

关于我国人工防雷效果的统计分析

黄美元 亢雪巧*

冰雹是我国农业上重大自然灾害之一,也是一种很典型的强对流云活动的产物。因此开展冰雹和人工防雷的试验研究是十分必要的。十多年来我国各地广泛开展了这项工作,取得了许多成绩。为了进一步总结经验,搞好人工防雷,需要科学地分析人工防雷的效果。但防雷效果的检验是一个复杂而尚未完全解决的课题,看法也并不一致。在这里,我们想从所收集到的国内实际资料出发,用统计方法讨论有关人工防雷的效果问题。

一、降雹资料的一些统计特点

为了进行防雷效果的统计检验,有必要了解降雹资料的统计特征。例如,降雹资料是否随机变量?服从何种统计分布?它的标准离差和变异系数多大等等,从简便和能够说明一定问题出发,各地在分析防雷效果时,经常采用年降雹日数和年受灾面积数作为统计的变量,我们也采用这两个量来分析讨论。年降雹日数我们抄自各地气象站的观测记载,从中选择观测年代较长(不少于30年)并能代表我国不同地区特点的资料。雹灾面积的资料获得较困难,我们只能从几次全国人工防雷会议上所提供的材料中得到。一般来说,这些资料年代都不长,准确程度也不如年降雹日数,对此只能从实际情况出发,进行初步讨论。

先分析年降雹日的资料。图1是国内几个典型地方年降雹日数随年代的变化情况,年代最长的地方是上海,有81年资料。从图中可以看到各地降雹日数的年变化还是比较大的,而且看不出有什么类似周期性的变化规律。我们还计算了这些地方降雹日数在不同间隔年代的相关系数(见图2)。从图看到,各年之间的相关系数基本上都在0.3—0.2之间,就是说年降雹日数在各年之间的相关是很小的,可以把它看作为一种随机变量。

表1给出了国内不同地区21个地方年降雹日数的一些统计情况;如平均值、标准离差和变异系数(即离差与平均值之比)。可以看到,各地变异系数的变化范围一般在0.49—2.36之间,而绝大部分落在0.6—1.1之间(占16/21)。并且有这样一个特点,即对于年降雹日数的平均值在1以下的地方,变异系数一般都较大,大部分在1.0以上,而年降雹日数在1.1以上的,其变异系数绝大多数在1.0以下(占12/13)。

