

华北区域一九七四年
汛期降水预报会议
技术材料选编

一九七四年七月

毛主席语录

思想上政治上的路线正确与否是决定一切的。

我们现在思想战线上的一个重要任务，就是要开展对于修正主义的批判。

现在的社会主义确实是前无古人的。社会主义比起孔夫子的“经书”来，不知道要好过多少倍。

科学技术这一仗一定要打，而且必须打好。不搞好科学技术，生产力无法提高。

人们为着要在自然界里得到自由，就要用自然科学来了解自然，克服自然和改造自然，从自然里得到自由。

目 录

关于夏季天气时段的划分及其预报的初步探讨

.....山西省晋中地区气象台 (1)

海、滦河流域年降水量的周期和趋势

——用时间序列方法对降水的分析.....华北地理所气候组
计算技术所概率组 (5)
天津计量所计算机组

适合手算的回归分析预报法

.....山西省岚县气象站 (12)

组合因子予告图

.....河北省保定地区气象台 (20)

一九七四年汛期降水预报的几点依据

.....河北省涿县气象站 (26)

试用周期迭加外延法作逐季逐月降水预报

.....山东省泰安地区气象台 (33)

单站长期预报几种工具

.....山东省沂源县气象站 (39)

利用自然正交函数制作大~暴雨日数预报

.....山东省气象台 (44)

用多因子同向百分率的回归曲线作降水定量预报法

.....天津市气象局气象台 (54)

汛期降水的统计转移预报

.....天津市气象局
南开大学数学系 (65)

冬季阶段主要环流系统与夏季主要雨带关系的初步探讨	河南省气象局气象台 科研所 (78)
用网格法制作夏季降水过程预报	河南省新乡地区气象台 (81)
对南阳6—8月份降水序列的初步分析	河南省南阳地区气象台 (87)
汛期开始及雨量集中时段预报方法	河南省栾川县气象站 (97)
地球自转速度变化和地极移动对天气气候的影响	中国科学院地理研究所气候室一组 (102)
试用平稳独立过程制作长期降水预报	内蒙古包头市气象台 (123)
冬春季节转换特征与六至八月降水关系	内蒙古商都县气象站 (129)
试用冬季东亚中纬地带环流突变预报北京汛期大~暴雨天气过程	北京市气象台 (133)

* 按收到稿件的时间先后排列

关于夏季天气时段的划分及其预报的初步探讨

山西省晋中地区气象台

为了更好地落实“两个服务”的方针，做好为农业生产服务工作，我们对影响我区大秋作物收成较大的“三伏”降水天气进行了初步分析，并对其长期预报进行了一些探索。现在谈谈我们的粗浅的认识和体会。

一、天气时段的由来及其划分

我区广大群众中流传着山西有“十年九旱”的气候特点。经过对历史资料分析，我们认为多数年分是旱中有涝，涝中有旱，旱涝交错，以旱为主。这种情况在夏季尤为明显。在预报工作实践中，我们体会到，在一段较长的时间内，冷暖空气活动的规律以及由它们构成的天气过程和气象要素的变化其有一定的阶段性和相似性。遵照毛主席“研究事物发展过程中的各个发展阶段上的矛盾的特殊性”的教导。我们对榆次地区历年七至八月分的降水资料进行了具体分析，发现七、八月分的降水分布具有明显的阶段性。

例如：七三年七月三至十一日无降水，十二至十八日连日有雨，十九日至八月十八日晴旱少雨，八月十九日至三十日转为多雨时期。

从天气图上可以看出，造成这种阶段性的主要因素是受西太平洋付高位置移动所制约。当付高西进北上（或南退）脊线位于 25°N — 30°N 附近时，我区处在付高边缘、自海洋源源不断地输送来不稳定的暖湿气流与西风槽东移相迁，造成我区多降水天气，并常有暴雨天气产生。这种情况一般可维持7—15天。我们称之为多雨时段。当付高脊线北进至 35°N 附近或南退至 25°N 以南时，我区多是晴旱少雨天气，常维持10—20天以上。

通过对历史资料分析，我们确定多雨时段的标准是：7—15天内总降水量大于30—50mm，其中有2—3天日降水量大于10mm，并有1—2天日降水量大于25mm。而小于1mm或无降水日数不超过3—5天。按此标准将61—73年的历史资料共划分出23个多雨时段。

