

利用卫星资料和实时天气资料预测玉米产量

M. J. Hayes

(美国国家干旱减灾中心)

W. L. Decker

(密苏里大学土壤和大气科学系)

摘要 在作物收获之前进行大范围作物状况评价,对于提前估测产量是非常关键的。卫星和天气资料为近实时的作物监测提供了条件。研究的目的是利用这些信息建立作物产量业务评估系统,在占全美国玉米产量 60% 的玉米带,对预测 42 个作物报告区的玉米生产进行评估。收集了 1985 ~ 1992 年共 8 年的卫星、气候和农业资料,建立了每个作物报告区标准化玉米产量预测模式,其自变量包括:卫星资料-植被状况指数,气候资料-作物水分指数。该模式大约解释了被观测到的标准化产量变化的 3/4,对其估产准确性和时效性亦进行了检验。将整个玉米带研究区域预测的玉米产量与美国农业部在玉米收获后几个月的统计结果相比较,发现 8 个生长季的预测误差低于 10%。此外,该预测可在玉米收获前 2 个月完成。该系统具有为人类经济活动提供及时的区域或全球农业生产信息的潜力。

关键词 遥感 玉米 区域生产预测 卫星 作物水分指数

引言

粮食生产在国际经济中起着重要的作用。对于任何国家,地区性和世界性商品粮的生产对价格和国际贸易都有很大的影响。在发展中国家,农业减产将产生灾难性的后果,可能导致粮食短缺,人民饿死,发达国家同样受此影响。1988 年,美国政府为遭受严重干旱的农民提供了 40 亿美元的资助。所以,不管一个国家是进口粮食还是出口粮食,在作物收获前 1~2 个月做出产量预测是非常重要的。及时的评价可以帮助国家制定相应的政策,改善作物生长条件。

本世纪 60 年代,随着卫星技术的出现,人们立即认识到卫星资料具有监测植被状况的潜力。80 年代,NOAA 极轨卫星的发射,提供了每天对地球的两次观测。截至目前,已积累了长达 15 年、具有相对连续性和经济性的卫星资料。

根据卫星资料得出的植被指数和专为监测植被设计的植被指数均已建立,而且在表

征区域植被状况和分布方面比从现有天气站点得出的气候变量效果更好(Kogan, 1995)。过去,在统计模型中,根据气候变量预测作物产量取得了一些成功,这些模型一般可解释作物产量变化的 50% 左右。Hayes 和 Decker (1996)最近发现,在美国玉米带,包含卫星资料变量的回归模型同样可解释玉米产量变化的 50%。

本研究的目的是建立一个利用卫星资料和气候资料的作物产量业务评估系统。采用 Kogan 提出的植被状况指数(VCI)和作物水分指数(CMI),及时做出美国玉米带大部分地区的玉米产量评估。一个模型的实用性取决于评估的准确性和及时性,本文对这两个方面均做了研究。作者本意是检验包含气候资料的玉米估产模型较单独使用卫星资料的模型准确性提高的程度(Hayes 和 Decker, 1996)。

美国玉米带的产量几乎占全世界玉米产量的 36%,而 1983 年和 1988 年的干旱对该区玉米生产和全球玉米供求产生了相当大的

影响。在这两年里,玉米总产比 1982 年和 1985 年减少 1 亿吨,产量下降对该区的经济以及国家经济和国际经济均产生了明显的影响。

1 方法

本研究所用资料是美国农业部评估作物状况的资料,将其集成到作物报告区。共选择了 42 个作物报告区,一般这些地区玉米面积超过 15% (图 1, 略)。伊利诺伊、南达科他和俄亥俄州例外,伊利诺伊州西南部的作物报告区(玉米面积少于 15%)包括在内,这样伊利诺伊州作物报告区的完整资料就可应用。由于本研究也包括气候资料,所以与 NOAA 气候分区不完全一致的南达科他州和俄亥俄州中部,虽然其面积超过 15%,但不包括在模式中。

