

提高冬春月降水量预测质量方法的探讨

李承宏 周佑云

(湖南省郴州市气象台, 423000)

提 要

月降水量预测是短期气候预测的业务之一,提高预测质量对防灾抗灾具有重要意义。运用数理统计原理,在因子选择标准和在各种资料中提取因子方面作了些探索,月降水量预测质量有了较大提高。

关键词: 月降水量 回归 滑动相关系数 因子选择

引 言

目前,数理统计仍然是短期气候预测的重要方法之一。前年,我们编写了一套气候预测软件,用于月降水量预测的有:逐步回归、最优子集回归、在各种历史资料库中挑选预测降水量因子的计算程序和气候资料库。气候资料有:历年北半球 500hPa 高度、北半球海平面气压、太平洋海温、74 种环流特征量(资料来源:湖南省气象台)、郴州站各气象要素的旬、月平均值(资料来源:郴州气象观测站)。经过不断的改进和完善,特别是对预测因子的精细选择,预测质量有了较大提高,取得了较为满意的效果。

1 选择因子的方法

在预报实践中发现,回归计算,常常拟合效果好,预测效果不稳定。有时复相关系数达到 0.9 以上,预测效果仍然不好。怎么样

才能拟合、预测效果都好呢?近年来,用最优子集回归取代逐步回归,预测效果有一定的提高,这已得到共识^[1,2]。从理论上讲,选择天气学意义好的因子,是解决这个问题的根本途径。但是,天气过程十分复杂,短期气候过程就更为复杂,从前期出现的气候资料中选出的因子,很难说出其天气、气候学意义的好坏。经过两年探索,我们有了从另外角度考虑的选择因子的方法。

1.1 单个因子的选择

设有预测量 Y_i , 因子 X_i , $i = 1, 2, \dots, n$; 设 t 为 $1 \sim n$ 之间的量。求 Y_{i1} 与 X_{i1} ($i1 = 1, 2, 3 \dots t$) 的相关系数 R_1 , 求 Y_{i2} 与 X_{i2} ($i2 = 2, 3, 4 \dots t + 1$) 的相关系数 $R_2 \dots$ 求 $Y_{i(n-t+1)}$ 与 $X_{i(n-t+1)}$ ($i(n-t+1) = t, t + 1, t + 2 \dots n$) 的相关系数 R_{n-t+1} , 这 $n - t + 1$

个相关系数叫滑动相关系数, t 为滑动长度^[3], 滑动步长为 1。如有 50 年资料, 一般可取滑动长度为 30 年; 若要更精细分析, 则取滑动长度为 20、10、5 年。

滑动相关系数一般是不相等的, 有以下三种型式:

(1)前后符号一致, (2)前后符号相反, (3)前后符号交错。对第(2)、(3)种绝对不能选取。

第(1)种, 其绝对值的大小又有以下 6 种亚型: 前后大小基本稳定、前小后大、前大后小、两头大中间小、两头小中间大、前后波动。

首选的是前后大小基本稳定型。当然, 滑动相关系数不可能完全一样大小, 但绝对值相差不超过 0.1, 可以认为基本稳定。

次选的是前小后大型。可以在作回归计算时, 只取后部若干年的资料, 而将前部一些年的资料舍弃。这样做减少了样本容量, 但如果某个因子后部滑动相关系数绝对值特别大, 还是可取的。

反复计算发现, 这样选择因子, 预测效果有较大提高。另外提及的是, 备选因子的相关系数要能通过显著性检验。

1.2 因子群体的选择

进行回归计算时, 参加计算回归系数的

那些样本回代出的是拟合值, 而没有参加计算回归系数的那些样本算出的才是预测值。如果将预测年之前的一些年的样本不参加计算回归系数, 作出这些年的预测, 再与实况对照就能看出这个因子群体预测效果。

例如要计算 2003 年的某个预测值, 可以将 1954~1996 年的资料计算, 得出 1997 年的预测值; 将 1954~1997 年的资料计算, 得出 1998 年的预测值……将 1954~2001 年的资料计算, 得出 2002 年的预测值, 再与实况对照。满意, 则用此因子群体; 不满意, 则再挑选另外的一个或几个因子加入。逐步回归、最优子集回归都会自动选择另外的因子组合, 这时又有了另外的上述几年的预测值……反复这样的过程, 直至满意为止。表 1 是郴州 1997 年 1~4 月至 2003 年 1~4 月降水量预测与实况的对比。表中括号内是按湖南省气象台定义的月降水量 6 个量级。降水量距平百分率 $p \leq -50\%$ 为特少、 $-50\% < p \leq -20\%$ 为偏少、 $-20\% < p < 0$ 为略少、 $0 \leq p < 20\%$ 为略多、 $20\% \leq p < 50\%$ 为偏多、 $p \geq 50\%$ 为特多。郴州 1~5 月降水量 1971~2000 年平均值分别为: 79.8、115.1、147.5、194.8、206.0mm。

表 1 郴州 1~4 月降水量预测与实况的对比

(单位: mm)