在人工防雷效果统计分析中常常用到参量的假设检验,为此首先要求了解统计量服从何种统计分布。我们分别采用了 χ^2 检验和柯尔莫哥洛夫检验方法对年降雹日进行了是

1977年6月30日收到。

* 北京大学地球物理系大气物理专业1976年工农兵毕业学员。

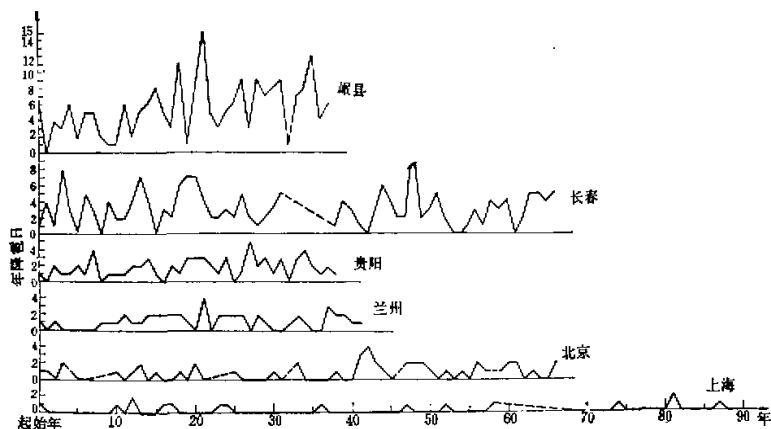


图 1 几个典型地方年降雪日随年代的变化。

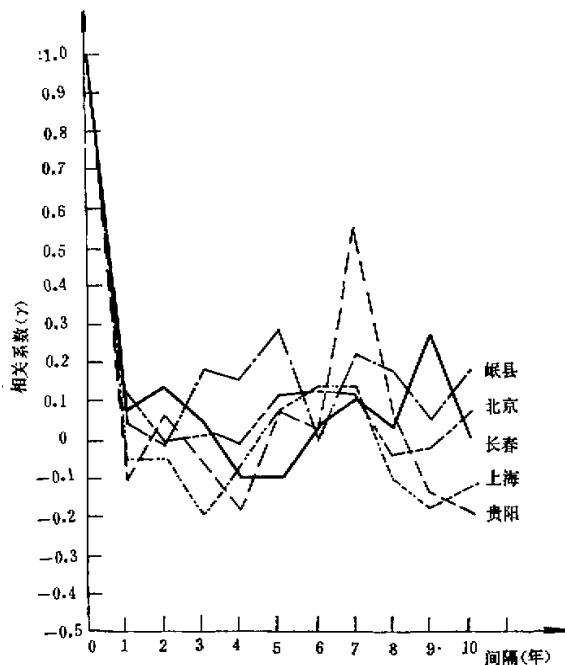


图 2 年降雪日数逐年相关系数。

否符合正态分布的检验。另外,我们考虑到,对于年降雪日数这个变量是一系列不连续的正整数,用泊松(Poisson)分布表达可能更好些,所以我们用 χ^2 检验法对年降雪日数进行了是否符合泊松分布的检验。以长春为例,把长春历年年降雪日数分组,分别统计它们出现的频率,相应地算出正态分布的理论频率。

表 1 国内若干地方年降雪日数的一些统计特征

地 点	起迄年代	有效年代	平 均 值	标 准 差	变异系数	是否符合泊松分布	是否符合正态分布
上 海	1883—1972	81	0.19	0.45	2.36	—	—
南 京	1929—1975	32	0.25	0.26	1.03	—	—
武 汉	1907—1974	59	0.51	0.25	0.49	—	—
酒 泉	1935—1974	40	0.68	0.69	1.01	—	—
靖 远	1937—1969	32	0.75	0.81	1.08	—	—
北 京	1875—1975	75	0.75	0.93	1.24	是	否
天 津	1905—1975	57	0.90	0.96	1.07	是	否
大 连	1905—1975	59	1.10	1.19	1.08	是	否
兰 州	1933—1974	42	1.14	0.98	0.86	是	是
太 原	1926—1975	33	1.36	1.31	0.96	是	是
西 峰	1938—1969	30	1.60	1.89	1.18	是	是
营 口	1905—1975	63	1.73	1.36	0.78	是	是
贵 阳	1935—1974	39	1.77	1.22	0.69	是	是
满洲里	1909—1974	41	1.78	1.24	0.70	是	是
海拉尔	1909—1974	50	1.96	1.44	0.74	是	是
临 夏	1943—1974	32	2.06	1.56	0.76	是	是
平 凉	1937—1975	39	2.15	1.82	0.85	是	是
沈 阳	1906—1975	57	2.22	1.23	0.56	是	是
博克图	1914—1953	35	2.80	1.77	0.63	是	是
长 春	1909—1974	61	3.11	2.26	0.72	是	是
岷 县	1937—1975	38	5.30	3.54	0.67	是	是

可得, $\sqrt{n} D_n = \sqrt{61} \times 0.1305 = 1.019$

取置信水平 $1 - Q(\lambda) = 5\%$, 查柯尔莫哥洛夫检验表, 可得 $\lambda = 1.36$, $\therefore \sqrt{n} D_n < \lambda$.

所以没有理由认为长春的年降雪日数不服从正态分布 $N(3.11, 2.26^2)$.

同样把长春历年年降雪日数分组, 统计各组的频数, 计算出各组泊松分布的理论频数。得到;

$$\sum \frac{(nP_{x_i} - f_i)}{nP_{x_i}} = 9.081,$$

查 χ^2 检验表, 当自由度 $r - m - 1 = 8 - 1 - 1 = 6$, 置信水平 $\alpha = 0.05$ 时, 得

$$\chi_{0.05}^2 = 12.592, \quad \therefore \sum \frac{(nP_{x_i} - f_i)}{nP_{x_i}} < \chi_{0.05}^2.$$

所以没有理由认为长春的年降雪日数不服从泊松分布 $P(3.11)$.

