

基于支持向量机回归的冬小麦叶面积指数遥感反演

梁 栋¹, 管青松^{1,2}, 黄文江^{2*}, 黄林生¹, 杨贵军³

(1. 安徽大学计算智能与信号处理教育部重点实验室, 合肥 230039;
2. 中国科学院遥感与数字地球研究所, 数字地球重点实验室, 北京 100094;
3. 北京农业信息技术研究中心, 北京 100097)

摘 要: 利用单一植被指数反演叶面积指数(LAI)时, 存在不同程度的饱和性且每种指数只能包含部分波段的信息, 该文提出利用支持向量机回归的方法进行叶面积指数的反演, 可以用更多的波段信息作为输入参数以提高LAI反演精度。选取冬小麦起身期、拔节期和灌浆期的实测光谱和叶面积指数数据, 用统计回归的方法分别建立NDVI-LAI和RVI-LAI模型, 用支持向量机回归(SVR)方法分别建立以NDVI、RVI以及蓝、绿、红和近红外4个波段数据作为输入参数的回归预测模型, 即NDVI-SVR、RVI-SVR和NRGB-SVR模型。上述5个模型分别利用对应时期的环境星HJ-CCD数据进行验证。结果表明: NDVI和RVI与叶面积指数(LAI)的回归模型预测的结果与实测值的RMSE分别为0.98与0.97; 预测精度分别为59.2%与59.3%。以NDVI和RVI结合实测叶面积指数(LAI)训练并预测的结果与实测值的均方根误差RMSE分别为0.71与0.83 预测精度分别为70.4%与67.1%。以蓝(B)、绿(G)、红(R)以及近红外(NIR)波段作为输入参数回归并预测的RMSE值为0.42, 预测精度为81.7%。通过支持向量机回归预测具有更好的拟合效果, 可以输入更多波段信息, 提高了叶面积遥感反演精度, 对冬小麦的多个生育期均具有较好的适用性。

关键词: 遥感, 光谱分析, 支持向量机, 反演, 叶面积指数, 植被指数

doi: 10.3969/j.issn.1002-6819.2013.07.015

中图分类号: S127

文献标志码: A

文章编号: 1002-6819(2013)-07-0117-07

梁 栋, 管青松, 黄文江, 等. 基于支持向量机回归的冬小麦叶面积指数遥感反演[J]. 农业工程学报, 2013, 29(7): 117-123.

Liang Dong, Guan Qingsong, Huang Wenjiang, et al. Remote sensing inversion of leaf area index based on support vector machine regression in winter wheat[J]. Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE), 2013, 29(7): 117-123. (in Chinese with English abstract)

0 引 言

叶面积指数 LAI (leaf area index) 是指单位土地面积上植物叶片单面总面积占土地面积的倍数。在生态学、大气和生态交互系统、全球气候变化研究中, 叶面积指数是一个重要结构参数^[1-3], 用来反映植物叶面积、冠层结构变化, 对植被的光合作用和生物物理过程起着重要的作用^[4-7]。LAI 在农业、生态、气候变化等领域已得到广泛应用^[8-9]。传统获取 LAI 的方法主要依靠地面测量, 只能获得地面有限点的 LAI 值, 不能满足作物长势和生态监测方面的需求^[10]。遥感具有大面积、全天候、多波段、简单无破

坏的特点, 为大范围反演 LAI 提供了可能^[11]。

叶面积指数的遥感反演方法主要有 2 类: 经验公式法和物理模型法^[12-14]。物理模型方法对植被的生理生化过程有较好的解释, 普适性较好, 但应用时需要输入植被的多种结构参数, 模型复杂, 难以推广使用。经验公式法主要是在植被指数与 LAI 之间建立线性或非线性模型, 获取大范围的 LAI 值。虽然对研究的背景、地区、植被类型具有较大的依赖性, 不具有很好的普适性, 但这种方法原理简单, 计算方便, 易于推广使用。陈雪洋等^[15]基于环境星 CCD 数据, 分析了 4 种植被指数与 LAI 的关系, 选择了 RVI 作为反演冬小麦 LAI 反演的最优模型。张瀛等^[16]通过 PROSAIL 模拟模型, 研究发现引入绿色波段的增强植被指数 (enhanced vegetation index, EVI) 能更好的用于 LAI 的遥感反演。郭琳等^[17]用环境星 CCD 数据, 通过支持向量机的方法建立归一化植被指数 (normalized differential vegetation index, NDVI) 与 LAI 的关系用于估算甘蔗的 LAI 值。宰松梅等^[18]利用支持向量机模型的方法

收稿日期: 2012-10-16 修订日期: 2013-02-27

基金项目: 国家自然科学基金 (61172127, 41201354); 高等学校博士学科点科研基金 (20113401110006)。中国科学院百人计划项目“植被定量遥感参数反演与真实性检验”项目资助。

