damage with MODIS data[J]. *Journal of Southwest Forestry College*, 2010, 30(1): 42–46, 54. (in Chinese)

9 常红, 王序铨. 应用微机对遥感 TM 图象进行森林资源分类的研究[J]. *西北师范大学学报: 自然科学版*, 1998, 34(1): 69–74.
Chang Hong, Wang Xuquan. Application of remote sense TM image to forest resource classification by computer[J]. *Journal of Northwest Normal University: Natural Science*, 1998, 34(1): 69–74. (in Chinese)

10 韩涛. 用 TM 资料对祁连山部分地区进行针叶林、灌木林分类研究[J]. *遥感技术与应用*, 2002, 17(6): 317–321.
Han Tao. Study on classification between coniferous forest and shrubbery based on TM data in the part of Qi Lian Mountainous area[J]. *Remote Sensing Technology and Application*, 2002, 17(6): 317–321. (in Chinese)

11 王蕾, 黄华国, 张晓丽, 等. 基于知识规则的马尾松林遥感信息提取技术研究[J]. *北京林业大学学报*, 2007, 29(3): 124–130.
Wang Lei, Huang Huaguo, Zhang Xiaoli, et al. RS information extraction techniques of *Pinus massoniana* forests based on knowledge rules[J]. *Journal of Beijing Forestry University*, 2007, 29(3): 124–130. (in Chinese)

12 Khalyani A H, Falkowski M J, Mayer A L. Classification of Landsat images based on spectral and topographic variables for land-cover change detection in Zagros forests[J]. *International Journal of Remote Sensing*, 2012, 33(21): 6956–6974.

13 Balthazar V, Vanacker V, Lambin E F. Evaluation and parameterization of ATCOR3 topographic correction method for forest cover mapping in mountain areas[J]. *International Journal of Applied Earth Observation and Geoinformation*, 2012, 18: 436–450.

14 Evrendilek F, Gulbeyaz O. Boosted decision tree classifications of land cover over Turkey integrating MODIS, climate and topographic data[J]. *International Journal of Remote Sensing*, 2011, 32(12): 3461–3483.

15 余坤勇, 刘健, 许章华, 等. 南方地区竹资源专题信息提取研究[J]. *遥感技术与应用*, 2009, 24(4): 449–455.
Yu Kunyong, Liu Jian, Xu Zhanghua, et al. Study on bamboo resources thematic information extraction in the south of China[J]. *Remote Sensing Technology and Application*, 2009, 24(4): 449–455. (in Chinese)

16 余坤勇, 许章华, 刘健, 等. “基于片层-面向类”的竹林信息提取算法与应用分析[J]. *中山大学学报: 自然科学版*, 2012, 51(1): 89–95.
Yu Kunyong, Xu Zhanghua, Liu Jian, et al. Algorithm and application analysis of film-based & class-oriented for bamboo forest information extraction[J]. *Acta Scientiarum Naturalium Universitatis Sunyatseni*, 2012, 51(1): 89–95. (in Chinese)

17 尚忠海, 光增云, 李世海, 等. 北带马尾松毛虫灾害发生级别的气候因子预报模型研究[J]. *生物数学学报*, 1997, 12(5): 606–610.
Shang Zhonghai, Guang Zengyun, Li Shihai, et al. Study on the model of climatic factor forecast occurring in difference of north *Dendrolimus punctatus* damage[J]. *Journal of Biomathematics*, 1997, 12(5): 606–610. (in Chinese)

18 Coops N C, Warning R H. A process-based approach to estimate lodgepole pine (*Pinus contorta* Dougl.) distribution in the Pacific Northwest under climate change[J]. *Climatic Change*, 2011, 105(1–2): 313–328.

19 Boden S, Pyttel P, Eastaugh C S. Impacts of climate change on the establishment, distribution, growth and mortality of Swiss stone pine (*Pinus cembra* L.) [J]. *iForest-Biogeosciences and Forestry*, 2010(3): 82–85.

20 Gomez-Mendoza L, Arriaga L. Modeling the effect of climate change on the distribution of oak and pine species of Mexico[J]. *Conservation Biology*, 2007, 21(6): 1545–1555.

21 陈绘画, 朱寿燕, 崔相富. 基于人工神经网络的马尾松毛虫发生量预测模型的研究[J]. *林业科学研究*, 2003, 16(2): 159–165.
Chen Huihua, Zhu Shouyan, Cui Xiangfu. A study on the forecast model of *Dendrolimus punctatus* occurrence based on artificial neural network[J]. *Forest Research*, 2003, 16(2): 159–165. (in Chinese)

22 朱建华, 陈顺立, 张再福. 森林病虫害预测预报[M]. 厦门: 厦门大学出版社, 2002.

