

基于时间序列 NDVI 相似性分析的棉花估产

高中灵^{1,2}, 徐新刚^{1*}, 王纪华¹, 靳华安², 杨浩¹

(1. 北京农业信息技术研究中心, 北京 100097; 2. 北京师范大学地理学与遥感科学学院, 北京 100875)

摘要: 多时相遥感数据能比单一时相反映更多的作物产量信息, 挖掘多时相遥感信息以提高作物估产精度具有重要意义。该文以新疆生产建设兵团农一师一团的棉花为研究对象, 提出了一种融合分区概念和时间序列 NDVI (归一化植被指数) 相似性分析的棉花估产方法。首先, 通过植被指数与产量的相关性比较分析, 确定 NDVI 为棉花估产因子, 在此基础上根据棉花品种和土壤条件的差异, 将研究区棉田划分为不同类型的生长区; 然后, 结合每个生长区获取的样点产量数据, 确定各生长区时序 NDVI 估产模型的拟合系数; 最后, 融合距离与角度相似性算法, 对各生长区内所有棉花像元的时序 NDVI 数据构成的向量与产量样点对应的时序 NDVI 向量进行相似性分析, 确定待测棉田像元最佳的估产模型, 实现对整个棉田区域棉花产量的遥感估测。结果表明, 基于分区和时序 NDVI 相似性分析的棉花产量预测值与实测值决定系数达到 0.77, 该方法具有较好的操作性和适用性。

关键词: 遥感, 估产, 时间序列分析, 分区, NDVI, 相似性分析, 棉花

doi: 10.3969/j.issn.1002-6819.2012.02.026

中图分类号: TP79; S127

文献标志码: A

文章编号: 1002-6819(2012)-02-0148-06

高中灵, 徐新刚, 王纪华, 等. 基于时间序列 NDVI 相似性分析的棉花估产[J]. 农业工程学报, 2012, 28(2): 148-153.
Gao Zhongling, Xu Xingang, Wang Jihua, et al. Cotton yield estimation based on similarity analysis of time-series NDVI[J]. Transactions of the CSAE, 2012, 28(2): 148-153. (in Chinese with English abstract)

0 引言

新疆是中国重要的棉花生产区, 棉花产量是影响区域经济发展的重要因素, 及时、准确地对棉花产量进行估算, 可以为管理部门制定相关政策提供重要依据。传统的产量调查方法只能得到离散的产量数据, 而且需要耗费大量的人力和财力, 遥感技术因具有大区域、实时和连续动态观测的优势, 在大区域作物估产方面得到了广泛应用。

基于空间遥感技术的作物估产方法可大致分为经验、半经验及理论模型三类。经验模型不涉及复杂的作物单产形成机理, 主要通过建立作物产量与敏感光谱波段或植被指数间的统计关系式来实现, 其中应用两个或多个波段反射率线性或非线性的组合的植被指数, 进行作物遥感估产的方法取得了较大进展^[1]。如应用 NOAA/AVHRR 和 MODIS 卫星遥感影像提取的植被指数, 结合气象、农学等相关知识进行的区域作物估产应用研究, 就取得了很好的效果^[2-5]。近年来, 神经网络与权重最优组合等非线性方法被用于玉米与大麦的遥感估产中, 也获得了较高的精度, 为经验性遥感估产方法的研究提供

新的活力^[6-7]。半经验模型是将作物生长模型简化, 利用遥感信息反演影响产量的指标因子, 在模拟作物干物质质量的动态变化的基础上, 通过作物干物质分配函数模型, 得到最终的作物产量信息。目前, Monteith^[8]提出的以光能效率模型概念为基础的半经验模型应用最为广泛^[9-11], 但模型中的辐射利用效率与收获系数难以实现准确计算, 从而限制了其在区域作物单产估测中的推广应用。作物估产的理论模型是通过作物生长模型动态模拟土壤-作物-大气间的相互作用关系, 来描述作物的生长、发育、产量形成的过程。尽管将遥感技术应用在作物生长模型估产研究中取得了一定进展^[12-15], 然而遥感数据不同尺度数据转换、区域作物模型参数难以获取等因素仍然是作物估产机理模型亟待解决的问题。总之, 虽然经验作物估产模型没有描述作物产量形成的内在机理, 但其方法简单, 不论在大区域或者小尺度范围内均可应用, 仍然是目前应用最为广泛的估产方法。

