

小麦单产的遥感-气象综合模式研究

李付琴 田国良

中国科学院遥感应用研究所

1992年3月23日收稿

摘 要

本文以北京顺义县为例,以气象因子与垂直植被指数(PVI)作为参数,用灰色模型G(0,2)和逐段订正模型即阶乘模型,建立冬小麦遥感信息-气象因子综合模型。计算结果表明,改进后的综合模型其平均精度比单纯的遥感信息模型提高近7%,个别年份达到10%以上。

关键词 遥感 冬小麦估产 灰色系统模型 阶乘模型

一、引 言

粮食产量的丰欠对我国这样一个拥有11亿人口的农业大国来说影响很大,因而及早地估测粮食产量具有重要的意义。

农业总产量等于面积乘单产,故产量预测可分为单产预测和面积预测两部分,种植面积可直接用遥感信息进行估测,而影响作物单产的因子很多,有农学、生物学、生产技术以及外界自然条件等,产量的年际波动很大,对预测模型的建立带来一定的困难。

国内外对单产模式研究得较多,常规的方法主要有农学预报、统计预报、气象统计方法等^[1-2],大部分已使用。但事实上,产量是各种因素综合反映作物的结果,因而各种方法均未能准确地反映作物的情况,尤其是产量波动较大时,遥感技术的发展,为能方便准确地进行作物估产提供了依据。

用遥感信息来建立作物单产模式国内还是近几年的事,但绝大部分是用单一的植被指数如比值植被指数,归一化植被指数等(可见光和近红外各波段的不同组合)与作物单产建立统计模型^[3-5]。为了消除土壤背景对植被指数的影响及从产量形成的生物学与农学观点对遥感单产模式进行研究以消除上述统计模型带来的误差,田国良等在研究水稻产量模式时,通过叶面积指数(LAI)与截获光合有效辐射的关系和LAI与卫星光谱计算得到的垂直植被指数(PVI)的关系,使气象因子的光温作用与作物光谱反射特征,统一于一个估产模式中,此模式是用卫星资料求出作物的LAI值,再用LAI求出单位面积的产量^[6-7]。北京大学遥感技术所在建立河南省冬小麦单产模式时,从作物产量构成三要素即亩穗数、穗粒数、千粒重出发,用可控样地建立三要素与PVI的统计模式,最后推广到大面积的单产模式中^[8]。由于PVI具有较好的消除土壤背景干扰的能力和对大气效应反应不灵敏等优点,且用农学方法建立模型可以避免单纯的统计模型带来的盲目性,因而这两种方法比前面的几种方法确具有一定的科学性。但植被指数仅仅反映了作物抽穗前后绿色作物的长势,其对作物产量的最终形成尤其灌浆期光谱资料不能完全反映出

来,有相当多的资料与现象表明,高植被指数、高叶面积指数并不代表高产,作物后期的外界条件,尤其灌浆期的气象条件对产量影响很大。本文在前人研究的基础上,以北京市顺义县冬小麦为例,用逐段订正的阶乘模型,选择小麦抽穗期的 *PVI* 和 5 月下旬的平均气温分别作为两阶段的预报因子,以 1986—1990 年 5 年资料为序列,用灰色预测的 $G(0, h)$ 模型建立冬小麦单产模型。

二、资料来源与方法介绍

1. 资料来源

顺义县遥感资料取自国家气象局气象卫星中心 NOAA 卫星 AVHRR 第 1 和第 2 通道的反射率数据,气象资料抄自北京市气象局,5 年的产量资料来自顺义县统计局及各样点所在乡。

顺义县为京郊县,经营规模,机械化程度都比较高,已进行了多年人工实割实测,且顺义县属“京津冀统一网络冬小麦遥感估产协作组”的一个抽中样县,在全县不同产量等级区都有样点,故可以认为产量资料来源基本可靠。