二、分析多雨时段前期天气变化特点，寻找预报规律

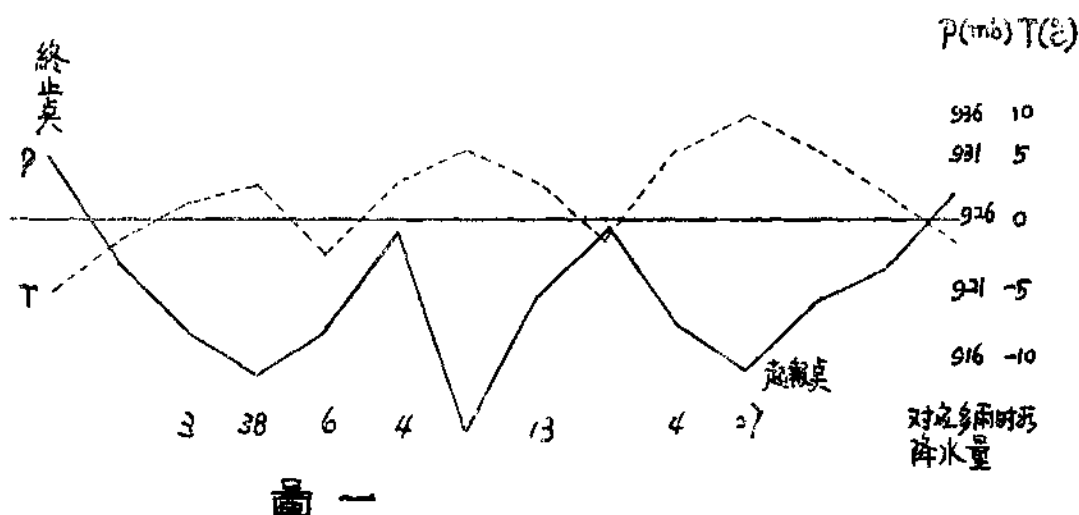
遵照毛主席“按照唯物辩证法的观点，自然界的变化的主要地是由于自然界内部矛盾的发展”的教导，根据几年来我们在长期预报中运用天气韵律关系的体会，将23个多

雨时段个例对应前六个月的天气及气象要素变化进行分析,发现一、二月分的压、温曲线特征与当年七、八月分多雨时段的出现有较好的对应关系。我们对其中21个多雨时段个例前期的压、温曲线特征进行归纳,可分为两大类。

(一) 低压高温类: 分主付两类

主类:

基本特征: 气压在较长的时间内(大于七天)维持在平均线以下并有3—4个峰点出现,其中最小峰谷差大于5毫巴,最大峰谷差大于8毫巴,在同一时间内温度曲线在平均线以上与气压呈反位相演变,并有谷点与气压峰点对应,最大峰谷差大于 10°C 。如图一:



过程日期: 气压曲线由平均线以上降至平均线以下, 出现第一个谷点的日期为起报点。气压曲线升至平均线以上出现第一个峰点的日期为终止点, 起报点到终止点的时段对应六个月(181天)以后的日期即为多雨时段。

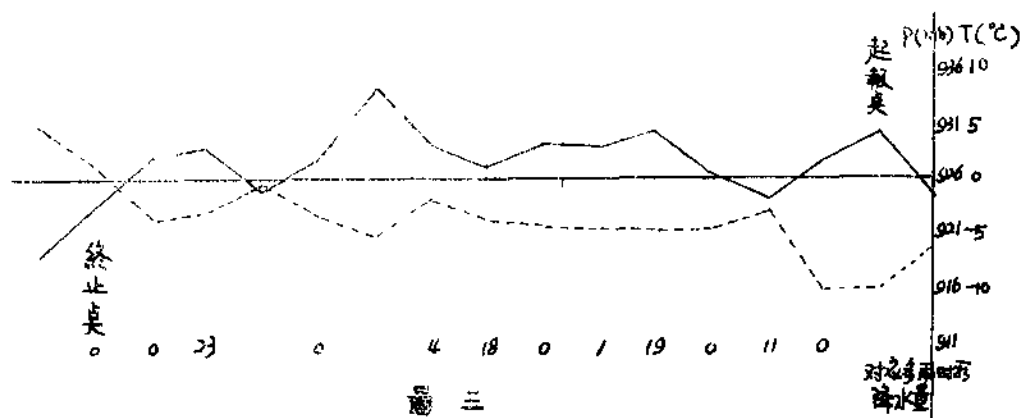
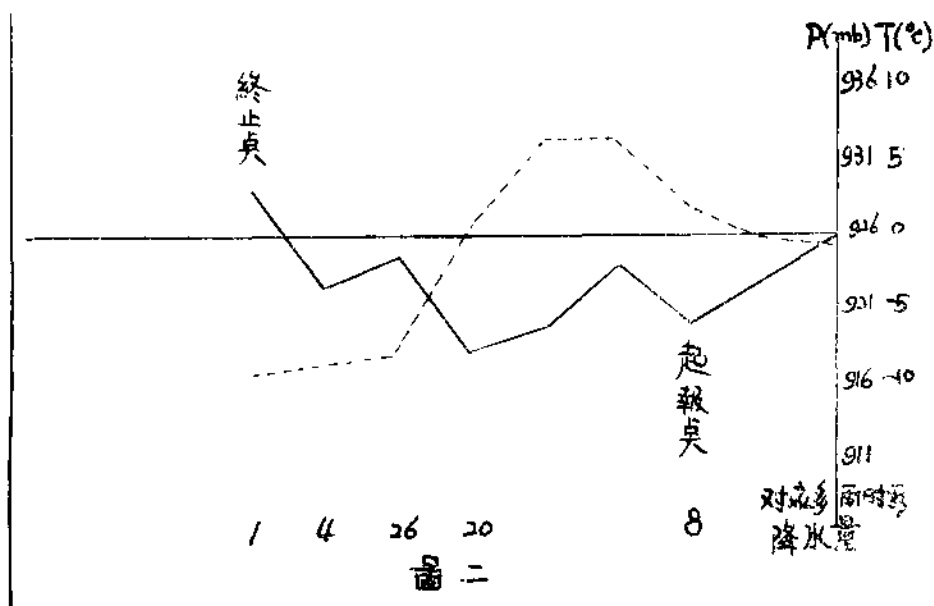
降水特点: 本类对应的多雨时段过程总降水量较大, 降水日数多在十天以上, 且总降水量大于50毫米。少数降水日数在十天以下者, 降水量为30—50毫米。在这类过程中, 常有2—3天日降水量大于25毫米。多出现在前期气压曲线谷点后1—2天对应的日期内。

