

基于多源数据的新疆棉田螨害大范围监测研究

杨丽丽, 王振鹏, 吴才聪*

中国农业大学信息与电气工程学院, 北京 100083

摘要 针对新疆棉田传统螨害监测方法耗时低效的问题, 提出了一种基于冠层高光谱、近地多光谱、环境数据与地面调查相结合的多源数据棉田螨害大范围监测方法。首先, 分别采集地面尺度的棉花冠层 350~2500 nm 高光谱遥感数据和不同时期低空尺度的棉田无人机多光谱遥感影像数据, 通过分析高光谱的原始光谱和一阶微分光谱特征, 提取出了 4 个螨害敏感波段, 分别为: 绿光波段 553 nm 附近、红光波段 680 nm 附近、红边波段 680~750 nm、近红外波段 760~1350 nm, 这几个波段同时包含在无人机搭载的多光谱传感器波段范围内, 验证了低空尺度下无人机遥感螨害监测的可行性。其次, 初选 23 种植被指数和 13 种田间环境数据, 结合地面调查的螨害发生情况做相关性分析。其中, SAVI、OSAVI、TVI、NDGI、平均湿度、温湿系数和 10 cm 土壤平均温度均与螨害发生达到极显著相关水平 ($\text{sig} \leq 0.01$); RDVI、RVI、MSR、最高温度、平均温度、积温、10 cm 土壤最高温度和 10 cm 土壤平均湿度均与螨害发生达到显著相关水平 ($\text{sig} \leq 0.05$)。选取 sig 值在 0.05 以下的 15 种特征值, 分别建立基于单一环境数据、单一植被指数、环境数据与植被指数相结合的 3 种支持向量机(SVM)棉田螨害发生监测模型。最后, 根据评价效果最优的监测模型, 绘制不同时期的螨害遥感监测空间分布图, 通过统计分布图中螨害和健康像元数计算出螨害面积占比, 将螨害占比与同时期田间环境数据进行相关性分析, 筛选出显著特征值, 再通过多元逐步回归分析法确定出与螨害面积值关系最密切的环境因子, 建立棉田螨害面积预测模型。结果表明: 基于单一环境数据的棉田螨害发生监测模型准确率为 62.22%, 基于单一植被指数的棉田螨害发生监测模型准确率为 75.56%, 基于环境数据与植被指数相结合的棉田螨害发生监测模型效果最优, 准确率为 80%。螨害面积预测模型的决定系数 $R^2 = 0.848$, 模型拟合度较好。本研究基于多源数据建立的棉田螨害发生监测模型和螨害面积预测模型, 可以为新疆地区棉田螨害的大范围监测和趋势预警提供参考。

关键词 无人机遥感; 棉叶螨; 环境数据; 监测; 预测

中图分类号: TP751

文献标识码: A

DOI: 10.3964/j.issn.1000-0593(2021)12-3949-08

引言

新疆作为我国最重要的产棉基地, 保障棉花产业健康发展, 对促进新疆经济具有重要意义^[1]。病虫害的大范围爆发严重影响了棉花的产量和品质, 其中棉叶螨因其隐蔽性强、个体微小的特点, 导致监测难度大, 危害程度深^[2]。传统的螨害监测方式主要是由当地植保人员定期在棉田内凭经验进行巡查, 手工标记棉叶螨受害中心植株位置。这种监测方式范围小、人力消耗大且时效性差, 不能准确评估大范围棉田的受害情况。

国内外学者在作物病虫害遥感监测方面进行了大量研

究。兰玉彬等^[3]通过获取柑橘果园的无人机高光谱影像建立柑橘黄龙病判别模型, 实现了低空高光谱遥感对柑橘黄龙病的大面积监测。黄焕华等^[4]利用小型固定翼无人机同时采集枯死松树的可见光和近红外航摄影像, 对松材线虫病的监测准确率可达到 80% 以上。田野^[5]利用多时相 HJ 小卫星遥感影像实现棉叶螨害监测, 但因其光谱分辨率难以达到要求还需结合地面光谱数据进一步研究。王守会等^[6]基于棉花冠层高光谱数据和 Landsat8 卫星遥感数据对区域螨害发生进行预测研究, 模型预测效果较好, 但卫星遥感影像分辨率过低使得监测样点之间的光谱差异大大降低。崔美娜等^[7]基于无人机多光谱影像建立棉田螨害监测识别模型, 实现区域范围内螨害识别, 但并未考虑棉田气象和土壤环境对螨害的影

收稿日期: 2020-11-03, 修订日期: 2021-02-19

基金项目: 国家重点研发计划项目(2016YFB0501805)资助

作者简介: 杨丽丽, 女, 1974 年生, 中国农业大学信息与电气工程学院副教授 e-mail: llyang@cau.edu.cn

* 通讯作者 e-mail: wucc@cau.edu.cn

响。赵亮等^[8]通过获取不同螨害程度叶片的地面高光谱数据,筛选出螨害敏感波段,建立了棉叶螨叶片光谱识别的最佳模型。

本研究利用无人机遥感的高时效性、低成本和高时空分辨率,弥补了地面光谱难以实现以点带面以及卫星遥感空间分辨率低、时效性差的问题,并综合考虑不同环境因素对于螨害发生的影响,建立了基于多源数据的棉田大范围螨害监测模型和螨害面积预测模型,可在实际应用环节完成快速大范围的监测需求,为棉田变量施药提供依据,有效减少损失。

