

多元回归积分及其在农业气候分析中的应用

许尊伍¹ 魏淑秋² 程林仙¹

(¹ 陕西省气象科学研究所, 西安 710015)

(² 中国农业大学, 北京 100094)

摘要: 本文根据对一元回归积分原理进一步理解的基础上, 结合大量的实践与试算, 提出了多元回归积分的概念和计算方法。利用上方法计算了汉中地区农科所的水稻产量及光温资料, 所得光温回归积分函数曲线与汉中盆地水稻生产实际相吻合。文章最后在综合国内外生产潜力方法的同时, 给出了利用光温回归积分函数曲线与各时段的偏差计算光温生产潜力的公式。

关键词: 回归积分; 光温生产潜力; 正交函数

中图分类号: S165.27; S162.3

1 问题的提出

1924年英国学者 R. A. Fisher 在研究罗萨姆斯特德(Rothamster)试验站历年冬小麦产量与降水量的关系时, 提出了回归积分(regression integral)的概念^[1]。这一方法以后在许多国家得到应用。J. W. Hopkins 研究加拿大西部降水量与温度对春小麦的影响以及松田昭美在研究日本福冈地区日照时数与水稻产量关系时, 都曾应用过这个方法, 并得出了很好的结果^[2,3]。在国内, 1963年卢其尧在《华北平原降水量对冬小麦产量的影响》一文中, 首先使用了这个方法。70~80年代初, 在农业气候分析及区划中已有广泛采用^[4,5]。1982年许尊伍在《陕西汉中盆地水稻冷害初探》一文中, 对这个方法的原理进一步做了详细论述。综上所述, 截止目前, 回归积分还仅限于单因子与作物产量形成的关系, 在农业气候分析和实际问题中, 往往有一定的局限性。我们在对这一方法原理的进一步理解基础上, 结合大量试验, 提出了多元回归积分的概念和计算方法。本文仅作为一个尝试, 可望从研究方法的角度做进一步的探讨。

2 原理概述

如果将作物的生育期分为 ρ 个小时段, 按照多元回归积分的概念(以二元为例), 作物的产量可用下式来表达^[1,5]:

$$\begin{aligned} \hat{Y}_c &= c + \int_0^{\rho} a_j(t) Q_j(t) dt \\ &= c + \int_0^{\rho} a_j(t) T_j(t) S_j(t) dt \end{aligned} \quad (1)$$

式中 \hat{Y}_c 是作物气候产量的估值, c 为待定常数, $a_j(t)$ 为两因子(如温度与日照时数)对于产量同时影响的时间函数, 它表示温度变化 1°C , 日照变化 1h , 作物产量的平均增量, $Q_j(t)$ 是 $t+dt$ 时段内温度值 T_p 与日照时数 S_p 对应之积的时间序列:

$$Q(t) = \begin{bmatrix} T_1 & S_1 \\ T_2 & S_2 \\ \dots & \dots \\ T_p & S_p \end{bmatrix}$$

由于 $a_j(t)$ 通常是一个相互独立、变化缓慢的时间函数, $Q_j(t)$ 为温度与日照时数对应之积的时间函数, 因而总可以用一个有限个项数 K 的正交标准多项式进行逼近:

$$a_j(t) = a_0\psi_{j0} + a_1\psi_{j1} + \dots + a_k\psi_{jk} \quad (2)$$

$$Q_j(t) = \rho_0\psi_{j0} + \rho_1\psi_{j1} + \dots + \rho_k\psi_{jk} \quad (k=K-1) \quad (3)$$