作者简介: 梁栋 (1963-), 男, 主要研究方向为图像处理、计算信号处理、模式识别等。安徽大学计算智能与信号处理教育部重点实验室, 230601。Email: dliang@ahu.edu.cn

*通信作者: 黄文江 (1972-), 博士, 研究员。北京 中国科学院遥感与数字地球研究所, 100094。Email: yellowstar0618@163.com

法测定甜椒的叶面积,取得了较好的效果。Shaban Shataee 等^[19]使用了不同机器学习的方法研究了森林的一些属性参数,取得了较好结果。Surya S. Durbha 等^[20]使用 PROSAIL 模型模拟植被冠层光谱,利用核主成分分析(kernel principal component analysis, KPCA)与支持向量机回归(support vector regression, SVR)相结合的方法建立模型,并使用多角度数据反演 LAI。当前的研究主要只针对植被某个生育时期建立模型,很少考虑到植被生育期的多样性,同时建模时所用的植被指数单一,包含的波段信息较少,本文考虑使用支持向量机的方法反演 LAI,使其适用于冬小麦的整个生育期,同时考虑反演时融合更多的波段信息,提高反演精度。

本文以北京通州地区冬小麦作为研究对象,利用 2009 年冬小麦起身期、拔节期和灌浆期 3 个时期的地面实测光谱和叶面积数据,用统计回归的方法分别建立 NDVI-LAI 和 RVI-LAI 模型,用 SVR 方法分别建立以 NDVI、RVI 以及蓝、绿、红和近红外 4 个波段数据作为输入参数的回归预测模型,即 NDVI-SVR、RVI-SVR 和 NRGB-SVR 模型。最后利用对应时期的环境星 HJ-CCD 数据结合建立的模型反演 LAI 并用实测数据验证。

1 材料与方 法

1.1 研究区和试验方案

试验区位于 115°25'~117°35'E, 39°28'~41°05'N,地处华北平原与太行山脉、燕山山脉的交接部位,与天津一起被河北省环绕,气候类型为典型的暖温带半湿润大陆性季风气候。本研究试验于 2009 年在北京顺义和通州 2 个郊区开展。顺义地区试验时间为 2009 年 3 月 31 号(起身期)、4 月 16 号(拔节期)、5 月 18 号(灌浆期),观测的内容为冠层光谱和叶面积指数,测量的数据用于建立预测模型;通州地区试验时间为 2009 年 4 月 1 日(起身期)、4 月 17 日(拔节期)、5 月 17 日(灌浆期),观测的内容是叶面积指数,测量的数据用于验证试验。

1.2 数据的获取

1.2.1 小麦地面光谱和叶面积数据

光谱采集使用美国 ASD 光谱仪。测定选择在少云无风条件下进行,每个样点测定 20 次,每次测量前后均用参考版进行校正。测完光谱后进行取样操作,取样面积为 50 cm×4 行(行距 15~20 cm),将取样样本密封带回实验室用比叶重法测定叶面积指数。同时每个样点同步采集 GPS 坐标数据。

1.2.2 遥感影像获取

试验从中国资源卫星应用中心网站上下载与

实测验证数据准同步的 HJ-CCD 影像数据,选择云覆盖量较少的影像进行下载,空间分辨率为 30 m。

1.3 数据的预处理

1.3.1 实测光谱数据处理

环境星 CCD 影像数据所包含的波段分别为:蓝波段(430~520 nm)、绿波段(520~600 nm)、红波段(630~690 nm)和近红外波段(760~900 nm)。实测光谱数据与 CCD 数据波段并没有一一对应,因此需根据环境星 CCD 传感器的 4 个波段响应函数对实测光谱数据进行模拟,得到与环境星 HJ-CCD 数据波段相对应的光谱反射率。

1.3.2 遥感数据预处理

环境星 HJ-CCD 数据的预处理主要包括辐射定标、几何校正和大气校正,最后得到 4 个波段的反射率数据。辐射定标的定标系数可以从影像说明文件中获取,几何校正采用校正好的 Landsat TM 影像作为参考图像进行校正,误差控制在 0.5 个像元以内,大气校正采用 ENVI4.7 的 FLAASH 大气校正模块来处理。

1.4 植被指数选择

经验模型法反演叶面积指数使用的植被指数有很多种^[21-22],本文选择常用的 2 种指数:归一化植被指数 NDVI 和比值植被指数 RVI (ratio vegetation index)。用这 2 种指数与 LAI 在不同的时期分别建立回归模型,同时对模型进行验证并且与支持向量机回归方法反演 LAI 作对比分析。NDVI 和 RVI 的计算式如下

$$NDVI = (R_{NIR} - R_R) / (R_{NIR} + R_R) \quad (1)$$

$$RVI = R_{NIR} / R_R \quad (2)$$

式中, R_{NIR} 和 R_R 分别为环境星 CCD 近红外波段和红波段的亮度值。

1.5 支持向量机回归

近年来,支持向量机(SVM)在各种分类和回归问题中的应用越来越多,在分类和回归问题中均能得到较好的应用^[23-24]。在分类问题中,试图寻找最优超平面来分离 2 类样本,也就是寻找到 2 类边缘距离最大的超平面。