1 实验部分

1.1 研究区概况

研究区位于新疆生产建设兵团第八师 145 团六分场六连,中心位置为 86°3′30.85″E, 44°32′30.23″N,属中高纬度的典型大陆性气候,光照充足,雨量稀少,气候干燥,冬季严寒,夏季严热,研究区气候条件适宜棉叶螨正常生长^[9]。研究区棉田面积 34.7 万 m²,约 520 亩,种植品种为新陆早 166 号和新陆早 163 号,采用 1 膜 3 管 6 行种植模式,蕾期一般处于当地的 6 月上中旬至 7 月上旬。

1.2 数据获取与处理

1.2.1 棉花冠层高光谱数据采集与处理

2018 年于北京时间 7 月 12 日 12:00—14:00 对蕾期棉花进行不同螨害程度的冠层光谱采集,采集设备为 ASD Field-Spec Pro FR2500 型背挂式野外高光谱辐射仪,光谱波长范围为 350~2 500 nm,采样间隔为 1 nm,光谱仪视场角为 25°。采集时保持探头垂直向下,光谱仪距离棉花冠层顶垂直高度约为 0.5 m,每次测量前后均进行标准白板校正,每个采样点采集 5 组光谱数据,取平均值作为该采样点的光谱值。

1.2.2 无人机多光谱数据采集与处理

2019 年棉花蕾期内,选择晴朗无云、无风或风力较小的天气,分别于 6 月 22 日、27 日、29 日和 7 月 3 日、9 日、13 日在研究区内进行 6 次多光谱数据采集,采集设备为瑞士 eBeeSQ 固定翼无人机搭载 Parrot Sequoia 五通道多光谱相机。该相机镜头焦距为 3.98 mm,分辨率为 1 280×960,可同时获取 4 个多光谱波段和可见光数据。4 个多光谱波段分别为绿光(中心波长 550 nm,波段宽度 40 nm)、红光(660 nm, 40 nm)、红边光(735 nm, 10 nm)和近红外光(790 nm, 40 nm)。

无人机起飞前用 Sequoia 多光谱相机配备的参考板进行辐射定标处理,根据预设航线垂直拍摄,影像范围覆盖整个研究地块。设置无人机飞行航线为 S 型,飞行高度为 120 m,飞行速度为 13 m·s⁻¹,旁向重叠度为 60%,航向重叠度为 75%,影像空间分辨率为 0.12 m。数据采集后,将影像数据导入 Pix4Dmapper 软件进行辐射定标和拼接处理。

1.2.3 螨害实地调查数据采集

在每期无人机多光谱影像采集的时间点附近由人工进行实地叶片螨害抽样调查。在研究区棉田内使用 VIVO X20 手

机拍摄健康和螨害叶片图像,采样比例 1:1,共 135 个采样点。采样时记录叶片受害程度并根据手机自带的 GPS 精确定位功能记录采样点经纬度信息。螨害等级参照国家标准《棉花叶螨测报技术规范》进行划分为 4 级:0 级为完害;1 级为叶面有零星黄斑;2 级为红斑占叶面积 1/3 以下;3 级为红斑占叶面积 1/3 以上。4 种不同螨害程度棉叶示例图像如图 1。



图 1 不同螨害程度棉叶图像

(a): 0 级; (b): 1 级; (c): 2 级; (d): 3 级

Fig. 1 Images of cotton leaves with different degree of mite damage

(a): Level 0; (b): Level 1; (c): Level 2; (d): Level 3

1.2.4 环境数据采集与处理

无人机多光谱数据采集期内,通过棉田研究区搭建的气象监测系统和土壤墒情设备,获取 2019 年 6 月—2019 年 7 月期间逐日每小时的气象数据和土壤温湿度数据共 13 种田间环境数据。再以 4 d 作为一个周期进行统计,计算出 9 种气象数据,即每周期最高温度、平均温度、最低温度、平均湿度、累计降雨量、平均风力、温湿系数、温雨系数、积温(逐日平均温度 ≥ 25 °C); 4 种土壤数据,即 10 cm 土壤最高温度、10 cm 土壤最低温度、10 cm 土壤平均温度和 10 cm 土壤平均湿度。同时考虑到农户施药对螨害的影响,在施药日期(6 月 24 日和 7 月 8 日)节点附近将周期长度细化为 2 d。

2 研究方法

首先,通过采集棉花冠层高光谱数据及无人机多光谱影像数据,筛选敏感波段位置,构建螨害监测植被指数;其次,通过相关性分析筛选相关性显著的螨害监测植被指数和田间环境数据,建立基于支持向量机(support vector machine, SVM)的棉田螨害发生监测模型;最后,基于最优模型分类结果生成螨害发生空间分布图,通过统计健康点和螨害点的像元数计算螨害面积占比值,并结合田间环境数据采用多元逐步回归分析方法建立棉田螨害面积预测模型。总体技术路线如图 2 所示。

