

# 基于植被指数的作物产量监测方法研究

焦险峰<sup>1,2</sup>, 杨邦杰<sup>1</sup>, 裴志远<sup>1</sup>, 王 飞<sup>1</sup>

(1. 农业部规划设计研究院, 北京 100026; 2. 中国农业大学工学院, 北京 100083)

**摘 要:** 在作物收获以前进行大范围的作物长势评价与作物产量估测, 对粮食供需平衡、贸易、农业政策制定非常重要。该文收集了 1984 年到 2002 年的 NOAA 卫星和农业统计资料, 计算耕地范围的植被状态指数 VCI、温度状态指数 TCI 和植被生长状态指数 VHI, 分析了遥感植被指数与作物产量间的相关关系, 分别建立了基于植被指数的线性回归模型和非线性回归模型。结果表明, 遥感植被指数与作物产量间存在较好的相关性, 其非线性回归模型在拟合精度上高于线性回归模型。研究目的是利用卫星资料得出应用于监测作物长势的植被指数, 建立作物产量监测模型, 应用于农作物遥感监测业务化运行系统。

**关键词:** 区域作物产量预测; 非线性回归模型; 归一化植被指数; 植被状态指数(VCI); 温度状态指数(TCI)

**中图分类号:** TP7.5; S127

**文献标识码:** A

**文章编号:** 1002-6819(2005)04-0104-05

焦险峰, 杨邦杰, 裴志远, 等. 基于植被指数的作物产量监测方法研究[J]. 农业工程学报, 2005, 21(4): 104-108

Jiao Xianfeng, Yang Bangjie, Pei Zhiyuan, et al. Monitoring crop yield using NOAA/AVHRR-based vegetation indices [J]. Transactions of the CSAE, 2005, 21(4): 104-108 (in Chinese with English abstract)

## 0 引 言

中国农业灾害频繁, 气候灾每年都在一定范围内发生。受气候因素的影响, 粮食产量逐年波动。在作物收获以前进行大范围的作物长势评价, 提前估测作物产量逐渐受到政府各部门的重视, 及时准确预测一个地区或全国的作物产量, 对粮食供需平衡、贸易、农业政策制定等非常重要。

作物产量预测有许多方法, 主要有 3 种模型: 气象模型、农学模型和遥感模型。由于影响作物产量的因素复杂, 气象模型和农学模型在大范围作物产量预测中精度波动较大。遥感模型是通过建立遥感测得的作物信息(光谱信息)与产量间的关系来估算作物产量。卫星遥感数据有更高度的概括性, 而且在表征区域植被状况和分布方面比从现有气象站点得到的气候变量效果更好(Kogan, 1995), 是农学参数和气象因子的综合反映。遥感模型的优点是适于监测大范围的作物产量。

大尺度的农作物产量遥感监测使用的数据主要是 NOAA/AVHRR 数据。20 世纪 80 年代, NOAA 极轨气象卫星系统投入业务应用, 提供每天对地球的两次观测, 自 NOAA-6 开始, AVHRR 具有了适合研究植被的通道。气象卫星的可见光和红外波段, 包含 90% 以上的植被信息, 且波段间的不同组合进行植被研究效果好<sup>[1]</sup>。这些波段间的不同组合方式被称为植被指数。植被指数比单波段用来探测生物量有更好的灵敏性。20 多年来, 已研究发展了 40 多个植被指数<sup>[2]</sup>。大量研究结果表明, 植物的叶面积系数、生物量、干物重与植被指数间存在较好的相关关系<sup>[3]</sup>。遥感手段预测作物产量就是

根据植被指数与作物产量间的相关关系构建模型, 监测作物产量。NOAA/AVHRR 数据重复探测周期短、覆盖范围大, 可以形成作物生长期间的植被指数过程, 到目前已积累了长达 21 年具有相对连续性和经济性的资料, 为遥感模型的建立提供了良好的基础。

本研究目的是利用卫星资料得出应用于监测作物长势的植被指数, 建立作物产量监测模型, 应用于农作物遥感监测业务化运行系统。采用 Kogan (2001) 提出的植被状态指数 VCI (Vegetation Condition Indices)、温度状态指数 TCI (Temperature Condition Indices) 和植被生长状态指数 VHI (Vegetation Health Indices) 作为建模因子, 建立统计模型, 监测作物产量。