近年来随着遥感的发展, 各种模型与新型传感器的出现, 逐步开展了多时相作物估产方法研究。作物在每个生育时期的生长状况均会影响产量的形成, 因此构建多时相遥感估产模型估算作物产量与单一时相相比更具有潜在的优势。尽管作物产量的形成受多种因素综合影响^[16-18], 如作物品种、气象因素、土壤因素、生物因素(病虫害等)以及田间管理与栽培方式等等, 但这些因素综合控制着作物的色素组分、含水率、细胞结构、叶面积指数以及冠层结构等植株理化参数, 这些参数的变化往往会引起作物光谱反射特性相应变化, 并对最终产量的形成产生影响, 这就为作物遥感估产经验模型的构建提供了一定的解释基础。此外, 气象因素对产量形成具

收稿日期: 2011-01-10 修订日期: 2011-09-20

基金项目: 国家科技支撑计划项目(2007BAH12B02); 国家 973 项目(2007CB714406)

作者简介: 高中灵(1980-), 女, 辽宁昌图人, 博士生, 主要从事定量遥感研究。北京 北京师范大学地理学与遥感科学学院, 100875。

Email: gzhongling@126.com

*通信作者: 徐新刚(1976-), 男, 湖北鄂州人, 副研究员, 主要从事农业遥感应用研究。北京 北京农业信息技术研究中心(国家农业信息化工程技术研究中心), 100097。Email: xxgpaper@126.com

有重要影响，但影响过程非常复杂，引入气象数据进行作物估产不但增加了模型的复杂性，反而可能会降低精度，而且，气象因素能够在植被指数上反映出来，土壤条件、作物品种以及田间管理方式等也是影响产量的主要因素，同样可以利用植被指数的时空变化反映。因此，本研究拟根据这些因素，应用分区概念将棉田区域划分为几个生长区，每个生长区是有相同品种、土壤条件以及田间管理方式相同或相近的集合，然后利用测产数据建立每一类别生长区的多元线性回归方程，通过向量之间的夹角公式与距离公式共同判断每个待估产像元对应的拟合系数，最终得到研究区棉花产量。

1 材料与方法

1.1 研究区域概况

新疆生产建设兵团一师一团（阿克苏地区）耕地面积约 130 km²，主要作物包括棉花、水稻、红枣套种辣椒、红枣、辣椒等，其中棉花种植面积所占比例最大。阿克苏位于天山以南，东望塔里木河，西临中吉边境，南邻塔里木盆地，属暖温带大陆性气候，气候干燥，降雨量少，平均降水约 75 mm/a，蒸发量大，光热资源丰富，平均气温约 10℃，是新疆棉花的高产区。

1.2 遥感数据及预处理

本文利用棉花苗期到吐絮期的 4 个时相遥感影像进行棉花种植面积提取与产量估算研究，包括环境卫星 CCD 影像、SPOT4 影像以及 TM5 影像，具体详见表 1。采用中光谱分辨率大气辐射传输模式（MODTRAN）实现遥感影像大气校正，该模型是由美国光谱科技公司、空军物理实验室联合开发的适用范围为 0.2 μm~∞ 的大气校正模型，计算精度达到 2 cm⁻¹，是目前应用最为广泛精度较高的大气辐射传输模型。利用 1:50 000 地形图对遥感影像进行几何校正，利用地面测量 GPS 坐标数据对几何校正结果验证，误差在一个像元之内。

1.3 产量数据测定

棉花测产：棉花产量在吐絮期已基本形成，测产一般在该时期内进行，产量调查结果数据部分用于建模，部分用于验证。按 5 行 3 m 长的取样面积进行棉花产量调查，数出棉花总株数与总成铃数，然后根据收获密度、平均单株成铃数以及单铃质量计算籽棉产量：籽棉产量=收获密度×平均单株成铃数×单铃质量。

表 1 遥感数据

Table 1 Remote sensing image used in the paper

传感器	轨道号	成像时间	分辨率/m	对应棉花生育期
环境卫星 1A	50-64	2009-05-28	30	苗期
TM5	147-32	2009-6-26	30	蕾期
SPOT4	205-268	2009-7-26	20	铃期
环境卫星 1A	46-64	2009-8-25	30	吐絮期