式中 ψ_{jk} 为时间的正交标准函数; a_0, a_1, \dots, a_k 为偏回归系数; $\rho_0, \rho_1, \dots, \rho_k$ 为温度和日照时数的复分配系数, 可依据下式:

$$\rho_k = \int_0^p \psi_{jk} T_j(t) S_j(t) dt \quad (k=0, 1, \dots) \quad (4)$$

求出。将公式(2)代入(1)得:

$$\begin{aligned} \hat{Y}_c &= c + \int_0^p (a_0\psi_{j0} + a_1\psi_{j1} + \dots + a_k\psi_{jk}) T_j(t) dt \\ &= c + a_0 \int_0^p \psi_{j0} T_j(t) S_j(t) dt + a_1 \int_0^p \psi_{j1} T_j(t) S_j(t) dt \\ &\quad + \dots + a_k \int_0^p \psi_{jk} T_j(t) S_j(t) dt \end{aligned} \quad (5)$$

由于求取 ψ_{jk} 较麻烦, 为简化运算, 用时间的正交非标准函数 ξ_{jk} 代替 ψ_{jk} , 利用公式:

$$\rho'_k = \int_0^p \xi_{jk} T_j(t) S_j(t) dt \quad (6)$$

将公式(3)变为^[3-6]:

$$Q_j(t) = \frac{\rho'_0 \xi'_{j0}}{\sum (\xi'^2_{j0})} + \frac{\rho'_1 \xi'_{j1}}{\sum (\xi'^2_{j1})} + \dots + \frac{\rho'_k \xi'_{jk}}{\sum (\xi'^2_{jk})} \quad (7)$$

$$\text{由于: } a_k = \frac{a'_k}{\sum (\xi'^2_{jk})}$$

$$\text{则: } a_k \psi_{jk} = a'_k \sqrt{\sum (\xi'^2_{jk})} \cdot \frac{\xi'_{jk}}{\sqrt{\sum (\xi'^2_{jk})}} = a'_k \xi'_{jk} \quad (8)$$

将公式(8)代入(5)得:

$$\begin{aligned} \hat{Y}_c &= c + a'_0 \int_0^p \xi'_{j0} T_j(t) S_j(t) dt + a'_1 \int_0^p \xi'_{j1} T_j(t) S_j(t) dt \\ &\quad + \dots + a'_k \int_0^p \xi'_{jk} T_j(t) S_j(t) dt \end{aligned} \quad (9)$$

将公式(6)和(8)分别代入公式(9)和(2)得:

$$\hat{Y}_c = c + a'_0 \rho'_0 + a'_1 \rho'_1 + \dots + a'_k \rho'_k \quad (10)$$

$$\text{和 } a_j(t) = a'_0 \xi'_{j0} + a'_1 \xi'_{j1} + \dots + a'_k \xi'_{jk} \quad (11)$$

这样,就把时间的正交标准函数化为正交非标准函数。由于 ξ_{jk} 可在正交多项式表中查到,使计算大大减化,于是求 $a_j(t)$ 函数的问题就归结为求解多元回归方程的问题了。

3 计算实例

利用汉中地区农科所 1960~1979 年水稻气候产量和 5 月 21 日~9 月 27 日每 10 d 平均气温和平均日照时数资料,将水稻生育期分为 13 个小时段,计算得 a'_0, a'_1, \dots, a'_6 , 经回归效果检验 ($F=5.00, F(5.14)=2.96$ 和系数检验 (ρ'_5 和 ρ'_6 被剔除) 得:

$$\hat{Y}_c = 395.1522 + 2.5629 \rho'_0 + 0.7933 \rho'_1 + 0.1480 \rho'_2 - 0.0586 \rho'_3 - 0.0236 \rho'_4 \quad (12)$$

利用公式(12)可求出水稻产量的估值。将气候产量的实际值与估值点绘如图 1。

由公式(13)得到图 2。

将 $a'_k (k=0, 1, \dots, 4)$ 之值代入(11)式得:

$$a_j(t) = 2.5629 \xi'_{j0} + 0.7933 \xi'_{j1} + 0.1480 \xi'_{j2} - 0.0586 \xi'_{j3} - 0.0236 \xi'_{j4} \quad (13)$$