## 1 数 据

本项研究收集 1984 年到 2002 年的 NOAA/AVHRR 气象卫星数据和新疆玉米单产数据, 在此基础上计算耕地范围的植被指数, 建模分析预测新疆玉米产量。

### 1.1 植被指数

NOAA/6 装载 AVHRR。通道 1 (CH1) 可见光波段 (VIS), 波长 0.58~0.68 μm, 典型的正常植被在 CH1 具有强吸收特点, 由通道 1 探测的地面植被反射率非常低; 通道 2 (CH2) 近红外波段 (NIR), 波长 0.725~1.1 μm, 正常植被在 CH2 具有强反射特点, 通道 2 对地面植被的响应非常强; 通道 4 (CH4) 热红外波段 (TIR), 波长 10.3~11.3 μm, 可用查找表转换为亮度温度 (BT)。通道 1、2 合成的植被指数反映了植被的生长状况, 在已研究发展的 40 多个植被指数中应用最广的是归一化植被指数 (NDVI)。NDVI 的计算公式是

$$NDVI = (CH2 - CH1) / (CH2 + CH1) \quad (1)$$

为了消除部分云和大气的干扰, 用最大值合成法, 每个像元取该像元每 7 天的最大值, 生成 7 天合成的 AVHRR/NDVI 数据, 并对 NDVI 和 BT 数据进行滤波

收稿日期: 2004-07-12 修订日期: 2004-10-08

基金项目: 国家自然科学基金资助 (30471009)

作者简介: 焦险峰, 高级工程师, 北京朝阳区麦子店街 41 号 农业部规划设计研究院, 100026



处理消除高频噪音。本研究所用数据来自 NASA 地球观测系统数据与信息中心 (EOS DIS) Global Vegetation Index 数据集。由此得到的一年 52 周的 NDVI 数据以时间为横坐标排列形成的时间序列数据能够描述作物从出苗、到成熟的周期性生长态势。

气候和生态系统不同地区的 NDVI 和 BT 值不具有可比性, 气候和管理措施较好的地区, 生物量和产量较高, 不能与生产潜力低的地区进行比较。为此 Kogan (1995) 提出用各地区每周的植被状态指数 VCI、温度状态指数 TCI 和 VCI、TCI 的合成值 VHI 来监测气候对植被的影响<sup>[4]</sup>。

植被状态指数 VCI 反映水分对植被影响的程度, 定义为

$$VCI = (NDV_i - NDV_{min}) / (NDV_{max} - NDV_i) \times 100 \quad (2)$$

温度状态指数 TCI 反映温度对植被影响的程度, 定义为

$$TCI = (BT_{max} - BT) / (BT_{max} - BT_{min}) \times 100 \quad (3)$$

植被生长状态指数 VHI 反映温度和水分条件联合作用下对植被影响的程度, 定义为

$$VHI = a(VCI) + (1 - a)(TCI) \quad (4)$$

式中 a——控制 VCI 和 TCI 对 VHI 影响程度的调节系数。

新疆地域广阔, 区内山地、平原、沙漠交错分布, 耕地面积只占土地总面积的 2%。因此, 植被指数的计算必须考虑耕地分布对植被指数的影响。本研究在 GIS 背景数据库的支持下, 提取新疆耕地范围内的 VCI、TCI 和 VHI 值, 图 1 是 2001 年 9 月第 1 周新疆耕地范围 VCI 值分布图, 图 2 是 2001 年全年 1~52 周新疆耕地范围 VCI、TCI 和 VHI 值曲线图, 横坐标 1~52 周, 纵坐标 VCI、TCI 和 VHI 值。

