

# 均生函数预报模型的改进实验

唐毓勇<sup>1</sup> 蒋国兴<sup>2</sup> 周礼杏<sup>2</sup> 黄淑娟<sup>1</sup>

(1. 广西百色市气象局, 533000; 2. 广西气象局)

**提 要:** 为了提高对百色市主汛期月降雨量的预报准确率, 基于均生函数预报方法, 应用百色市 1951—2006 年 6、7、8 月份的月降雨资料, 通过“短序列建模”和“残差订正”方法进行改进, 得到原序列、短序列预报模型及其相应的残差订正预报模型。各改进模型均使预报精度和准确率得到了不同程度的提高, 尤其是 MGFAFM<sub>s</sub> 模型的预报  $T_s$  比 MGFFM 模型提高了 25%。研究表明: 在月雨量预报业务实践中采取均生函数 4 模型集合预报方式将是一种很好的选择。

**关键词:** 均生函数 短序列建模 残差订正 集合预报

## Experimental Investigation on the Improved Prediction Models with Average-Generation Functions

Tang Yuyong<sup>1</sup> Jiang Guoxing<sup>2</sup> Zhou Lixing<sup>2</sup> Huang Shujuan<sup>1</sup>

(1. Baishe Meteorological Office, Guangxi 533000; 2. Guangxi Autonomous Region Meteorological Bureau)

**Abstract:** In order to increase the forecasting accuracy, techniques of short-series modeling and residual correction are improved and the prediction models of original series, short-series and corresponding residual correction are developed on the basis of prediction technique of averaged-generation functions. The used data are the precipitation data of Baishe City from June to August in 1951-2006. The results show that the forecasting accuracy and precise of these improved models have raise to some extent, especially the  $T_s$  of MGFAFM model is 25% higher than MGFFM. The operational use of a consensus forecast with four models to the monthly rainfall amount is a good choice.

**Key Words:** average-generation function short series modeling residual correction consensus forecast

### 引 言

广西百色市的年降水量主要集中在每年的 6—8 月份, 而且年际变化很大。本文的目

的在于寻求一种比较稳定可靠的月雨量客观预报方法。

在现有的时间序列预测模型中, 如自回归(AR)、自回归滑动(ARMA)和门限自回归(TAR)模型中, 在制作多步预测时, 预测值

会趋于平均值,且往往对极值的拟合效果欠佳。指数平滑模型和灰色模型等可以制作多步预测,但它们表示的是一种指数增长,对于呈起伏型变化的气候序列不适用。目前短期气候预测方法大多是以统计方法加经验模型为主,由于气候因子周期性变化的不确定性,使我们对于转折性变化的预测能力较差,依据气候时间序列蕴涵不同时间尺度振荡的特征,魏凤英<sup>[1]</sup>等人拓展了数理统计中算术平均值的概念,定义了时间序列的均生函数,提出了视均生函数为原序列生成的、体现各种长度周期性的基函数的新构思。均生函数预测模型既可以作多步预测,又可以较好地预测极值,为长期预报和短期气候预测开辟了一条新途径。均生函数预报模型已在气象预测中得到了广泛的应用<sup>[2-4]</sup>。

在应用均生函数原序列模型进行百色市降水集中期月雨量预报的业务试验中发现:最近4年的平均预报准确率只有67%。因此,本文拟对原预报模型进行适当改造以求进一步提高预报准确率。

## 1 预报模型建模

### 1.1 均生函数预报模型

均生函数模型基于系统状态前后记忆的基本思想,构造一组周期函数,通过建立原时间序列与这组函数间的回归,建立预测方程。本文采用的资料为1951—2002年百色市6、7、8月的月雨量实况资料按月分别建立预报模型(简称“原序列模型”:MGFFM),具体建模过程可参考文献<sup>[1]</sup>。

### 1.2 均生函数短序列预报模型

通过对各月降水量原序列进行突变性检验,可考察分析降水的气候变化特征。截取发生气候突变后不少于20年的实况序列进行建模(即“短序列模型”:MGFFMs),用于考察分析气候态(或气候阶段)对预报模型预

报能力的影响。

用Mann-Kendall方法,可以明确突变开始的时间,指出突变区域。为检验突变点的真假性,用Yamamoto法进行检验。定义信噪比 $B_{SN}$ :

$$B_{SN} = \frac{|\bar{X}_1 - \bar{X}_2|}{S_1 + S_2} \quad (1)$$

其中 $\bar{X}_1(S_1)$ 和 $\bar{X}_2(S_2)$ 分别是突变点两端的2个时期的平均值(均方根差),在 $B_{SN} \geq 1$ 时,认为检测的点为气候真突变点。