1.1 卫星资料

NOAA 极轨卫星平台装载遥感器之一是甚高分辨率辐射计(AVHRR)。虽然其原设计是为了探测云以及水、雪和地球表面的冰,但后来发现该仪器也有监测植被的特性。此特性来自于植被在 AVHRR 的两个波长通道中具有唯一的光谱特征。1 通道相应于可见光波段(0.58~0.68 μm),典型的正常植被在 1 通道波段具有强吸收特点,所以由 1 通道探测的地面植被反射辐射就非常低。同时,2 通道探测近红外波段辐射(0.72~1.10 μm),正常的植被在这个波段具有强反射性,所以 2 通道对地面植被的响应比较强。两通道合成的植被指数反映了植被的生长状况。本研究应用的归一化植被指数(NDVI),根据下式计算:

$$I_{NDV} = (V_2 - V_1) / (V_2 + V_1) \quad (1)$$

其中 I_{NDV} 为 NDVI, V_2 为第 2 通道值, V_1 为第 1 通道值,使用 1985~1992 年共 8 年的卫星资料,该资料是全球卫星资料的一部分,分辨率为 16 km,称为全球植被指数(GVI),来自 NOAA-9 和 NOAA-11。

Kidwell(1990)对 GVI 作了全面的解释。

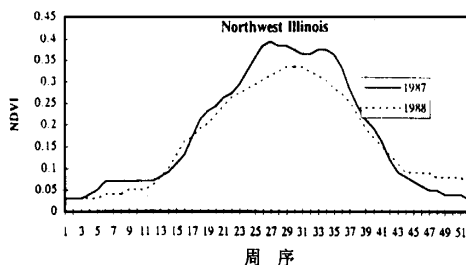


图 2 伊利诺伊州西北部作物报告区 1987 年和 1988 年平滑 NDVI 季廓线

每年对这些资料进行周的合成,采用周合成方法是为了,去掉云的影响。然后把所有周卫星资料集成到作物报告区一层,每个作物报告区包括 22~113 个 GVI 像元,据此确定了 1985~1992 年每个作物报告区每周的 NDVI 平均值。

把资料集成到作物报告区一层,可以减少周 NDVI 值的附加噪音,此种方法已在美国玉米带、全美和苏丹的类似研究中使用。

对本研究中的每一作物报告区,将 1985~1992 年的周 NDVI 值每年按“季节”排列绘出廓线。为了减少由云、水汽和气溶胶波动造成的大气影响,以及太阳-目标物-感应器间的几何位置等引起的噪音对资料的影响,采用 Van Dijk(1986)提出的复合平滑技术对生长季廓线进行了处理。处理后的生长季廓线准确地反映了区域植被变化动态。伊利诺伊州西北部作物报告区的平滑廓线表明了该区典型的植被动态,同时表明了在每个特殊生长季中天气因素的作用(图 2),1988 年伊利诺伊州西北部遭受了严重的干旱,这一结果由该年廓线中较低的 NDVI 值明显地反映出来。

生态因子不同的作物报告区的 NDVI 值不具有可比性,如土壤、气候和管理措施较好,生物量和产量较高的地区不能与生产潜力较低的地区进行比较。为此,Kogan(1990)对 NDVI 进行了修正,即根据区域植被状况对 NDVI 进行标准化,并在 1995 年提出用作物报告区的已知周 VCI 值来衡量天气对植被的影响。

每个作物报告区 8 个生长季的周 VCI 采用下式计算:

$$I_{vc} = \frac{I_{NDV} - I_{min}}{I_{max} - I_{min}} \times 100 \quad (2)$$

式中: I_{max} 是 8 年资料中某周最大的 NDVI, I_{min} 是 8 年资料中相应周中最小的 NDVI.

在本研究中, 生长季中每周 VCI 值其中一年为 100, 一年为 0, 其他 6 年某周的 VCI 值介于 0 和 100 之间. 如果又有新的资料追加, 那么就要建立新的 I_{max} 、 I_{min} , 相应地每周的 VCI 都要重新计算. 最终, 当资料序列包括足够多年的时候, 半永久性的 I_{max} 和 I_{min} 值就确定了. 图 3 给出了伊利诺伊州西北部作物报告区 1987 年(产量相对高的年份)和 1988 年(严重干旱年)的资料及生长季廓线, 并与数据库中所有生长季资料中的最大和最小值进行了比较.