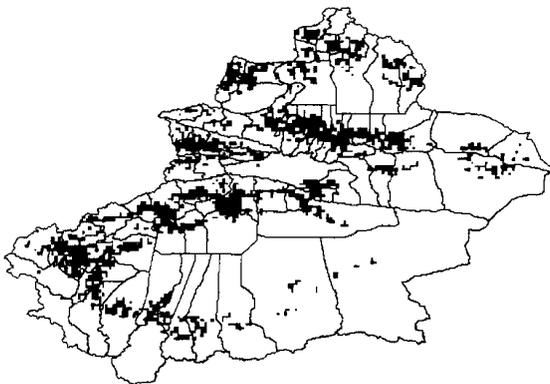


图 1 2001 年 9 月第 1 周新疆 VCI 值分布图

Fig 1 Color-coded map of VCI distribution on arable land at the first week of September 2001, Xinjiang

### 1.2 农业资料

一个地区粮食产量受气候、经济、科技等多种因素影响, 经济、科技对粮食产量的影响是一个长期的逐渐递增的渐变过程, 而气候是粮食产量年际间波动的关键因素。分析 1984 到 2001 年新疆维吾尔自治区玉米单产数据

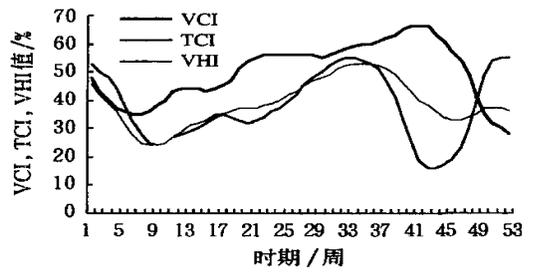


图 2 2001 年 1~52 周新疆耕地范围 VCI、TCI 和 VHI 值曲线图

Fig 2 Variations of the values of TCI, VCI, VHI distribution on arable land Xinjiang Region from the first to fifty fifth week of 2001

(Y), 通过线性拟合得到一条递增的趋势线表达经济、科技等因素决定的玉米单产 (Y<sub>t</sub>) (图 3), 该因素使玉米单产从 1984 年的 3.1 t/hm<sup>2</sup> 增加到 2001 年的 7.1 t/hm<sup>2</sup>, 17 年间增加了 129%。回归分析模型如下式, 用于预测分析 Y<sub>t</sub> 值。

$$Y_t = 0.236x - 465.16 \quad (5)$$

在经济和科技进步因素使玉米产量逐年递增的同时, 气候因素决定玉米产量的年际间波动 (dY) (图 4)。dY 可用下式表示。

$$dY = Y/Y_t * 100 \quad (6)$$

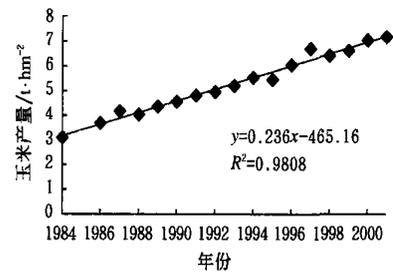


图 3 玉米产量 (Y<sub>t</sub>) 一元线性回归模型

Fig 3 Linear regression model of corn yield

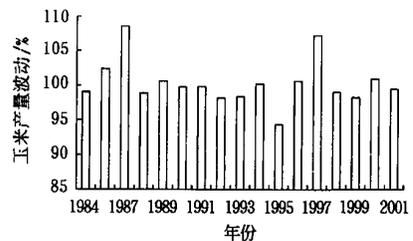


图 4 逐年波动的玉米产量 (dY)

Fig 4 Corn yield ratio (%) variation from year to year

## 2 方法

### 2.1 相关分析

在科学技术进步因素使玉米产量逐年增加的背景下, 玉米产量的波动 dY 与 VCI、TCI、VHI 之间存在相关性和规律性。

1984 年至 2001 年, 每周的 VCI、TCI 和 VHI 值与

玉米产量波动  $dY$  做相关分析, 选择相关系数高的周作为 VCI TCI 和 VHI 的代表建立模型。相关系数分布见图 5。

在第 20 周到 40 周 (5~ 10 月) 的作物生长期, VCI 与  $dY$  的相关系数由小逐渐变大, 达到最大值后, 由大逐渐变小。表明作物生长初期, 产量与植被状态指数的相关性很小, 随着作物的生长, 叶面积指数的增大, 产量与植被状态指数的相关性逐渐增加, 达到最大值, 随后相关性逐渐减少。TCI VHI 与  $dY$  的相关系数在作