图 2 即为温度、日照时数对于水稻产量形成同时影响的时间函数曲线。它是光温双因子与水稻产量形成的动态模式,表示水稻生育期各阶段温度变化 1°C 、日照时数变化 1h ,影响产量形成的数值。从图 2 明显看出, $a_j(t)$ 函数有 2 个峰点和 1 个谷点。6 月上旬和 9 月上旬产量的平均增量最大。前者正值水稻插秧初期,晴朗的天气有利于缓秧和壮秧。此时,10 d 的平均气温上升 1°C ,日照时数增加 1h ,产量平均增加 $45\text{ kg}\cdot\text{hm}^{-2}$ 左右。后者正值水稻盛花期至灌浆阶段,是水稻全生育期对光温变化最敏感的时段,10 d 平均气温上升 1°C ,日照时数增加 1h ,产量平均增加 $150\text{ kg}\cdot\text{hm}^{-2}$ 左右。而 7 月中旬是全年气温最高、日照时间较长且强度也是较强的阶段,水稻正处在幼穗分化阶段,光温偏高对穗分化有抑制作用,10 d 平均气温上升 1°C ,日照时数增加 1h ,每公顷水稻平均减少 60 kg 左右。就生物学意义而言,是符合汉中盆地水稻生产实际的。

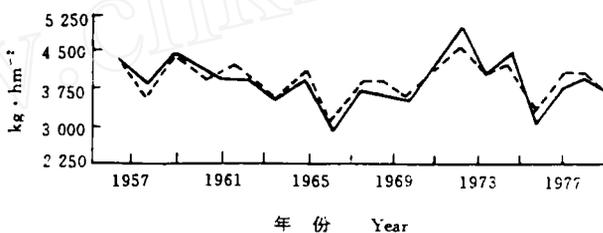


图 1 历年产量的拟合曲线

Fig. 1 Curve fitting of crop yield

----计算产量 Computational yield;
——实际产量 Practical yield

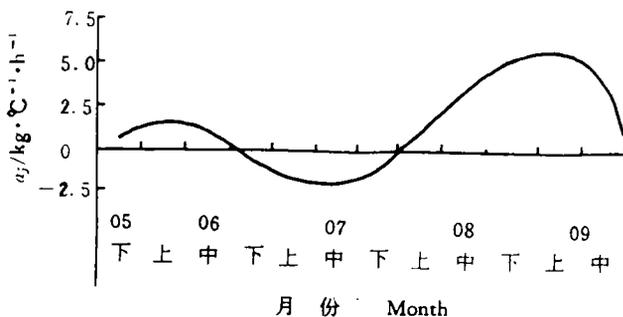


图 2 $a_j(t)$ 函数分布曲线

Fig. 2 Functional curve of $a_j(t)$

4 多元回归积分方法求算水稻最大光温生产潜力

一元回归积分从60年代初至今已有近30年的历史。它之所以能够被广泛采用,是因为其应用价值和特点。其对于产量形成和某一因子分布的动态关系是有意义的,但对于研究一个地区气候条件对某一作物的栽培以至产量形成的综合影响,就有一定的局限性,这就要借助于多元回归积分方法。因此,本文提出多元回归积分方法,并以求算汉中盆地水稻最大光温生产潜力为例。

4.1 平均生产潜力

$$\hat{Y} = c + at + \sum_{j=1}^p a_j(t) \Delta T_j \Delta S_j \quad (14)$$

式中 \hat{Y} 为平均生产潜力, at 为趋势项, $\Delta T_j, \Delta S_j$ 分别为不同时段温度、日照时数的平均偏差:

$$\Delta T_j = \frac{1}{N} \sum_{i=1}^N |T_{ij} - \bar{T}_j|$$

$$\Delta S_j = \frac{1}{N} \sum_{i=1}^N |S_{ij} - \bar{S}_j|$$

其中 N 为资料年数。

根据图2或公式(13)与(14)算得的平均光温生产潜力为 $7000 \text{ kg} \cdot \text{hm}^{-2}$ 。

4.2 最大生产潜力

$$\hat{Y}_{\max} = c + at + \sum_{j=1}^p a_j(t) \Delta T'_j \Delta S'_j \quad (15)$$