1.4 估产因子确定

在棉花不同生育时期，棉花光谱信息与其他地物光谱之间的差异不断变化，不同时期棉花生长状况对产量

的敏感程度也明显不同。因此，棉花种植面积提取与遥感估产需要根据实际情况选择最佳时相，这是利用经验模型进行棉花遥感估产的重要环节。棉花生长过程大致可以分为苗期、蕾期、铃期、吐絮期等关键生育期，从现蕾到吐絮阶段是营养生长与生殖生长的重叠时期，是影响棉花产量和质量形成的关键生育期，也是棉花估产的最佳时期。因此分别在蕾期（6 月）、铃期（7 月）和吐絮期（8 月）分别获取遥感影像用于棉花单产估算。植被指数与产量有明显的相关性，但是如何确定最佳植被指数用于估产研究是需要解决的问题。绿色植物在红光波段具有很强的吸收特性，在近红外波段具有很强的反射特性，当植被受到水分或营养胁迫时，红光波段的吸收特性与近红外波段的反射特性均减弱，因此基于红光与近红外波段反射率的植被指数经常用于作物长势监测与估产研究，而且这两个波段数据较容易获取，能够满足不同时间尺度与空间尺度业务应用的需求。最具有代表性的植被指数除 NDVI 外，还包括差值植被指数（DVI）、比值植被指数（RVI）、垂直植被指数（PVI）、调整土壤植被指数（SAVI）等，都可以作为监测作物长势与产量估算的依据。本文根据几个常用植被指数（表 2）与产量数据进行相关性分析，进而确定用于估产的植被指数。

表 2 基于红光与近红外波段的植被指数

Table 2 Vegetation index based on red and near infrared reflectance

缩写	名称	英文	公式
NDVI	归一化植被指数	Normalized difference vegetation index	$NDVI = (\rho_{NIR} - \rho_R) / (\rho_{NIR} + \rho_R)$
DVI	差值植被指数	Difference vegetation index	$DVI = \rho_{NIR} - A \times \rho_R$
RVI	比值植被指数	Ratio vegetation index	$RVI = \rho_{NIR} / \rho_R$
PVI	垂直植被指数	Perpendicular vegetation index	$PVI = (\rho_{NIR} - A \times \rho_R - B) / \sqrt{(1 + A^2)}$
SAVI2	调整土壤植被指数	Soil adjusted vegetation index	$SAVI_2 = \rho_{NIR} / (\rho_R + I/S)$

注： ρ_R 、 ρ_{NIR} 分别为红、近红外波段的反射率；S、I 分别为土壤线的斜率与截距，S=1.07，I=0.06。

棉花苗期主要以营养生长为主，而且该时期植被光谱受到地膜影响较大，所以选择蕾期、花铃期、吐絮期 3 个棉花关键生育期遥感数据用于估产研究。对这 3 个时期的 5 种植被指数与产量调查数据进行相关性分析，得到每个时期 5 种植被指数与产量的相关系数（表 3），可以发现 NDVI 在每个时期与产量的相关系数均最大，因此确定其为用于估产的因子。

表 3 棉花关键生育期各植被指数与棉花产量的相关系数

Table 3 Correlation coefficient between VI and cotton yield at different key growth period

植被指数	蕾期	花铃期	吐絮期
DVI	0.431	0.403	0.477
NDVI	0.513	0.532	0.500
RVI	0.461	0.492	0.351
PVI	0.429	0.454	0.381
SAVI2	0.386	0.491	0.337

1.5 时间序列 NDVI 数据相似性分析

作物分区对于估产研究具有重要的意义，分区的本质就是减少估产模型的异质性，增加分模型的同一性，是提高估产精度的重要表现。本研究提出了一种基于时间序列 NDVI 相似性分析的分区方法。首先确定影响棉花单产差异的主要环境因子，根据每个生长区的棉花具有相似的生长环境的原则确定生长区的类别。每一类别生长区获取一组产量调查数据，将每个待测像元对应的时间序列 NDVI 向量与所有调查点对应的时间序列 NDVI 向量进行相似性分析，将相似性最高的时序 NDVI 向量对应的生长区类别确定为待测像元对应的生长区类别，最终实现研究区所有棉花像元生长区的划分。

由于棉花各个生育期的 NDVI 都能在一定程度上反映产量信息，因此估算棉花产量要综合考虑 NDVI 的动态变化趋势。NDVI 数据的动态变化趋势因生长区的不同而不同，即使在同一个生育期，不同生长区的 NDVI 对产量的贡献率并不相同，而同一生长区内各因素对棉花产量具有相似的影响。利用每个生长区的棉花产量调查数据，利用多元线性拟合方法构建每个生长区的产量与 NDVI 之间的多元关系模型（公式（1））。

$$Y = A \times N^T + b \quad (1)$$

式中， Y 表示棉花产量， kg/hm^2 ， $N = (N_1, N_2, N_3, \dots, N_n)$ 表示时间序列 NDVI 数据， n 表示用于估产的遥感时相， A 表示时间序列 NDVI 数据对应的多元线性拟合系数， b 表示常数项， A 与 b 利用每个生长区的产量调查数据计算得到。然后，获取每个待测像元对应的时间序列 NDVI 向量，根据两个向量之间的夹角与距离公式，逐个计算这些向量与产量调查点对应的时序 NDVI 向量之间的相似性，确定每个像元所属的生长区类别，将对应生长区的多元线性拟合系数赋值给该像元，最终实现棉花单产估算。本研究中两个向量之间的相似性分析采用了夹角公式与距离公式联合的方法， P 、 Q 是两个任意的具