物生长期也是由小逐渐变大, 达到最大值后, 由大逐渐变小有规律的变化过程。在作物生长期,  $dY$  分别与第 33 周 VCI (VCB3)、第 31 周 TCI (TCB1)、第 32 周 VHI (VHB2) 的相关系数最大。

确定了玉米产量与植被指数 VCB3、TCB1 和 VHB2 的相关性后, 通过散点图 (图 6) 可进一步分析数据间的相关关系。从图 6 可以看出, 逐年波动的玉米产量随着植被指数的增加而减小, 植被指数与玉米产量的波动之间存在明显的负相关关系。

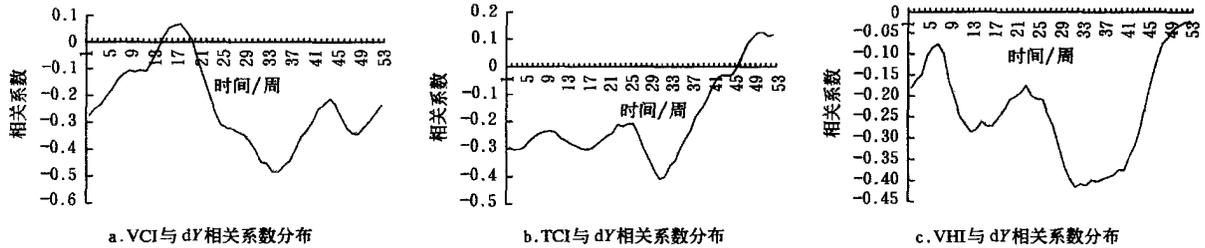


图 5 1984~ 2001 年玉米产量波动 ( $dY$ ) 和每周植被指数相关系数分布图

Fig 5 Diagram of correlation coefficient of the vegetation indices versus corn yield ratio (% of 1984~ 2001)

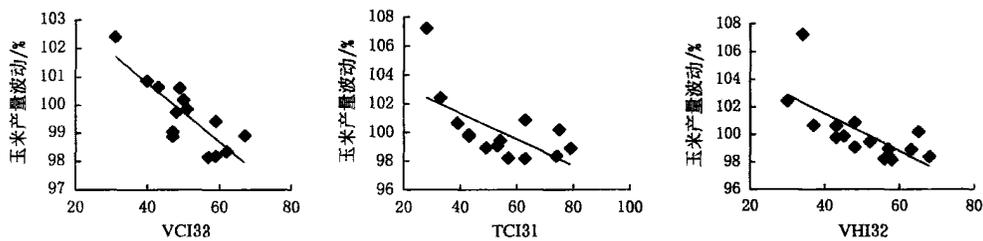


图 6 植被指数与玉米产量波动  $dY$  关系散点图

Fig 6 Scatter plot corn yield ratio (% of 1984~ 2001) versus VCI TCI and VHI at week of peak correlation

2.2 模型

根据相关分析, 可进一步建立基于植被指数的作物产量监测模型, 本研究分别建立了植被指数与玉米产量的线性和非线性回归模型。

1) 线性回归模型

分别用 VCB3、TCB1、VHB2 值与  $dY$  做一元线性回归 (表 1), 表 1 中  $dY$  为逐年波动的气候产量,  $R^2$  是回归方程相关系数的平方, 它的大小反映样本数据配合回归方程的紧密程度,  $R^2$  越接近 1, 表明方程拟合精度越高, RMSE 是剩余标准差。从表 1 可看出, VCB3 的拟合精度最好,  $R^2$  达到 0.67, TCB1 和 VHB2 的拟合精度不理想,  $R^2$  都小于 0.5。用第 31 周的 TCI 和第 33 周的 VCI 做二元线性回归分析,  $R^2$  仅为 0.47 拟合精度也不理想。

2) 曲线回归模型

为了建立精度更高的回归模型, 本研究对不同的植被指数与作物产量间进行非线性回归。从散点分布图不能清楚的得出植被指数与作物产量间符合那种曲线走向, 因此对两者分别进行了二次方程、三次方程、对数方程、指数方程、乘幂等多种曲线回归分析。经过比较研