其中 $\Delta T'_j$ 和 $\Delta S'_j$ 分别表示不同阶段温度和日照时数的最大偏差,其值是从历史资料中按时段挑取最大值,与累年平均值求差。由于各阶段两个因子的极值往往不出现在同一年,因此只限于挑取一个因子的极值,而另一个因子的值则取对应年份之值。为了确定挑取哪个因子的极值,可由与产量较大的相关系数来定。由公式(13)和(15)计算出最大光温生产潜力为 $8000 \text{ kg} \cdot \text{hm}^{-2}$ 。

5 讨论

目前,国内外关于光温生产潜力的计算方法很多。最常用的有:利用总辐射量、最大光能利用率以及光合有效辐射、反射、漏射、光饱和点、量子效率等有关的系数之乘积来求算^[5];布德科用温度订正系数订正光合生产力^[6];黑岩(S. Duroiwa)根据光强、日照长度和时间、温度以及群体结构(叶的透光度、消光系数、累积叶面积系数等)计算光合生产力^[5],等等。总括以上各类光温生产潜力的计算方法,均涉及较多的系数测定。而在目前仪器设备的条件下,是很难测定准确的。如光强的水平、垂直分布、透光率、漏光率等。而叶面积系数的测定又耗资费力;现有的系数又有地区、季节和生长状况的局限性。有的生产力计算还要进行温度、叶面积系数等订正。因此,以上方法仍属理论计算,一般来说,和实际有相当的差距,难以实现。我们认为,利用多元回归积分曲线,结合各时段出现的平均偏差与最大偏差可

求出平均生产潜力与最大生产潜力。由于农业气象资料均取自所讨论地区, 求算结果比较接近实际, 也易于实现。当然针对不同品种、不同要求可选择不同要素进行计算。本文仅以光温生产潜力为例。

参 考 文 献

- 1 Fisher R. A. The influence of rainfall on the yield of wheat at Rothamsted. Roy Soc T [London] Phil Trans Ser B, 213, 89~142
- 2 Hopkins J W. Weather and wheat yield in Western Canada; I. Influence of rainfall and temperature during the growing season on plot yield. Canad J Res, 1935, (12)
- 3 松田昭美. 暖地(における)水稻栽培と气象要因の统计学研究(1). 农业气象, 1959, 15(1):15~20
- 4 魏淑秋. 栾城冬小麦产量与气候条件的统计学分析. 作物学报, 1980, 6(3):161~170
- 5 王世普等. 北京地区冬小麦产量和气象要素的统计学分析. 中国农业科学, 1979, (1):10~18
- 6 许尊伍. 陕西汉中盆地水稻冷害初探. 气象学报, 1982, 40(1):89~95
- 7 北京农业大学农业气象专业编. 农业气象学. 北京: 科学出版社, 1984, 175~192

The Multiple Regressive Integration and Its Application in Analysis of Agricultural Climate

Xu Zunwu¹ Wei Shuqiu² Cheng Linxian¹

(¹Shaanxi Institute of Meteorology, Shaanxi 710015)

(²China Agricultural University, Beijing 100094)

Abstract: Based on further comprehension of single regressive integration and a large number of calculations, the concept and computation method of multiple regressive integration were proposed. Regressive integration curve based on thermal and solar light data and rice yield from Han-Zhong Institute Agricultural Science is coincident with the actual rice production of Hanzhong basin. Finally, we present the formula for potential productivity analysis using the photo-thermal regressive integration curve and its difference between each interval while synthesizing various computing methods of the potential productivity.

Key words: regressive integration; light and thermal potential productivity; orthogonal functions