在回归中,目标是在有限的噪声样本基础上预测出未知的连续函数。支持向量机回归有线性回归和非线性回归^[25-27],非线性回归只是在线性回归的基础上引入核函数^[28]得到的。下面对线性回归作简单介绍,支持向量机用于回归时需引入损失函数的概念,Vapnik^[29]提出了如式(3)所示的 ε -不敏感损失函数

$$L_\varepsilon(f(x), y) = \begin{cases} 0 & |f(x) - y| < \varepsilon \\ |f(x) - y| - \varepsilon & \text{其他} \end{cases} \quad (3)$$

式中, ε 为不敏感系数, 用于控制拟合精度。 $L_\varepsilon(f(x), y)$ 是损失函数, $f(x)=\omega x+b$ 是回归函数, 用于拟合数据 $\{x_i, y_i\}, i=1, \dots, m, x_i \in R^d, y_i \in R$ 。其中 ω 、 b 为系数。

当使用线性回归函数拟合样本数据时, 假设所有训练数据的拟合误差精度为 ε , 即

$$\begin{cases} y_i - \omega x_i - b \leq \varepsilon \\ \omega x_i + b - y_i \leq \varepsilon \end{cases} \quad i=1, \dots, n \quad (4)$$

根据结构风险最小化准则, 应使目标函数 $\frac{1}{2}\|\omega\|^2$ 最小。考虑到在 ε 误差下不能估计的数据,

引入松弛因子 $\xi_i \geq 0, \xi_i^* \geq 0$, 则标准的 ε 不敏感支持向量机可以表示为

$$\begin{aligned} \min \quad & \frac{1}{2}\|\omega\|^2 + C \sum_{i=1}^n (\xi_i + \xi_i^*) \\ \text{s.t.} \quad & \begin{cases} y_i - \omega x_i - b \leq \varepsilon + \xi_i \\ \omega x_i + b - y_i \leq \varepsilon + \xi_i^* \\ \xi_i, \xi_i^* \geq 0 \end{cases} \end{aligned} \quad (5)$$

常数 $C>0$ 为平衡因子, 用于控制对超出误差范围的样本的惩罚程度。最后, 通过引入拉格朗日乘数, 函数 $f(x)$ 可以表示为:

$$f(x) = \sum_{nSV} (\alpha_i - \alpha_i^*) \cdot x_i \cdot x + b \quad (6)$$

式中, nSV 是支持向量 (SVs) 的个数, α_i, α_i^* 是拉格朗日乘子。

对于非线性回归, 只需要引入核函数, 通过非线性映射将输入空间映射到高维的特征空间, 在高维空间上进行线性回归。核函数有很多种, 也可以根据需要自己设计, 本文采用径向基核函数(RBF):

$$k(x, x_i) = \exp(-\|x - x_i\|^2 / p^2) \quad (7)$$

式中, x_i 为核函数中心, p 为函数的宽度参数, 控制了函数的径向作用范围。引入核函数后, 式 (6) 将变为

$$f(x) = \sum_{nSV} (\alpha_i - \alpha_i^*) \cdot k(x, x_i) + b \quad (8)$$

本文分别以 NDVI、RVI 以及蓝、绿、红和近红外 4 个波段反射率作为输入参数即自变量, LAI 为因变量, 运用 SVR 进行训练和预测。SVR 中的参数用交叉验证法获得, 同时考虑到样本较少, 为防止“过学习”, 对参数 C 进行适当调整。使用 LIBSVM 软件包^[30]实现回归和预测。

2 结果与分析

2.1 LAI 反演模型的建立

2.1.1 LAI 的植被指数反演模型

本文利用实测的顺义试验数据进行建模, 利用

2009 年 3 月 31 日 (起身期)、4 月 16 日 (拔节期) 和 5 月 18 日 (灌浆期) 冠层光谱模拟出的环境星 HJ-CCD4 个波段数据计算归一化植被指数 (NDVI) 与比值植被指数 (RVI), 结合同时期实测的 LAI 进行统计分析。由于在不同生育时期植被指数 (VI) 与叶面积指数 (LAI) 统计模型有所不同, 因此对 3 个时期数据分别建立 NDVI-LAI 与 RVI-LAI 回归方程。图 1 为 3 月 31 日 (起身期) 植被指数与叶面积指数的相关关系。