2.1 螨害监测植被指数计算

植被指数(vegetation index, VI)通过不同波段反射率的线性或非线性组合变化,在一定程度上可以突显植被的不同特征,削弱背景等因素对植被光谱特征的干扰。工作中根据地面尺度的棉花冠层高光谱数据提取出多个螨害敏感波段,从中选取与低空尺度的无人机多光谱波段位置相对应的 4 个敏感波段,初步构建 23 种螨害监测植被指数。植被指数及其计算公式如表 1 所示。

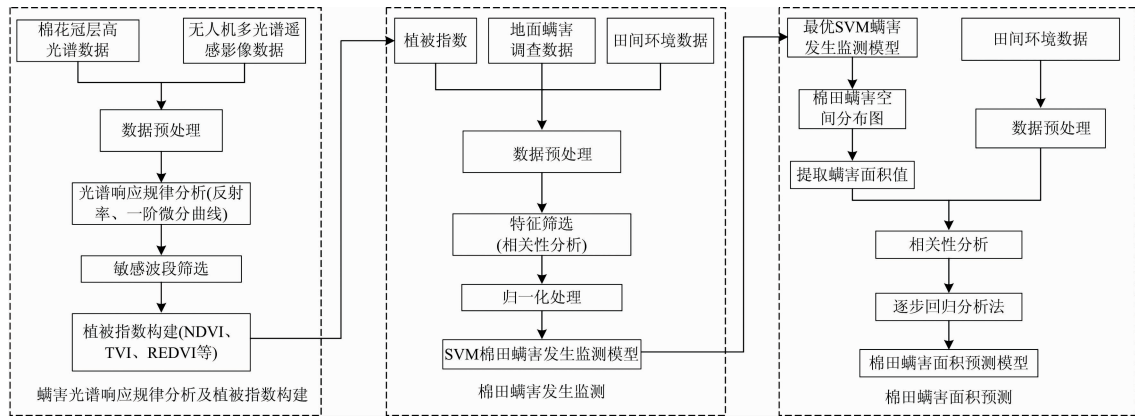


图 2 棉田螨害监测技术路线

Fig. 2 Technical route of cotton field spider mite monitoring

表 1 植被指数及其计算公式

Table 1 Calculation formulas for vegetation index

植被指数	公式
NDGI	$(G - R) / (G + R)$
TVI	$0.5 \times [120 \times (NIR - G) - 200 \times (R - G)]^{[10]}$
MSR	$(NIR / R - 1) / [(NIR / R)^{0.5} + 1]^{[10]}$
MSAVI	$0.5 \times \{ (2 \times NIR + 1) - [(2 \times NIR + 1)^2 - 8 \times (NIR - R)]^{0.5} \}^{[11]}$
RDVI	$(NIR - R) / [(NIR + R)^{0.5}]^{[10]}$
RVI	$NIR / R^{[11]}$
REWDRVI	$(0.15 \times NIR - RE) / (0.15 \times NIR + RE)^{[11]}$
DVI	$NIR - R^{[10]}$
GNDVI	$(NIR - G) / (NIR + G)^{[11]}$
SAVI	$1.5 \times (NIR - R) / (NIR + R + 0.5)^{[12]}$
ARI	$(1/G) - (1/RE)$
OSAVI	$1.16 \times (NIR - R) / (NIR + R + 0.16)^{[12]}$
NDVI	$(NIR - R) / (NIR + R)^{[12]}$
EVI	$2.5 \times [(NIR - R) / (NIR + 6 \times R - 7.5 \times G + 1)]^{[13]}$
GRVI	$NIR / G^{[13]}$
GDVI	$NIR - G^{[13]}$
GOSAVI	$(1 + 0.16) \times (NIR - G) / (NIR + G + 0.16)^{[13]}$
REDVI	$NIR - RE^{[13]}$
RERV	$NIR / RE^{[13]}$
RESAVI	$1.5 \times [(NIR - RE) / (NIR + RE + 0.5)]^{[13]}$
RERDVI	$(NIR - RE) / [(NIR + RE)^{0.5}]^{[13]}$
RENDVI	$(NIR - RE) / (NIR + RE)^{[13]}$
REOSAVI	$(1 + 0.16) \times (NIR - RE) / (NIR + RE + 0.16)^{[13]}$

注: G, R, RE, NIR 分别为绿, 红, 红边和近红外波段光谱反射率
 Note: G, R, RE, NIR represent reflectance of green, red, red-edge and near-infrared bands respectively

2.2 模型建立

通过相关性分析筛选反映螨害最敏感的植被指数和田间环境数据作为建模特征值, 分别用 0(无螨害)和 1(有螨害)作为目标值, 将 135 个样本数据集输入 SVM 分类器进行训练和分类, 采用单一环境数据、单一植被指数、环境数据与植被指数相结合的 3 种方法建立基于 SVM 棉田螨害发生监

