

江苏省单季晚稻产量预报的分段加权动态模式

汤志成 高苹

(江苏省气象科学研究所)

提 要

本文利用灰色系统GM(1, 1)模型逐段滑动平均模拟趋势产量 y_t , 根据单季晚稻的穗数、粒数和粒重对最终产量的贡献确定出生育三个阶段的权重, 算出各阶段的气象产量。通过MAICE逐步回归方法分别建立各个生育阶段末气候产量的预测模式, 分段加权动态地进行单季晚稻的产量预测。最后还作了1986—1988年产量的试报。

一、引言

单季晚稻的大多数品种对温、光反应都很敏感, 属于感温性、感光性均强的类型(1)。在我省一般5月份播种, 10月份收获, 总生育期较长。大致可将其全生育期划分为三个阶段: 穗数决定阶段(即营养生长期), 大致为5月上旬—6月下旬; 粒数决定阶段(即营养生长、生殖生长并进期),

为7月上旬—8月下旬; 粒重决定阶段(即生殖生长阶段), 为9月上旬—10月下旬。而这三个阶段的生长状况对最终产量的形成贡献也各不相同。

为了动态地按作物生育阶段逐段进行产量预测, 通常采用阶乘模型或累加模型(2)。这两种模型均是以订正残差逐段逼近的原理来提高模拟精度的, 虽然其拟合效果较好, 但预报结果并不一定理想, 并不能

真正体现各生育阶段对产量的贡献。本文试图考虑作物三个生育阶段的生长状况对产量的不同贡献，分段建立动态的产量预报模式，作出单季晚稻的产量预报。

本文将实际产量（全省平均亩产） y 分为依社会生产水平的变化而变化的趋势产量 y_i 和随历年气象环境条件而变化的气候产量 y_w 及随机误差产量 ε 。 ε 在一般情况下，可以忽略不计。其比值模型为

$$y = y_i \cdot y_w \quad (1)$$

二、趋势产量的模拟

根据省统计局提供的1964年以来我省单季晚稻产量资料，作出历年产量分布曲线（附图）。由图可见，产量波动较大。尤其是从1982年以来，产量猛增，难以用一种简单的函数模拟趋势产量 y_i 。故本文先将产量序列逐步滑动分段，对每段使用灰色系统GM(1,1)模型模拟，然后再对产量序列相应的各年之多个模拟值进行平均，以模拟趋势产量和进行趋势产量的预测。这个方法，我们简称为灰色系统GM(1,1)模型逐段滑动平均。即设有一产量原始序列

$$y^{(0)} = \{y^{(0)}(j)\} \\ (j=1, 2, \dots, n)$$

按步长 q ($q < n$) 进行逐步滑动分段，则每段 (i) 原始序列为

$$y^{(0)}(t) = \{y^{(0)}(i), y^{(0)}(i+1), \dots, y^{(0)}(i+q-1)\} \\ \begin{matrix} (i=1, 2, \dots, n-q+1) \\ (t=i, i+1, \dots, i+q-1) \end{matrix}$$

对每段 (i) 的数据序列根据GM(1,1)模型的数学方法，作累加生成序列

$$y^{(1)}(t) = \sum_{m=i}^t y^{(0)}(m) \quad (2)$$

建立白化形式的微分方程

$$\frac{dy^{(1)}}{dt} + a_i y^{(1)} = u_i \quad (3)$$

用最小二乘法求解参数 a_i 、 u_i ，再求取时间

响应函数

$$y^{(1)}(t+1) = (y^{(0)}(i) - (u_i/a_i) e^{-a_i t} + u_i/a_i) \quad (4)$$

就可得到 $\hat{y}^{(1)}(t+1)$ ，经累减生成

$$\hat{y}^{(0)}(t+1) = \hat{y}^{(1)}(t+1) - \hat{y}^{(1)}(t) \quad (5)$$

获取原始序列的拟合值 $\hat{y}^{(0)}(t+1)$ ，即为经过GM(1,1)模型处理后第 i 段的数据序列。

显然对于 $t+1$ 时刻有 P_{t+1} 个 $\hat{y}^{(0)}$

$(t+1)$ 值。则 $t+1$ 时刻的 $\hat{y}^{(0)}(t+1)$ 平均数为

$$\hat{y}_t = 1/P_{t+1} \cdot \sum_{i=1}^{P_{t+1}} \hat{y}^{(0)}(t+1) \\ \begin{matrix} (t=0, 1, 2, \dots, n-1) \\ (i=1, 2, 3, \dots, n-q+1) \end{matrix} \quad (6)$$

其中

$$P_{t+1} = \begin{matrix} t+1 & t+1 < q \\ q & q \leq t+1 \leq n-q+1 \\ n-(t+1)-1 & t+1 > n-q+1 \end{matrix}$$