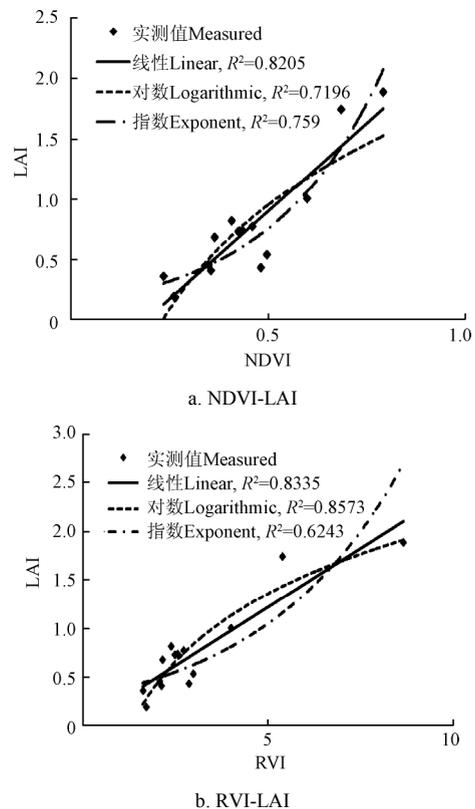


图 1 起身归一化植被指数 NDVI、比值植被指数 RVI 与叶面积指数 LAI 的相关关系

Fig.1 Relationship between vegetation index and LAI at erecting stage

3 个不同时期建立的 NDVI-LAI 和 RVI-LAI 估算模型如表 1 所示。

由表 1 可以看出, 在各个生育时期, 以归一化植被指数 (NDVI) 建模时, 一般指数模型的相关性较好, 在起身期和拔节期, 冬小麦快速生长, 叶面积指数增长迅速, 然而 NDVI 变化范围在 0~1 之间, 从而表现为叶面积指数随 NDVI 呈指数型增长, 在灌浆期, 小麦生长速度减缓, 这种增长模型已不明显, 从表 1 中可以看出, 灌浆期 3 种模型的相关系数相近, 此时的 NDVI 也已达到饱和。灌浆期仍然表现为指数模型也体现了 NDVI 自身的缺陷: 容易受到土壤背景和大气的影响, 在植被覆盖度较大

的地区容易发生饱和现象。当选择比值植被指数 (RVI) 建模时, 对数模型和线性模型的相关系数较好, 从表 1 中可以看出这 2 个模型差别很小。RVI 是近红外波段与红光波段反射率的比值, 并没有进行归一化, 变化范围相对较大, 冬小麦生长阶段, RVI 变换明显, 建模时与叶面积指数得到了较好的线性关系或者对数关系, 从这点来看, RVI 在一定程度上可以克服饱和性的问题。以上的分析仅是根据冬小麦的生长特点以及植被指数的本身特点进行的定性分析, 对于在不同时期建模时所具有的差异应在大量数据的基础上进行系统的分析, 本文所用的试验数据有限难以达到要求, 在后续的研究中可以对这个问题有个较为深入的分析。本文对不同时期和不同植被指数均选择相关系数最好的回归方程作为最后的 LAI 反演模型, 例如 3 月 31 日 (起身期) 的 NDVI-LAI 模型选择 $R^2=0.8605$ 的指数模型作为最后的 LAI 反演模型。模型建立完成后使用影像数据反演通州地区的叶面积指数 (LAI), 同时利用 2009 年 4 月 1 日 (起身期)、4 月 17 日 (拔节期)、5 月 17 日 (灌浆期) 实测叶面积指数 (LAI) 数据对反演结果进行验证。

表 1 不同时期 NDVI 和 RVI 的 LAI 估算模型

Table 1 LAI estimation models by NDVI and RVI

日期 Date	植被指数 Vegetation indices	关系模型 Relation models	R^2	RMSE
起身期 3 月 31 日 (erecting stage March 31)	NDVI	$y=2.9051x-0.5537$	0.8205	0.2016
		$y=1.2527\ln(x)+1.8194$	0.7196	0.2521
	RVI	$y=0.1344e^{3.4522x}$	0.8605	0.1778
		$y=0.2415x+0.0133$	0.8335	0.1942
		$y=1.0105\ln(x)-0.2688$	0.8573	0.1798
		$y=0.2833e^{0.2582x}$	0.5711	0.3074
拔节期 4 月 16 日 (elongation stage April 16)	NDVI	$y=3.0338x-0.964$	0.6131	0.3829
		$y=1.5957\ln(x)+1.8242$	0.5113	0.4303
	RVI	$y=0.1257e^{2.956x}$	0.7071	0.3331
		$y=0.0682x+0.4788$	0.5872	0.3759
		$y=0.7396\ln(x)-0.4096$	0.5934	0.3730
		$y=0.5886e^{0.0558x}$	0.3848	0.4591
灌浆期 5 月 18 日 (filling stage May 18)	NDVI	$y=19.867x-13.845$	0.6348	0.6525
		$y=16.073\ln(x)+5.6626$	0.6242	0.6620
	RVI	$y=0.0029e^{8.0912x}$	0.6708	0.6258
		$y=0.3058x-0.7877$	0.7846	0.5078
		$y=3.4684\ln(x)-5.6091$	0.7836	0.5091
		$y=6339e^{0.1193x}$	0.6612	0.6371