付类: 气压曲线演变与主类相同。温度曲线, 过程前期在平均线以上, 后期降到平均线以下, 并有 $8-10^{\circ}$ 的负距平出现。如图二

降水特点: 本类对应的多雨时段过程前期降水量不大, 并有2—3天无降水或降水日数小于1毫米。后期降水量较大, 在温度出现最大负距平的日期对应的日期内常有降水量大于25毫米的天气过程出现。

(二) 高压低温类:

基本特征: 气压曲线在较长的一段时间内(大于7天)维持在平均线以上, 并有3—4个峰点出现, 其中最小峰谷差大于5毫巴, 最大峰谷差大于8毫巴。温度曲线在平均线以下与气压曲线呈反位相演变, 并有谷点与气压峰点对应。如图三



过程日期：气压曲线由平均线以下升到平均线以上出现第一个峰点的日期为起报点。降到平均线以下出现第一个谷点的日期为终止点。起报点到终止点的日期对应六个月后的时段即为多雨时段。

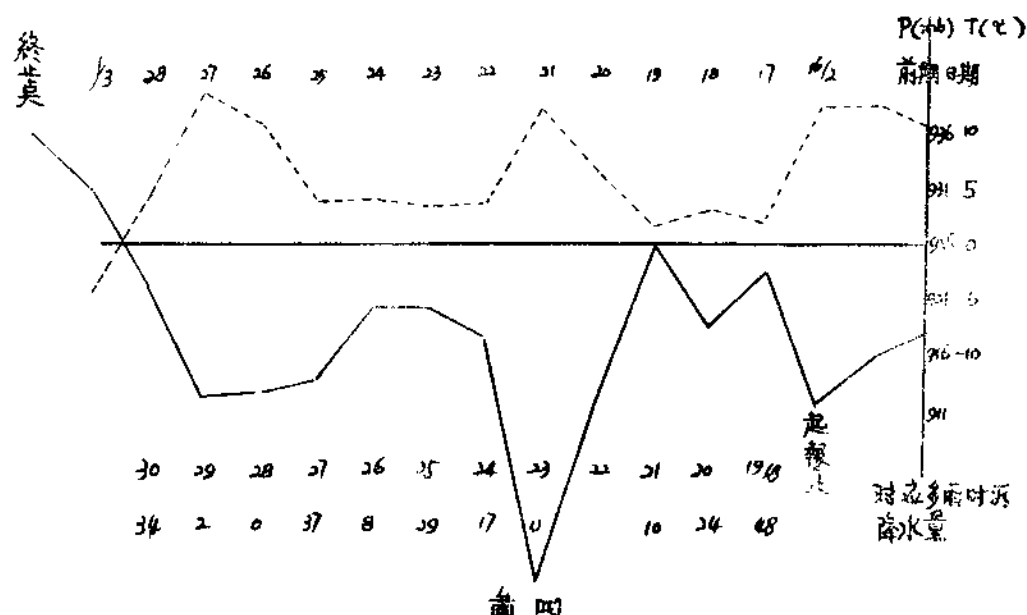
降水特点：本类对应的降雨时段过程总降水量不大，一般在30—50毫米。日降水量达到30毫米以上的日数很少。