测模型。模型训练和测试使用 Matlab R2019b 和 LIBSVM 3.24 工具箱, 并选取准确率(Accuracy)、精确率(Precision)、召回率(Recall)和对精确率和准确率整体评价的 F_1 值作为模型精度的评价指标, 确定出最优螨害发生监测模型。

使用田间环境数据分别与螨害面积占比值作相关性分析, 筛选出显著特征值, 再基于多元逐步回归分析法选出与螨害面积值关系最密切的环境特征建立棉田螨害面积预测模型, 选取决定系数 R^2 作为模型评价指标。

3 结果与讨论

3.1 地面冠层光谱数据反射率分析

棉花叶片在遭受不同程度的棉叶螨危害后, 棉花冠层光谱反射率曲线呈现规律性变化, 呈现明显的波峰和波谷等特征。图 3 为不同螨害程度的棉花冠层光谱反射率曲线。

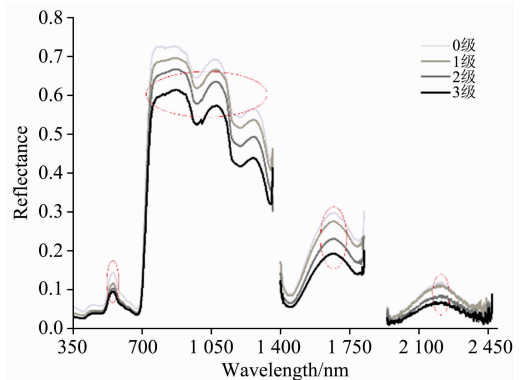


图 3 不同螨害程度地面冠层光谱反射率曲线

Fig. 3 Spectral reflectance curves of ground canopies with different degrees of mite damage

由图 3 可见, 在绿光波段范围内 553 nm 波长附近出现第一个反射波峰后反射率逐渐下降, 至红光波段范围内 680 nm 波长附近出现反射波谷-红谷, 此时具有较低的反射率。反射率于 705 nm 波长处开始呈现加速上升趋势, 且在近红外波段 760~1 350 nm 范围内反射率较高, 同时在此范围内

光谱反射率随着螨害程度的逐渐加重表现出较为明显的下降趋势,在中红外波段 1 550~1 800 nm 范围内的光谱反射率也表现出了同样的规律。

红边是绿色植物最明显的光谱特征之一。为分析不同螨害程度棉叶冠层光谱的红边特征,对红边(680~750 nm)范围内的光谱反射率求一阶微分,如图 4 所示。可以看出 680~750 nm 范围内的一阶微分光谱曲线图像呈现“双峰”现象,且随着螨害加重,反射率一阶微分值逐渐减小,表明红边波段对不同程度的螨害较为敏感。

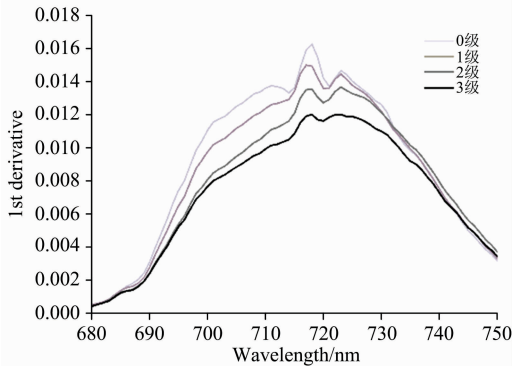


图 4 不同螨害程度地面冠层一阶微分光谱曲线

Fig. 4 First derivative spectral curve of ground canopy with the different degrees of spider mites damage

分析结果表明,棉花叶片在遭受不同程度的棉叶螨为害后,可通过棉花冠层光谱反射率及反射率一阶微分的变化进行表征,绿光波段 553 nm 附近、红光波段 680 nm 附近、红边波段 680~750 nm 范围、近红外波段 760~1 350 nm 范围都是与螨害发生相关的敏感波段,与文献[8]研究结果一致。

3.2 无人机多光谱数据反射率分析

3.1 节中高光谱的螨害敏感波段分析结果显示,本实验中无人机携带的多光谱相机波段范围属于螨害敏感波段,为进一步确定无人机携带多光谱相机进行虫害监测的可行性,对无人机多光谱遥感影像进行了不同螨害程度采样点的反射率提取,图 5 为反射率曲线。

由图 5 可见,绿光波段 550 nm 附近、550~660 nm 波段

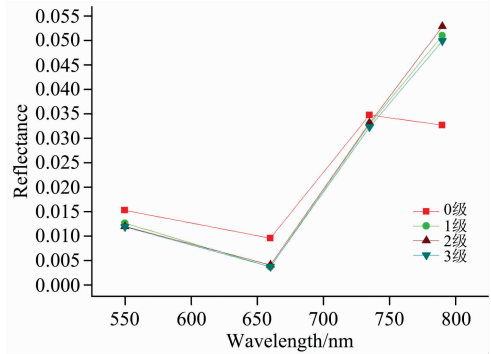


图 5 不同螨害程度的多光谱影像反射率曲线

Fig. 5 Reflectance curve of multispectral image with different degrees of spider mite damage