2.1.2 支持向量机回归 (SVR) LAI 反演模型

由于每个时期的 LAI 观测数据较少, 不同时期分别进行支持向量机训练样本数过少, 因此将 3 个时期数据放在一起作为训练样本, 训练时植被指数和波段数据作为自变量, LAI 作为目标值。本文建立 3 个 SVR 回归模型, 即分别以 NDVI、RVI 为自

变量的 NDVI-SVR、RVI-SVR 回归模型, 以及以模拟的环境星 HJ-CCD 4 个波段数据作为自变量的 NRGB-SVR 模型, 与统计模型不同, 在同一自变量条件下不同时期的 SVR 模型均相同。

2.2 验证结果

在样本数较少时, 建立 LAI 估算模型往往会呈现“过度拟合”现象^[31], 使模型存在误差, 有必要结合野外实测数据对模型进行检验, 验证模型的有效性。本文使用通州地区 2009 年 4 月 1 日、4 月 17 日和 5 月 17 日地面实测 LAI 数据和与实测数据准同步的环境星 HJ-CCD 数据对建立的模型进行验证。利用环境星 HJ-CCD 反射率数据结合前文建立好的模型反演 LAI。对统计回归模型而言, 不同时期不同植被指数所使用的回归方程均不同, 对 SVR 模型而言, 自变量相同时均使用同一 SVR 模型反演 LAI。最后将同一模型反演所得的 3 个时期 LAI 数据放在一起与实测数据比较, 分析不同模型的反演精度。本文选取预测精度和均方根误差 (RMSE) 作为模型的评价指标, 并对反演数据和实测数据进行了相关性分析, 分析结果和 3 个时期的总体验证结果见表 2。

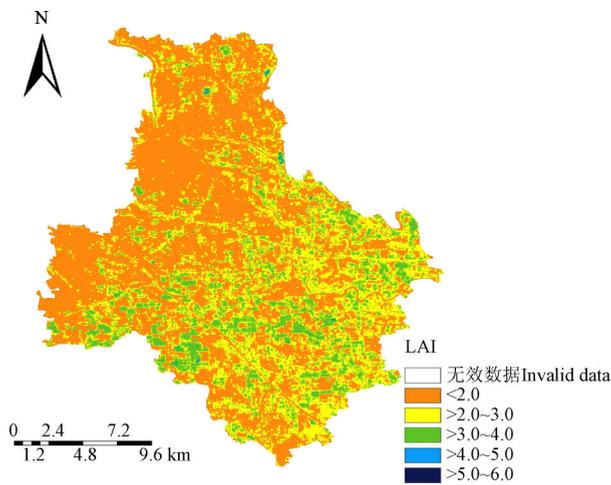
表 2 3 个时期不同模型总体验证结果

Table 2 Overall verification results of different models

模型 Models	RMSE	预测精度 Accuracy	R^2	p-value
NDVI-LAI	0.9853	0.5913	0.507	<0.0001
RVI-LAI	0.9747	0.5914	0.371	0.001
SVR	NDVI	0.7073	0.7045	<0.0001
	RVI	0.8302	0.6707	<0.0001
	R、G、B、NIR 4 波段	0.4256	0.8172	0.792

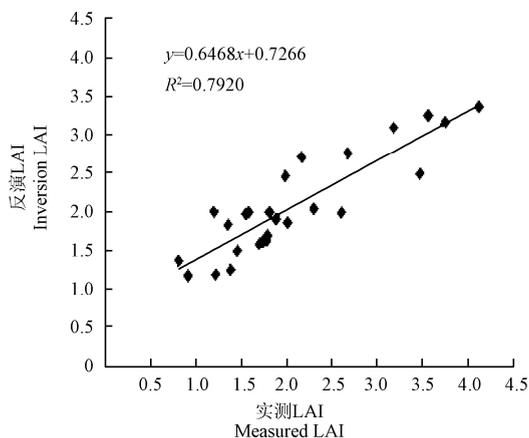
由表 2 可以看出, 5 种模型得到的反演结果与实测数据相关分析结果全部表现为极显著相关 ($p\text{-value}<0.01$), 且相关系数均高于 0.6, 以 R、G、B 和 NIR4 个波段为输入参数的 SVR 预测模型得到的相关系数最高, 为 0.89。以 NDVI 为自变量, 支持向量回归 (SVR) 方法得到的 RMSE 比 NDVI-LAI 统计回归方法得到的 RMSE 降低了 28%, 预测精度从 60% 左右提高到了 70% 左右, RVI 也得到了类似的结果。说明基于结构风险最小化原理的支持向量回归 (SVR) 拟合效果要比普通的回归方法优越, 用该方法训练预测叶面积指数 (LAI) 是可行的。同时从表 2 可以看出以 R、G、B 和 NIR 4 个波段为输入参数的 SVR 预测模型得到的 RMSE 最小, 为 0.4256, 预测精度 81% 也达到了最高。上表中前 4 个模型的输入参数只包含了红波段和近红外波段 2 个波段, 包含的信息相对较少, 融合了 4 个波段的模型, 包含的信息最多, 相比于单波段已受到外界