少雨时段对应前六个月压温曲线的特点是在较长的一段时间内，呈同位相演变、或在平均线附近振动、且振幅不大、本文对此不做详细讨论。

三、予报实例

七三年二月1—12日压温曲线都在平均线附近振动，符合少雨时段的特点。十三日

以后气压曲线转入平均线以下且连续下降，到十六日出现第一个谷点三月一日升至平均线以上，二日达到峰点。中间十八日、二十一日，二十六日至二十七日分别出现谷点。符合多雨时段的低压高温主类的特点。如图四：



根据此类降水特点应予以八月十六日以前为少雨时段，十六日以后转为多雨时段。同时参考其它方法，最后定出八月二十日以前为少雨时段，有严重伏旱。八月二十日以后转为多雨时段，旬总降水量为100毫米（占八月预报总量的80%）二十日前后，二十五至二十六日，三十日前后分别有中——大雨天气出现。实况是八月一日至十八日总降水量9.3毫米，十九日至三十一日总降水量210.6毫米。并在十九日、二十日，二十五日，二十七日，三十日分别出现了大于25毫米的降水天气过程。对整个过程预报来说基本正确，但总降水量的预报偏少了110.6毫米。

四、问题讨论

1. 效果检验：从61—73年一、二月分压温曲线来看，共有22次符合上述多雨时段的前期特征，其中21次对应七、八月分有多雨时段出现，空报一次。13年中七、八月分共出现了23个多雨时段，用上述压温曲线特征可报出21次，漏报了2次。准确率为 $21/24 = 87.5\%$ 。

2. 对发生在2—3天的短期内的中到大雨天气过程，我们也进行了初步分析，可以看出六个月内压温曲线变化的特点是：压温曲线越过平均线急升急降，并常伴大风、风沙天气现象出现。同时我们还试用多种韵律迭加，和多因子相关概率统计方法进行综合分析，其具体预报指标还有待进一步探讨。

3. 以上我们对多雨时段的预报问题做了初步分析。由于我们水平不高，工作做得还不够细致，有待今后在实际工作中继续充实提高。

海、滦河流域年降水量的周期和趋势

——用时间序列方法对降水的分析

华北地理所气候组

计算技术所概率组

天津计量所计算机组

一、引言

本文是用时间序列分析方法对海、滦河流域23个地点的年降水量周期及趋势作一初步分析,目的是希望能为根治海河水利工程设计及长期天气预报提供参考依据。

本文采用的大多是水文资料,各地点均有50年以上,其中个别年代短缺者,用邻近地点资料插补。

资料的整理、分析及执笔工作,由华北地理所负责。计算工作由中国科学院计算技术所及天津计量管理所负责。

二、方法⁽¹⁾⁽²⁾

1. 周期提取

用方差分析方法提取周期,是利用 F 检验来判断系统性差异的显著程度。

对时间序列 $X(t)$ 的 N 个观测数据,按 $1, 2, 3, \dots, \frac{N}{2}$ 排列周期,当 N 为奇数时,

取 $\frac{N}{2} = \frac{N-1}{2}$ 。此处只取三年以上的周期。

$$F = \frac{\sigma_1/r_1}{\sigma_2/r_2} \quad (1)$$

式中: σ_1 ——“组间”离差平方和

r_1 ——“组间”自由度

σ_2 ——“组内”离差平方和

r_2 ——“组内”自由度

分别计算 $3, 4, \dots, \frac{N}{2}$ 各周期的 F 值,并挑选 F_{max} 进行检验。给出信度 α , 若

$$F_{max} \geq F_{\alpha}$$

则周期显著, 并以此为第一周期。若 $F_{max} < F_{\alpha}$, 则周期不存在。

第一周期选出后, 将原序列去掉第一周期各值, 得到一个新序列, 按上述步骤选出第二周期, 余类推。

2. 趋势

将时间序列 $X(t)$ 看作为随机过程的一个现实, 假定它由趋势 $f(t)$ 、周期 $S(t)$ 、平稳随机部分 $\eta(t)$ 和白噪声 $\zeta(t)$ 四部分组成, 构成一个加法模型, 即

$$X(t) = f(t) + S(t) + \eta(t) + \zeta(t) \quad (2)$$

白噪声 $\zeta(t)$ 可以略去。用右端前三项统计计算, 给出 $X(t)$ 的预报值。

(1) 趋势分析

对时间序列 $X(t)$, 取 $t_0 = 1$, $t_n = 1 + \frac{n}{N}$

用函数

$$\begin{aligned} f(t) = & a_0 + a_1 t + a_2 t^2 + a_3 t^3 + a_4 t^4 + a_5 \frac{1}{t} + a_6 \frac{1}{t^2} \\ & + a_7 \sqrt{t} + a_8 \frac{1}{\sqrt{t}} + a_9 e^t + a_{10} \ln t \end{aligned} \quad (3)$$