23 魏初奖, 庄晨辉, 谢大洋, 等. 福建省马尾松毛虫灾区区划及其应用[J]. *南京林业大学学报: 自然科学版*, 2002, 26(3): 35–39.
Wei Chujiang, Zhuang Chenhui, Xie Dayang, et al. Plague division and application of *Dendrolimus punctatus* Walker in Fujian province[J]. *Journal of Nanjing Forestry University: Natural Science Edition*, 2002, 26(3): 35–39. (in Chinese)

24 Weiss A. Topographic position and landforms analysis[R]. San Diego, CA: ERSI User Conference, 2001.

25 韦金丽, 王国波, 凌子燕. 基于高分辨率 DEM 的地形特征提取与分析[J]. *测绘与空间地理信息*, 2012, 35(1): 33–36.
Wei Jinli, Wang Guobo, Ling Ziyang. The extraction and analysis of landform characters based on high-resolution DEM[J]. *Geomatics & Spatial Information Technology*, 2012, 35(1): 33–36. (in Chinese)

26 段霞. 人工模拟摘叶对马尾松单株承受力的测定[J]. *河南林业科技*, 2006, 26(4): 11–12.

27 Alonzo M, Roberts K, Roberts D. Identifying santa barbara’s urban tree species from AVIRIS imagery using canonical discriminant analysis[J]. Remote Sensing Letters, 2013, 4(5): 513 – 521.

28 Liao W Z, Pizurica A, Scheunders P, et al. Semisupervised local discriminant analysis for feature extraction in hyperspectral images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2013, 51(1): 184 – 198.

29 Wang Y W, Dunn B L, Arnall D B. Assessing nitrogen status in potted geranium through discriminant analysis of ground-based spectral reflectance data[J]. Hortscience, 2012, 47(3): 343 – 348.

30 Prades C, Gomez-Sanchez I, Garcia-Olmo J, et al. Discriminant analysis of geographical origin of cork planks and stoppers by near infrared spectroscopy[J]. Journal of Wood Chemistry and Technology, 2012, 32(1): 54 – 70.

31 Chen X H, Wang S Q, Chen J, et al. New algorithm for spectral mixture analysis based on Fisher Discriminant Analysis: evidence from laboratory experiment[J]. Journal of Infrared and Millimeter Waves, 2009, 28(6): 476 – 480.

32 Du Qian. Modified Fisher’s linear discriminant analysis for hyperspectral imagery[J]. IEEE Geoscience and Remote Sensing Letters, 2007, 4(4): 503 – 507.

33 李秀珍, 王成华, 宋刚. 基于 Fisher 判别分析法的潜在滑坡判识模型及其应用[J]. 中国地质灾害与防治学报, 2009, 20(4): 23 – 26.

Li Xiuzhen, Wang Chenghua, Song Gang. Potential landslide identification model based on Fisher discrimination analysis method and its application[J]. The Chinese Journal of Geological Hazard and Control, 2009, 20(4): 23 – 26. (in Chinese)

34 朱永生. 实验数据多元统计分析[M]. 北京: 科学出版社, 2009.

35 王文川, 吴海波, 赵晓慎, 等. 基于 Fisher 判别分析法的泥石流预报模型研究[J]. 水土保持通报, 2011, 31(6): 104 – 107.

Wang Wenchuang, Wu Haibo, Zhao Xiaoshen, et al. Forecasting model of debris flow based on Fisher discrimination analysis method[J]. Bulletin of Soil and Water Conservation, 2011, 31(6): 104 – 107. (in Chinese)

Fisher Discriminant Analysis of *Dendrolimus punctatus* Walker Pest Levels

Xu Zhanghua¹ Li Conghui² Liu Jian^{3,4} Yu Kunyong^{3,4} Gong Conghong^{3,4} Tang Mengya^{3,4}
(1. College of Environment and Resources, Fuzhou University, Fuzhou 350108, China
2. College of Nature Conservation, Beijing Forestry University, Beijing 100083, China
3. Institute of Geomatics Application, Fujian Agriculture and Forestry University, Fuzhou 350002, China
4. College of Forestry, Fujian Agriculture and Forestry University, Fuzhou 350002, China)

Abstract: The *Dendrolimus punctatus* Walker is the most serious forest defoliator in China. The study of pest remote sensing prediction is helpful to the comprehensive and timely identification of possible damage regions, and can provide the technical support for the direct and indirect losses reduction. Taking Fujian Province as the study area, the *Dendrolimus punctatus* Walker overwintering generation was investigated in 13 monitoring stations, Sanming City, Jiangle County, Shaxian County, Nanping City, Huaan County, Yunxiao County, Nan’an City, Anxi County, Putian City, Changting County, Jianyang City, Ningde City, and Fuqing City, successively from February to May 2012. The investigation indices contained GPS coordinates, pine forests canopy spectrum, climate, terrain, stand, pest source, human environment and so on. After the analysis and extraction of these factors above, the sub-factors used to predict the *Dendrolimus punctatus* Walker pest levels were obtained with correlation analysis. With the help of Fisher discriminant analysis, the discriminant functions of 4 pest levels, non-damage, mild damage, moderate damage and severe damage, were constructed, and the self-test result with samples of modeling group showed that the discriminant accuracies were 71.0% , 84.9% , 87.2% , 100% respectively, with total discriminant accuracy of 80.8% . The test result with samples of validation group showed that the discriminant accuracy and accurate rate were 73.00% and 75.99% respectively, which indicated that there was good feasibility to use Fisher discriminant analysis to predict the *Dendrolimus punctatus* Walker pest levels.

Key words: *Dendrolimus punctatus* Walker Pest levels Fisher discriminant analysis Correlation analysis