因素影响,多波段可以更好的反映地物的信息,得到的结果也最好,说明融合更多波段信息的 LAI 反演方法能提高反演精度,得到更好的效果。支持向量回归(SVR)为多波段协同反演叶面积指数(LAI)提供了新方法,而且本文的训练和验证数据均选用了3个不同的生长时期,说明SVR方法不受生育期的限制,可以在整个生育期内推广使用。图2为用环境星影像数据得到的LAI反演结果,其中图2a是灌浆期通州地区LAI反演结果的空间分布情况,反演的模型用的是NRGB-SVR回归模型。从LAI反演结果分布来看,通州地区小麦主要分布在中部和东部地区,与实际情况基本相符。在整个区域,LAI取值有较大变化,说明了NRGB-SVR模型的有效性。图2b为3个时期地面GPS点影像反演LAI与实测LAI数据的散点图,实测值和反演值之间具有较好的一致性,决定系数达到了0.792,说明NRGB-SVR模型具有较高的精度。



a. 灌浆期通州地区 LAI 反演分布图

a. Inversion LAI distribution of winter wheat at filling period in Tongzhou



b. NRGB-SVR 模型反演 LAI 与实测 LAI 散点图

b. Scatter diagram of measured LAI and inversion LAI with NRGB-SVR model

图2 环境星 CCD 数据的反演叶面积指数 LAI 结果

Fig.2 HJ-CCD images inversion results of LAI

3 结论

本文通过支持向量机回归的方法(SVR)建立叶面积指数(LAI)反演模型,模型的输入参数选择了2种植被指数NDVI和RVI,以及利用实测数据模拟的环境星HJ-CCD4个波段的反射率数据,训练样本选择了冬小麦起身期、拔节期和灌浆期3个时期的数据。同时建立了NDVI、RVI与LAI的经验模型。最后利用环境星HJ-CCD数据结合模型反演LAI,并用实测数据进行验证。结果表明:

1) 冬小麦不同生育期NDVI、RVI与LAI的关系呈较明显的正相关关系,这与前人研究结果相一致,但是在冬小麦的不同生长时期,NDVI、RVI和LAI的关系模型有所不同,因此需在不同生育时期建立不同的反演模型,以达到更好的预测结果。

2) 基于植被指数建立的经验模型与SVR方法建立的模型相比,SVR方法预测精度比NDVI-LAI和RVI-LAI经验模型分别提高了10%和8%。说明SVR方法具有更好的拟合效果,对于反演LAI具有可行性。同时在冬小麦的不同生育时期,SVR方法可以用统一训练好的模型进行预测,说明SVR方法适用于冬小麦的主要生育期。

3) 使用环境星HJ-CCD4个波段作为输入参数的SVR模型所得到的预测效果最好,精度达到了81%。与单波段数据易受环境因素影响相比,多波段数据包含了更为丰富的信息,支持向量机回归方法(SVR)在输入参数上可以融合更多的波段信息,也提高了LAI反演精度。本文的方法为多波段协同反演叶面积指数提供了新的思路。

[参 考 文 献]

- [1] Kouiti Hasegawa, Hiroshi Matsuyama, Hayato Tsuzuki, et al. Improving the estimation of leaf area index by using remotely sensed NDVI with BRDF signatures[J]. Remote Sensing of Environment, 2010, 114: 514-519.
- [2] Zhiqiang Xiao, Shunlin Liang, Jindi Wang, et al. Real-time retrieval of Leaf Area Index from MODIS time series data[J]. Remote Sensing of Environment, 2011, 115: 97-106.
- [3] Darvishzadeh R, Atzberger C, Skidmore A K, et al. Leaf Area Index derivation from hyperspectral vegetation indices and the red edge position[J]. International Journal of Remote Sensing, 2009, 30(23): 6199-6218.
- [4] Chen J M, Cihlar J. Retrieving leaf area index of boreal conifer forests using landsat TM images[J]. Remote Sensing of Environment, 1996, 55(2): 153-162.
- [5] Rasmus Houborg, Martha Anderson, Craig Daughtry. Utility of an image-based canopy reflectance modeling tool for remote estimation of LAI and leaf chlorophyll content at the field scale[J]. Remote Sensing of Environment, 2009, 3(1): 259-274.