来拟合 $X(t)$ 。用逐步回归算法, 通过 F 检验, 舍选(3)中的系数 a_i , 给出趋势项 $f(t)$ 。

若得 $a_i = 0$ ($i = 1, 2, \dots, 10$) 则认为该时间序列 $X(t)$ 无趋势项 $f(t)$ 。

(2) 周期分析

将 $X(t)$ 减去 $f(t)$, 则有 $y(t) = X(t) - f(t)$, 用周期图分析方法提取周期

取周期函数

$$S(t) = \sum_{i=1}^l (b_{i1} \cos w_i t + b_{i2} \sin w_i t) \quad (4)$$

拟合 $y(t)$, 其中 l, b_{i1}, b_{i2}, w_i 为待定参数。

对有 N 个数据的 $y(t)$, 可能周期有,

$$\frac{N}{1}, \frac{N}{2}, \dots, \frac{N}{\left[\frac{N}{2}\right]}$$

这里

$$\left[\frac{N}{2}\right] = \begin{cases} \frac{N}{2} & \text{当 } N \text{ 为偶数时} \\ \frac{N-1}{2} & \text{当 } N \text{ 为奇数时} \end{cases}$$

对这组可能周期 $T_i = \frac{N}{i} (i = 1, 2, \dots, \left[\frac{N}{2}\right])$, 有

$$b_{i,1} = \frac{2}{N} \sum_{j=1}^N y_j \cos \frac{2\pi i j}{N}$$

$$b_{i,2} = \frac{2}{N} \sum_{j=1}^N y_j \sin \frac{2\pi i j}{N}$$

一般称统计量

$$g_i = b_{i,1}^2 + b_{i,2}^2$$

为随机过程 $X(t)$ 的周期图, 为了从这些可能周期中选取 $X(t)$ 的真正周期, 给下面的筛选方法, 取

$$G = \sum_{i=1}^{\left[\frac{N}{2}\right]} g_i$$

$$g_{ij} = \max \left[g_1, g_2, \dots, g\left(\frac{N}{2}\right) \right]$$

在 $\eta(t)$ 为白噪声的假定下, 统计量

$$Q_i = \frac{g_{ij}}{G}$$

服从 Fisher 分布给定显著水平 α , 经统计检验, 得到 $y(t)$ 的 l 个周期 $T_{i1}, T_{i2}, \dots, T_{il}$, 取频率

$$\omega_i = \frac{2\pi}{T_{il}}$$

给出周期函数 $S(t)$ 。

(8) 平稳时间序列分析

时间序列 $X(t)$ 经过趋势 $f(t)$ 和周期 $S(t)$ 的提取

$$Z(t) = X(t) - f(t) - S(t)$$

一般假定为一个平稳时间序列。

设 $Z(t)$ 是一个各态历经的平稳时间序列, 用 n 阶自回归模型

$$Z_{n+1}^* = \sum_{i=1}^n C_i Z_{n+1-i} \quad (5)$$

对平稳时间序列进行预报, 选取 (5) 中的待定系数 C_i , 解 n 阶代数方程组

$$\begin{pmatrix} P_{(0)} & P_{(1)} & \dots & P_{(n-1)} \\ P_{(1)} & R_{(0)} & \dots & P_{(n-2)} \\ \dots & \dots & \dots & \dots \\ R_{(n-1)} & R_{(n-2)} & \dots & P_{(0)} \end{pmatrix} \begin{pmatrix} C_1 \\ C_2 \\ \dots \\ C_n \end{pmatrix} = \begin{pmatrix} R_{(n)} \\ P_{(n-1)} \\ \dots \\ R_{(1)} \end{pmatrix}$$

其中 R 为相关函数。得 C_t , 给出预报值 Z'_{n+1} 。

综合上述三项统计计算结果, 给出结果预报值, 并作修匀处理, 得最终预报结果。

三、结 果

(一) 周 期

降水周期, 目前国内外都在进行研究⁽³⁾⁽⁴⁾⁽⁵⁾⁽⁶⁾。这些研究认为降水周期与太阳活动周期有某种程度的相关性。例如: 80至90年的世纪周期、35年左右的布吕克纳周期、22至23年的海尔周期(亦称太阳黑子磁周)、11年左右的太阳活动周期以及5至6年的半太阳活动周期等, 都在降水序列中有某种程度的反映。由于我们使用资料年限较短, 所以没能算出较长的周期。但仅从这些已算出的周期也可看出确实与太阳活动的周期性有关系。

