

# 用模糊均生函数作区域夏季旱涝预测

俞炳启 胡洛林

(江苏省气象台, 南京 210008)

## 提 要

用模糊均生函数(FMGF)建模方案建立宁镇(南京、镇江)地区夏季旱涝趋势预测模型,其拟合效果,特别是对大旱大涝的拟合效果十分理想。1994年用该模型预测宁镇地区梅雨量,结果与实况较为接近。

关键词: 模糊均生函数 旱涝 预测

## 引 言

1991年梅雨期,江苏淮河以南频繁出现暴雨和大暴雨,形成百年未遇的大范围洪涝灾害。1994年夏季,又遇罕见的严重干旱,6—8月的总雨量比常年偏少4—8成,梅雨量仅20—45mm,比常年偏少8—9成。4年内二次严重的旱涝气象灾害,给江苏省的工农业生产 and 人民生活带来巨大损失。尤其是南京、镇江等地区,洪涝、干旱的影响更为严重。本文用南京、镇江地区(以下简称宁镇地区)逐年梅雨量,分析近40年来该地区夏季旱涝趋势,并用对极值有较好拟合和预测效果的模糊均生函数模型作旱涝预测,试报结果令人满意。

## 1 宁镇地区夏季旱涝概况

宁镇地区夏季的主要降水是梅雨,梅雨量的多寡直接影响夏季旱涝。我们以宁镇地区1954—1993年逐年梅雨量分析该地区夏季旱涝概况。

1954—1993年宁镇地区梅雨量的平均值为210mm,最多636mm,出现在1954年。最少仅8mm,出现在1978年。极差628mm。在划分旱涝等级时,涝的各个级别跨度取100mm,正常及早的各个级别跨度取50mm<sup>[1]</sup>。40年来旱涝情况见表1。

考虑到目前宁镇地区已具备一定抗旱排涝能力,偏旱和偏涝都已不会形成严重灾害,在讨论旱涝情况时,可把偏旱和偏涝视为正常。

从表1可以看到:

表1 1954—1993年宁镇地区旱涝级别及频数

级别	梅雨量/mm	频数	百分比/%
大旱	≤85	5	30
旱	86—135	7	
偏旱	136—185	10	50
正常	186—235	6	
偏涝	236—335	4	20
涝	336—435	2	
大涝	≥436	6	

(1)40年中,正常和偏旱或偏涝的年份占50%,旱年和大旱年占30%,涝年和大涝年占20%。也就是说,有一半的年份非涝即旱。

(2)大旱年和大涝年出现的频数分别为5和6,总的几率为27.5%。这说明,在宁镇地区夏季,平均每隔不到4年就会出现大旱或大涝灾害。

## 2 模糊均生函数模型

曹鸿兴等拓广了数理统计中算术平均值的概念,定义了时间序列的均值生成函数(mean generation function,缩写为MGF),简称均生函数,以建立具有多步预测能力的数学模型<sup>[2]</sup>。并进一步将均生函数的概念推广到模糊集中,定义了模糊均生函数(FMGF),给出相应的建模方案<sup>[3]</sup>,使模型的拟合效果,尤其是对原序列极值的拟合效果更为理想。因此,比较适用于预测干旱、洪涝等异常现象。

设一等间隔时间序列

$$X(t) = \{X(1), X(2), \dots, X(N)\} \quad (1)$$

式中  $N$  为样本量。定义式(1)的均生函数

$$\bar{X}_l(i) = \frac{1}{n_l} \sum_{j=0}^{n_l-1} X(i+j) \quad (2)$$

$i = 1, 2, \dots, l; 1 \leq l \leq M; n_l = \text{INT}(N/l)$   
 $M = \text{INT}(N/2)$  或  $\text{INT}(N/3)$ ;  $\text{INT}$  表示取整。对  $\bar{X}_l(i)$  作周期性延拓, 则可得到式(2)的外延序列

$$f_l(i) = \bar{X}_l\left(i - l \text{INT}\left(\frac{i-1}{l}\right)\right) \quad (3)$$

把上述均生函数的概念推广到模糊集中。

设论域  $U = \{u_i | i = 1, 2, \dots, N\}$ , 在  $U$  上构造模糊子集  $A$  和  $B$ 。定义具有周期性, 且随起报时刻的远近以指数形式下降的隶属函数

$$\mu_A(i) = \begin{cases} re^{-\beta(N-i)} \sin \frac{2\pi}{l}(N-i) & i < N \\ 1 & i \geq N \end{cases} \quad (4)$$