给定显著性水平 $\alpha = 0.05$ ,即 $\mu_{0.05} = \pm 1.96$ ,对1951—2002年百色市6、7、8月月雨量序列进行突变性检验。计算结果绘成图1。

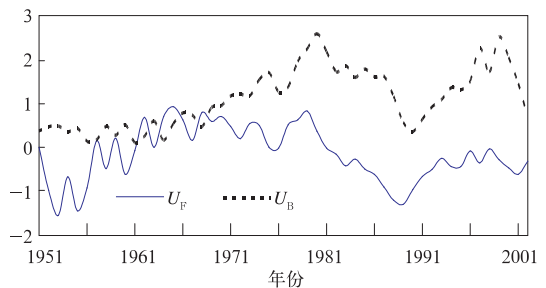


图1 1951—2002年百色6月月降雨量M-K统计量曲线

在6月M-K曲线中(图1),1962—1969年间有6个交叉点,各交叉点位于信度线之间,说明这一时段百色6月份降水处于一种显著的高频振动,其中1968年的 $B_{SN} = 1.02$ ,说明从1968年开始百色6月份降水过渡到另一个平均状态,具有下降趋势,但不显著。7月M-K曲线(图略)只有一个交叉点即1957年,其 $B_{SN} = 1.23$ ,但 $U_F$ 线都在信度直线之间,说明其存在突变,但不显著。8月的M-K曲线(图略)有2个交叉点:1954、2001年,两年的 $B_{SN}$ 值分别为0.87、0.62,均小于1.0,但在1973—1981年间 $U_F$ 曲线有8年突破信度直线,降水表现为显著减少趋势。

通过突变性检验,6月份降水的短序列模型选用1969—2002年月雨量进行建模;7月份选用1958—2002年月雨量进行建模;8月份选用1982—2002年月雨量进行建模。

### 1.3 均生函数残差订正模型

由均生函数原序列模型(或短序列模型)可得到拟合(预报)数据序列:

$$X_1 = (x_1(1), x_1(2), \dots, x_1(N)), \quad (2)$$

定义实测值  $X$  的残差为:

$$q(t) = x(t) - x_1(t) \quad (3)$$

用式(2)得到拟合残差数据序列  $q(t) = (q(1), q(2), \dots, q(N))$ , 然后利用均生函数方法对残差序列建立新的模型, 称此模型为残差模型  $q(t)$ 。由均生函数预报模型与残差模型构成新的预测模型, 称为均生函数残差订正预报模型(MGF AFM 或 MGF AFMs), 即:

$$x_2(t) = x_1(t) + q(t) \quad (4)$$

## 2 预报模型检验

用 2003—2005 年的月雨量对上述模型进行独立样本预报试验, 并应用于 2006 年百色市 6、7、8 月的月降水量预报业务实践。

采用距平百分率的平均绝对误差( $S_s$ )和距平倾向准确率( $T_s$ )检验预报模型的拟合预报水平:

$$S_s(i) = \frac{100}{N} \sum \left| \frac{x_i(t) - x(t)}{\bar{x}} \right| \quad (5)$$

$i = 1, 2; t = 1, 2, \dots, N;$   
 $\bar{x}$  = 气候平均值

$$T_s(i) = \frac{100}{N} \sum [IF((\frac{x_i(t) - \bar{x}}{\bar{x}} \times \frac{x(t) - \bar{x}}{\bar{x}}) > 0, 1, 0)] \quad (6)$$

分别统计均生函数原系列(1951—2002 年)预报模型(MGFFM)、短序列预报模型(MGFFMs)及其残差订正预报模型(MGF AFM 和 MGF AFMs)的输出数据(见表 1 和表 2)。

从表 1 来看, 4 个预报模型对各月降水的拟合预报能力差异比较明显。总体而言,

残差订正预报模型可以使拟合  $S_s$  明显减小; 短序列模型的拟合  $S_s$  最大, 但预报  $S_s$  较小; 经过残差订正的短序列预报模型(MGF AFMs)的预报  $S_s$  最小。

表 1 均生函数预报模型拟合预报距平的平均绝对误差统计表

	拟合				预报			
	6月	7月	8月	模型平均	6月	7月	8月	模型平均
MGFFM	16%	27%	18%	20%	68%	44%	39%	50%
MGFAFM	12%	16%	8%	12%	63%	35%	47%	48%
MGFFMs	24%	21%	27%	24%	62%	28%	31%	40%
MGFAFMs	22%	14%	17%	18%	54%	28%	9%	30%
月平均	19%	20%	18%		62%	34%	31%	

