

# 基于 PLS 算法的棉花黄萎病高空间分辨率遥感监测

竞霞<sup>1</sup>, 黄文江<sup>2\*</sup>, 琚存勇<sup>3</sup>, 徐新刚<sup>2</sup>

(1. 西安科技大学测绘科学与技术学院, 西安 710054; 2. 国家农业信息化工程技术研究中心, 北京 100097;  
3. 东北林业大学林学院, 哈尔滨 150040)

**摘要:** 棉花黄萎病危害程度大, 发生范围广, 已成为中国乃至世界上棉花主要病害之一。论文基于野外定位调查数据及高空间分辨率遥感影像, 利用变量投影重要性 (VIP) 准则筛选最优变量, 用偏最小二乘回归 (PLS) 方法建立棉花黄萎病病情严重度的定量估测模型, 并利用已建立的估测模型和高分辨率 IKONOS 影像获取了不同病情严重度的空间分布图。研究表明: 在所分析的 13 个遥感因子中, 增强植被指数 (EVI)、再归一化植被指数 (RDVI)、全球环境监测指数 (GEMI)、差值植被指数 (DVI)、修改型土壤调整植被指数 (MSAVI)、归一化植被指数 (NDVI) 为棉花黄萎病病情严重度遥感估测的敏感因子, 能够有效估测棉花黄萎病病情严重度, 其模型预测值与实测值间的  $R^2$ 、RMSE 和 RE 分别为 0.78、0.45、9.2%。论文利用 PLS 算法和高分辨率卫星影像实现了棉花黄萎病病情严重的遥感监测, 研究结果对实现大范围农作物病虫害的遥感监测具有重要的参考价值。

**关键词:** 遥感监测, 偏最小二乘法, 高空间分辨率, 棉花黄萎病, 病情严重度

doi: 10.3969/j.issn.1002-6819.2010.08.039

中图分类号: S127

文献标识码: A

文章编号: 1002-6819(2010)-08-0229-07

竞霞, 黄文江, 琚存勇, 等. 基于 PLS 算法的棉花黄萎病高空间分辨率遥感监测 [J]. 农业工程学报, 2010, 26(8): 229-235.  
Jing Xia, Huang Wenjiang, Ju Cunyong, et al. Remote sensing monitoring severity level of cotton verticillium wilt based on partial least squares regressive analysis[J]. Transactions of the CSAE, 2010, 26(8): 229-235. (in Chinese with English abstract)

## 0 引言

棉花黄萎病危害程度大, 发生范围广, 已成为中国乃至世界范围棉花主要病害之一, 严重影响棉花生产效益的提高, 尤其是落叶型黄萎病对棉花产量和品质所造成的损失更具毁灭性, 少则减产 10%~20%, 严重时高达 50%~70%。如何及早发现黄萎病害并控制其蔓延已成为当前研究的难点和重点。棉花黄萎病的传统监测主要由植保专家等通过田间调查取样的方法判断其病害严重度, 费时费力, 时效性差, 难以在大范围内展开, 而且其严重度评价受观测者的主观因素影响较大。棉花受黄萎病菌感染后, 叶肉细胞被破坏、水分含量下降、叶片变黄干枯、叶绿素降低、叶面积指数下降, 这些理化参数的变化会影响棉花反射辐射或发射辐射的总量, 从而为遥感技术快速、大面积监测棉花黄萎病的发生、发展提供了可能。

对棉花胁迫的遥感监测目前主要集中于利用近地高光谱数据监测棉花病虫害的发生情况<sup>[1-5]</sup>, 近地高光谱数据可以为不同病虫害的光谱特征识别提供理论依据, 是棉花病虫害监测的基础性工作。但近地高光谱数据具有

空间上的随机性, 难以反映病虫害发生的空间特征, 而航天航空遥感影像具有空间上的连续性, 结合 GIS 可以较容易的实现灾害发生范围的监测和预测。

多变量回归分析技术是农作物病虫害定量遥感监测的主要方法之一, 国内外学者利用普通最小二乘算法的回归分析技术在农作物病虫害遥感监测方面做了许多研究<sup>[6-8]</sup>并取得了一定的成果。利用普通最小二乘算法进行多元回归分析时, 若自变量间存在严重多重相关性则会使普通最小二乘算法失效, 使所建模型的准确性、可靠性难以得到保证, 而 PLS 算法则可以消除各变量间多重共线性对估测模型的影响<sup>[9-10]</sup>。基于此本文在遥感因子与黄萎病病情严重度相关分析基础上, 采用 PLS 算法和高分辨率 IKONOS 影像建立棉花黄萎病病情严重度的多变量估测模型。

## 1 研究区概况

石河子垦区地处天山北麓中段, 古尔班通古特大沙漠南缘 (东经 85°59'12"~86°08'13", 北纬 44°15'13"~44°19'13"), 面积 7 529 km<sup>2</sup>, 是新疆自治区和建设兵团棉花生产基地。石河子垦区属典型温带大陆性气候, 冬季长而严寒, 夏季短而炎热, 年平均气温 7.5~8.2℃, 日照 2 318~2 732 h。