用上面类似的检验计算方法, 可以对各台站的年降雪日数进行分布检验, 但是对上海, 酒泉, 南京, 靖远, 武汉等五个地方的资料没有检验, 主要是因为这些地方的年降雪日数只在 0—2 之间变动, 分挡太少, 不符合检验的条件。对其余 16 个地方的年降雪日数的分布都进行了统计检验, 其结果列在表 1 的最后二项中, 从这结果中可以看到, 16 个台站的年降雪日数都符合泊松分布, 对绝大部分 (13/16) 台站来说, 年降雪日数也可以用正态分布来近似。对于年降雪日数平均值较小的地方, 用正态分布来近似偏离较大, 显然这是由于年降雪日不可能是负值的缘故。

对于受灾面积我们可以进行类似的统计分析。表 2 给出了几个地方的资料, 从表中看到, 对于统计单位较小的大队和公社来说, 变异系数较大(约为 1.0 左右), 以省和县为单位来统计, 变异系数较小(约在 0.5 左右), 这实际上反映了代表性的问题。变异系数较大的代表性差, 要从统计上说明问题更困难些。

表 2 受灾面积的统计特征

地 点	有效年代	平 均 值	标 准 差	变 异 系 数
甘 肃 省	19	152.4	93.5	0.61
岷 县	15	18.1	9.4	0.52
昔 阳 县	9	13.3	5.0	0.38
麻迷图公社	10	17.2	17.9	1.04
黄茂营大队	8	7.4	8.1	1.10
老羊圈大队	6	26.1	25.6	0.98

对于雹灾面积的分布检验, 想直接用各地雹灾面积资料是有困难的, 因为各地资料的年代太短, 不符合分布检验的要求, 为此, 我们把年代较长的地方的资料综合起来进行统计检验。具体进行是先把每一个地方的资料标准化, 即求 $x'_i = \frac{x_i - \bar{x}}{\sigma}$, 其中 $x_i, \bar{x}, \sigma$ 是原来的雹灾面积、平均值和离差。 x'_i 为标准化后的雹灾面积。然后把几个地方标准化后的数据分组, 进行正态分布的检验, 我们利用的资料有: 甘肃省、岷县和麻迷图公社, 总共有 44 年资料。用柯尔莫哥洛夫法和 χ^2 法检验结果表明, 雹灾面积的分布可以用正态分布来近似, 当然这只是一初步的计算, 严格的检验还要等到收集到较多较完整的资料之后。

二、人工防雷效果的统计分析

在介绍了国内一些地区年降雹日和雹灾面积的一些统计特性之后, 可以讨论一下人工防雷效果估计的问题。我们分析的基本思想是: 统计对比防雷前后的年降雹日或雹灾面积, 并进行 t 检验。根据二组变量的 t 检验, 有,

$$t_a = \frac{\bar{x} - \bar{y}}{\sqrt{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}} \cdot \sqrt{\frac{n_1 n_2 (n_1 + n_2 - 2)}{n_1 + n_2}}$$

把上式变换一下, 并考虑到效果检验时, 只讨论减少的一方, 所以只是单边检验, 因此, t_a 要用 t_{2a} 代替, 上式变为:

$$\frac{\bar{x}}{\bar{y}} - 1 = t_{2a} \cdot \frac{s_2}{\bar{y}} \cdot \sqrt{\frac{(n_1 - 1)s_1^2/s_2^2 + (n_2 - 1)}{n_1 + n_2 - 2}} \cdot \left(\frac{1}{n_1} + \frac{1}{n_2} \right)$$