有 n 个元素的向量，计算两个向量之间的相似性通常采用距离公式（2）或者角度公式（3）

$$\|P - Q\| = \sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \dots + (p_n - q_n)^2} \quad (2)$$

$$\theta = \arccos \frac{[P, Q]}{\|P\| \|Q\|} = \arccos \frac{\sum_{i=1}^n p_i q_i}{\sqrt{\sum_{i=1}^n p_i^2} \sqrt{\sum_{i=1}^n q_i^2}} \quad (3)$$

然而仅利用距离公式计算两个向量的相似性，可能会发生模值接近但是夹角很大的情况；而利用角度公式计算两个向量的相似性时，则有可能发生角度接近，但模值却差别很大的情形。因此，本文采用将两者联合的方式判断两个向量的相似程度：

1) 用角度公式计算向量夹角 θ ，把 3 个最接近于 0 的向量保留；

2) 对于保留下来的向量再用距离公式，距离最小的两个向量才是相似性最高的两个向量。

1.6 数据分析与计算方法

利用 excel 软件对棉花产量调查数据进行整理，获取每个生长区的多元线性拟合系数 A 和常数项 b ，本研究所用环境卫星 CCD 影像和 TM5 影像的空间分辨率为 30 m，SPOT4 的空间分辨率为 20 m，要得到估产区域每个像元对应的时间序列 NDVI 向量，需要有一致的空间分辨率，因此将 20 m 的 SPOT4 影像重采样为 30 m 分辨率的数据，最后基于 IDL 平台实现向量之间的比较分析与产量计算。

2 结果与分析

2.1 棉花种植区域提取

估产作物种植面积提取是估产的一个重要内容。本研究利用多时相遥感数据，通过目视解译、野外调查与历史资料，基于地块单元的决策树方法提取棉花种植面积。目视解译是遥感影像分类的基本前提，图 1 是研究区主要作物的目视判读结果。

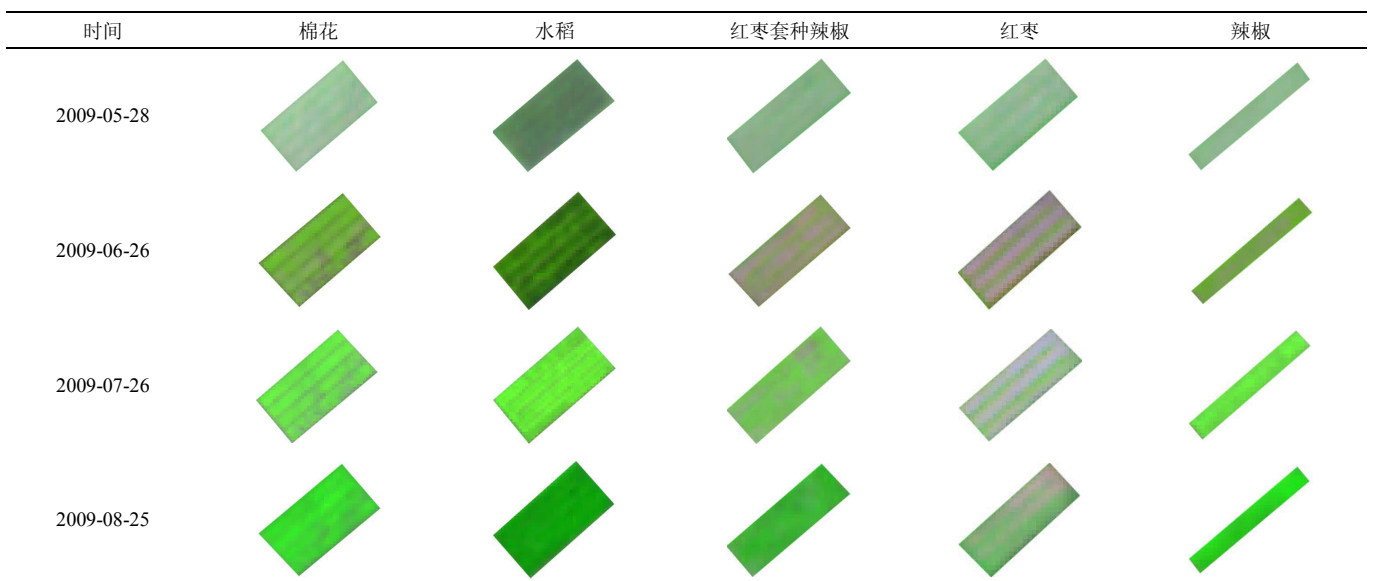
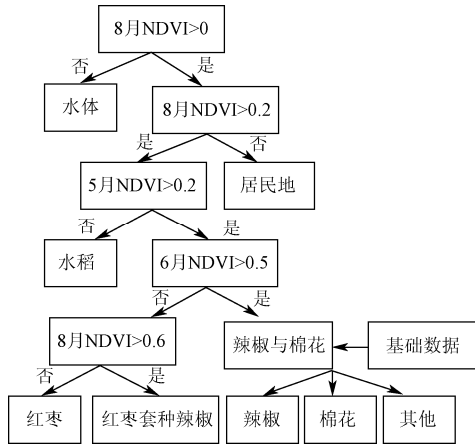