## 2 材料与方法

### 2.1 病情指数调查

根据研究区域内遥感影像的覆盖范围和棉花黄萎病的发病情况, 准同步在研究区域内相对均匀的调查了 70

收稿日期: 2009-12-04 修订日期: 2010-04-09

资助项目: 农业部行业科技项目 (200903010); 西安科技大学培育基金项目 (2009013)

作者简介: 竞霞 (1978-), 女, 博士, 研究方向为农业遥感。西安 西安科技大学测绘科学与技术学院, 710054

\*通信作者: 黄文江 (1974-), 男, 博士, 研究员, 研究方向为农业遥感。北京 国家农业信息化工程技术研究中心, 100097。

Email: huangwj@nrcita.org.cn

个点并利用 GPS 记录下相应点的位置, 每点选取 10 株棉花, 分别调查其发病情况, 计算测试群体的病情指数 (disease index,  $DI$ )<sup>[1]</sup>

$$DI = \frac{\sum(X \cdot f) \times 100}{n \cdot \sum f} \quad (1)$$

式中:  $X$ ——各级病级值;  $n$ ——最高病级值;  $f$ ——各级株数。

根据测试群体  $DI$  的不同将黄萎病冠层病情严重程度分为 5 级, 即: 正常 ( $b_0$ ):  $DI=0$ ; 轻度 ( $b_1$ ):  $DI$  在 0~25% 之间; 中度 ( $b_2$ ):  $DI$  在 25%~50% 之间; 严重 ( $b_3$ ):  $DI$  在 50%~75% 之间; 极严重 ( $b_4$ ):  $DI$  在 75%~100% 之间。

## 2.2 IKONOS 影像获取与预处理

选用 2008 年 7 月 25 日新疆石河子 ( $86.02875^\circ \sim 86.16871^\circ$ ,  $44.51535^\circ \sim 44.66714^\circ$ ) 的 IKONOS 影像监测棉花黄萎病病情严重程度, 并对该景影像进行了大气校正、几何校正和图像融合的预处理工作。

大气校正: 采用大气辐射传输模型中的 6S (the Second Simulation of the Satellite Signal in the Solar Spectrum) 模型<sup>[11]</sup>进行 IKONOS 影像的大气校正。6S 模型首先计算大气校正参数  $x_a$ ,  $x_b$  和  $x_c$  (表 1), 然后利用式 (2) 计算校正后的反射率。

$$\begin{aligned} \rho &= y / (1 + x_c \cdot y) \\ y &= x_a \cdot L_i - x_b \end{aligned} \quad (2)$$

式中:  $\rho$ ——校正后的二向性反射率因子;  $L_i$ ——第  $i$  波段辐射亮度。

表 1 IKONOS 影像 6S 模型大气校正参数

Table 1 6S model atmospheric correction parameters of IKONOS image

Band	Blue	Green	Red	NIR
$x_a$	0.00296	0.00304	0.0035	0.00459
$x_b$	0.10554	0.06403	0.03393	0.01858
$x_c$	0.16384	0.12075	0.08369	0.05752

几何校正: 在 IKONOS 影像研究区域内利用高精度 GPS 均匀布设 27 个控制点, 为提高定位的准确性将控制点选取在靠近主要交通干道的田间小路交叉处, 经过仔细核对确定其中 24 个点作为地面有效控制点, 用来进行 IKONOS 图像的几何精校正与精度检验, 确保几何校正误差控制点在 0.5 个像元内。

图像融合: 选用基于平滑滤波的强度调整法 (smoothing filter-based intensity modulation, SFIM)<sup>[12]</sup>对 IKONOS 影像多光谱和全色波段进行融合以提高其多

光谱影像的空间分辨率 (图 1), 其计算公式为

$$IMAGE_{SFIM} = \frac{IMAGE_{low} \times IMAGE_{high}}{IMAGE_{mean}} \quad (3)$$

式中:  $IMAGE_{SFIM}$ ——融合后的图像;  $IMAGE_{low}$ ——低分辨率的图像对高分辨率的图像重采样后的图像;  $IMAGE_{mean}$ ——对高分辨率的图像通过邻域平滑的卷积运算后的图像;  $IMAGE_{high}$ ——高分辨率的图像。对高分辨率图像进行邻域平滑卷积运算时模板的大小为  $2a+1$ , 其中  $a$ ——低分辨率图像的分辨率/高分辨率图像的分辨率。



a. 原始多光谱 IKONOS 影像 (4 m)



b. SFIM 变换后的图像 (1 m)

注: b 显示的是 a 方框中的部分

图 1 IKONOS 多光谱与全色波段融合结果

Fig.1 IKONOS panchromatic and multi-spectral image fusion

## 2.3 变量设置

棉花受黄萎病菌侵染后会在外部形态及内部生理上发生一系列的变化, 可见光和近红外波段以及它们的不同组合方式对这些变化的监测具有重要指示意义。根据国内外相关研究结果, 本文选用 9 种常用植被指数 (表 2) 作为偏最小二乘回归分析的入选变量。