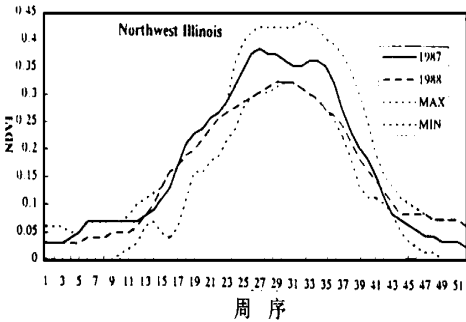


图 3 伊利诺伊州西北部作物报告区 1987 年和 1988 年 NDVI 生长季廓线与 1985~1992 年间最大和最小 NDVI 廓线的比较

1.2 气候资料

本研究中的气候变量选择了作物水分指数(CMI). Palmer(1968)提出用 CMI 监测一周的区域作物状况. 用气温和降水资料计算 CMI, 并考虑前一周的状况. CMI 的效应具有累积性, 能够快速反映降水对土壤水的补充(Hamidi, 1978), 在美国的每一个气候区都计算周的 CMI 值. 收集了 42 个作物报告区的 1985~1992 年每周的 CMI 资料, 由 NOAA 的国家气候资料中心和农业部与 NOAA 的联合农业天气局提供.

如同 VCI 一样, CMI 的权重是由地域和

季节确定的, 因为它是依赖于气温和降水的周标准值, 因此对内布拉斯加州东部和俄亥俄州西部的作物报告区而言, 相似的 CMI 值具有相同的解释, 其与作物和作物发育阶段有关. 由于这是 CMI 和 VCI 的共同特征, 所以在玉米带的区域评价模式中采用这两个变量是可取的.

1.3 农业资料

美国农业部提供了 42 个作物报告区的所有产量和面积的统计资料. 对每个作物报告区的产量进行标准化处理, 即每年的产量与 8 年平均产量之比再乘 100. 标准化产量就是本研究中建立模式的因变量. 进行标准处理是为了从不同的作物报告区获取的产量能在代表大多数玉米区域的同样模式中被检验. 例如, 在伊利诺伊州西北部的作物报告区和内布拉斯加州的作物报告区, 虽然其实际平均产量相差 1800 kg/hm², 但其标准化产量 100% 均表示两个作物报告区的“平均”产量. 每个生长季与玉米发育相关的物候资料来自各州的农业统计办公室. 只有印地安那、明尼苏达和内布拉斯加 3 个州(的办公室)可提供每个作物报告区的完整的物候资料系列, 艾奥瓦州可提供一部分. 而其他州只能提供州范围的物候期资料, 所以伊利诺伊、威斯康星、密歇根、俄亥俄和艾奥瓦州的部分资料是利用州的平均资料和作物报告区周边的物候资料进行估计的. 密苏里州也提供了每一作物报告区的玉米物候资料, 对于估计艾奥瓦和伊利诺伊州的物候资料非常有用. 并整理了每一作物报告区的玉米吐丝、蜡熟的物候期资料, 由于在玉米生殖生长和早期灌浆期内, 天气在最终产量形成中起着重要作用, 因此选择这些发育期的出现时间.

1.4 建立模式

预测作物产量的统计模式通常包括一个或几个反映气象要素和技术变化的变量, 多数模式并不考虑生产区域的自然地理变异. 本研究通过 VCI 和 CMI 预测玉米带 42 个作物报告区的标准化玉米产量时, 考虑了地

理差异和气候差异。与本世纪 60 年代和 70 年代相比, 由于技术进步而增加的玉米产量不多, 加之较短的 8 年研究周期, 所以本模式不包括技术趋势变化。

对于采用 VCI 建立模式, 并在一定时间内做出预测, 进行了详细的调研分析, 计算了生长季内每周的 VCI 与玉米产量的相关性, 最后选择相关系数最高的时段作为 VCI 的代表。加拿大统计局采用此项技术预测加拿大的春小麦和其他 5 种作物的产量。图 4 给出了在伊利诺伊州西北部作物报告区 8 个生长季 r^2 每周的变化情况, 除 31~36 周, VCI 可解释标准化产量变化近 2/3 外, r^2 值都相对小。