范围、660~735 nm 波段范围处不同等级螨害发生点的光谱反射率基本重叠在一起,可能的原因是无人机的飞行高度影响影像分辨率,造成不同螨害程度的采样点在光谱响应上的变化差异较小,光谱差异很难区分,不能进行受害等级区分,但有害和无害的光谱反射率差异比较明显。故利用无人机多光谱遥感进行螨害监测模型构建时,只构建基于 SVM 的二分类模型,将螨害 1 级及以上均视为螨害发生,即未发生螨害记为 0,发生螨害记为 1。

3.3 螨害监测模型特征筛选

将选取的 23 种植被指数和预处理后的 13 种田间环境数据分别与地面采样点的螨害发生数据做相关性分析,用显著性(sig)作为评价指标,当 $\text{sig} \leq 0.05$ 时,植被指数与螨害发生达到显著相关水平,当 $\text{sig} \leq 0.01$ 时,植被指数与螨害发生达到极显著相关水平。从表 2 看出,所选的植被指数中,SAVI, OSAVI, TVI 和 NDGI 与螨害发生呈极显著相关, RDVI, RVI 和 MSR 与螨害发生呈显著相关。从表 3 可以看出,平均湿度、温湿系数和 10 cm 土壤平均温度与螨害发生呈极显著相关,最高温度、平均温度、积温、10 cm 土壤最高温度和 10 cm 土壤平均湿度与螨害发生呈显著相关。研究中选取 $\text{sig} \leq 0.05$ 的植被指数及环境特征作为螨害监测建模的特征值。

表 2 植被指数与螨害发生之间相关性

Table 2 Correlation between the occurrence of spider mite and vegetation indices

植被指数	RDVI	SAVI	OSAVI	TVI	NDGI	RVI	MSR
相关系数	0.407*	0.222**	0.304**	0.492**	-0.304**	0.210*	0.197*

注: ** 和 * 分别表示 0.01, 0.05 的显著水平

Note: ** and * represent significant at the 0.01 and 0.05 levels of probability respectively

表 3 环境数据与螨害发生之间相关性

Table 3 Correlation between the occurrence of spider mite and environmental data

环境数据	最高温度	平均温度	平均湿度	温湿系数	积温	10 cm 土壤最高温度	10 cm 土壤平均温度	10 cm 土壤平均湿度
相关系数	-0.218*	-0.178*	0.221**	0.243**	-0.213*	-0.201*	-0.258**	0.213*

注: ** 和 * 分别表示 0.01, 0.05 的显著水平

Note: ** and * represent significant at the 0.01 and 0.05 levels of probability respectively

3.4 基于 SVM 算法的棉田螨害监测模型

选取表 2 及表 3 中的 7 种植被指数和 8 种田间环境数据作为建模特征值, 在 135 个地面采样点中, 随机选取出总样本数的 2/3 作为训练集, 其余的 1/3 作为测试集进行模型构建, 分别基于单一环境数据、单一植被指数、环境数据与植被指数相结合的 3 种方法建立 SVM 棉田螨害发生监测模型

M1, M2 及 M3。SVM 模型的精度主要与核函数和 SVM 内的参数选取有关, 研究中选择径向基函数(radial basis function, RBF)作为 SVM 训练模型的核函数, 使用交叉验证的方法调整 SVM 内的惩罚参数 c 和 γ 参数, 得到各类评价指标效果最优的监测模型。模型结果如表 4 所示。

表 4 不同螨害监测模型分类结果对比

Table 4 Comparison of classification results of the different degrees of spider mite monitoring models

建模方式	样本	训练样本			测试样本					
		健康	螨害	准确率/%	健康	螨害	准确率/%	精确率/%	召回率/%	F_1 值/%
单一环境数据 M1	健康	26	23	64.44	13	8	60	56.52	61.9	59.09
	螨害	9	32		10	14				
单一植被指数 M2	健康	27	9	76.67	15	6	68.89	65.22	71.43	68.18
	螨害	12	42		8	16				
环境数据与植被指数结合 M3	健康	26	13	82.22	12	5	80	75	70.59	72.73
	螨害	3	48		4	24				