表 2 部分常用植被指数  
Table 2 Some vegetation indices

植被指数名称	缩写	公式	参考文献
比值植被指数(ratio vegetation index)	$RVI$	$NIR/R$	Person et al. <sup>[13]</sup>
归一化植被指数(normalized difference vegetation index)	$NDVI$	$(NIR - R) / (NIR + R)$	Rouse et al. <sup>[14]</sup>
差值植被指数(difference vegetation index)	$DVI$	$NIR - R$	Jordan et al. <sup>[15]</sup>
再归一化植被指数(renormalized difference vegetation index)	$RDVI$	$\sqrt{NDVI * DVI}$	Roujean et al. <sup>[16]</sup>
变换植被指数(transformed vegetation index)	$TVI$	$0.5 * [120(NIR - G) - 200(R - G)]$	Broge et al. <sup>[17]</sup>

归一化差异绿度指数(normalized difference greenness index)	NDGI	$(G-R)/(G+R)$	Chamard et al. [18]
转下页			
接上页			
植被指数名称	缩写	公 式	参考文献
修改型土壤调整植被指数(modified soil adjusted vegetation index)	MSAVI	$\frac{1}{2}[(2NIR+1)-\sqrt{(2NIR+1)^2-8(NIR-R)}]$	Qi et al. [19]
全球环境监测指数(global environment monitoring index)	GEMI	$\frac{\eta(1-0.25\eta)-(R-0.125)/(1-R)}{\eta=[2(NIR^2-R^2)+1.5NIR+0.5R]/(NIR+R+0.5)}$	Pinty et al. [20]
增强植被指数(enhanced vegetation index)	EVI	$\frac{NIR-R}{NIR+C_1R-C_2B+L}(1+L)$	Liu et al. [21]

### 2.4 变量投影重要性 (VIP)

自变量  $x_j$  在解释因变量  $Y$  的重要性时, 可以用变量投影重要性指标  $VIP_j$  (variable importance in projection) 来测度, 其计算公式为

$$VIP_j = \sqrt{\frac{p}{Rd(Y;t_1,t_2,t_3,\dots,t_m)} \sum_{h=1}^m Rd(Y;t_h)w_{hj}^2} \quad (4)$$

式中:  $VIP_j$ ——第  $j$  个自变量的投影重要性指标;  $m$ ——从原变量中提取的成分个数;  $p$ ——自变量个数;  $t_h$ ——第  $h$  个成分;  $Rd(Y, t_h)$ ——成分  $t_h$  对因变量  $Y$  的解释能力, 为二者相关系数的平方;  $Rd(Y;t_1, t_2, t_3, \dots, t_m)$ ——成分  $t_1, t_2, t_3, \dots, t_m$  对因变量  $y$  的累计解释能力;  $w_{hj}$ ——轴  $w_h$  的第  $j$  个分量, 用于测量  $x_j$  对构造成分  $t_h$  的边际贡献, 并且对于任意  $h=1, 2, 3, \dots, m$  均有

$$\sum_{j=1}^p w_{hj}^2 = w_h^T w_h = 1 \quad (5)$$

对于  $p$  个自变量  $x_j (j=1, 2, \dots, p)$ , 如果它们在解释  $Y$  时的作用都相同, 则所有的  $VIP_j$  均等于 1; 否则, 对于  $VIP_j$  很大 (大于 1) 的  $x_j$ , 它在解释  $Y$  时就有更加重要的作用 [10]。

### 2.5 偏最小二乘回归 (PLS) 方法

偏最小二乘回归 (partial least squares regression, PLS) 是一种先进的多元分析方法, 于 1983 年由伍德 (S. Wold) 和阿巴诺 (C. Albano) 等人首次提出。它集多元线性回归分析、主成分分析和典型相关分析的基本功能为一体, 可以较好地解决以往用普通多元线性回归分析方法无法解决的问题。特别是当自变量集合内部存在较程度的相关性时, 其结论比普通多元回归更加可靠。PLS 通过提取概括原数据信息的综合变量 (新成分), 建立新变量与因变量的回归关系, 最后再表达成原变量的回归方程 [9-10]。

设  $E_0 (n \times p)$  为标准化的自变量数据矩阵,  $F_0 (n \times 1)$  为对应的因变量向量, 则成分  $t_i$  的计算公式为

$$t_i = E_{i-1} W \quad (6)$$

式中:  $W_i$ ——矩阵  $E_{i-1}^T F_0 F_0^T E_{i-1}$  最大特征值所对应的特征

向量, 计算式为  $W_i = \frac{E_{i-1}^T F_0}{\|E_{i-1}^T F_0\|}$ ,  $T$ ——转置矩阵;  $E_{i-1}$ ——