2. 穗帽 (*tasseled cap*) 变换和 *PVI* 的计算

Kauth 等人描述作物在土壤中开始生长到枯萎在几个波段 (MSS) 的反射特征如图 1 所示。作物反射光谱大部分在一个被称为土壤平面的上方,其形状象帽子,故称为穗帽。各地土壤背景不同,作物在各波段的反射特征也不同,而从这几维空间中提取反映土壤特征的亮度 (*brightness, BR*), 作物绿色特征的绿度 (*greenness, GN*), 反映作物枯萎的黄色度 (*yellowness, YE*) 和干扰项 (*nonsense, NS*) 称为穗帽变换^[9]。

kauth-Thomas 和 Jackson^[10] 等人用 Gram-Schmidt 计算 n 维的变换系数。

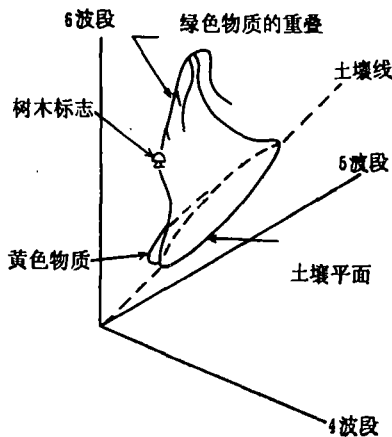


图 1 穗帽

Fig.1 The Tasseled Cap

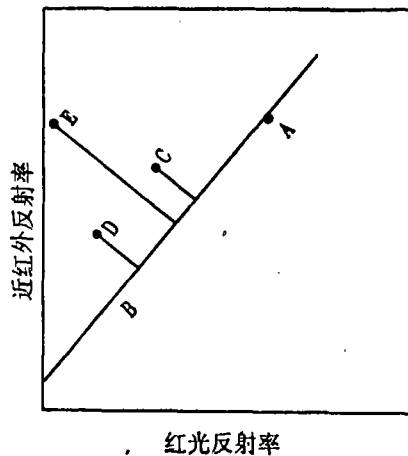


图 2 土壤线和植被

Fig.2 Diagram of the Soil Line and Vegetation

二维空间中,主要是在可见光和近红外波段,植被在土壤中生长时,其可见光(红光)反射率减小,近红外反射率增加,植被点将远离土壤线,而其到土壤线的垂直距离即是 PVI ,它可以排除由于土壤湿度、土壤质地不同而带来的误差。其原理见图 2。

图中 A 、 B 分别代表干土和湿土点, C 、 D 点表示有部分植被覆盖的点,但植被量相同, C 点表示土壤表面干燥的情况, D 点则是土壤表面潮湿的情况。从图中可以看出, C 、 D 点在两波段中的反射率是不同的,但其到土壤线 AB 的垂直距离是相等的。点 E 是植被全部覆盖地面的情况,故其到土壤线的垂直距离最大。这些差别,其它植被指数一般很难能反映出来,因而,从这里我们也可以看出 PVI 的优越性。

3. 灰色预测模型 $G(0, h)$ 介绍

部分信息已知,部分信息未知的系统称为灰色系统。灰色系统有一种新观点,即认为任何随机过程都是在一定幅值范围,一定时区内变化的灰色量,称这随机过程为灰色过程。因而它可以通过原始数据的整理来寻找数的规律,建立预测微分方程。灰色预测模型能克服回归分析存在的缺陷,对样本量的多少没有要求,也不需要典型的分布规律,计算量小,尤其适合我国的产量预测。

灰色模型有 $G(0, h)$ 和 $G(1, h)$ 等。 $G(1, h)$ 为动态模型,它反映了其他 $h-1$ 个变量对某一变量的 1 阶导数的影响,是个动态变化过程。而 $G(0, h)$ 模型则是 $1=0$ 的特殊情况,不考虑变量的导数,是静态模型。根据资料所处年代为 1986—1990,此期间,我国农业政策相对来说比较平稳,基本上是一个静态过程,故选择 $G(0, h)$ 模型作为预测模型。

$G(0, h)$ 模型是一个有 h 变量(包括自变量和因变量)的线性模型。它的建模步骤为:

(1) 原始数据的累加生成

原始数据也许是杂乱无章的,建立灰色模型要求先将原始数列 $x_i^{(0)}$ 进行累加生成,得到生成数据 $x_i^{(1)}$,二者满足下列关系:

$$x_i^{(1)}(k) = \sum_{n=1}^k x_i^{(0)}(n) \quad (1)$$

式中: $x_i^{(0)}(n)$ ——第 i 个原始数据列的第 n 时刻之值, $x_i^{(1)}(k)$ ——第 i 个原始数据列经累加生成后生成数列的第 k 时刻之值。

(2) 模式系数的计算

$G(0, h)$ 模型指 0 阶的 h 变量的模型,是指无动态分量,指系统的静态。由 $G(1, h)$ 推导得到它的一般形式为:

$$x_1^{(1)} = b_1 x_2^{(1)} + b_2 x_3^{(1)} + \dots + b_{h-2} x_{h-1}^{(1)} + a \quad (2)$$

系数矩阵为:

$$\hat{b} = [b_1, b_2, \dots, b_{h-2}, a]^T = (B^T B)^{-1} B^T Y_N(\cdot)$$

其中:

$$B = \begin{bmatrix} x_2^{(1)}(2), \dots, x_{h-1}(1)(2), 1 \\ x_2^{(1)}(3) \dots x_{h-1}(1)(3), 1 \\ \dots \dots \dots \\ x_2^{(1)}(N), \dots, x_{h-1}(1)(N), 1 \end{bmatrix}$$

$$Y_N(\cdot) = [x_1^{(1)}(2), x_1^{(1)}(3), \dots, x_1^{(1)}(N)]^T$$

(3) 数据还原

由于模型是累加生成数列建立起来的,其计算得到预测值也是累加值,因而应用到实际情况时应对数据进行还原,即通过累减生成来得到,令 $x^{(r)}$ 为 r 次生成数列, $x^{(r)}$ 作 i 次累减生成记为 $\alpha^{(i)}$, 则其一般形式为:

$$\begin{aligned} \alpha^{(0)}(x^{(r)}(k)) &= x^{(r)}(k) \\ \alpha^{(1)}(x^{(r)}(k)) &= \alpha^{(0)}(x^{(r)}(k)) - \alpha^{(0)}(x^{(r)}(k-1)) \\ \alpha^{(2)}(x^{(r)}(k)) &= \alpha^{(1)}(x^{(r)}(k)) - \alpha^{(1)}(x^{(r)}(k-1)) \\ &\dots \dots \dots \\ \alpha^{(i)}(x^{(r)}(k)) &= \alpha^{(i-1)}(x^{(r)}(k)) - \alpha^{(i-1)}(x^{(r)}(k-1)) \end{aligned} \tag{3}$$

式中, $\alpha^{(0)}$ 为 0 次累减,即无累减。

4. 阶乘模型

作物从播种到收割是个连续的过程,任何时期作物生长不好都会影响其最终产量,但很难将整个过程统一建立模型。如遥感信息只能观测到抽穗前作物生长的综合反映,只能作长势监测,对后期的气象条件尤其灌浆期的外界因子影响不能反映出来,这样,即便前期长势很好遇后期的不利气象条件作物产量仍就不高;反之,如前期作物生长不好,那么后期环境条件再好也难高产。因而它们不能作为互补因子用在同一灰色模型中,这样就需要逐段订正的模型,再阶乘模型正好能解决上面提出的问题。

阶乘模型的基本方程是:

$$y_i = (y_{i-1}) \cdot \prod_{i=1}^p x_i \tag{4}$$

式中, y_i 是作物某一发育阶段的产量, y_{i-1} 是前一发育阶段预测的产量, x_i 是预测因子。

用植被指数 dVI 作为预测因子,用 $G(0, h)$ 模型得到抽穗前的产量模型为:

$$y_1 = v_1(x_{1i})$$

再将原始产量序列用 y_1 序列除,得到新序列为 y'_i 即:

$$\frac{y_0}{y_1} = y'_i$$

然后,用 y'_i 序列用气象因子且以 $G(0, h)$ 为模型得到第二步预测模型为

$$y_2 = v_2(x_{2i})$$

预测产量: $y = y_1 * y_2$

如果对作物各个发育阶段都建立模型,则可以继续建立第三次,第四次预测模型。模型中预测因子也可以从一个因子到几个因子。而预测的产量即是几次预测值之积。

三、计算结果与分析

1. 顺义县土壤线的计算

根据历年的物候资料我们知道，北京地区小麦返青一般在 3 月下旬，故用顺义 1986 年 3 月 1 日，1989 年 3 月 7 日，1990 年 3 月 15 日 3 天的 NOAA-AVHRR 光谱资料代表裸土的光谱值制作土壤线，顺义县土壤比较类似，好多数据有重复，为简便起见，在 3 天资料中任取 272 点(也有重复)，在第 1 和第 2 通道 CH₁-CH₂ 反射率坐标平面中用统计回归方法得到一条土壤回归线：

$$CH_2 = -1.6354 + 1.2730 CH_1 \tag{5}$$

$$R = 0.9602 \quad N = 272$$

图 3 表示了它们的统计关系。

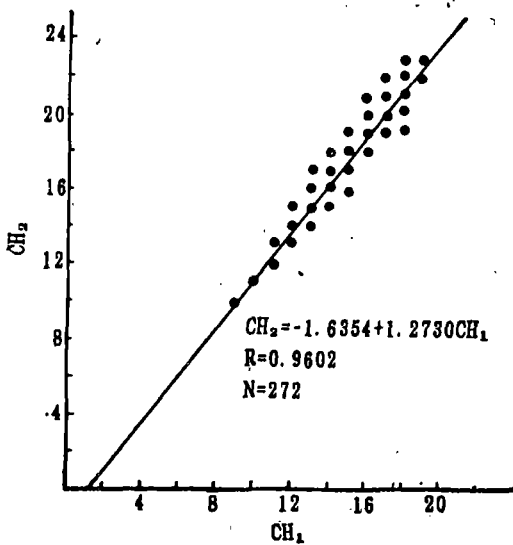


图 3 可见光与近红外波段的土壤线

Fig.3 soil line (for visible and near-infrared wavebands)

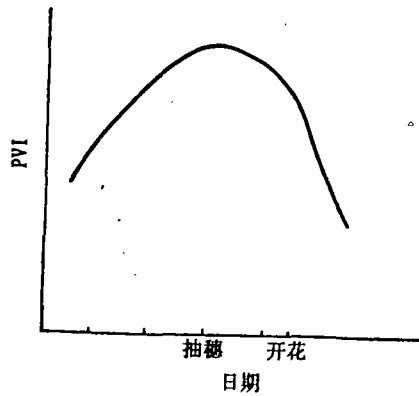


图 4 PVI 随时间变化图

Fig. 4 change of PVI with time

2. PVI 的计算和积温订正

据北京大学遥感技术所在河南省做的冬小麦田间实验知道，返青以后，冬小麦 PVI 值随着作物的增长其值增加，到抽穗期达到峰值，而后到扬花期间稍有减弱，在扬花后迅速减少直到收割止^[8](图 4)。

顺义县平原面积有 145.8 万亩，在 NOAA 卫星上有 1 千多个像元点，逐个像元点建立的模型没有大量对应的产量资料，也难定位，而且工作量大，计算烦。考虑到顺义县以农业为主，主要作物是小麦，全县规模经营，机械化程度高，生产水平比较平均，因而对计算得到各像元点的 PVI 值进行全县平均，用平均值与全县平均产量建立模型。

我们所得到的 NOAA 卫星资料由于各年的天气气候条件、卫星轨道等影响,很难能在同一生育期内得到各年的资料,因而必须对上述计算得到的 PVI 值进行订正,使它们都处在相同的生育期。我们知道,影响作物发育的主要因子是积温,从图 4 中可以看出,冬小麦在拔节到抽穗阶段, PVI 随时间基本上是线性递增,而抽穗至扬花,扬花至成熟,又是线性递减,因而可以通过算出单位积温的 PVI 递增、递减率来订正我们所算出的资料。