图 1 研究区作物目视解译结果

Fig.1 Result of visual interpretation in experimental area

基于像元尺度的遥感影像分类通常得到离散的分类结果，对于具有空间连续性的土地利用区域中出现的误分像元难以实现纠正，因此本文利用已有的地块边界数据裁切遥感影像，提取不同时期每个地块单元 NDVI 的平均值，根据遥感影像的目视判读结果和经验知识，利用决策树方法获取研究区的土地利用情况，这样提取的棉花种植区域在空间上具有连续性。具体分类过程与分类结果分别见图 2 与图 3。



注：NDVI 代表地块单元 NDVI 平均值

图 2 实验区土地利用分类流程

Fig.2 Flow chart of land-use classification in experimental field

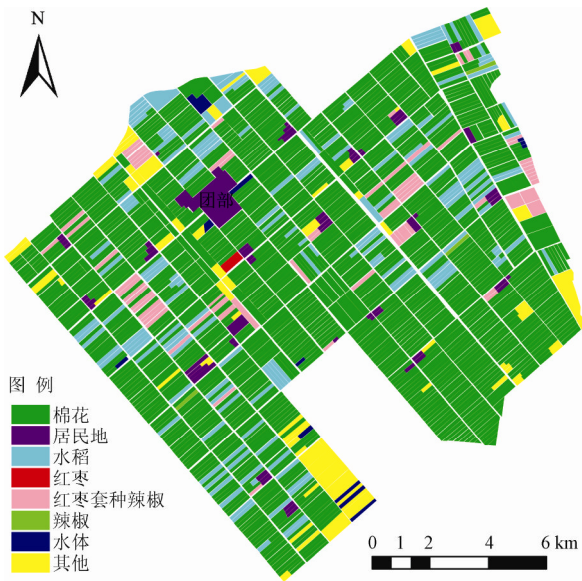


图 3 实验区土地利用分类结果

Fig.3 Result of land-use classification in experimental field

2.2 棉花生长区确定

本研究中影响棉花产量差异的主要因素有土壤条件和棉花品种（采用统一的田间管理方式），所以生长区的划分主要依据这两个要素。研究区土壤类型包括壤土与粘土；氮素含量和有机质含量是土壤环境中影响作物长势和产量的两个主要方面，所以根据土壤类型、土壤氮含量和有机质含量作为生长区划分的依据。研究区内

棉花种类较多，本研究将棉花分为三种类型：杂交棉、长绒棉和普通棉。杂交棉一般指杂交生产的棉花种子，有时也指用杂交种子种植的棉花，其产量较高；长绒棉叶片形状和株高与其他棉花有所差别，其生长期比其他棉花长 10~15 d；其他品种棉花本文统称为常规棉或普通棉。因此，根据土壤条件与棉花种类将研究区划分为 10 类生长区， A_1 、 A_2 、 A_3 、 b 表示每个生长区不同时期 NDVI 对应的拟合系数与常量，见表 4。

表 4 棉花生长区划分

Table 4 Determination of cotton growing area

棉花种类	土壤类型	土壤氮	土壤有机质	棉花主要生育期对应的拟合系数与常数项/(kg·hm ⁻²)			
				A_1 (2009-06-26)	A_2 (2009-07-26)	A_3 (2009-08-25)	b (常量)/10 ⁴
普通棉	壤土	低	低	0.17	1.76	0.18	-0.98
普通棉	壤土	低	正常	0.02	2.41	0.03	-1.30
杂交棉	壤土	低	正常	0.22	2.33	0.15	-1.45
普通棉	壤土	正常	正常	0.05	2.50	0.05	-1.40
普通棉	壤土	正常	高	0.01	2.47	0.01	-1.33
杂交棉	壤土	正常	高	0.07	2.07	0.33	-1.27
普通棉	粘土	正常	正常	0.04	2.17	0.06	-1.13
杂交棉	粘土	正常	正常	0.13	1.37	0.11	-0.58
长绒棉	粘土	正常	正常	0.04	1.03	0.40	-0.46
普通棉	壤土	正常	正常	0.02	2.20	0.02	-1.12