自变量阵  $E_{i,2}$  对成分  $t_{i-1}$  回归得到的残差阵, 表达式为

$$E_i = E_{i-1} - t_i P_i^T, \quad \text{而 } P_i = \frac{E_{i-1}^T t_i}{\|t_i\|^2}。$$

具体选取几个成分, 可用交叉验证法 (cross validation) [22] 确定, 当增加新的成分对减少方程的预测误差没有明显的改善作用时, 就停止提取新的成分。假如共有  $k$  个成分入选, 建立  $F_0$  与  $t_i$  的回归方程, 得

$$\hat{F}_0 = r_1 t_1 + r_2 t_2 + \dots + r_k t_k \quad (7)$$

由于  $t_i$  均是  $E_0$  的线性组合, 所以  $\hat{F}_0$  可以写成  $E_0$  的线性表达形式

$$\hat{F}_0 = E_0 \beta \quad (8)$$

其中:  $\beta = \sum_{i=1}^k r_i W_i^*$ , 而  $r_i = \frac{F_0^T t_i}{\|t_i\|^2}$ ;  $W_i^* = \prod_{j=1}^{i-1} (I - W_j P_j^T) W$ ;  $I$ ——单位阵。

最后根据标准化的逆运算变换成因变量  $Y$  对原始自变量  $X$  的回归方程。

### 2.6 模型精度评价

用未参与建立模型的样本评价模型的精度, 评价指标采用决定系数 (determination coefficient,  $R^2$ )、均方根误差 (root mean square error,  $RMSE$ ) [23] 和相对误差 (relative error,  $RE$ ) [24]。  $R^2$  越接近于 1, 而  $RMSE$  和  $RE$  越接近于 0, 说明模型预报值与实测值越吻合, 模型的精度越高。3 个评价指标的计算公式分别为

$$R^2 = \frac{(\sum (y_i - \bar{y}_i)(\hat{y}_i - \bar{y}_i))^2}{\sum (y_i - \bar{y}_i)^2 \sum (\hat{y}_i - \bar{y}_i)^2} \quad (9)$$

$$RMSE = \sqrt{\frac{1}{n} \sum (y_i - \hat{y}_i)^2} \quad (10)$$

$$RE = \frac{|\sum y_i - \sum \hat{y}_i|}{\sum y_i} \times 100\% \quad (11)$$

式中:  $n$ ——检验样本个数;  $y_i$ ——黄萎病病情严重度实测值;  $\bar{y}_i$ ——病情严重度实测值的平均值;  $\hat{y}_i$ ——模型预测值;  $\bar{y}_i$ ——模型预测值的平均值。

## 3 结果与分析

### 3.1 遥感因子与病情严重度的相关分析

由于研究对象、研究目的及应用环境的差异, 不同

的植被指数具有不同的优势和局限性, 本文对部分常用植被指数(表2)、IKONOS 影像各波段(B1: 蓝光, B2: 绿光, B3: 红光, B4: 近红外)光谱反射率与黄萎病病情严重度的相关分析结果如表3所示。

由表3可知, 可见光波段光谱反射率与棉花黄萎病病情严重正相关, 随黄萎病病情严重度增加, 可见光波段光谱反射率增大。棉花受黄萎病病菌侵染后, 病菌不断产生分生孢子破坏叶片叶绿素。在可见光波段内植物的光谱特性主要受叶片内各种色素含量的影响, 特别是叶绿素含量对可见光波段植物的光谱特性起着重要的作用。棉花受黄萎病病菌危害后正常生长发育被阻止, 叶绿素含量降低, 叶绿素在蓝红波段的吸收减少、反射增强, 特别是红光波段反射率升高, 叶片变为黄色(红色+绿色=黄色), 这与黄萎病发生后棉花叶片褪绿、发黄症状吻合。随棉花黄萎病病情严重度增加, 近红外波段光谱反射率减小, 且棉花黄萎病病情严重度与近红外波段冠层光谱反射率值的相关性远大于可见光波段。这与陈兵等利用近地高光谱数据监测棉花黄萎病不同病情严重度下冠层光谱特征的变化规律相一致<sup>[1]</sup>。

大丽轮支菌侵入棉花根表层后, 病原菌在木质部产

生分生孢子并刺激邻近薄壁细胞产生胶状物质等堵塞导管, 使水分和养分运输发生困难, 叶肉细胞被破坏, 水分含量下降, 棉株萎蔫死亡, 近红外波段光谱反射率下降。由植被指数的计算公式以及棉花受黄萎病危害后可见光波段光谱反射率增加而近红外波段光谱反射率下降可知: 随黄萎病病情严重度增加, 植被指数值下降, 病情严重度与植被指数均达到了极显著负相关(表3)。由于植被指数能够有效地综合各有关的光谱信号, 在增强植被信息的同时使非植被信号最小化, 通常情况下它们比用单波段来探测绿色植被更具有灵敏性, 因此本研究中大部分植被指数与棉花黄萎病病情严重的相关性优于单波段。