月	年						
	1997	1998	1999	2000	2001	2002	2003
1 月	预测 91.3(略多)	132.8(特多)	74.9(略少)	67.0(略少)	85.4(略多)	71.3(略少)	75.9(略少)
	实况 117.5(偏多)	136.3(特多)	45.9(偏少)	45.1(偏少)	116.6(偏多)	63.7(略少)	78.2(略少)
2 月	预测 107.7(略少)	119.9(略多)	52.1(特少)	64.5(偏少)	107.8(略少)	119.8(略多)	51.2(特少)
	实况 145.4(偏多)	147.2(偏多)	15.4(特少)	66.8(偏少)	92.2(偏少)	81.6(偏少)	58.0(特少)
3 月	预测 223.5(特多)	244.0(特多)	130.7(略少)	163.3(略多)	120.6(略多)	75.4(特少)	111.8(偏少)
	实况 235.1(特多)	224.6(特多)	53.5(特少)	178.7(偏多)	121.6(略多)	92.6(偏少)	111.4(偏少)
4 月	预测 247.0(偏多)	86.9(特少)	164.2(略少)	208.3(略多)	317.6(特多)	205.4(略多)	108.4(偏少)
	实况 262.5(偏多)	74.9(特少)	214.0(略多)	256.7(偏多)	265.4(偏多)	251.8(偏多)	119.4(偏少)
5 月	预测 189.0(略少)	188.6(略少)	250.7(偏多)	188.8(略少)	111.2(偏少)	194.2(略少)	232.2(略多)
	实况 183.2(略少)	208.4(略多)	236.8(略多)	160.1(偏少)	114.5(偏少)	120.6(偏少)	197.3(略少)

从表中可见, 2003 年 1~4 月降水量预测与实况都在同一个量级, 特别是 3 月降水量预测与实况仅差 0.4mm。前几年的检验性预测, 就正负距平趋势而言, 除 1997 年 2 月、1998 年 5 月、1999 年 4 月、2002 年 2 月

趋势报错外, 其余都对, 准确率 86%。不少特多、特少的都在同一个量级, 同一级的占 43%。

2 从各种资料中挑选因子

2.1 从北半球 500hPa 高度、北半球海平面

气压、太平洋海温中挑选因子

这些资料是网格点数据。先求出前期某月每个网格点数值与某月降水量的相关系数,选取相关系数符号相同、绝对值大且能通过0.01显著性检验、相互靠近的点4~5个,求其平均值;再求这样的历年值与降水量的滑动相关系数。前后大小基本稳定、前小后大的可入选,否则舍弃。

2.2 从环流特征量、气象要素旬、月平均值中挑选因子

求这些前期气候资料与降水量的滑动相关系数,在一种资料内取滑动相关系数符号

相同、绝对值较大、前后大小基本稳定的月或旬相加,再平均。这样处理后,滑动相关系数同样稳定且绝对值增大。例如副热带高压北界头年6、10、11月平均值和6、10、11月平均值相加再平均的值(表中简写:平均)与郴州4月降水量的滑动相关系数如表2。(资料年代1954~2002年、滑动长度30年、滑动步长1年。)表中滑动相关系数各行左起第一个为1954~1983年因子与预测值的相关系数,第二个为1955~1984年因子与预测值的相关系数,余类推。

表2 副高北界与郴州4月降水量滑动相关系数

月	滑动相关系数																				
6月	0.3	0.3	0.4	0.4	0.4	0.4	0.4	0.4	0.4	0.3	0.3	0.3	0.3	0.3	0.4	0.4	0.4	0.4	0.5	0.5	
10月	0.3	0.3	0.3	0.3	0.3	0.2	0.2	0.2	0.2	0.3	0.3	0.3	0.3	0.3	0.2	0.2	0.2	0.2	0.3	0.3	0.3
11月	0.4	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.3	0.4	0.4	0.3	0.3	0.3	0.3	0.4	0.4	0.3	0.3	0.3	0.3
平均	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.6	0.6	0.6

从表2可见滑动相关系数不大的旬、月平均值,再经平均后滑动相关系数绝对值增大而且更稳定。有时滑动相关系数前小后大与前大后小的几个、两头大中间小与两头小中间大的几个再经平均后滑动相关系数绝对值增大而且变得稳定。

3 小结

在大气环流、海温场、当地气象要素历史资料中选取滑动相关系数前后大小基本稳定、前小后大的预测因子;在一种资料内取滑动相关系数符号相同、绝对值较大、前后大小基本稳定的月或旬相加,再平均,滑动相关系数同样稳定且绝对值增大,这样来选择单个因子。将最近几年以前的资料参与回归系数的计算,求出最近几年的预测值,再与实况对

比,就能看出因子群体的优劣。如果对预测效果不满意,再选因子加入,回归程序会选择另外的因子群体,又可求出最近几年的预测值,对比预测效果是否提高,这样反复进行来选择因子群体。通过以上方法,郴州冬春月降水量预测质量有了较大提高。这对提高短期气候预测整体质量不失为有益的探索。

参考文献

- 1 施能. 气象科研与预报中的多元分析方法. 北京: 气象出版社, 1995: 89~90.
- 2 罗伯良等. 湖南省夏季旱涝预报模型研究及试报. 气象, 1999, 25(10): 23~26.
- 3 施能. 多元分析和概率统计预报. 广西壮族自治区河池地区气象学会编, 1983: 81~83.

Research on Improvement Quality of Monthly Rainfall Forecast

Li Chenghong Zhou Youyun

(Chenzhou Meteorological Observatory, Hunan Province 423000)

Abstract

Monthly rainfall forecast is one of the short-term climate prediction operations, It is important for disaster prevention to improve the quality of forecast. By using statistic principles, some improvement on factors selection standard and factors pick-up from data have been made, to improve monthly rainfall forecast greatly.

Key Words: monthly rainfall regression moving correlation factors selection