预报: 2003—2006 年

表 2 均生函数预报模型拟合预报距平倾向准确率统计表

	拟合				预报			
	6月	7月	8月	模型平均	6月	7月	8月	模型平均
MGFFM	90%	84%	67%	80%	50%	75%	75%	67%
MGFAFM	94%	90%	80%	88%	100%	75%	75%	83%
MGFFMs	76%	78%	76%	77%	75%	75%	100%	83%
MGFAFMs	82%	87%	86%	85%	100%	75%	100%	92%
月平均	86%	84%	77%		81%	75%	88%	

预报: 2003—2006 年

从表 2 来看也有类似的特征, 4 个预报模型对各月降水的拟合预报能力可以出现比较明显差异。总体而言, 残差订正预报模型可以使拟合  $T_s$  明显提高; 残差订正短序列模型对拟合  $T_s$  提高的幅度最大(比原模型提高了 9%), 预报  $T_s$  也最高(高达 92%); 原序列预报模型(MGFFM)的预报  $T_s$  最小。

从表 1 和表 2 还可以发现: 6 月和 8 月的多模型同月平均拟合  $S_s$  差异很小, 但是他们的同月平均预报  $S_s$  却相差 1 倍; 6 月和 8 月的多模型同月平均拟合  $T_s$  差异最大, 然而他们的同月平均预报  $T_s$  却相差较小。由此可见, 并非拟合效果越好, 预报能力就越大。

上述特点也可以从图 2 中得以充分的体现。

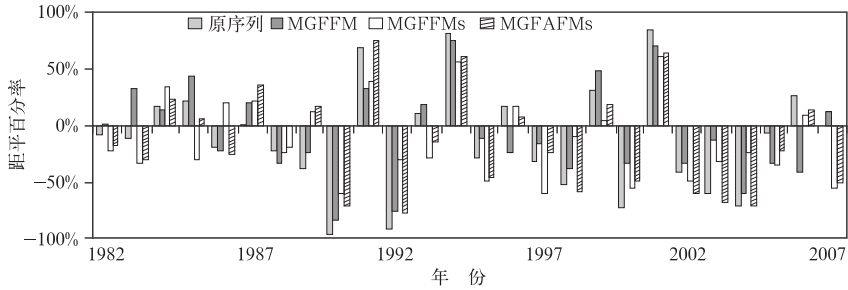


图 2 百色 8 月份雨量均生函数预报模型拟合预报图

虽然 MGFAFMs 模型的预报  $T_s$  比 MGFFM 模型提高了 25% 之多;但是,采用单一模型将无法避免未来气候突变所带来的

风险。因此,在业务实践中采取“多模型集合预报”的方式也许是一种更好的选择。表 3 给出了 4 模型集合预报的实验结果:

表 3 均生函数预报模型集合预报表

	2003 年			2004 年			2005 年			2006 年		
	6 月	7 月	8 月	6 月	7 月	8 月	6 月	7 月	8 月	6 月	7 月	8 月
MGFFM	1	0	1	0	1	1	0	1	1	1	1	0
MGFAFM	1	1	1	1	0	1	1	1	1	1	1	0
MGFFMs	0	1	1	1	1	1	1	0	1	1	1	1
MGFAFMs	1	0	1	1	1	1	1	1	1	1	1	1
集合预报	1	—	1	1	1	1	1	1	1	1	1	—

注:表中“1”表示预报正确,“0”表示预报错误

由表 3 可见,即使按照简单多数的原则,集合预报的成功率也可以达到 92% (11/12)。

### 3 结 论

均生函数预报模型、尤其是经过残差订正的短序列预报模型,对百色降水集中期的月雨量预报表现出很强的预测能力。

在月雨量预报业务实践中采取均生函数

4 模型集合预报方式将是一种很好的选择。

### 参考文献

- [1] 魏凤英. 现代气候统计诊断与预测技术[M]. 北京:气象出版社,1999:128-134.
- [2] 俞炳启,胡洛林. 用模糊均生函数作区域夏季旱涝预测[J]. 气象,1995,21(12):32-34.
- [3] 张德宽,杨贤为,邹旭恺. 均生函数—最优子集在高温极值预测中的应用[J]. 气象,2003,29(4):44-47.
- [4] 钟利华. 均生函数模型在广西 1998 年春播低温阴雨结束期预报中的应用[J]. 广西气象,1998,19(4):12-15.