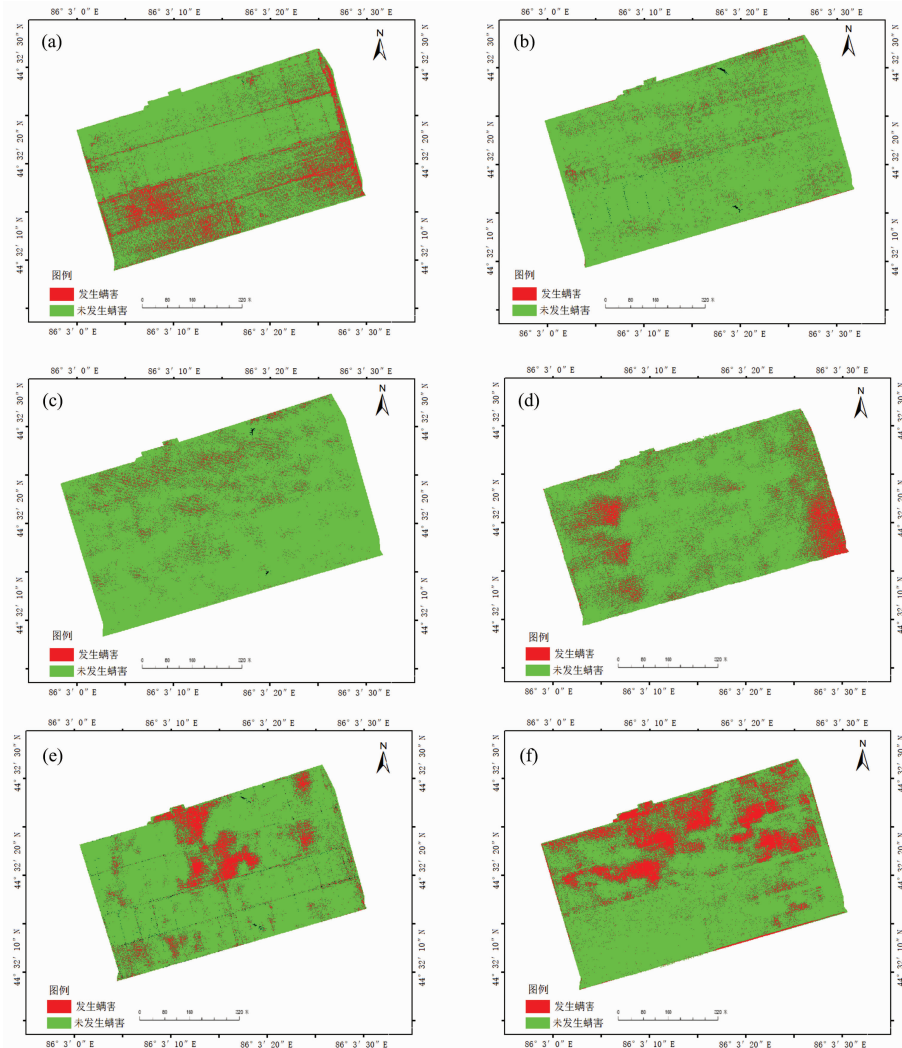


图 6 不同时期(6月 22 日, 27 日, 29 日, 7 月 3 日, 9 日, 13 日)棉田螨害监测空间分布图(a, b, c, d, e, f)

Fig. 6 Spatial distribution map (a, b, c, d, e, f) of cotton field spider mite monitoring in (June 22, 27, 28, July 3, 9, 13) different periods

表 4 中, M3 效果最好, 测试集分类准确率为 80%, 精确率为 75%, 召回率为 70.59%, F_1 值为 72.73%; M2 次之, M1 分类效果最差。表明, 同时使用环境数据和植被指数综合考虑了影响螨害发生的主要因素, 从而使模型具有更高的精度和监测效果。

3.5 棉田螨害严重度空间分布

基于棉田螨害监测模型 M3, 分别绘制 6 月 22 日、27 日、29 日和 7 月 3 日、9 日、13 日共计 6 景无人机遥感螨害监测空间分布图[图 6(a—f)]。

图 6 中可见, 6 月 22 日开始, 研究区螨害成片发生且基本分布在棉田东部以及中路以南区域[图 6(a)], 棉田实际调查发现棉田东部及南部区域靠近路边, 地边杂草较多且农户活动多聚集于两地, 同时发现监测期间研究区中部以南地区因棉田灌溉设施漏水, 棉花长势较弱, 这些因素综合导致了棉田边缘易发生的螨害扩散到了棉田中部及以南。根据 6 月 22 日棉田螨害监测情况, 农户于 6 月 24 日使用 25%阿维乙螨唑悬浮剂对叶螨进行施药防治, 从图 6(b)可以看出, 6 月 27 日分布图中螨害面积逐渐减小, 叶螨数量较之前有所下降。但由于棉田内施药时间不同及叶螨抗药性强, 观察 6 月 29 日棉田螨害监测图[图 6(c)]发现棉田北部螨害扩散较其他地区略微严重。

随着 7 月新疆地区温度逐渐升高, 高温条件对棉叶螨扩散极其有利, 同时药剂对叶螨的防治效果下降, 导致 7 月 3 日螨害逐渐加重, 呈现局部棉田叶螨暴发状态且多集中于东、西部地块边缘区域[图 6(d)]。因此, 农户于 7 月 8 日再次对棉叶螨进行施药防治, 但随着植保人员和农户在田中不断移动进行实地调查等操作, 导致叶螨随之迁移扩散, 通过图 6(e)可以看出, 中部以北区域出现局部田块爆发成灾现象, 防治难度增加, 易由逐渐扩散转变成迅速扩散。7 月 13 日再次对棉田进行监测, 发现螨害中部以北区域已达到大面积爆发状态, 而中部以南区域螨害防治程度较好, 无局部爆发状态[图 6(f)]。