单变量对信息的压缩易导致单变量模型的估测精度低于多元回归模型<sup>[25]</sup>, 而变量间存在的严重多重相关性, 又会使普通最小二乘算法失效, 破坏参数估计, 扩大模型误差, 丧失模型稳健性, 使所建模型的准确性、可靠性难以得到保证<sup>[10]</sup>, 因此本文选择 PLS 算法研究建立棉花黄萎病病情严重度估测模型。

表3 遥感因子与病情严重度相关系数 ( $n=46$ )

Table 3 Correlation between remote sensing factor and severity levels

相关系数	B1	B2	B3	B4	TVI	RVI	RDVI	NDVI	NDGI	MSAVI	GEMI	EVI	DVI
DI	0.351	0.441	0.547	-0.656	-0.700	-0.652	-0.698	-0.671	-0.667	-0.695	-0.692	-0.708	-0.695
B1	1.000	0.965	0.937	-0.296	-0.534	-0.767	-0.635	-0.780	-0.549	-0.518	-0.714	-0.636	-0.518
B2		1.000	0.972	-0.380	-0.611	-0.836	-0.711	-0.845	-0.582	-0.597	-0.784	-0.722	-0.597
B3			1.000	-0.524	-0.738	-0.908	-0.820	-0.927	-0.752	-0.722	-0.881	-0.831	-0.722
B4				1.000	0.961	0.798	0.917	0.803	0.776	0.968	0.855	0.905	0.968
TVI					1.000	0.924	0.991	0.935	0.868	0.999	0.962	0.985	0.999
RVI						1.000	0.961	0.983	0.851	0.918	0.966	0.969	0.918
RDVI							1.000	0.974	0.873	0.988	0.989	0.998	0.988
NDVI								1.000	0.865	0.927	0.993	0.977	0.927
NDGI									1.000	0.853	0.872	0.884	0.853
MSAVI										1.000	0.956	0.981	1.000
GEMI											1.000	0.988	0.956
EVI												1.000	0.981
DVI													1.000

### 3.2 PLS 估测模型

利用 PLS 算法建模与其它建模方法一样, 成分个数的增加会降低误差, 提高模型的预测精度, 但成分过多时又会发生过拟合现象, 使预测误差增加, 因此成分个数的确定是 PLS 算法的关键之一<sup>[26]</sup>。

#### 3.2.1 PLS 成分确定

确定 PLS 成分数时, 既要保证所提取的成分对系统解释能力最强, 又要克服变量之间的多重共线性问题。本文采用目前广泛使用的交叉验证法确定 PLS 成分数, 计算成分  $t_h$  对因变量  $y$  的交叉有效性(表4), 根据增加一个新的成分后能否对模型的预测功能有明显改进确定成分个数。

表4 成分  $t_h$  对因变量  $y$  的交叉有效性

Table 4 Cross validation between competent and dependent variable

成分个数	$Q_h^2$	$Q_h^2(\text{cum})$
1	0.7299	0.7299
2	-0.1249	0.7029

注:  $Q_h^2$ ——成分  $t_h$  对因变量  $y$  的交叉有效性;  $Q_h^2(\text{cum})$ ——使用前  $k$  个成分建模的累计交叉有效性。

由表4可知:  $Q_2^2 = -0.1249 < 0.0975$  (95%的置信区间), 因此认为  $t_2$  成分的边际贡献不显著, 引进新的成分  $t_2$  对减少方程的预测误差无改善作用, 本文采用一个成分进行 PLS 建模。

#### 3.2.2 VIP 准则筛选变量

利用 VIP 准则进行变量筛选时, 若自变量的 VIP 值大于 1.0, 说明自变量在解释因变量时具有比较重要的作

用；若自变量的 *VIP* 值在 0.5~1.0，说明自变量对因变量解释作用的重要性还不很明确，需增加样本或根据别的条件进行判断；若自变量 *VIP* 值小于 0.5，则自变量对因变量的解释基本没有意义<sup>[27]</sup>。根据式 (4)，计算了所有变量的 *VIP* 值 (图 2)，本文选取 *VIP* 值大于 1 的变量建立棉花黄萎病病情严重度估测的 PLS 模型。

### 3.2.3 研究区病情严重度估测

以 *VIP* 值大于 1 的 6 个变量即 *EVI*、*RDVI*、*GEMI*、*DVI*、*MSAVI*、*NDVI* 为自变量，以棉花黄萎病病情严重

度为因变量 *y* 进行 PLS 迭代运算得到的棉花黄萎病病情严重度 PLS 估测模型为

$$y = -1.6801RDVI - 1.5828NDVI - 0.4062MSAVI - 1.7838GEMI - 1.5315EVI - 1.6248DVI + 8.159 \quad (12)$$

以未参与建模的 24 个野外调查点进行模型精度检验，*R*<sup>2</sup>、*RMSE* 和 *RE3* 个指标分别为 0.78、0.45、9.2%，说明利用 PLS 算法建立的多元回归分析模型能较好的估测棉花黄萎病病情严重度。