2.3 结果分析

本文融合分区概念与时间序列 NDVI 数据相似性分析方法实现棉花估产研究。首先根据研究区概况确定生长区的类别，利用产量调查数据确定每个生长区对应的多元线性拟合系数，然后通过待测像元对应的 NDVI 向量与产量调查点对应像元的 NDVI 向量的相似性分析，实现所有像元生长区的确定，然后将生长区对应的拟合系数赋值给该生长区所有待测像元的时间序列 NDVI 数据，实现研究区棉花产量估算。基于时间序列 NDVI 数据相似性分析的棉花估产方法与单时相经验估产模型相比，能够包含棉花关键生育期的产量信息，并且考虑了棉花生长环境的空间异质性问题，因此产量估算精度明显高于单时相遥感估产结果。图 4a、b 分别表示利用单时相 NDVI 数据和时间序列 NDVI 数据相似性分析方法得到的棉花单产估测值与实测值的比较结果。

其中，单时相遥感估产，利用的是与产量相关性最高的铃期 NDVI 数据（表 3），利用时间序列 NDVI 数据相似性分析方法实现估产时，铃期对应的 NDVI 系数最大（表 4），这说明铃期 NDVI 反映的棉花产量信息量最大。铃期是棉花营养生长和生殖生长的旺盛时期，既有根、茎、叶等营养器官的生长，又有现蕾、开花、结铃等生殖器官的发育，当棉花植株营养充足，生长状态良好，生物量迅速积累，产量逐渐形成；反之，当营养供应不足时，植株代谢过程受阻甚至停止生长引起早衰现象，蕾铃等还未发育成熟就已脱落，使棉花产量严重降低，因此这个时期的棉花田间管理至关重要。

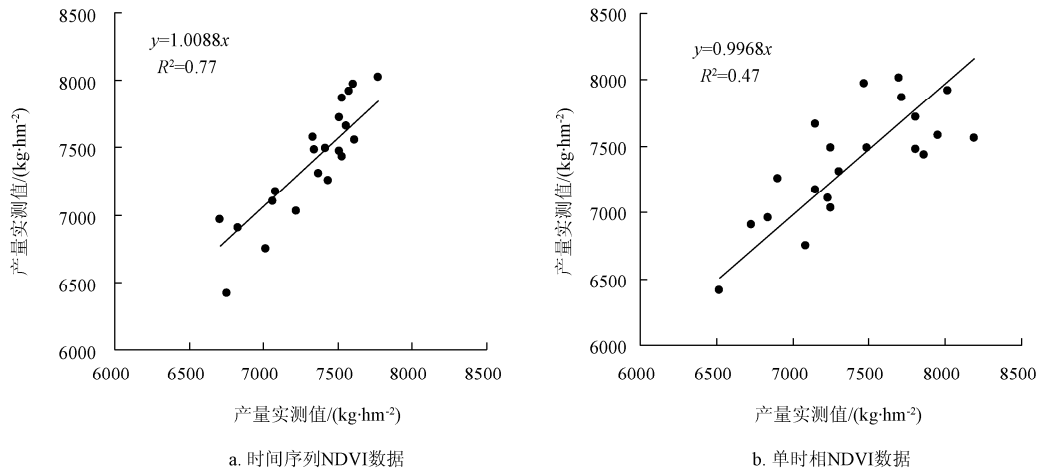


图4 时间序列 NDVI 与单时相 NDVI 数据棉花单产预测值与实测值比较

Fig.4 Relationship between estimated cotton yield and measured value utilizing time-series and mono temporal NDVI data

图5是利用时间序列NDVI向量分析方法得到的实验区棉花单产结果,从中可以看出实验区东南部棉花产量偏低,该区域主要种植普通棉花,土壤含氮量较低,盐碱程度严重,因此棉花产量偏低。基于时间序列NDVI数据相似性分析方法得到的实验区棉花单产平均值为 $6\,825\text{ kg/hm}^2$,与实际单产平均值 $6\,375\text{ kg/hm}^2$ 基本一致。