究, 认为一元二次方程最适合于研究区的曲线回归模型 (表 2)。

表 1 玉米产量线性回归模型

Table 1 Linear regression models of corn yield ratio

植被指数	回归分析模型	R	R <sup>2</sup>	RMSE
VCB3	$dY = -0.1043(VCB3) + 104.95$	0.820	0.673	0.719
TCB1	$dY = -0.0912(TCB1) + 104.99$	0.631	0.398	1.829
VHB2	$dY = -0.1363(VHB2) + 106.96$	0.686	0.470	1.716
TCB1 \ VCB3	$dY = -0.0979(TCB1) + 0.06632(VCB3) + 101.928$	0.689	0.474	1.779

表 2 玉米产量非线性回归模型

Table 2 Non-linear regression models of corn yield ratio

植被指数	回归分析模型	R	R <sup>2</sup>	RMSE
VCB3	$dY = 0.0028(VCB3)^2 - 0.3867(VCB3) + 111.73$	0.862	0.743	0.667
TCB1	$dY = 0.006(TCB1)^2 - 0.7399(TCB1) + 121.21$	0.873	0.761	1.197
VHB2	$dY = 0.0062(VHB2)^2 - 0.748(VHB2) + 121.3$	0.768	0.590	1.570

为了更直观的显示回归拟合的结果, 建立了非线性

回归模型与实际值的拟合图(图7)。

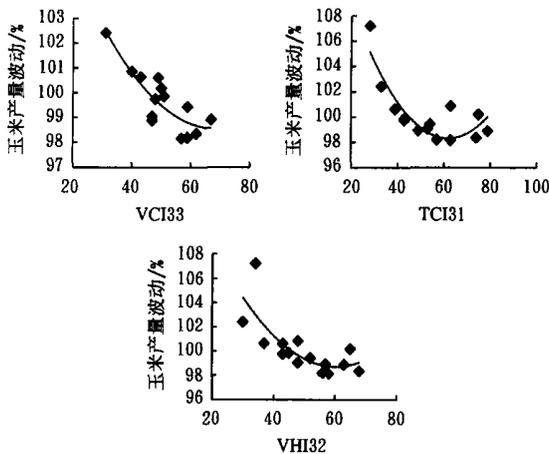


图7 基于植被指数玉米产量波动dY监测模型

Fig 7 Diagrams of crop yield ratio monitoring models based on the vegetation indices

从表2和图6可以看出,非线性回归得到的监测模型的拟合精度比一元线性回归模型的精度有显著提高。在本研究比较分析的多种非线性回归模型中,一元二次方程更适合研究区的非线性回归模型。所建模型中,TCI与作物产量的拟合精度最高, $R = 0.87$ ,RMSE(剩余标准差) $= 0.667$ ,VCI次之, $R = 0.86$ ,但RMSE稍大于1,RMSE $= 1.197$ 。

模型建成后,用2002年的植被指数和统计数据对模型进行了检验。以2002年第33周VCI(VCI33)值,用公式: $dY = 0.0028(VCI33)^2 - 0.3867(VCI33) + 111.73$ ,计算产量波动dY,用公式5计算趋势产量 $Y_t$ ,最后用公式6计算2002年新疆玉米产量。计算结果与农业统计数据比较,得到满意的预测结果。

### 3 结论

全球植被指数GVI数据集自1982年至今已有20多年的历史,由此计算的植被指数TCI、VCI、VHI有很好的时间序列用于统计分析模型。植被指数VCI、TCI和VHI反映了作物叶绿素、形体和温度特征,物理意义明确,与作物产量间存在着较好的相关关系,可用于估测农作物产量。

本研究在GIS背景数据库支持下,计算新疆耕地范围内的植被指数TCI、VCI、VHI值,通过分析该数据与玉米产量间的相关关系,应用统计模型,建立了新疆粮食产量预测模型。在作物收获前6周就可以预测作物产量。

统计模型操作简便,不讨论作物产量形成的复杂过程,可用于运行化的遥感估产方法,统计模型的局限性是物理意义不够明确,模型直接跨年和跨地区应用效果不好,但随着逐年建模研究和运行工作的深入,累计的资料和经验将使模型有更好的解释和发展。