根据历年的物候资料统计分析知道,顺义县冬小麦从返青至抽穗一般需 $\geq 0^{\circ}\text{C}$ 的积温为 750°C , 因此,在计算时,认为小麦从返青期向后累加 $\geq 0^{\circ}\text{C}$ 积温达到 750°C 时为抽穗期,如果从返青后到资料日期日 $\geq 0^{\circ}\text{C}$ 的累计积温小于 750°C 则将该日的 PVI 订正增加到 750°C 积温时的 PVI , 反之,如果资料日期日的累计积温大于 750°C , 则用递减率订正到积温 750°C 时的 PVI 。

经过积温订正后我们得到顺义县各年抽穗期的 PVI 值(表 1)。

表 1 顺义县各年抽穗期的 PVI 值
Table 1 PVI of the heading period in Shunyi

年代	1986	1987	1988	1989	1990
PVI	4.3130	5.9908	5.7965	5.0234	6.2365

3. 气象因子的选择

影响小麦后期生长的主要因子是气温和降水,前面已谈到,北京顺义县是规模经营,机械化程度较高,农田以喷灌为主,因而一般情况下不缺水,所以温度是影响小麦的主要因子,因此经过筛选,选择 5 月下旬小麦灌浆时期的平均气温作为影响它的气象因子。

顺义县历年 5 月下旬的平均气温列于表 2。

表 2 顺义县各年 5 月下旬平均气温
Table 2 Average air temperature in last ten days of May in Shunyi

年代	1986	1987	1988	1989	1990
气温($^{\circ}\text{C}$)	23.4	20.7	19.1	23.3	19.7

4. 阶乘模型的建立

顺义县历年的产量资料见表 3

表 3 顺义县各年平均产量
Table 3 Average yield in Shunyi

年 件	1986	1987	1988	1989	1990
产量(斤/亩)	595.1	637.3	674.2	733.9	754

用 $G(0,2)$ 模型建立产量与 PVI 的第一步模型为

$$Y_1(K) = 120.3212PVI(K) - 108.9294 \quad (6)$$

用还原模型计算得到的预测值和相对误差列于表 4

表 4 用 PVI 作为预测因子得到的预测值和误差

Table 4 Predicted yield and relative errors using PVI as a predicting factor

年 代	1986	1987	1988	1989	1990
$y_1(k)$	595.1	1213.3	1957.1	2601.7	3402.0
\hat{y}_1	595.1	618.2	743.8	644.6	800.3
$(\hat{y} - y)/y\%$	0	-3.00	10.33	-12.17	6.14

用原始产量除以前面的预测值得到的新序列与 5 月下旬平均气温建立第二步 $G(0,2)$ 模型为

$$y_2(k) = 0.048196T_5 - 0.099143 \quad (7)$$

同样用还原模型和前面预测模型得到第二步预测值为表 5 表示:

表 5 用 5 月下旬平均气温作为预测因子得到的预测值

Table 5 Corrected value using temperature as a correcting factor

年 代	1986	1987	1988	1989	1990
y/\hat{y}_1	1.0000	1.0310	0.9064	1.1385	0.9422
$y_2(k)$	1.0000	2.0263	2.9468	4.0698	5.0193
\hat{y}_2	1.0000	1.0263	0.9205	1.1230	0.9495

代入阶乘模型得综合模型的产量预测值 $y = \hat{y}_1 * \hat{y}_2$ 和误差列于表 6:

表 6 综合模型产量预测值与误差

Table 6 Final predicted yield and relative errors using a comprehensive model

年 代	1986	1987	1988	1989	1990
y	595.1	637.3	674.2	733.9	754.0
\hat{y}	595.1	634.4	684.7	723.9	759.8
$(\hat{y} - y)/y\%$		-0.45	1.56	-1.37	0.77