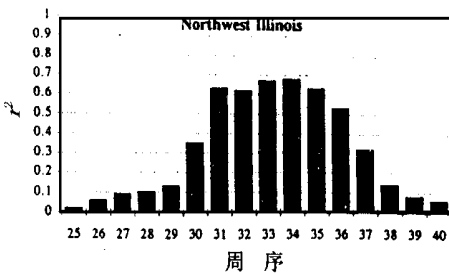


图 4 伊利诺伊州西北部作物报告区 1985~1992 年用周植被状况指数解释的玉米产量变化的相关系数(r^2)

由图 4 可知, 伊利诺伊州西北部与所有 42 个作物报告区相似, 在每一作物报告区 r^2 值的升高与玉米吐丝、蜡熟等主要发育阶段都密切相关。可见对每一作物报告区采用的物候期确定并计算模式变量 VCI 资料的时段是合理的, 所选时段为蜡熟期前后的 7 周。每一作物报告区的生长季平均 VCI 采用蜡熟前后各 3 周的 VCI 计算。

采用相同的方法确定了模式中 CMI 变量的时段, 同样确定为 7 周, 但集中于玉米吐丝期。就 CMI 而言, 其解释玉米标准化产量最大变化的时段较 VCI 的时段约偏早 2~3 周, 由于植被生物量对降水量的响应较快, 因此 CMI 变量的临界期较 VCI 偏早出现是合

理的。

一旦确立了每个作物报告区 8 个生长季的 VCI 和 CMI, 那么模式就可建立。首先从所有资料中取出 1/4 (84 例) 用作模式的检验。理论做法是从数据库中随机选取 2 年的资料建立模式, 然后利用这些资料作模式的独立检验。但是由于总共只有 8 年的资料, 取出 2 年的资料可能导致建模时丢掉一些关键信息。另一种独立检验模式的方法是在模式中用 1993 年和 1994 年的资料, 但遗憾的是由于气候应用分局(CAB)的取消, 该数据库中缺乏 1993 年和 1994 年玉米带的卫星资料。由于上述两种方法都不可取, 因此试用另一种方法对模式的有效性进行检验。该方法不提供完全独立的测试, 通过随机分层从三个地理区域, 即玉米带的东部、中部和西部选取 84 个个例。与此类似, 各年作物报告区的选取也是一致的。

以标准化玉米产量作为因变量的模式是应用其余 252 例资料, 采用一系列统计分析方法建立的, 这些分析方法包括自变量 VCI 和 CMI 的单一或共同的线性和曲线关系。本研究的所有统计分析均采用统计分析系统 (SAS) 程序完成。

2 结果

最终模式选择的变量包括 VCI 和 CMI, 以及二者的相互作用项(VCI/CMI), 几乎解释了玉米产量变化的 75% ($r^2 = 0.73$)。模式如下:

$$\text{标准化产量} = 96.8 + 0.12\text{VCI} + 16.1\text{CMI} - 0.13\text{VCI/CMI} \quad (3)$$

当自变量与因变量相互作用时, 相互作用项是必需的, 当降水量过多并和作物相互

线性模式将导致过高地估计产量, 在这种情况下, 当 CMI 达到表示过湿条件时, 相互作用项的负系数将降低预测的玉米产量。

为了“检验”模式的有效性, 采用作物报告区的 VCI 和 CMI 代入模式预测作物报告

区的玉米标准化产量。并计算了实际产量和预测标准化产量的差异,以确认模式是否合适。表 1 给出了残差结果。理想模式的残差平均应为 0,本模式 84 例的残差是 1.23%。表 1 说明,预测产量与实际产量相差在 5% 以内的占 32% (27 例),在 10% 以内的占 74% (62 例)。在残差大于 5% 的例子中,模式预测低于实际产量的为 29 例,高于实际产量的为 28 例。

表 1 模式检验: 预测与实测标准化产量残差分布

残差 %	个例(共 84 个)	百分比
0.0~ 4.9	27	32.1
5.0~ 9.9	35	41.7
10.0~ 14.9	13	15.5
15.0~ 19.9	7	8.3
20.0~ 24.9	1	1.2
≥ 25.0	1	1.2