根据(6)式便可算得历年趋势产量拟合值，作出趋势产量曲线(图1)。

当 $i=n-q+1$ 时，由(4)式即可算出 $y^{(1)}(n+1)$ ，再经累减生成可得到

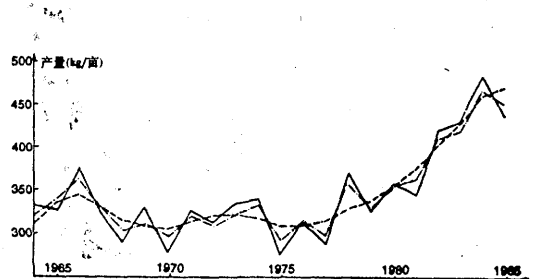


图1 历年单季晚稻产量和趋势产量及拟合值的变化曲线

社会产量，-----趋势产量，.....产量拟合值

$y^{(0)}(n+1)$ ，这便是趋势产量预报值。

本文比较了滑动步长 q 的变化，选择 $q=7$ ，即滑动步长为7年所分离的趋势产量比较好。如附图中虚线即为趋势产量的变化曲线，与我省历年单季晚稻产量年景基本一致。

三、气候产量与气象因子的相关分析

1. 气候产量的获取

根据多年农气观测资料，用产量结构三要素的穗数、粒数和粒重与产量作相关分析，用以反映单季晚稻全生育期中三个主要生育阶段的生长状况对最终产量的贡献。

经计算，各阶段的权重值分别为

$$\lambda_1 = 0.3662, \lambda_2 = 0.3900, \lambda_3 = 0.2438.$$

由(6)式计算出历年趋势产量 y_t ，再由(1)式变换

$$y_w = y/y_t \times 100\%$$

计算出 y_w ，那么便可得到历年的 y_{w1} 、 y_{w2} 和 y_{w3} 。

2. 相关普查

选择全省6个代表站1964—1985年5—10月的旬平均温度、旬平均最高气温、旬平均最低气温、旬降水量、旬降水日数、旬日照时数，以及它们的组合光温积，降水强度(降水量/降水日数)和温度旬较差，经过膨化处理⁽²⁾，按上述三个时段与各自相关的气候产量 y_w 作相关普查。除进行直线相关外，还同时作了平方、指数、对数及导数等多种曲线变形处理，寻求最佳相关关系。

3. 因子的稳定性检验

为了避免因子的相关关系的不稳定，本文采用滑动相关⁽³⁾对因子进行稳定性检验。然后，根据滑动相关系数的变化特征，可按下列三个指标⁽⁴⁾筛去一些滑动相关系数序列 r_k 波动较大或近期有下降的不稳定因子：

(1) 平均滑动相关系数 $\bar{r}_{q,k} \geq 0.40$;

(2) 滑动相关系数的均方差 $\sigma_r < 0.2$ 或 < 0.15 ;

(3) 近期三个滑动相关系数 r_{n-q-1} 、 r_{n-q} 、 $r_{n-q+1} \geq 0.35$ 。

对于20年左右样本的因子，根据我们的工作经验，滑动步长一般以10年(即 $q=10$)左右为佳。太大则难以通过检验，选不到因子；过小则易于通过，失去检验的意义。

四、气候产量模式的建立

常用的逐步回归方法所建立的回归模型往往随给定的 F 检验水平的不同而不同。因此，本文采用比较客观的MAICE逐步回归分析⁽⁵⁾方法建模。所谓MAICE逐步回归就是用一个极大似然函数定义的赤池信息量准则 $AIC = -2\log L(\hat{\theta}, x) + 2p$ 代替一般逐步回归的 F 检验。从一组可供选择的模型中选出 AIC 最小(即MAICE)的模型就是最佳模型。即

(1) 引入某一个未进入方程的变量后的所有回归模型中，引入该变量后的回归方程的 AIC 值最小的变量；

(2) 在某一个已进入方程的变量后的所有回归模型中，剔除该变量后的回归方程的 AIC 值最小的变量。

根据以上选出的比较稳定而相关又好的因子，应用MAICE回归分析方法建立了三个阶段的回归方程如下：

$$y_{w1} = -0.0801x_1 - 0.5707x_2 + 41.2235 \\ F = 16.9563 \quad R^{***} = 0.8006$$

$$y_{w2} = 4.7385x_3 + 18.9297x_4 + 36.6325 \\ F = 6.9670 \quad R^{**} = 0.6504$$