在公式中, $\bar{y}$ 和 $\bar{x}$ 以及 s_2 和 s_1 分别为防雷前后统计量的平均值和标准差。 n_2 和 n_1 是防雷前后统计量具有的年代数。 $s_2/\bar{y}$ 是防雷前统计量的变异系数。 $\frac{\bar{x}}{\bar{y}} - 1$ 说明防雷后比防雷前统计量减少的百分数。

根据上述公式和前面分析的降雹统计量变化的范围, 可以计算出一些表供分析查用。表 3a 和 3b 就是其中之二(取置信水平 $\alpha = 0.05$)。这样当我们知道了防雷前后雹灾的统

表 3a 防雷效果的检验表

$n_2 = 10$																			
$-(\bar{x}/\bar{y} - 1)$	$s_2/\bar{y}$ s_1/s_2	0.25			0.5			0.57			1.0			1.5			2.0		
		0.5	1.0	2.0	0.5	1.0	2.0	0.5	1.0	2.0	0.5	1.0	2.0	0.5	1.0	2.0	0.5	1.0	2.0
1		48.0	48.0	48.0	96.0	96.0	96.0												
2		34.5	35.1	40.0	69.0	70.2	80.0												
3		25.6	29.6	36.7	51.2	59.2	73.4	76.8	88.0										
4		21.4	26.3	34.8	42.8	52.6	69.6	64.2	76.9	85.6									
5		18.7	24.3	33.7	37.4	48.6	67.4	56.1	72.9	74.8	97.2								
6		16.6	22.7	32.7	33.2	45.4	65.4	49.8	68.1	98.1	66.4	90.8	99.6						
7		15.1	21.6	32.0	30.2	43.2	64.0	45.3	64.8	96.0	60.4	86.4	90.6						
8		13.0	20.8	31.6	26.0	41.6	63.2	39.0	62.4	94.8	52.0	83.2	78.0						
9		12.9	19.9	30.9	25.8	39.8	61.8	38.7	59.7	92.7	51.6	79.6	77.4						
10		12.1	19.4	30.7	24.2	38.8	61.4	36.3	58.2	92.1	48.4	77.6	72.6				96.8		
15		9.5	17.5	29.4	19.0	35.0	58.8	28.5	52.5	88.2	38.0	70.0	57.0				76.0		
20		8.1	16.4	28.6	16.2	32.8	57.2	24.3	49.2	85.8	32.4	65.6	48.6	98.4			64.8		
25		7.2	15.9	28.4	14.4	31.8	56.8	21.6	47.7	85.2	28.8	63.6	43.2	95.4			57.6		
30		6.5	15.3	27.7	13.0	30.6	55.4	19.5	45.9	83.1	26.0	61.2	39.0	91.8			52.0		
40		5.8	14.9	27.6	11.6	29.8	55.2	17.4	44.7	82.8	23.2	59.6	34.8	89.4			46.4		
60		5.0	14.3	27.1	11.0	28.6	54.2	15.0	42.9	81.3	22.0	57.2	33.0	85.8			44.0		
120		4.2	13.7	26.7	8.4	27.4	53.4	12.6	41.1	80.1	16.8	54.8	25.2	82.8			33.6		

注：表中空白区均大于 100。

计量和资料年代数,就可以从表中查到一个降雹减少的临界百分数;当实际减少的百分数大于此临界值时,就认为防雷效果是经得起检验的,可以认为防雷有效果。如果实际减小百分数小于此临界值,则认为防雷前后雹灾情况没有显著差别,统计到的变化是属于资料的随机起伏,因此就没有理由认为是人工防雷的效果,现举例说明;如昔阳防雷前后有关数据是: $n_2 = 9$, $n_1 = 8$, $s_2/\bar{y} = 0.4$, $s_1/s_2 = 0.7$, 查表 3a, 当 $n_2 = 10$ 时, $\frac{\bar{x}}{\bar{y}} - 1 = -41\%$, 而实际上防雷后雹灾面积比防雷前减少了 46.6%, 所以实际减少值大于表中临界值,因此认为,昔阳在开展人工防雷前后雹灾面积有显著差别,其防雷效果是有统计意义的。