- [6] Jianxi Huang, Yuan Zeng, Andres Kuusk et al. Inverting a forest canopy reflectance model to retrieve the overstorey and understorey leaf area index for forest stands[J]. *International Journal of Remote Sensing*, 2011, 32(22): 7591—7611.
- [7] 王纪华, 赵春江, 黄文江, 等. 农业定量遥感基础与应用[M]. 北京: 科学出版社, 2008.
- [8] Thorp K R, Wang G, West A L, et al. Estimating crop biophysical properties from remote sensing data by inverting linked radiative transfer and ecophysiological models[J]. *Remote Sensing of Environment*, 2012, 124: 224—233.
- [9] 赵虎, 裴志远, 马尚杰, 等. 利用 HJ-1-A/B CCD2 数据反演冬小麦叶面积指数[J]. *农业工程学报*, 2012, 28(10): 172—176.
Zhao Hu, Pei Zhiyuan, Ma Shangjie, et al. Retrieving leaf area index of winter wheat using HJ-1-A/B CCD2 data[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2012, 28(10): 172—176. (in Chinese with English abstract)
- [10] 刘晓臣, 范闻捷, 田庆久, 等. 不同叶面积指数反演方法比较研究[J]. *北京大学学报: 自然科学版*, 2008(2): 57—64.
Liu Xiaochen, Fan Wenjie, Tian Qingjiu et al. Comparative analysis among different methods of leaf area index inversion[J]. *Journal of Peking University: Natural Science Edition*, 2008(2): 57—64. (in Chinese with English abstract)
- [11] 李鑫川, 徐新刚, 鲍艳松, 等. 基于分段方式选择敏感植被指数的冬小麦叶面积指数遥感反演[J]. *中国农业科学*, 2012, 45(17): 3486—3496.
Li Xinchuan, Xu Xingang, Bao Yansong, et al. Retrieving LAI of winter wheat based on sensitive vegetation index by the segmentation method[J]. *Scientia Agricultura Sinica*, 2012, 45(17): 3486—3496. (in Chinese with English abstract)
- [12] Vuolo F, Dini L, D'Urso G. Retrieval of leaf area index from CHRIS/PROBA data: an analysis of the directional and spectral information content[J]. *International Journal of Remote Sensing*, 2008, 29 (17/17): 5063—5072.
- [13] 蒙继华, 吴炳方, 李强子. 全国农作物叶面积指数遥感估算方法[J]. *农业工程学报*, 2007, 23(2): 160—167.
Meng Jihua, Wu Bingfang, Li Qiangzi. Method for estimating crop leaf area index of China using remote sensing[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2007, 23(2): 160—167. (in Chinese with English abstract)
- [14] 方秀琴, 张万昌. 叶面积指数(LAI)的遥感定量方法综述[J]. *国土资源遥感*, 2003(3): 58—62.
Fang Xiuqin, Zhang Wanchang. The application of remotely sensed data to the estimation of the leaf area index[J]. *Remote Sensing For Land and Resources*, 2003(3): 58—62. (in Chinese with English abstract)
- [15] 陈雪洋, 蒙继华, 杜鑫, 等. 基于环境星 CCD 数据的冬小麦叶面积指数遥感监测模型研究[J]. *国土资源遥感*, 2010, 21(2): 55—58.
Chen Xueyang, Meng Jihua, Du Xin, et al. The monitoring of the winter wheat leaf area index based on HJ 1 CCD data[J]. *Remote Sensing for Land and Resource*, 2010, 21(2): 55—58. (in Chinese with English abstract)
- [16] 张瀛, 孟庆岩, 武佳丽, 等. 基于环境星 CCD 数据的环境植被指数及叶面积指数反演研究[J]. *光谱学与光谱分析*, 2011, 31(10): 2789—2793.
Zhang Ying, Meng Qingyan, Wu Jiali, et al. Study of environmental vegetation index based on environmental satellite CCD data and LAI inversion[J]. *Spectroscopy and Spectral Analysis*, 2011, 31(10): 2789—2793. (in Chinese with English abstract)
- [17] 郭琳, 裴志远, 张松龄, 等. 基于环境星 CCD 图像的甘蔗叶面积指数反演方法[J]. *农业工程学报*, 2010, 26(10): 201—205.
Guo Lin, Pei Zhiyuan, Zhang Songling, et al. Estimation method of sugarcane leaf area index using HJ CCD images[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2010, 26(10): 201—205. (in Chinese with English abstract)
- [18] 宰松梅, 温季, 郭冬冬, 等. 基于支持向量机模型和图像处理技术的甜椒叶面积测定[J]. *农业工程学报*, 2011, 27(3): 237—241.
Zai Songmei, Wen Ji, Guo Dongdong, et al. Determination of leaf area of sweet pepper based on support vector machine model and image processing[J]. *Transactions of the Chinese Society of Agricultural Engineering (Transactions of the CSAE)*, 2011, 27(3): 237—241. (in Chinese with English abstract)
- [19] Shaban Shataee, Syavash Kalbi, Asghar Fallah, et al. Forest attribute imputation using machine-learning methods and ASTER data: comparison of k-NN, SVR and random forest regression algorithms[J]. *International Journal of Remote Sensing*, 2012, 33(19): 6254—6280.
- [20] Surya S Durbha, Roger L King, Nicolas H Younan. Support vector machines regression for retrieval of leaf area index from multiangle imaging spectroradiometer[J]. *Remote Sensing of Environment*, 2007, 107(1): 348—361.
- [21] Rouse J W, Haas R H, Schell J A, et al. Monitoring vegetation systems in the great plain with ERTS[C]// *Proc. of the 3rd ERTS Symposium*, 1973(1): 48—62.
- [22] Jordan C F. Derivation of leaf area index from quality of light on the forest floor[J]. *Ecology*, 1969, 50(4): 663—666.
- [23] 丁世飞, 齐丙娟, 谭红艳. 支持向量机理论与算法研究综述[J]. *电子科技大学学报*, 2011, 40(1): 2—10.
Ding Shifei, Qi Bingjuan, Tan Hongyan. An overview on theory and algorithm of support vector machines[J]. *Journal of University of Electronic Science and Technology of China*, 2011, 40(1): 2—10. (in Chinese with English abstract)