$$y_{w3} = -0.0002x_5 - 0.0107x_6 \\ + 25.8289$$

$$F = 6.9437 \quad R^{**} = 0.6498$$

式中 x_1 为5月上旬降水日数的平方； x_2 为6月下旬降水日数； x_3 为8月上旬降水量； x_4 为7月上旬降水强度； x_5 为9月中旬降水量；

x_0 为9月中旬—10月中旬降水日数的累加值的平方。

上述方程的因子意义比较明确，基本反映了影响江苏省单季晚稻产量的主要因子。5月上旬正值单季晚稻播种期间，在此期间，如果天晴少雨，就不会发生烂种烂秧现象，是单季晚稻培育壮秧的良好条件。6月下旬—7月上旬正是单季晚稻分蘖期间，如梅雨长而明显，则不利于光合物质的积累，影响分蘖，从而导致亩有效穗数不足而影响产量。8月初是单季晚稻幼穗分化期间，雨水少，日照充足，有利于单季晚稻由营养生长向生殖生长的过渡，可形成大穗；9月中旬—10月中旬处于单季晚稻的生育后期，晴天少雨多日照，则有利于晚稻的充分灌浆，增加千粒重。

五、动态预测及试报

1. 动态预测

根据上述生育期分段，可以动态的在每个生育阶段末进行产量预测，即：

第一阶段（穗数形成期）末的预报方程为

$$\hat{y} = \hat{y}_1 \cdot \hat{y}_w = \hat{y}_1 \cdot (\hat{y}_{w1} / \lambda_1) \quad (7)$$

第二阶段（粒数形成期）末的预报方程为

$$\hat{y} = \hat{y}_1 \cdot (\hat{y}_{w1} + \hat{y}_{w2}) / (\lambda_1 + \lambda_2) \quad (8)$$

第三阶段（粒重形成期）末预报方程为

$$\hat{y} = \hat{y}_1 \cdot (\hat{y}_{w1} + \hat{y}_{w2} + \hat{y}_{w3}) \quad (9)$$

于是我们就可在6月下旬、8月下旬及10月下旬分别逐段地动态地进行单季晚稻产量预测。

2. 预报方程拟合效果和近3年的试报利用以上所建立之模式作出历年产量的

拟合值（如图1之点虚线），拟合效果较佳，趋势一致，其平均拟合误差仅稍高于3%（表1）。

另外，用上述模式试作了1986年、1987年和1988年的3年单季晚稻的产量预报。结果也较为满意。相对误差在±3%以内，完全可以投入业务使用。

表1 回归方程的平均拟合误差(%)

$$\left(\frac{1}{N} \sum_{i=1}^N |\hat{y}_i - y_i| / y_i \right)$$

第一阶段末	第二阶段末	第三阶段末
3.32	3.25	3.09

表2 近三年单季晚稻产量预测模式试报一览表

项目 年代	实际值 y(公斤/亩)	预测值 \hat{y} (公斤/亩)	相对误差 $(y - \hat{y})/y \times 100\%$
1986年	475	472.5	0.51
1987年	468	473.6	-1.20
1988年	476	488.0	-2.52

六、结论

1. 用GM(1, 1)模型逐步滑动平均方法作趋势产量模拟和预测是可取的，其计算结果与当地各个时段的生产力水平基本符合，趋势一致。

2. MAICE逐步回归分析方法建立的方程，复相关系数R的值均达到0.01信度的显著性水平，近三年的预测效果也较好。用AIC代替F值进行检验是成功的。

3. 所建方程中的因子都是几个关键期的重要因子，说明分段加权模式吸取了作物的生育特点，既能够确实地反映出各阶段的生长状况对最终产量形成的贡献大小，又避免了如累加和阶乘模式中第一阶段所占的

权重特别大,后面各阶段的权重太小的弊病,所以它的各阶段的拟合误差比较稳定。

4. 该模式属动态统计模式,难以避免小概率事件的发生。

参考文献

(1)南京农学院等,全国高等农林院试用教材,作物栽培学(南方本),上海科学技术出版社,1979年。

(2)汤志成等,江苏省双季早稻产量预报的累加型模式,江苏农业学报,1986年第3期。

(3)朱盛明,相关系数稳定性分析方法及其应用,气象学报,1982年第4期。

(4)汤志成等,海温、大气环流在产量预报中的应用,气象科学,1988年第3期。

(5)毛宗英,应用MAICE的逐步回归分析,数学的实践与认识,1987年第4期。

The piecewise weighing dynamic model for yield prediction of the middle-season rice in Jiangsu

Tang Zhicheng Gao Ping

(The Meteorological Research Institute of Jiangsu Province)

Abstract

Using the stagewise moving average of the Grey System model GM (1,1), the trend yield Y_t is simulated. According to the number of the ears and grain as well as the weight of a grain, especially their contribution to the yield, the weight coefficients of three growing stages are developed and the climate-related yield of each stage is also found out. The predicting models for climate-related yield each stage are established with MAICE successive regression method. And the rice yield is predicted dynamically by piecewise weighting. Finally, the yields of 1986—1988 are estimated.