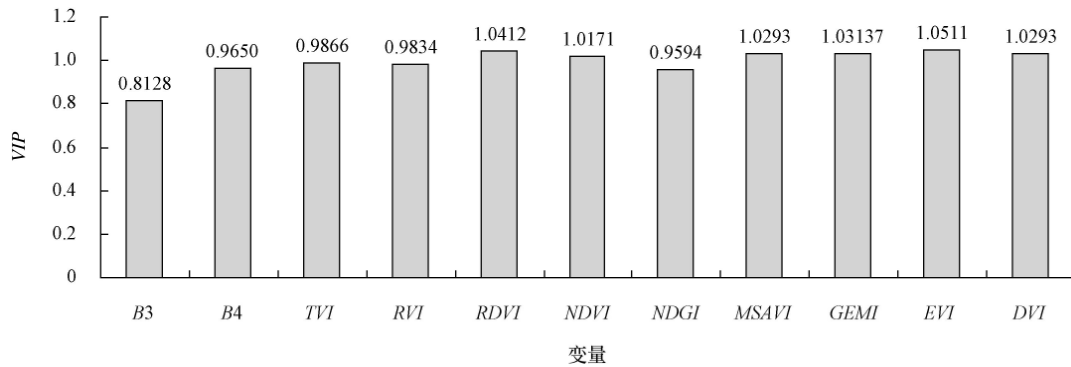


图 2 变量投影重要性

Fig.2 Variable importance in projection

在利用遥感影像监测棉花正常生长区域及受胁迫区域 (包括黄萎病发病区域及其它胁迫区域) 的基础上，进一步利用生境条件和时相信息提取棉花黄萎病发病区域<sup>[25]</sup>，按像元分别计算 PLS 估测模型中各个变量的值，代入公式 (12) 即可得到棉花黄萎病发病区域内每个像元所对应的病情严重度 (图 3)。研究区域内棉花黄萎病的病情严重度等级主要是 1 级和 2 级，说明截止到 2008 年 7 月 25 日研究区域内棉花黄萎病发病较轻，这与野外调查研究区域内棉花黄萎病实际发病情况是相符的。

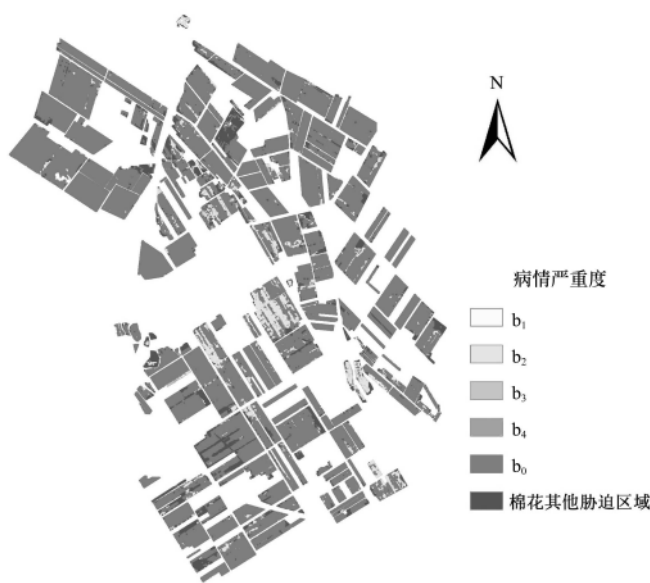
Fig.3 Severity spatial distribution maps of cotton verticillium wilt

## 4 结 论

棉花黄萎病病情严重度与相应位置的遥感影像数据存在较强的相关性，这为遥感估测棉花黄萎病发病情况提供了可能。然而遥感数据构造的变量之间存在的多重相关性，会导致普通最小二乘算法难以建立有效的目标参数估测模型。本文采用的 PLS 算法能够降低甚至消除变量共线性对遥感估测模型的不利影响，建立了精度可靠的棉花黄萎病病情严重度遥感估测模型。另外变量投影重要性准则作为量化自变量对目标参数贡献率的一种方法，可以优化估测模型，使其符合参数节俭的原则，提高模型的运算效率和可移置性。研究结果为航空航天遥感影像监测农作物病虫害提供了一种方法参考和技术支持，对实现农作物病虫害的大面积遥感监测具有重要的借鉴意义。

利用 *VIP* 准则和 PLS 算法建立了黄萎病病情严重度估测模型，取得了较高的监测精度。但本文的估测模型是针对棉花黄萎病，以高空间分辨率 IKONOS 影像为例建立的，因此该模型是否适用于其它分辨率的遥感影像，能否外推到其它病虫害的遥感监测还有待进一步验证。

高空间分辨率影像和高光谱分辨率影像在农作物病虫害遥感监测中各有其优势和局限性，究竟是高空间分辨率影像还是高光谱分辨率影像更适合于本研究区域的棉花黄萎病遥感监测还有待进一步研究。限于篇幅，关于棉花黄萎病发病区域的遥感监测本文没有详细阐述。