从表 3 中我们可以得到一些有意义的概念,以防雷前有 6 年历史资料,变异系数为 0.5 为例,在这样的地方搞人工防雷,要使防雷效果在统计上有意义,最少要搞二年,如果只搞一年,即使降雹减少到零,也没有统计意义,仍然可认为是随机起伏。如雹灾只减少 60%,就必须搞 5 年以上,如雹灾只减小 40%,就要搞 20 年以上。这就是说,为了使防雷效果在统计上可信,人工防雷工作要坚持一定的年代,开展的时间太短不能很好说明问题,搞的年代长一些,对防雷效果容易作出判断。

表 3b 防雷效果的检验表

$-(\bar{x}/\bar{y}) - 1)$	n_2		3		6		10		15		20		50	
	$s_2/\bar{y}$													
	n_1		0.5	1.0	0.5	1.0	0.5	1.0	0.5	1.0	0.5	1.0	0.5	1.0
1			168	336	109	218	96.0	192	91.0	182	88.8	198	86.0	172
2			107	214	79.2	158	70.2	140	65.8	132	64.2	128	60.6	121
3			85.8	172	67.2	134	59.2	118	55.4	111	53.2	106	49.6	99.2
4			77.6	155	60.2	120	52.6	105	49.0	98.0	47.4	94.8	43.4	86.8
5			71.6	143	55.6	111	48.6	97.2	44.6	89.2	42.8	85.6	39.2	78.4
6			68.2	136	52.4	105	45.4	90.8	41.8	83.6	39.8	79.6	36.0	72.0
7			65.2	130	50.4	101	43.2	86.4	39.6	79.2	37.6	75.2	33.6	67.2
8			62.6	125	48.0	96.0	41.6	83.2	37.6	75.2	35.8	71.6	31.8	63.6
9			60.0	120	46.4	92.8	39.8	79.6	36.2	72.4	34.2	68.4	30.2	60.4
10			59.0	118	45.4	90.8	38.8	77.6	35.0	70.0	33.0	66.0	29.0	58.0
15			55.5	111	41.2	82.4	35.0	70.0	31.2	62.4	29.0	58.0	24.6	49.2
20			53.4	107	39.8	79.6	32.8	65.6	29.0	58.0	26.6	53.2	22.0	44.0
25			52.0	104	38.4	76.8	31.8	63.6	27.4	54.8	25.2	50.4	20.4	40.8
30			51.0	102	38.0	76.0	30.6	61.2	26.6	53.2	24.2	48.4	19.2	38.4
40			50.0	100	36.8	73.6	29.8	59.6	25.4	50.8	22.8	45.6	17.8	35.6
60			49.2	98.4	35.8	71.6	28.6	57.2	24.2	48.4	21.4	42.8	16.0	32.0
120			48.2	96.4	34.8	69.6	27.4	54.8	22.8	45.6	19.8	39.6	13.8	27.6

从表 3 中也可以看到, 防雷前历史资料的代表性和起伏变化程度, 对分析防雷效果有很大影响, 如在相同的 $c = 0.75$, $s_1/s_2 = 1$, 雹灾减少 80% 时, 对于 $n_2 = 3$ 时, 需要搞人工防雷 20 年才能说明问题, 而对于 $n_2 = 20$ 时, 只要搞 3 年防雷就能说明问题, 同样, 对于不同的变异系数, $\frac{\bar{x}}{\bar{y}} - 1$ 的变化也是很大的。以 $n_2 = 10$, 雹灾减少 80% 为例, 当 $s_2/\bar{y} = 0.25$ 时, 搞一年防雷就能说明问题, 当 $s_2/\bar{y} = 0.75$ 时, 要搞 4 年防雷, 当 $s_2/\bar{y} = 1.0$ 时, 需要搞 9 年, 而当 $s_2/\bar{y} = 1.5$ 时, 需要搞 120 年以上才能说明问题, 这在实际中是不可能的, 因此, 对于降雹统计量的变异系数很大的地方, 用一般的统计方法来检验其防雷效果是很困难的。