式中,  $r$  为由经验或试算确定的常数,  $l$  为周期长度,  $\beta$  按对过去观测值重视程度事先给定。定义  $\mu_B(i) = X(i)/X_{\max}$ ,  $X_{\max} = \max X(i)$ 。

用模糊向量的内积定义模糊均生函数

$$\bar{X}_l(i) = \frac{c}{n_l} \sum_{j=0}^{n_l-1} \mu_A(i+jl) \cdot \mu_B(i+jl)$$

式中  $c$  为给定常数, 使  $\bar{X}_l(i)$  与原序列量级相同。若取  $c = X_{\max}$ , 则上式变化为

$$\bar{X}_l(i) = \frac{1}{n_l} \sum_{j=0}^{n_l-1} \mu_A(i+jl) X(i+jl) \quad (5)$$

$i = 1, 2, \dots, l$

用式(3)作周期性延拓, 则可得到式(5)的外延序列。在实际计算中, 可对原序列作一、二阶差分运算, 得到一阶和二阶差分序列。对原序列和一阶、二阶差分序列分别用式(5)求出模糊均生函数, 并用式(3)作周期性延拓, 得到3组外延序列  $f_l^{(0)}(t)$ 、 $f_l^{(1)}(t)$ 、 $f_l^{(2)}(t)$ 。

另外构造一组累加延拓序列

$$f_l^{(3)} = X^{(0)}(1) + \sum_{i=1}^{l-1} f_l^{(1)}(i+1)$$

$t = 2, 3, \dots, N, \quad l = 1, 2, \dots, M$

这样, 共得到  $4 \times M$  个外延序列。

用兼顾数量评分和趋势评分的双评分准则, 筛选  $4 \times M$  个外延序列。若引入  $k_0$  个外延序列, 则模糊均生函数预测模型为

$$\hat{X}(t) = a_0 + \sum_{i=1}^{k_0} a_i f_i(t)$$

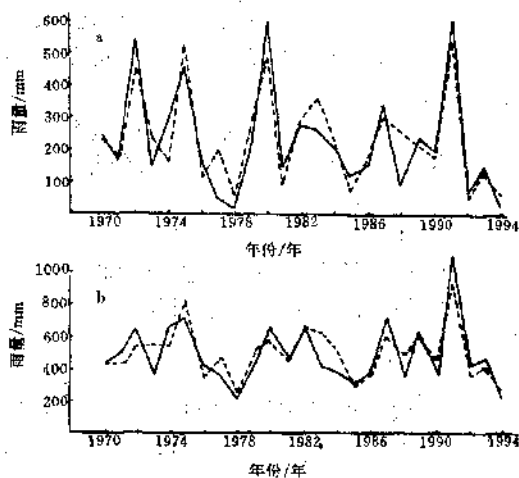
式中,  $a_0$  为常数项,  $a_1, a_2, \dots, a_{k_0}$  为回归系数。

### 3 宁镇地区夏季旱涝预测

用宁镇地区 1954—1993 年逐年梅雨量建立基于模糊均生函数的预测模型。经试算, 在该序列中, 用样本量  $N = 24$  (即用 1970—1993 年梅雨量所建序列), 周期最大长度  $M = \text{INT}(N/3) = 8, \beta = 0.01$  时, 所建模型对原序列的拟合效果最为理想。用上述方法建立的预测模型如下:

$$\begin{aligned} \hat{X}(t) = & 242.71 - 1.24f_7^{(0)}(t) \\ & + 1.95f_7^{(1)}(t) + 3.50f_7^{(2)}(t) \\ & - 2.43f_8^{(3)}(t) - 0.14f_8^{(3)}(t) \end{aligned}$$

式中,  $f_7^{(0)}(t)$  为原序列周期长度为 7 年的模糊均生函数外延序列。  $f_7^{(1)}(t)$  为一阶差分序列、周期长度为 7 年的模糊均生函数外延序列。  $f_8^{(3)}(t)$  表示周期长度为 8 年的模糊均生函数累加延拓序列。其余项, 依此类推。预测模型的拟合值与实况值曲线如附图 a 所示。预测模型的拟合级别与实况级别见表 2。