- [24] James Tin-Yau Kwok. Support vector mixture for classification and regression problems[C]//Proc of 14th International conference on Pattern Recognition, Brisbane, Australia, 1998, 1: 255—258.
- [25] Gulsen Taskin Kaya, Okan K. Ersoy, Mustafa E. Kamasak. Support vector selection and adaptation for classification of earthquake images[C]//Proc. IGARSS, 2009, 2: 851—854.
- [26] James Tin-Yau. Support vector mixture for classification and regression problems[C]//Proc of 14th International Conference on Pattern Recognition, Brisbane, Australia, 1998, 1: 255—258.
- [27] 王定成, 方廷健, 唐毅, 等. 支持向量机回归理论与控制的综述[J]. 模式识别与人工智能, 2003, 16(2): 192—197.
- [28] Wang Dingcheng, Fang Tingjian, Tang Yi, et al. Review of support vector machines regression theory and control[J]. Pattern Recognition and Artificial Intelligence, 2003, 16(2): 192—197. (in Chinese with English abstract)
- [29] Si Wu, Shun-ichi Amari. Conformal transformation of kernel functions: a data-dependent way to improve support vector machine classifiers[J]. Neural Processing Letters, 2002(15): 59—67.
- [30] Vapnik V N. The Nature of Statistical Learning Theory[M]. New York: Springer-Verlag, 1995.
- [31] Chang C C, Lin Chih-Jen. 2001. LIBSVM: A library for support vector machines[EB/OL]. Software available at <http://www.csie.ntu.edu.tw/~cjlin/libsvm>.
- [32] 浦瑞良, 宫鹏. 高光谱遥感及其应用[M]. 北京: 高等教育出版社, 2000.

Remote sensing inversion of leaf area index based on support vector machine regression in winter wheat

Liang Dong¹, Guan Qingsong^{1,2}, Huang Wenjiang^{2*}, Huang Linsheng¹, Yang Guijun³

(1. Key Laboratory of Intelligent Computing & Signal Processing, Ministry of Education, Anhui University, Hefei 230039, China;

2. Laboratory of Digital Earth Sciences, Institute of Remote Sensing and Digital Earth, Chinese Academy of Sciences, Beijing

100094, China; 3. Beijing Agriculture Information Technology Research Center, Beijing 100097, China)

Abstract: The method of inverting leaf area index (LAI) using a single vegetation index (VI) was influenced by different degrees of saturation and each index could contain in general two bands of information. This paper proposed the method of using support vector machine regression (SVR) for leaf area index inversion, which could use more band information as input parameters in order to improve LAI inversion accuracy. Using the winter wheat's actual spectra measurement and leaf area index data in the period of erecting stage, elongation stage and filling stage, we established a NDVI-LAI and RVI-LAI model with the statistical regression method respectively, and established regression prediction model using NDVI, RVI, as well as blue, green, red and near-infrared four-band data as input parameters with the support vector machine regression (SVR) method, namely the NDVI-SVR, RVI-SVR and NRGB-SVR model. The above five models used the corresponding period environment HJ-CCD data for validation respectively. The results showed that: the RMSE of 0.98, 0.97 with the prediction accuracy value of 59.2%, 59.3% was obtained using the NDVI-LAI and RVI-LAI regression model respectively, and the RMSE of 0.71, 0.83 with the prediction accuracy value of 70.4%, 67.1% was obtained using NDVI-SVR and RVI-SVR regression model respectively. With blue (B), green (G), red (R) and near infrared (NIR) bands as input parameters of support vector machine regression and prediction, the RMSE value is 0.39, the prediction accuracy value is 81.7%. Support vector machine regression (SVR) prediction has a better fitting effect, and can input more band information to improve the leaf area index remote sensing inversion accuracy which is suitable for winter wheat's multiple birth period.

Key words: remote sensing, spectroscopic analysis, support vector machines, inversion, leaf area index, vegetation index