对照表 4 和表 6 可以看出,作物产量并不是随着 PVI 的增加而增加,所以如果单纯的以 PVI 作为预测因子而建立模型其误差很大,结果不够理想,但如果以后期的气象因子加以订正,其精度有明显的提高,说明后期气象条件对小麦单产的贡献是不能忽视的,因而在进行单产预测时,必须将后期的气象因子考虑进去。

四、结果与讨论

从上面的计算结果分析表明,把遥感信息与气象条件结合起来建立一个统一的模型

比单纯的遥感模型和气象模型都要优越得多,这与我们的理论分析是相符的,符合模型的平均精度比纯遥感模型提高达 7%,有个别年甚至在 10% 以上。说明外界环境对产量的构成影响还是很大的,在进行遥感作物估产时不能忽视后期气象条件的影响。由于这种模型采用阶乘模型形式,它可以根据作物生长不同时期影响产量的主要因子,随时加入新的参量,从而在实现动态估产中不断提高精度。

由于条件的限制,也由于顺义县所处的特殊环境,我们在选择气象因子方面只选择了 5 月下旬的平均气温作为因子,事实上,影响小麦灌浆到成熟其因子是复杂的,有光、温、水等,因而模型外推到别的地方应增加其它因子。再则,在作物生长期,有时会碰到灾害天气,如冰雹、干热风等,这些特大的灾害天气无法在气象或植被指数上用确切的数字反映出来。由于顺义县这五年中没有大灾害天气,故模型中没有用到,在实际预测时,就有可能出现这种情况,因而必须在模式中加入一个灾害修正因子。

参 考 文 献

- [1] 刘树泽、张宏名、蓝鸿弟,作物产量预报方法,气象出版社,1987。
- [2] 张宏名、王家圣,农作物遥感监测与估产,北京农业大学出版社,1989。
- [3] 赵玉金、井继东,用气象卫星资料试作山东省冬小麦估产预报,遥感信息,1988,3。
- [4] 肖乾广等,用气象卫星数据对冬小麦进行估产的试验,环境遥感, VOL.1, NO.4, 1986 年。
- [5] 京津冀冬小麦遥感估产试验研究论文集,北京科学技术出版社,1987。
- [6] 项月琴、田国良,遥感估算水稻产量——I. 产量与辐射截获量间关系的研究,环境遥感, VOL.3 NO.4, 1988。
- [7] 田国良、项月琴,遥感估算水稻产量——II. 用光谱数据和陆地卫星图像估算水稻产量,环境遥感, VOL.4, NO.1,1989。
- [8] 朱晓红等,冬小麦产量构成分析与遥感估产,环境遥感, VOL.4, NO.2, 1989。
- [9] Kauth, R.J., Lambeck, P.F., Richardson, W., Thomas, G.S., and Pentland A. P., Feature extraction applied to agricultural crop as seen by landsat, Pro. ceedings of the LACIE, Symposium, NASA/ Johnson Space center, Houston, TX, PP. 705—721, 1979。
- [10] R.D. Jackson, Spectral indices in n-space, Remote Sensing of Environment, 13 PP. 409—421, 1983。

A Research on Remote Sensing-Meteorological Model for Wheat Yield Estimation

Li Fuqin Tian Guoliang

(Institute of Remote Sensing Application, Chinese Academy of Sciences)

Abstract

In growing season, from planting to harvest of crop is a complicated process. There are many factors that influence the final yield which is a comprehensive response of various factors, such as environmental conditions, biological factors and agricultural management and technique etc. Thus, remote sensing information or meteorological data does not reflect the conditions of crop growth actually. In this paper, an example in Shunyi County, Beijing, is given, and a winter wheat yield comprehensive model is proposed. The model is combined Perpendicular Vegetation Index (PVI) with meteorological data. A grey-system model $G(0, 2)$ and successive corrected methods were used to establish a remote sensing-meteorological model for estimating winter wheat yield. By using multitemporal NOAA-AVHRR imageries and air temperature, the average accuracy of yield estimation is improved 7% in comprehensive model than in PVI model only.

Key words Remote sensing Winter wheat yield estimation Grey-system model Successive multiplied model