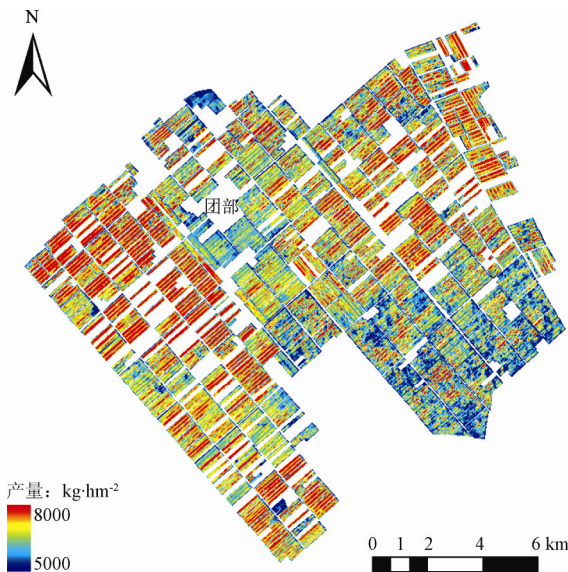


图5 实验区棉花单产估算结果

Fig.5 Result of cotton per unit area yield estimation in experimental field

3 结论

在棉田分区的基础上,应用时间序列NDVI数据相似性分析方法开展了棉花估产研究。由于棉花品种以及土壤因素等原因导致不同生长区内棉花的生长轨迹不同,同一生育期内不同生长区棉花携带产量信息量的多少不同,因此基于棉花品种和土壤条件因素对研究区利用时间序列NDVI数据相似性分析方法进行分区,实现棉花单产估产研究,避免了同一生育期内不同生长区NDVI对产量权重系数的差异引起的计算误差,对于提高

估产精度具有重要的意义。计算结果表明:利用时间序列NDVI数据相似性分析方法与单时相估产模型相比,由于前者充分利用了棉花各生育期包含的产量信息,而且引入了分区概念,是估产模型构建的合理性体现,棉花产量预测值与实测值决定系数达到0.77,具有较高的精度。

利用多光谱遥感影像,基于时间序列NDVI数据相似性分析方法进行棉花估产研究,从方法方面来讲,只用了两个向量的对比分析,采用了距离公式和角度公式联合的方法确定两个向量的相似性;从数据角度来说,只需要红光和近红外波段反射率,而且这两个波段数据较易获取,便于不同时空尺度棉花估产研究,因此,本文提出的估产方法具有一定的可操作性和适用性。

[参 考 文 献]

- [1] Wiegand C L, Richardson A J, Escobar D E, et al. Vegetation indices in crop assessments[J]. Remote sensing of Environment, 1991, 35(2): 105-119.
- [2] Van Dijk A. A Crop Condition and Crop Yield Estimation Method Based on NOAA/AVHRR Satellite Data [D]. Missouri: University of Missouri-Columbia, 1986.
- [3] Dalezios N R, Domenikiotis C, Loukas A, et al. Cotton Yield Estimation Based on NOAA/AVHRR Produced NDVI[J]. Physics and Chemistry of the Earth(B), 2001, 26(3): 247-251.
- [4] Domenikiotis C, Spiliotopoulos M, Tsiros E, et al. Remotely sensed estimation of annual cotton production under different environmental conditions in Central Greece[J]. Physics and Chemistry of the Earth, 2005, 30(1/3): 45-52.
- [5] 王长耀, 林文鹏. 基于MODIS-EVI的冬小麦产量遥感预测研究[J]. 农业工程学报, 2005, 21(10): 90-94.
Wang Changyao, Lin Wenpeng. Winter wheat yield estimation based on MODIS EVI[J]. Physics and Chemistry of the Earth, 2005, 21(10): 90-94. (in Chinese with English abstract)
- [6] Uno Y, Prasher S O, Lacroix R, et al. Artificial neural networks to predict corn yield from Compact Airborne Spectrographic Imager data[J]. Computers and Electronics in Agriculture, 2005, 47(2): 149-161.
- [7] 徐新刚, 王纪华, 黄文江, 等. 基于权重最优组合和多时