监测图整体趋势与实际调查情况较为相符, 通过基于环境数据与植被指数相结合建立的 SVM 棉田螨害监测模型可以反映田间螨害变化, 基于无人机的低空遥感可以辅助用来实现棉田大范围的螨害发生监测。

3.6 棉田螨害面积预测模型的建立

首先基于模型分类结果计算螨害面积占比值, 考虑到实际应用环节, 环境数据更易获取, 故只将初选的 13 种田间环境数据与棉田螨害面积值作相关性分析, 选取 $\text{sig} \leq 0.05$ 的田间环境数据, 结果如表 5 所示。

表 5 环境数据与棉田螨害面积之间相关性

Table 5 Correlation coefficient between environmental data and cotton field spider mite area

环境数据	最高温度	平均湿度	温湿系数	积温	10 cm 土壤最高温度	10 cm 土壤平均温度
相关系数	0.707*	0.844*	0.931**	0.837*	0.856*	0.974**

注: ** 和 * 分别表示 0.01, 0.05 的显著水平

Note: ** and * represent significant at the 0.01 and 0.05 levels of probability respectively

分别使用 SVM 回归和多元逐步回归方法构建螨害面积模型, 以决定系数 R^2 作为指标来衡量模型的性能。在 Matlab 下调用 LIBSVM 工具包建立 SVM 螨害面积预测模型 M_1 , 并通过交叉验证方式寻优参数, 得到最终模型 $R^2 = 0.796$ 。采用 SPSS 进行多元逐步回归分析, 从表 5 中逐步选出与螨害面积值关系最密切的田间环境数据作为建立预测模型 M_2 的特征值, 最终得到多元逐步回归方程为

$$y = 110.826 - 0.193x_1 - 0.089x_2 - 0.5x_3 + 0.423x_4 - 4.022x_5 \quad (1)$$

式(1)中, x_1 是最高温度, x_2 是平均湿度, x_3 是温湿系数, x_4 是 10 cm 土壤最高温度, x_5 是 10 cm 土壤平均温度, y 是预测的螨害面积占比值。模型评价结果见表 6。

表 6 模型评价结果

Table 6 Model evaluation results

模型	R	R^2	调整后 R^2	标准估算的误差
M_2	0.874	0.848	0.835	0.9149

由表 6 可知, 多元逐步回归预测模型的决定系数 $R^2 = 0.848$, 说明该模型拟合度优于 SVM 回归预测模型, 并对多元逐步回归预测模型进行显著性检验, 计算出 P 值为

0.001, 小于 0.05。

本研究最终使用多元逐步回归模型对测试样本进行预测, 将螨害面积占比值单位转换成亩数, 结果如表 7 所示。

结果表明模型对棉田螨害发生面积的预测效果较好。

表 7 棉田螨害面积预测模型预测结果对比

Table 7 Comparison of prediction results of cotton field spider mite area prediction models

日期	实际值/亩	预测值/亩
6.22	64.168	58.374
6.27	18.361	18.591
6.29	20.348	18.101
7.3	50.055	41.666
7.9	43.087	42.39
7.13	72.831	69.83

4 结 论

对棉田大范围螨害监测和预测方法进行研究, 通过地面尺度的棉花冠层高光谱数据结合低空尺度下的无人机多光谱遥感影像数据进行敏感波段筛选, 计算螨害相关植被指数,

并结合试验区的气象和土壤等田间环境数据,使用单一环境数据、单一植被指数以及环境数据与植被相结合的 3 种方式建立 SVM 蚜害发生监测模型并进行对比分析。结果表明,同时使用环境数据和植被指数建立的模型具有更高的准确度。基于已建立的 SVM 棉田蚜害监测模型,绘制不同时期的蚜害监测空间分布图,并分别统计健康点和蚜害点像元数计算蚜害面积占总面积比值,结合田间环境数据建立棉田蚜害面积预测模型,模型的决定系数 $R^2 = 0.848$,可以用来对棉田蚜害面积发生进行预测。

研究所建立的模型基于田间环境数据和植被指数,这两

个因素本身就具有时间和空间的变异性,对一段时间内的棉田进行监测和面积预测,验证了时间尺度的可移植性,对于区域上的验证,需要采集不同区域的田间环境数据和植被指数,理论上与时间上的验证同理。目前只针对棉田是否发生蚜害建立了监测模型,对不同蚜害等级的发生监测未进行研究,且模型虽然考虑了田间环境数据和施药防治对蚜害的影响,但是越冬基数、天敌等因素同样会影响蚜害的发生。本研究结果可为大范围棉田蚜害监测和面积预测提供方法参考,后续研究中应融合更多相关因素,提升蚜害监测与预测的精度。