附图 a: 历年梅雨量、模型拟合值、预测值(1994年)  
 b: 历年6—8月总雨量、模型拟合值、预测值(1994年)  
 实线: 实况值 虚线: 拟合值

表2 模型拟合、预测级别与实况对比

年份	1970	71	72	73	74	75	76	77	78	79	80	81	82
实况	偏涝	偏旱	大涝	偏旱	偏涝	大涝	偏旱	大旱	大旱	正常	大涝	旱	偏涝
拟合	正常	偏旱	大涝	偏涝	偏旱	大涝	旱	正常	大旱	偏涝	大涝	大旱	偏涝
年份	1983	84	85	86	87	88	89	90	91	92	93	1994年	
实况	偏涝	正常	偏旱	偏旱	涝	旱	偏涝	偏旱	大涝	大旱	偏旱	实况	大旱
拟合	涝	正常	大旱	偏旱	偏涝	偏涝	正常	正常	大涝	大旱	偏旱	预测	大旱

从附图 a 和表 2 中可以看到:

(1) 该模糊均生函数模型的拟合值曲线与实况值曲线十分接近。拟合级别与实况级别完全一致的有 11 年; 拟合与实况相差一个级别的有 8 年。两者共占 79.2%。

(2) 如前所述, 由于宁镇地区已具备一定的抗旱排涝能力, 故若把偏旱年和偏涝年均视为正常年, 则除 1977 年和 1988 年外, 其余 22 年的旱涝趋势拟合与实况完全一致, 拟合准确率达 91.7%。

(3) 该模型对极值的拟合较好。在参加统计的 24 年中, 共出现 1972、1975、1980 和 1991 年 4 个大涝年, 模型的拟合与实况级别完全一致, 无一空漏现象。在原序列中所出现的 3 个大旱年中, 除 1977 年拟合在正常级别, 属漏报外, 1978 年和 1992 年拟合和实况十分一致。

用上述模型对 1994 年梅雨量作试报, 计算结果为 83.4mm, 实况为 30mm。可见试报结果令人满意。

我们用同样的方法, 得到宁镇地区夏季 (6—8 月) 总雨量的模糊均生函数预测模型:

$$\hat{X}(t) = -272.76 + 0.91f_7^{(3)}(t) + 0.98f_8^{(0)}(t)$$

其拟合和预测效果也较为理想。预测模型的拟合曲线与实况值曲线见附图 b。

#### 4 讨论

4.1 由于引起气候变化的影响因素错综复杂, 而气候资料的时间序列却反映了曾经发生过的所有因果与结构关系的影响。因此, 利用时间序列自身演变规律进行气候预测已被广泛采用<sup>[4]</sup>。该方法就是通过寻找历年梅雨量的自身变化规律, 再用外延序列对梅雨量作出预测。

4.2 该方法既考虑了观测值随起报时刻远近效用逐渐下降, 又考虑了时间序列的周期性变化。经试算, 拟合效果, 特别是对极值量级的拟合效果较好, 但雨量值的拟合有一定差距。例如, 1991 年宁镇地区梅雨量为 618mm, 是原序列中雨量最多的一年, 模型拟合值为 523mm, 是拟合序列中数值最大的, 实况与拟合均属大涝年, 但拟合值误差较大, 达 95mm。

4.3 本文中所用样本长度及其它参数是通过试算确定的。如何选择最佳样本长度及其它参数, 使拟合和预测效果最好, 有待进一步探讨。

#### 参考文献

- 1 吴震, 陈家富. 江浦县夏季干旱的分析和预测. 江苏气象, 1994(5).
- 2 曹鸿兴, 魏凤英. 基于均值生成函数的时间序列分析. 数值计算与计算机应用, 1991, 12(12).
- 3 魏凤英, 曹鸿兴. 模糊均生函数及其应用. 气象, 1993, 19(2), 7—11.
- 4 周家斌. 一种新的时间序列预报方法. 科学通报, 1983, 28.

## Regional Prediction of Summer Floods/Droughts with Fuzzy Mean Generating Function Model

Yu Bingqi Hu Luolin

(Jiangsu Provincial Meteorological Service, Nanjing 210008)

#### Abstract

A prediction model for summer floods/droughts in the Nanjing-Zhengjiang region by fuzzy mean generating function is developed. The model is satisfactory, especially in cases of severe floods/droughts. The results of the precipitation prediction in this region during the Meiyu period in 1994 is close to actual precipitation.

**Key Words:** fuzzy mean generating function floods/droughts prediction