- 相遥感的作物估产[J]. 农业工程学报, 2009, 25(9): 137—142.
- Xu Xingang, Wang Jihua, Huang Wenjiang, et al. Estimation of crop yield based on weight optimization combination and multi-temporal remote sensing data[J]. Transactions of CASE, 2009, 25(9): 137—142. (in Chinese with English abstract)
- [8] Monteith J L. Climate and the Efficiency of crop production in britain[J]. Philosophical Transactions of the Royal Society of London. Series B, 1977, 281(980): 277—294.
- [9] Daughtry C S T, Gallo K P, Goward S N, et al. Spectral estimates of absorbed radiation and phytomass production in corn and soybean canopies[J]. Remote Sensing of Environment, 1992, 39(2): 141—152.
- [10] Lobell D B, Asner G P, Ortiz-Monasterio J I, et al. Remote sensing of regional crop production in the Yaqui Valley, Mexico: estimates and uncertainties[J]. Agriculture, Ecosystems and Environment, 2003, 94(2): 205—220.
- [11] Leblon B, Guerif M, Baret F. The use of remotely sensed data in estimation of PAR use efficiency and biomass production of flooded rice[J]. Remote Sensing of Environment, 1991, 38(2): 147—158.
- [12] Moulin S, Bondeau A, Delecolle R. Combining agricultural crop models and satellite observations: from field to regional scales[J]. International Journal of Remote Sensing, 1998, 19(6): 1021—1036.
- [13] Dorigo W A, Zurita-Milla R, de Wit A J W, et al. A review on reflective remote sensing and data assimilation techniques for enhanced agroecosystem modeling[J]. International Journal of Applied Earth Observation and Geoinformation, 2007, 9(2): 165—193.
- [14] 任建强, 陈仲新, 唐华俊, 等. 基于遥感信息与作物生长模型的区域作物单产模拟[J]. 农业工程学报, 2011, 27(8): 257—264.
- Ren Jianqiang, Chen Zhongxin, Tang Huajun, et al. Regional crop yield simulation based on crop growth model and remote sensing data[J]. Transactions of the CSAE, 2011, 27(8): 257—264. (in Chinese with English abstract).
- [15] 杨鹏, 吴文斌, 周清波, 等. 基于作物模型与叶面积指数遥感影像同化的区域单产估测研究[J]. 农业工程学报, 2007, 23(9): 130—136.
- Yang Peng, Wu Wenbin, Zhou Qingbo, et al. Assimilating remotely sensed LAI into GIS-based EPIC model for yield assessment on regional scale[J]. Transactions of the CSAE, 2007, 23(9): 130—136. (in Chinese with English abstract).
- [16] 徐新刚, 吴炳方, 蒙继华, 等. 农作物单产遥感估算模型研究进展[J]. 农业工程学报, 2008, 24(2): 290—298.
- Xu Xingang, Wu Bingfang, Meng Jihua, et al. Research advances in crop yield estimation models based on remote sensing[J]. Transactions of the CSAE, 2008, 24(2): 290—298. (in Chinese with English abstract)
- [17] 徐新刚. 农作物产量模型研究[D]. 北京: 中国科学院研究生院, 2007.
- Xu Xingang. Research of Crop Yield Models[D]. Beijing: Doctor Dissertation, Graduate School of Chinese Academy of Sciences, 2007. (in Chinese with English abstract)
- [18] Yuan Jinhua, Xu Renkou, Wang Ning, et al. Amendment of acid soils with crop residues and biochars[J]. Pedosphere, 2011, 21(3): 302—308.

Cotton yield estimation based on similarity analysis of time-series NDVI

Gao Zhongling^{1,2}, Xu Xingang^{1*}, Wang Jihua¹, Jin Huan², Yang Hao¹

(1. Beijing Research Center for Information Technology in Agriculture, Beijing 100097, China;

2. School of geography, Beijing normal university, Beijing 100875, China)

Abstract: Multi-temporal remote sensing images contain more crop yield information than mono temporal images, and it is very significant to mine information from multi-temporal remote sensing data for improving the precision of crop yield estimation. In the paper, taking the cotton field of First Regimental, Agriculture First Division, Xinjiang Production and Construction Corps as the studying area, a method of cotton yield estimation was proposed by integrating the concept of cotton growing area with similarity analysis of time-series NDVI data. Firstly, the NDVI was determined as the dominant factor of cotton yield estimation through correlation analysis between vegetation index and cotton yield from all sampled plots. Secondly, the whole studying area was divided into several cotton growing areas according to cotton variety and soil condition. And then the linear-fitting analyses were used to acquire the coefficient of yield model for each growing area. Finally, multiple linear regression coefficients for each cotton pixel were determined by similarity analysis between NDVI vectors from unknown-yield cotton pixels and all known ones as the investigated yields. Thus, cotton yield estimation throughout the whole studying area was realized by time-series NDVI data. The analyses show that the coefficient of determination (R^2) between the estimated and investigated yield can reach to 0.77, which indicates that the method is reasonable and adaptable.

Key words: remote sensing, estimation, time-series analysis, cotton growing area, NDVI, similarity analysis, cotton yield