表1所列周期是按方差分析挑选顺序先后排列。

周 期 表

表 1

地 点	第一周期		第二周期		第三周期		地 点	第一周期		第二周期		第三周期	
	T	F	T	F	T	F		T	F	T	F	T	F
浑 源	12	2.5	23	2.4	8	2.2	三 家 店	17	3.5	10	2.1	4	3.3
承 德	5	3.8	21	2.2	16	3.3	保 定	5	3.3	17	2.2	9	2.3
唐 山	5	2.5	16	3.2			献 县	7	2.4	3	3.5	5	2.5
玉 田	17	3.1	5	3.3	7	4.8	沧 州	3	2.8	5	2.2*	8	3.4
香 河	17	2.8	5	6.0			衡 水	10	4.4	18	2.0	14	2.6
杨 柳 青	17	3.2	5	4.4			石 家 庄	5	4.5	7	2.7	13	2.6
天 津	17	2.7	9	5.9	4	3.2	深 泽	10	2.3	16	2.0	12	2.2
通 州	21	2.1	16	2.3	9	4.0	临 清	8	3.3	7	2.4	19	2.3
芦 沟 桥	17	2.3	13	5.0	5	4.9	大 名	8	3.3	19	2.4		
北 京	17	4.1	5	2.4			张 家 口	3	3.1	5	2.8	11	2.2

* 者为信度 $\alpha=0.25$, 其它 $\alpha=0.05$ 。

从计算结果看: (1) 5年左右的周期, 占所选周期总数的25%。(2) 11年左右的周期占所选周期总数的25%。(3) 17至18年左右的周期占所选周期总数的25%稍多些, 这个长度恰好是布吕克纳周期的一半左右。这就说明, 海、滦河流域降水的周期, 大多数与太阳活动的周期是相近的。大致说来, 多雨年大多表现为5年周期, 少雨年表现为11年左右的周期。用群众的语言来说, 就是: “十年一早, 五年一涝”。

周期的地区分布, 大致可分为四个类型。即: 浑源、张家口、承德一带, 主要是5年和11年的周期; 北京、天津、保定一带, 主要是5年和17年的周期; 石家庄、衡水一带主要是5年和7年的周期; 沧州、临清、大名一带主要是8年和19年的周期。

在本次用时间序列分析计算降水趋势时,我们只选用一个周期,多数为5年左右的周期和两年多的周期,即准两年周期。

(二) 降水趋势

各地点降水趋势分级:正常~该点多年平均值 ± 50 毫米,偏多~正常上限+100毫米,多~大于偏多上限,偏少~正常下限-100毫米,少~小于偏少下限。

经过计算1974年~1982年海、滦河流域的降水趋势(见表2),大致可以认为是两

年 降 水 量 未 来 趋 势

表 2

地 点 \ 年	1974	1975	1976	1977	1978	1979	1980	1981	1982	备 注
大 同	偏多	偏少	正常	正常	正常	正常	偏少	正常	偏少	
浑 源	正常	偏少	偏多	偏少	偏多	"	偏少	正常	偏少	
朔 县	偏多	正常	正常	正常	偏多	"	偏多	多	少	
张家口	正常	正常	"	偏多	正常	"	正常	偏少	正常	
承 德	偏多	偏多	"	正常	偏多	"	正常	正常	偏少	
唐 山	多	正常	"	正常	偏少	多	偏少	正常	正常	
玉 田	偏多	"	"	偏少	少	正常	正常	偏多	正常	
香 河	正常	"	"	偏少	正常	偏多	正常	偏少	少	
北 京	正常	偏少	偏多	正常	偏少	正常	正常	正常	偏少	
三家店(北京)	正常	正常	"	正常	偏少	多	偏多	正常	正常	
芦沟桥(北京)	正常	偏多	"	正常	正常	偏多	正常	偏少	偏少	
保 定	偏多	正常	"	偏多	偏少	正常	多	正常	正常	
新 镇(文安)	偏多	正常	正常	"	偏少	正常	多	偏多	偏少	
献 县	多	正常	偏少	"	偏多	多	多	正常	正常	
沧 州	偏多	偏多	正常	"	正常	多	多	偏多	偏多	
衡 水	偏多	正常	偏少	"	正常	偏多	多	正常		
石家庄	多	偏少	偏少	正常	正常	偏少	偏多	正常	少	
深 泽	多	正常	偏少	偏多	偏多	多	多	正常	正常	
忻 县	偏多	正常	正常	正常	正常	少	偏少	偏多	少	
大 名	正常	正常	偏少	正常	"	偏多	偏多	偏多	偏少	
安 附	偏少	偏多	正常	偏多	"	偏少	偏少	正常	少	
济 南	正常	正常	偏少	"	"	正常	偏少	正常	少	
临 清	偏多	正常	少	"	"	正常	多	正常	偏少	

个波峰和两个波谷。第二个波的振幅比第一个波要来得大些。即:

1, 1974年本流域多数地点年降水为正常或偏多,是第一个波峰。

2. 1975~1978年全流域多表现为下降的趋势,其中一九七六年倾向是正常和偏少,是第一个波谷。

3. 1979~1980年本流域多雨和偏多地点增多,是第二个波峰,比第一波峰更高些。

4. 1982年本流域少雨和偏少地点居多数,是第二个波谷,是未来近十年的最低点。

从总的趋势看,未来近十年的年降水量是处于正常偏多状态。

在两个波动中,降水趋势在地区分布上又各有若干不同,分述如下:

1974年多雨和偏多主要出现在三个地区,即:保定、石家庄、衡水、沧州地区;唐山、玉田地区;大同、朔县地区。北京、天津地区正常。唯安阳一处偏少。

1975~1978年间,其中1976年北京、天津、保定地区偏多;石家庄、衡水以南大片地区偏少;唐山地区正常。1977年北京、天津地区正常,往北偏少,往南偏多;1978年玉田、唐山、北京地区偏少,往南正常,往北偏多。

1979~1980年间,北京、保定、衡水、沧州是大片多雨地区,其它地区亦多有上升趋势。

1981~1982年间,年降水倾向减少。1982年多数地区是少雨及偏少,即:大同、浑源、朔县、忻县地区;石家庄、安阳、济南、临清地区;北京、香河地区。其他地区正常。

(三) 预报精度

根据预报结果与实况相比,算出预报精度。以预报结果落在该级或邻近一级为对,落在以上范围之外为错。

从预报结果检验可以看出,总的成功率可以达到69%,报错约为31%。不同地区效果也不同,有的地区如:大同、张家口等成功率较高,保定、承德等成功率较低。多数是由于极值报错,其中少雨年80%报错,多雨年有50%报错(见表3)。

予 报 结 果 检 验

表 3

地 点	年数	报对	%	报错	%	七三年	地 点	年数	报对	%	报错	%	七三年	备 注
大 同	6	6	100			对	新 镇	5	4	80	1	20		
浑 源	6	4	67	2	33	对	献 县	6	4	67	2	33	对	
朔 县	6	4	67	2	33	错	沧 州	6	4	67	2	33	对	
张家口	6	6	100			对	衡 水	5	4	80	1	20	对	
承 德	6	3	50	3	50	错	石 家 庄	6	4	67	2	33	对	
唐 山	6	4	67	2	33	对	深 泽	4	2	50	2	50		
玉 田	5	4	80	1	20		忻 县	6	5	84	1	16	对	
香 河	5	4	80	1	20		大 名	6	6	100			对	
北 京	10	7	70	3	30	错	安 阳	6	4	67	2	33	错	
三家店	4	1	25	3	75		济 南	6	4	67	2	33	错	
芦沟桥	5	3	60	2	40		临 清	6	5	84	1	16	对	
保 定	6	2	33	4	67	错								

(四) 讨 论

1. 用福氏级数计算华北地区年降水量的周期, 不大理想。因为降水序列振幅较大, 跳动强烈, 不易达到检验标准。因而会漏选一些周期, 可选用的周期少, 代表性就不全面。本次计算降水趋势的只用一个周期, 是个缺陷。一般用方差分析效果较好。

2. 不同地点的预报精度不一, 所以使用本文结果时, 应予以注意。

3. 本文结果只是用一种数理统计方法计算得出, 只能作为一个参考, 应与其它方法结合起来考虑。对今后的预报要不断订正。

本文缺点错误欢迎批评指正。

本工作进行中, 得到中央气象局、大气物理所、河北气象局、天津气象局的很多帮助, 仅致谢意。

参 考 文 献

1. 计算技术所、地球物理所: 用《时间序列分析》方法预报天山地震 地震 1973. 2
2. 大气物理所: 概率统计和单站天气预报方法 1973. 2
3. 统计学长期天气预告方法的若干研究 气象学报34卷4期 史久恩等
4. 水文气象特征值的周期性 涅·弗·罗基诺夫 苏联科学院通报 地理学(俄) 1971. 4
5. 水文学的序列中周期性的分析 杰·安德尔 水文学(英) 1971. 1
6. 汛期降雨历史演变规律的分析 河北省气象局气象台 1972. 7

适合手算的回归分析预报法

山西岚县气象站

回归分析是数理统计制作天气预报方法中常用的一种，（见文献〔1〕）但计算量大，需电子计算机。目前县站只有算盘，开展此法困难较多，我站在“十大”精神鼓舞下，遵照毛主席理论联系实际的教导，以求适合手算，便于掌握，采取了以下措施：

- 一、用先粗筛后细选的方法找因子；
- 二、用线性变换减少计算量；
- 三、用单、复相关系数的起码值作显著性检验；
- 四、用简易一些方法作区间估计；
- 五、用简单代数式表明运算步骤。具体作法如下：