玉米产量估计

作物估产业务模式的成功与否取决于产量估测的有效性和准确性。在本研究中,应用方程(3)对整个研究区域的每个生长季作了玉米产量预测,共 336 例。8 个生长季的预测结果与美国农业部对该区玉米的最终估计进行了比较。

由于模式中使用的 VCI 变量的计算需要 7 周的卫星资料,其中包括蜡熟后 3 周,所以对某一生长季而言,预测的有效性依赖于所有作物报告区玉米蜡熟期的出现时间。此外,为了做出估测,处理资料也需要时间。在本研究中,配有计算机软硬件,该资料的处理大约在收集到最后一周卫星资料后的一周内完成。因此,估计大区的玉米生产,在作物报告区玉米达到蜡熟期后,大约需要 4 周时间。

表 2 给出在玉米带本研究区 8 个不同生长季内业务预测系统有效的估产日期。其中 7 个生长季的预测可在 10 月 1 日之前完成,只有 1992 年较晚,在 10 月 4 日,且不包括密歇根的中南部作物报告区,那里由于夏季低温而使玉米发育期延迟。正常情况下玉米带

的玉米收获期是从 10 月到 11 月,所以这一预测可在收获前的 1~ 2 个月内完成,为了解该区玉米的可能产量提供了重要信息。

玉米产量预测的有效性不仅取决于预测的时效性,而且取决于预测的准确性。本预测结果与美国农业部对该区生产的统计结果作了比较,图 5(略)给出了美国农业部统计的和预测的玉米总产量。从中可知,与统计结果相比:模式的预测结果 4 年偏高(1986,1988,1990 和 1991 年),4 年偏低(1985,1987,1989 和 1992 年)。

表 2 玉米估产模式的时效性

年代	有效估产日期	年代	有效估产日期
1985	9 月 30 日	1986	9 月 23 日
1987	9 月 16 日	1988	9 月 18 日
1989	9 月 17 日	1990	9 月 30 日
1991	9 月 15 日	1992	10 月 4 日*

* 不包括密歇根州中南部作物报告区

表 3 预测玉米带产量与美国农业部最终产量误差

年代	误差(10 ⁷ t)	误差百分比
1985	- 2.57	- 1.9
1986	+ 1.51	+ 1.2
1987	- 5.63	- 5.2
1988	+ 0.74	+ 1.1
1989	- 4.34	- 3.7
1990	+ 6.76	+ 5.4
1991	+ 4.15	+ 3.7
1992	- 8.99	- 6.0

表 3 给出了模式预测值与实际值的距平 and 距平百分比。在 8 年中,预测值与实际值相差在 6% 以内,模式预测低于最终产量的最大误差是 1992 年,为- 6% (875× 10⁴ t)。在严重干旱的 1988 年,虽然该年的玉米产量较次低年(1987)减少 40%,但模式预测误差仅为 1.1%。因此,模式能够反映该地区严重干旱的影响,其在 1988 年提供的信息具有较高的可靠性。

3 讨论

研究结果表明:应用卫星资料和天气气候资料建立的玉米产量预测业务模式,可在玉米收获前1~2月提供有益的信息。在作物估产模式中,增加了气候资料的模式较单一的卫星资料模式或气候资料模式效果更好。

虽然基于上述技术建立的业务模式不可能收集当年生长季的资料并做出产量预测,但在每一个生长季结束后这些资料是可以加入模式的。随着资料年限的增加,模式将趋于稳定,每增加一个生长季资料将带来很小的变化。同时利用本研究技术可建立用于其他

地区预测其他作物产量的模式。

此外,由于卫星资料接收和处理技术的不断改进,数据库存取更为方便,区域作物预测模式将更有价值。新建的探路者数据系列可提供以10天为周期的卫星资料,其像元为 8 km^2 。随着探路者数据系列的累积,本研究所介绍方法的应用和未来的研究方向均表明,卫星资料将成为大范围植被状况监测和潜在产量水平预测的一个有力手段。

李春强译自 *Int. J. Biometeorology*,
1998, Vol 42, 10~15, 李郁竹校