注：b<sub>1</sub>——轻度；b<sub>2</sub>——中度；b<sub>3</sub>——严重；b<sub>4</sub>——极严重；b<sub>0</sub>——正常

图 3 棉花黄萎病不同病情严重度空间分布图

## [参 考 文 献]

- [1] 陈兵, 王克如, 李少昆, 等. 棉花黄萎病冠层高光谱遥感监测技术研究[J]. 新疆农业科学, 2007, 44(6): 740-745. Chen Bing, Wang Keru, Li Shaokun, et al. Study on monitoring canopy spectrum of cotton verticillium wilt by remote sensing technique[J]. Xinjiang Agriculture Sciences, 2007, 44(6): 740-745. (in Chinese with English abstract)
- [2] 冯先伟, 陈曦, 包安明, 等. 水分胁迫条件下棉花生理变化及其高光谱响应分析[J]. 干旱区地理, 2004, 27(2): 250-255. Feng Xianwei, Chen Xi, Bao Aaming, et al. Analysis on the cotton physiological change and its hyperspectral response under the water stress condition[J]. Arid Land Geography, 2004, 27(2): 250-255. (in Chinese with English abstract)
- [3] 赵德华, 李建龙, 宋子健, 等. 不同施氮水平下棉花群体反射光谱的差异性分析[J]. 作物学报, 2004, 30(11): 1169-1172. Zhao Dehua, Li Jianlong, Song Zijian, et al. Difference of Canopy Spectral Reflectance to nitrogen nutrient in cotton with different nitrogen applications[J]. Acta Agronomica Sinica, 2004, 30(11): 1169-1172. (in Chinese with English abstract)
- [4] 田会东, 谢宝瑜, 赵永超, 等. 棉花枯萎病的光谱识别[J]. 棉花学报, 2008, 20(1): 51-55. Tian Huidong, Xie Baoyu, Zhao Yongchao, et al. Using spectral data to distinguish cotton fusarium wilt disease[J]. Cotton Science, 2008, 20(1): 51-55. (in Chinese with English abstract)
- [5] 陈鹏程, 张建华, 李眉眉, 等. 土耳其斯坦叶螨为害棉叶的生理变化及光谱特征分析[J]. 昆虫知识, 2007, 44(1): 61-65. Chen Pengcheng, Zhang Jianhua, Li Meimei, et al. Physiological change and hyperspectral character analysis of cotton leaves infested by tetranychus turkestani[J]. Chinese Bulletin of Entomology, 2007, 44(1): 61-65. (in Chinese with English abstract)
- [6] Steddom K, Bredehoeft M W, Khan M, et al. Comparison of visual and multispectral radiometric disease evaluations of cercospora leaf spot of sugar beet[J]. Plant Disease, 2005, 89(2): 153-158.
- [7] 蒋金豹, 陈云浩, 黄文江, 等. 冬小麦条锈病严重度高光谱遥感反演模型研究[J]. 南京农业大学学报, 2007, 30(3): 63-67. Jiang Jinbao, Chen Yunhao, Huang Wenjiang, et al. Study on hyperspectral remote sensing retrieval models about winter wheat stripe rust severity[J]. Journal of Nanjing Agricultural University, 2007, 30(3): 63-67. (in Chinese with English abstract)
- [8] 黄木易, 王纪华, 黄文江, 等. 冬小麦条锈病的光谱特征及遥感监测[J]. 农业工程学报, 2003, 19(6): 154-158. Huang Muyi, Wang Jihua, Huang Wenjiang, et al. Hyperspectral character of stripe rust on winter wheat and monitoring by remote sensing[J]. Transactions of the CSAE, 2003, 19(6): 154-158. (in Chinese with English abstract)
- [9] Naesset E, Bollandsas O M, Gobakken T. Comparing regression methods in estimation of biophysical properties of forest stands from two different inventories using laser scanner data[J]. Remote Sensing of Environment, 2005, 94: 541-553.
- [10] 王惠文, 吴载斌, 孟洁. 偏最小二乘回归的线性与非线性方法[M]. 北京: 国防工业出版社, 2006.
- [11] Tanre D, Holben B N, Kaufman Y J. Atmospheric correction against algorithm for NOAA-AVHRR products: theory and application[J]. 1992, 30(2): 231-248.
- [12] Liu J G. Smoothing Filter-based Intensity Modulation: a spectral preserve image fusion technique for improving spatial details[J]. International Journal of Remote Sensing, 2000, 21(18): 3461-3472.
- [13] Pearson R L, Miller, L D. Remote mapping of standing crop biomass for estimation of the productivity of the shortgrass prairie[C]. In: Proceedings of the Eighth International Symposium on Remote Sensing of Environment, 1972: 1357-1381.
- [14] Rouse J W, Haas R H, Schell J A, et al. Monitoring the vernal advancement of retrogradation of natural vegetation[M]. NASA/GSFC Type III Final Rep, 1974.
- [15] Jordan C F. Derivation of leaf area index from quality of light on the forest floor[J]. Ecology, 1969, 50: 663-666.
- [16] Roujean J L, Breon F M. Estimating PAR absorbed by vegetation from bidirectional reflectance measurements[J]. Remote Sensing of Environment, 1995, 51(3): 375-384.
- [17] Broge N H, Leblanc E. Comparing prediction power and stability of broadband and hyperspectral vegetation indices for estimation of green leaf area and canopy chlorophyll density[J]. Remote Sensing of Environment, 2001, 76(2): 156-172.
- [18] Chamard P, Courel M F, Ducouso M, et al. Utilisation des bandes spectrales du vert et du rouge pour une meilleure évaluation des formations végétales actives. Télédétection et cartographie, éd. Aupelf-uref, 1991: 203-209.
- [19] Qi J, Chehbouni A, Huete A, et al. A modified soil adjusted vegetation index[J]. Remote Sensing of Environment, 299, 48: 119-126.
- [20] Pinty B, Verstraete M M. GEMI: a non-linear index to monitor global vegetation from satellites[J]. Plant Ecology, 1992, 101: 15-20.
- [21] Liu H Q, Huete A R. A feedback based modification of the NDVI to minimize canopy background and atmospheric noise. IEEE Transactions on Geoscience and Remote Sensing, 1995, 33, 457-465.
- [22] 许和连, 赖明勇, 钱晓英. 外商直接投资影响因素的偏最小二乘回归建模分析[J]. 中国管理科学, 2002, 10(5): 20-25. Xu Helian, Lai Mingyong, Qian Xiaoying. The partial least-squares regression analysis of impact factor of foreign direct investment in china[J]. Chinese Journal of Management Science, 2002, 10(5): 20-25. (in Chinese with English abstract)
- [23] Scull P, Okin G, Chadwick O A, et al. 2005. A comparison of methods to predict soil surface texture in an alluvial basin[J].