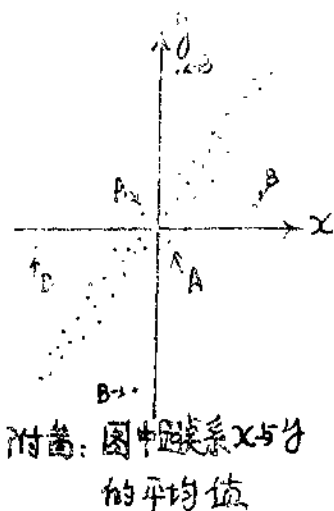
一、筛选因子

为广泛而严格地挑选因子，尽量减少计算量，首先以农谚、经验及气象学原理为线索，采用四步筛选。

1. 符号检验：对比予报量的距平与所考虑的因子的距平，看它们符号的异同情况，根据二项分布的符号检验表，可以查到在一定信度下作为相关因子的起码条件，从而判

断是否可以作为一个因子。此法较简单，便于广泛试查，但对定量予报而言，尚嫌粗糙，故需将合格者再作进一步检验。（有关符号检验法详见参考文献〔2〕）

2. 图象检验：将历史资料予报量 y 及因子 x_i 点在直角坐标图上，选择点子近似直线分布而又密集者作为予报因子。如附图。（注意附图中按符号检验法 A 处点子算反常而实际好， B 处点子算正常而实际太分散。所以为克服符号检验的错觉，不丢掉好的因子，通过“符号检验法”的标准可稍稍放宽。）



3. 相关系数检验：

计算 y 与 x 的（单）相关系数 r

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\left[\sum_{i=1}^n (x_i - \bar{x})^2 \right] \cdot \left[\sum_{i=1}^n (y_i - \bar{y})^2 \right]}} = \frac{S_{xy}}{\sqrt{S_{xx}S_{yy}}} \quad (1)$$

从表 1 中, 由资料 (样本) 总数 n 及因子数 $p=1$, 按一定信度 α , 可以查到 (单) 相关系数的起码值 r_α , 若 $|r| \geq r_\alpha$, 便认为 y 与 x 的相关是显著的。按照 r 绝对值从大到小的次序, 将入选的因子依次编号, 记作 $x_1 x_2 \dots x_p$ 。

复相关系数的起码值 P_α

表 1

α $n \backslash p$	5 %						1 %					
	1	2	3	4	5	6	1	2	3	4	5	6
10	632	758	839	898	942	973	765	855	911	949	975	991
12	576	697	777	838	886	925	708	800	860	904	938	963
14	532	648	726	786	835	877	661	753	814	861	898	928
16	497	608	683	741	790	832	623	712	773	821	859	891
18	468	574	646	703	751	792	590	677	737	785	824	857
20	444	545	615	670	717	757	561	647	706	752	791	825
22	423	520	587	641	686	726	537	620	677	724	762	796
24	404	498	563	615	660	699	515	596	652	697	736	769
26	388	479	542	592	636	674	496	574	630	674	712	745
28	373	461	523	572	614	651	479	555	609	652	690	722
30	361	446	506	553	595	631	463	538	590	633	670	701
35	334	414	469	514	552	586	429	500	550	591	626	657
40	312	387	439	482	518	550	403	469	517	556	589	618
45	294	365	414	455	490	520	380	444	489	526	558	587
50	279	346	394	432	465	495	361	422	465	501	533	559
55	266	330	375	412	444	473	344	403	445	479	509	535
60	254	316	359	394	425	452	330	386	426	459	488	514
70	235	292	333	366	395	419	306	358	396	427	454	477
80	220	273	312	342	370	393	286	336	371	401	426	448
90	207	258	294	323	349	371	270	317	351	378	402	424
100	197	245	279	307	331	352	256	302	333	359	382	403

说明: α 为信度, n 为资料 (样本) 总数, p 为因子数目。

当 $p=1$ 时, 复相关系数的起码值 P_α , 即为 (单) 相关系数的起码值 r_α 。

本表对 n 值可内插。

本表录自文献[3]。表中数字已放大1000倍, 如632即为0.632。

显然好的因子除与予报量相关密切外, 还应与其它因子互相独立, 此点在选因子时应注意。

为予报年总降水量 y , 在农谚的启发下, 寻找前期旱涝、风力等要素, 通过筛选得四个因子: x_1 (去年二月日照率), x_2 (去年十月平均气温), x_3 (去年年平均绝对湿度), x_4 (去年十二月平均风速), 用此试作回归方程。

4. 复相关系数检验。这步是在建立多因子予报之后, 用于验收予报公式的, 具体做法后边再谈。