### [参考文献]

- [1] Baret F, Guyot G, Marjor D J. TSAVI: A vegetation index which minimizes soil brightness effects on LAI and APAR estimation [C]. Processing of the 12<sup>th</sup> Canadian Symposium on Remote Sensing Vancouver, Canada, 1989: 1355- 1358
- [2] 田庆九. 植被指数研究进展[J]. 地球科学进展, 1998, (4)
- [3] 谭凯璇, 李郁竹. 遥感植被指数与冬小麦叶面积系数和麦土比的关系[J]. 遥感信息, 1989, 1
- [4] Felik N. Kogan. Droughts of the late 1980s in the United States as derived from NOAA polar orbiting satellite data [J]. Bulletin of the American Meteorological Society, 1995, 76, 655- 668
- [5] 中华人民共和国农业部编. 中国农业统计资料[M]. 北京: 中国农业出版社
- [6] 陈沈斌. 种植业可持续发展的支持系统—农作物卫星遥感估产[J]. 地理科学进展, 1998, (2).
- [7] 杨星卫. 全球稻谷生产国遥感估产可行性研究[J]. 应用气象学报, 1998, (2).
- [8] 陈沈斌. 建立我国主要农作物卫星遥感估产运行系统的主要技术环节及解决途径[J]. 自然资源学报, 1997, (4).
- [9] 黄敬峰. 基于GIS的水稻遥感估产研究[J]. 遥感技术与应用, 2002, (3).
- [10] Felik N Kogan. Operational space technology for global vegetation assessment [J]. Bulletin of the American Meteorological Society, 2001, 82(9).
- [11] K. Dabrowska-Zielinska. Modelling of crop growth conditions and crop yield in Poland using AVHRR-based indices [J]. International Journal of Remote Sensing, 2002, 23(6).
- [12] L U W T. Monitoring Brazilian soybean production using NOAA [J]. AVHRR-based vegetation condition indices [J]. International Journal of Remote Sensing, 2002, 23(6).
- [13] 杨邦杰. 国家级农情监测系统结构设计[J]. 农业工程学报, 1997, (3)
- [14] 牛志春. 青海湖环湖地区草地植被生物量遥感监测模型[J]. 地理学报, 2003, (5)
- [15] 李郁竹. 农作物气象卫星遥感监测和估产研究进展及前景探讨[J]. 气象科技, 1997, (3)
- [16] 黄敬峰. 冬小麦气象卫星综合遥感[M]. 北京: 气象出版社, 1996
- [17] 吴炳方. 全国农情监测与估产的运行化遥感方法[J]. 地理学报, 2000, (1)

## Monitoring crop yield using NOAA/AVHRR-based vegetation indices

Jiao Xianfeng<sup>1,2</sup>, Yang Bangjie<sup>1</sup>, Pei Zhiyuan<sup>1</sup>, Wang Fei<sup>1</sup>

(1. Chinese Academy of Agricultural Engineering, Beijing 100026 China;

2. Engineering College, China Agricultural University, Beijing 10083, China)

**Abstract:** Satellite remote sensing is the most effective means to monitor crop production on a regional scale. When figuring the vegetation growth condition and distribution on a large scale, vegetation indices derived from NOAA polar orbiting satellite work better than meteorological data derived from weather station. In this paper, vegetation indices including VCI (Vegetation Condition Index), TCI (Temperature Condition Index) and VHI (Vegetation Health Index) are extracted from 16 km seven-day composite NOAA AVHRR/NDVI time series images. Then, the values of VCI, TCI and VHI distribution on arable land were calculated. Based on the calculated vegetation indices and the crop yield statistical data, the linear regression models and the non-linear regression models were established, respectively, to express the relationships between the vegetation indices and crop yield. The major conclusions in this study are (1) the vegetation indices and crop yield have good correlation in a certain week of the crop growth season, (2) the fitting accuracies of the non-linear regression models are much higher than those of the linear regression models, namely, the results obtained from the non-linear regression models are more accordant with the agricultural statistic data in comparison with those from the linear regression models. Those methods can be used in the operational system of crop yield estimation on a national scale.

**Key words:** crop yield estimation; non-linear regression models; NDVI; VCI; TCI