

利用滚动式相关分析方法普查历史资料初探^①

侯淑梅 孙忠欣 孙秀忠 蔡冬梅 信志红

(山东省东营市气象局, 东营 257091)

提 要

介绍了一种新的普查历史资料的方法——滚动式相关分析方法,并将滚动相关分析法与一般相关分析法的相关分析结果进行了对比分析。结果表明,滚动式相关分析方法提高了相关稳定性,预报效果优于一般相关分析法。

关键词: 滚动相关分析法 稳定性 降水量预报

1 问题的提出

无论是查找预报指标,还是制作预报方程,只要涉及筛选预报因子的问题,就需要普查历史资料。目前比较常用的普查方法(简称一般法)是用相关系数作为其显著与否的评判标准^[1]。本文采用滚动相关分析法。普查因子时,在不改变样本容量的前提下(保证置信度的一致性),改变样本内容。也就是说,每次确定样本时,在原始样本的基础上,增加一个新样本,去掉一个老样本,然后用这组样本重新进行普查。根据实际情况,可以滚动 n 次(3次或5次)。将滚动 n 次得到的 n 个相关系数场迭套,选取 n 次相关系数均能满足置信标准的因子作为备选因子。我们将这种方法称为滚动式相关分析方法,简称滚动法。经过滚动法筛选的因子,提高了相关稳定性。

2 资料来源与方法

预报量为东营市垦利县 1963~1996 年 1~12 月的降水量,预报因子选取 1962~1996 年太平洋海温资料(SST)($10^{\circ}\text{S}\sim 50^{\circ}\text{N}$ 、 $120^{\circ}\text{E}\sim 80^{\circ}\text{W}$ 范围内 $5^{\circ}\times 5^{\circ}$ 286 个网格点资料)。降水量第一次取 1963~1992 年 30 年的资料进行相关普查,第二次取 1964~1993 年的资料,每次向后延续一年,滚动 5 次。因子选取与降水量月份相距两个月的前期 11

个月北太平洋月平均海温场资料,例如 1 月份降水用上年 1 月~上年 11 月份 SST 资料进行普查(本文所指上年均指降水量的上一年,当年是指降水量的同一年)。通过以上计算,各月降水量分别得到 5 个相关系数场,将 5 个相关系数场迭套,挑取 5 个相关系数场中相关系数均大于等于 0.35($\alpha=0.05$)的格点,并判断这些格点是否能连成一片(规定相邻的格点数大于等于 4 作为连成一片的入选条件)。将连成片的格点的海温值进行平均,作为一个备选因子。经过以上计算,得到各月降水量的 SST 因子库,留待备用。

3 两种普查方法入选因子的对比分析

比较一般相关分析法和滚动相关分析法所得的相关因子发现,滚动法的入选因子普遍少于一般法。比如,6 月份降水量与 SST 的相关普查中,一般法入选 27 个关键区,共 336 个格点,滚动法入选 10 个关键区,共 83 个格点。一般法在前 11 个月的 SST 场中都有关键区入选,而滚动法中前期第 5 个月和第 10 个月的 SST 场中都没有关键区入选,其它月份中,无论是关键区还是格点数,滚动法都比一般法少。表 1 列出了 1~12 月两种方法的入选因子情况。

从表 1 可见,一般法中每个月入选的关键区都在 10 个以上,最多的 6 月份有 27 个

^① 山东省气象局“九五”科研项目“东营市月、季、年降水预报自动化系统”课题资助

关键区被选中,滚动法入选的关键区都在10个以下。入选格点数一般法都远远大于滚动法,最多的6月份高达336个。滚动法入选最多的是12月份,选中格点168个。同样,其它月份降水量的普查中,滚动法比一般法入选因子明显减少,无形中起到了提高筛选标准的作用。

4 滚动法的相关显著性分析

通过滚动式相关分析方法得到了与每个月降水量相关显著的典型因子,每个月份入选因子的位置和相关的程度是不同的。表2列出了7月份降水量与前期北太平洋海温场的相关分析情况。

表1 1~12月一般法和滚动法入选因子情况

降水量月份	1	2	3	4	5	6	7	8	9	10	11	12
一般法关键区	12	17	4	18	8	27	8	13	16	10	10	15
滚动法关键区	3	11	4	5	3	10	7	10	5	3	8	10
一般法格点数	83	107	27	194	62	336	62	135	133	77	95	198
滚动法格点数	26	56	21	35	13	83	39	86	26	14	39	168

表2 东营市7月份降水量与前期北太平洋海温场的相关分析

海温场月份 ¹⁾	-7	-8	-9	-10	-11	-12	+1	+2	+3	+4	+5
关键区个数	0	0	1	1	2	1	1	0	1	0	0
相关格点总数	0	0	4	4	12	8	4	0	7	0	0
平均相关系数 ²⁾	/	/	0.426	0.429	0.389	0.396	0.362	/	0.391	/	/
最大相关系数 ³⁾	/	/	0.519	0.555	0.504	0.514	0.438	/	0.485	/	/

1)月份前的“-”表示上年,“+”表示当年;2)每月5个相关系数场中所有入选格点的相关系数的平均值;
3)每月5个相关系数场中所有入选格点的相关系数的最大值。

从表2可以看出,与7月份降水相关显著的海温场共有7个关键区,都是正相关区。各月关键区内所有格点的平均相关系数为0.398,其中上年10月份最大,其平均相关系数为0.429,入选的关键区位于北太平洋海流东端。各月关键区内入选格点最大相关系数的平均值为0.503,其中上年10月份最大相关系数为0.555,是所有入选格点中相关系数最大的格点,格点位置为(40°N、145°E)。