References

- [1] YAN Jian-feng(闫建峰). Village Science and Technology(乡村科技), 2020, 11(22): 47.
- [2] YANG Li-li, ZHANG Da-wei, LUO Jun, et al(杨丽丽, 张大卫, 罗 君, 等). Transactions of the Chinese Society for Agricultural Machinery(农业机械学报), 2019, 50(2): 14.
- [3] LAN Yu-bin, ZHU Zi-hao, DENG Xiao-ling, et al(兰玉彬, 朱梓豪, 邓小玲, 等). Transactions of the Chinese Society of Agricultural Engineering(农业工程学报), 2019, 35(3): 92.
- [4] HUANG Huan-hua, MA Xiao-hang, HUANG Hua-yi, et al(黄焕华, 马晓航, 黄华毅, 等). Journal of Environmental Entomology(环境昆虫学报), 2018, 40(2): 306.
- [5] TIAN Ye, ZHANG Qing, LI Xi-can, et al(田野, 张 清, 李希灿, 等). Arid Zone Research(干旱区研究), 2017, 34(2): 423.
- [6] WANG Shou-hui, DAI Jian-guo, LI Shuan-ming, et al(王守会, 戴建国, 李栓明, 等). Jiangsu Agricultural Sciences(江苏农业科学), 2018, 46(13): 96.
- [7] CUI Mei-na, DAI Jian-guo, WANG Shou-hui, et al(崔美娜, 戴建国, 王守会, 等). Xinjiang Agricultural Sciences(新疆农业科学), 2018, 55(8): 1457.
- [8] ZHAO Liang, CHEN Bing, XIAO Chun-hua, et al(赵 亮, 陈 兵, 肖春华, 等). Journal of Gansu Agricultural University(甘肃农业大学学报), 2015, 50(5): 94.
- [9] LIU Hong-xia(刘红霞). China Cotton(中国棉花), 2016, 43(1): 45.
- [10] MA Hui-qin, HUANG Wen-jiang, JING Yuan-shu, et al(马慧琴, 黄文江, 景元书, 等). Transactions of the Chinese Society of Agricultural Engineering(农业工程学报), 2017, 33(5): 162.
- [11] QU Sha, LI Zhen-hai, QIU Chun-xia, et al(屈 莎, 李振海, 邱春霞, 等). Transactions of the Chinese Society of Agricultural Engineering(农业工程学报), 2017, 33(12): 186.
- [12] Kanke Y, Tubana B, Dalen M, et al. Precision Agriculture, 2016, 17(5): 507.
- [13] Cao Qiang, Miao Yuxin, Shen Jianning, et al. Precision Agriculture, 2016, 17(2): 136.

Research on Large-Scale Monitoring of Spider Mite Infestation in Xinjiang Cotton Field Based on Multi-Source Data

YANG Li-li, WANG Zhen-peng, WU Cai-cong*

College of Information and Electrical Engineering, China Agricultural University, Beijing 100083, China

Abstract Xinjiang's traditional cotton field spider mite monitoring method is time-consuming and inefficient. The paper proposes combining ground hyperspectral, UAV multi-spectrum, environmental data and field survey for dynamic monitoring of large-scale spider mite damage. Firstly, cotton canopy hyperspectral data and low-altitude-scale UAV multi-spectral data in different cotton periods are collected separately by analyzing the original hyperspectral spectrum and first-order differential spectral characteristics, four sensitive bands of spider mite damage are extracted as below: green light band near 553 nm, red light band near 680 nm, red side band of 680~750 nm, and near infrared band of 760~1350 nm, which are also included in the multi-spectrum carried by UAV. Secondly, the correlation analysis among 23 vegetation indices, 13 field environmental data, and the occurrence of spider mites surveyed on the ground is done. SAVI, OSAVI, TVI, NDGI, average humidity, temperature-humidity coefficient and average soil temperature of 10 cm are all significantly correlated with spider mite occurrence ($\text{sig} \leq 0.01$); RDVI, RVI, MSR, maximum temperature, average temperature, accumulated temperature, the highest temperature of 10 cm soil and the average humidity of 10 cm soil all reach a significant correlation level with the occurrence of spider mite damage ($\text{sig} \leq 0.05$). 15 characteristic values with sig values below 0.05 were selected; cotton field mite monitoring models based on single environmental data, single vegetation indices, and a combination of environmental data and vegetation indices are established respectively using support vector machine (SVM). Finally, from the optimum model, we can draw the spatial distribution map of spider mite damage in different periods and calculate the proportion of spider mite damage are based on the number of spider mite damage and healthy pixels in the statistical distribution map. Then the field environmental data is analyzed for correlation, the environmental factors most closely related to the spider mite area value are determined by multiple stepwise regression analysis, and the cotton field spider mite area prediction model is established. The results show that the accuracy rate of the cotton field spider mite monitoring model based on a single environmental data is 62.22%, while the accuracy rate of the cotton field mite monitoring model based on a single vegetation index is 75.56%. Moreover, the most effective model is based on the combination of environmental data and vegetation indices with an accuracy rate of 80%. The coefficient of determination of the spider mite area prediction model is $R^2 = 0.848$. In this study, based on multi-source data, the cotton field spider mite occurrence monitoring model and spider mite area prediction model can provide a reference for the large-scale monitoring and trend warning of cotton field mite damage in Xinjiang.

Keywords Unmanned aerial vehicle remote sensing; Cotton spider mite; Environmental data; Monitoring; Prediction

(Received Nov. 3, 2020; accepted Feb. 19, 2021)

* Corresponding author