下面我们利用几次全国防雷会议上提供的一些资料, 应用表 3 来分析一下国内几个地方的防雷效果, 有关的实际数据列在表 4 中。

从表 4 可以看到, 大多数地方开展人工防雷后雹灾减少了, 并且还经得起统计检验, 特别是对于那些历史资料较长, 变异系数较小的地方, 说服力较强。另外, 有些地方虽然防雷后雹灾减少了 90% 以上, 甚至达 100%, 但仍然不能经得起检验, 主要因为变异系数较大, 开展防雷的年代又不够长。

表 4 国内若干地方防雹效果情况

防 雹 地 点	统 计 量	n_2	n_1	s_1/s_2	$s_2/\bar{y}$	$(\frac{\bar{x}}{y} - 1)_*$	$(\frac{\bar{x}}{y} - 1)_*$	是否通过 t 检验
察右前旗老羊圈大队	受灾面积	6	$\frac{10}{(1967-1976)}$	0.1	1.0	-96%	-91%	✓
察右前旗黄茂营大队	受灾面积	8	$\frac{14}{(1963-1976)}$	1.0	1.0	-61%	-70%	×
丰镇麻迷图公社	受灾面积	10	$\frac{5}{(1972-1976)}$	0.1	1.04	-92%	-87%	✓
义县红墙子公社	受灾面积	5	$\frac{9}{(1966-1974)}$	$\leq *1$	$\leq *1$	-99.4%	-93%	✓
灵邱史庄公社	受灾面积	8	$\frac{15}{(1960-1974)}$	$\leq *1$	$\leq *1$	-80%	-70%	✓
固阳大庙公社	受灾面积	5	$\frac{7}{(1968-1974)}$	0.25	0.67	-81%	-58%	✓
灵 邱 县	受灾面积	11	$\frac{13}{(1962-1974)}$	$\leq *1$	$\leq *1$	-63%	-63%	✓
岷 县	受灾面积	12	$\frac{5}{(1970-1974)}$	0.71	0.5	-51.5%	-46%	✓
昔 阳 县	受灾面积	9	$\frac{8}{(1969-1976)}$	0.7	0.43	-46.6%	-42%	✓
照 苏	年降雪日	18	$\frac{1}{(1974)}$	$\leq *1$	0.27	-65%	-48%	✓
民和试验 I 区	年降雪日	9	$\frac{1}{(1974)}$	$\leq *1$	0.35	-80.5%	-67%	✓
民和试验 II 区	年降雪日	8	$\frac{1}{(1974)}$	$\leq *1$	0.38	-89.2%	-73%	✓
冕 宁	年降雪日	15	$\frac{2}{(1973-1974)}$	$\leq *1$	0.66	-46.6%	-81%	×
冕 宁	年成灾降雪日	15	$\frac{2}{(1973-1974)}$	$\leq *1$	0.97	-100%	< -100%	×

注: * 表示估计值。

三、几点看法

通过以上分析讨论,可以得到以下看法:

(1) 降雪日数和雹灾面积可以看作是随机变量,其分布可用正态分布来近似,对于年降雪日来说,用泊松分布来近似更好些,因此用统计方法来分析防雹效果是适宜的,而且,人工防雹后实际的雹灾减少量一般比较大,可能较容易地用统计方法来区分自然变化和人工影响,而对于人工降水的增加来说,一般只有 10% 左右,要从统计上区分自然变化和人工影响就较困难。

(2) 通过资料分析和统计检验看到,国内不少地方的防雹效果在统计上是可信的,说明我国人工防雹是取得一定效果的(对于以县为单位来说,防雹后受灾面积减少约 50% 左右)。而对另外一些地方来说,虽然雹灾减少了,但在统计上还不能说明问题,需要进一步努力。

(3) 从时间序列上统计对比防雹前后的变化可以说明一些问题,但也有不足,特别是对于历史资料年代不长,而变异系数又较大时,要求搞防雹年代较长才行,有时甚至无法满足要求。因此,有条件时,应同时考虑随机作业方案,以提高统计的灵敏度和可信度。