5 两种因子的预报效果对比分析

针对滚动法与一般法普查所得因子,分别用最优子集回归方法对各月降水量建立统计方程。利用所建方程,回代历史样本,并分别对1997~2000年进行预报。结果表明,用滚动式相关分析所得因子建立的方程,其预报准确率高于一般相关分析法。图1绘制了两种方法的最优子集回归方程对7月份降水量的预报情况。从图1可以看出,两种方法对1963~1996年的拟合情况基本接近,但对1997~2000年的预报中,滚动法因子的预报

误差小于一般法,尤其是1997年7月份降水偏少,滚动法预报正确,一般法预报错误。本文还对每个月两种方法的预报结果进行了评分。评分标准根据距平符号来衡量,符号一致为正确。总的来看,多数月份的滚动法预报准确率高于一般法。表3列出了用两种方

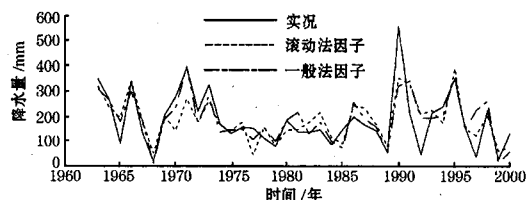


图1 一般相关分析法和滚动相关分析法的最优子集回归方程7月份降水量预报图

表3 滚动法与一般法相关普查所获因子的最优子集回归方程对1997~2000年4月份的预报及评定情况

年份	1997	1998	1999	2000
实况	11.3	28.9	8.8	3.3
滚动法	20.3	36.9	12.1	9.0
趋势评定	报对	报对	报对	报对
一般法	40.3	24.0	0.0	8.7
趋势评定	报错	报错	报对	报对

法所获得的 SST 因子建立的最优子集回归方法对 1997~2000 年 4 月份的预报及评定情况。4 月份气候平均值为 28.1mm。从表 3 可见,滚动法对 4 年的预报均正确,一般法只报对了两年。其它月份预报情况类似。

6 降水量预报

利用上述方法,普查了东营市各季和全年降水与 SST 的滚动式相关分析,得到东营市季、年降水量的海温备选因子场。利用最优子集回归方法分别建立东营市季、年降水量预报方程。降水量样本长度为 1963~1996 年共 34 年,1997~1998 年为试报年。夏季降水量预测方程为:

$$y = -6045.82 + 8.07x_1 + 5.86x_2 + 16.12x_3 \quad (1)$$

其中, y 为夏季降水量, x_1 为上年 9 月北太平洋海流区东半部的 4 个格点, x_2 为上年 10 月北太平洋海流区西半部的 4 个格点, x_3 为当年 3 月北赤道海流北部的 4 个格点平均海温。因篇幅所限,其它季节的降水预测方程略。图 2 给出了东营市夏季和全年降水量的逐年计算值(虚线)及实况值(实线)。从图 2 可见,各季和全年降水量的预报趋势基本与实况一致,几个异常旱涝年份,比如 1990 和 1995 年夏季降水异常偏多,拟合得很好。从 1999 和 2000 年的业务应用情况来看,1999 年各个季节和全年的预报都正确,效果很好。2000 年夏季降水量 222.9mm,较常年偏少 42.6%,模型预报 354.7mm,接近常年略偏少,虽然量级差别较大,但降水趋势预报是正确的。其它季节和全年降水量的预报与实况

差异也不是很大。由此可见,该预测工具实用准确率很好,有较高的应用价值。

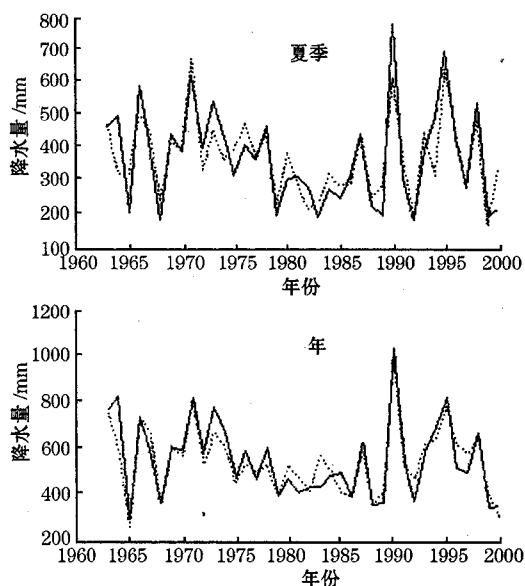


图 2 东营市夏季、年降水量实况(实线)与计算值(虚线)历史演变图

7 小结

通过以上分析表明,滚动式相关分析方法是一种较好的相关普查方法,可以较好地应用在科研和业务中应用。其筛选的显著因子的稳定性较好,对于基层台站开展预报工具的科研开发有较高的推广使用价值。

参考文献

- 1 谢考现,崔秀兰,刁秀广.短期气候预测因子的选取及利用.气象科技,1998,2.

Preliminary Study of Rolling Correlative Analysis Method for Investigating History Data

Hou Shumei Sun Zhongxin Sun Xiuzhong Xin Zhihong Wang Zonghai
(Dongying Meteorological Bureau, Shandong Province 257091)

Abstract

A new method, rolling correlative analysis, which can generally investigate the history data is introduced, and then the comparison analysis between rolling correlation and general correlation methods is carried out. The results show that the correlation of the factors which are gained from rolling correlation method is more stable, and its forecast accuracy is better than general correlation method.

Key Words: rolling correlative analysis method stability precipitation forecast