- The Professional Geographer, 57(3): 423-437
- [24] 李崇贵, 赵宪文. 森林郁闭度定量估测遥感比值波段的选择[J]. 林业科学, 2005, 41(4): 72-77.  
Li Chonggui, Zhao Xianwen. Selection of remote sensing ratio wave band for canopy density quantitative estimation[J]. 2005, 41(4): 72-77. (in Chinese with English abstract)
- [25] 竞霞. 基于多源多时相数据的棉花黄萎病遥感监测研究[D]. 北京: 北京师范大学, 2009.  
Jing Xia. Monitoring Cotton Verticillium Wilt Based on Multi-source and Multi-temporal Remote Sensing Data[D]. Beijing: Beijing Normal University, 2009. (in Chinese with English abstract)
- [26] 李民赞. 光谱分析技术及其应用[M]. 北京: 科学出版社, 2006.
- [27] Umetrics AB Inc., Tutorial SIMCA-P Version 11.5, Umetrics AB Company, Umea, Sweden, 2007.

## Remote sensing monitoring severity level of cotton verticillium wilt based on partial least squares regressive analysis

Jing Xia<sup>1</sup>, Huang Wenjiang<sup>2\*</sup>, Ju Cunyong<sup>3</sup>, Xu Xingang<sup>2</sup>,

(1. College of Geomatics, Xi'an University of Science and Technology, Xi'an 710054, China;

2. National Engineering Research Center For Information Technology In Agriculture, Beijing 100097, China;

3. College of Forestry, Northeast Forestry University, Harbin 150040, China)

**Abstract:** The objective of this study was to estimate the severity level of cotton verticillium wilt using high spatial resolution satellite data and partial least-squares regressive analysis. Firstly, remote sensing factors for monitoring cotton verticillium wilt were picked out by the pre-processed IKONOS image. Then the variable factors were selected according to an approach of Variable Importance in Projection, and used to establish the severity estimating model of cotton verticillium wilt using partial least squares (PLS) regression analysis. Finally the model was applied to calculate the severity level of each pixel in the region of cotton verticillium wilt. The results show that these vegetation indices, i.e. enhanced vegetation index (*EVI*), renormalized difference vegetation index (*RDVI*), global environment monitoring index (*GEMI*), difference vegetation index (*DVI*), modified soil adjusted vegetation index (*MSAVI*) and normalized difference vegetation index (*NDVI*), are sensitive factors for monitoring severity level of cotton verticillium wilt. The model based on those variables achieves better accuracy since precision assessment indices such as determination coefficient ( $R^2$ ), root mean square error (*RMSE*) and relative error (*RE*) which are 0.78, 0.45 and 9.2% respectively. Severity level of cotton verticillium wilt can be effectively estimated utilizing high spatial resolution image and partial least squares regression analysis and the result presents an important reference approach for further monitoring crop pests and diseases at large scale using airborne and airspace remote sensing data.

**Key words:** remote sensing monitoring, partial least squares, high spatial resolution, cotton verticillium wilt, severity level