

非随机化人工降雨试验效果 评价方法研究

曾光平 郑行照 方仕珍 李顺来

(福建省气象科学研究所, 福州 350001)

提 要

本文在分析各种非随机化人工降雨试验的基础上提出一种新的评价方案——区域控制模拟试验。统计数值模拟分析及采用本方案对古田人工降雨试验再评价、对河南、江西、山东等省飞机、高炮人工降雨作业效果评价均表明, 本方案比其它随机试验方案能较客观、准确评价作业效果。

文中还采用统计数值模拟方案探讨区域雨量相关、作业效果对区域控制模拟方案统计结果的影响。

关键词: 非随机化; 人工降雨; 区域控制模拟试验; 效果评价

一、引 言

随着生产发展, 人工影响天气学科越来越引人关注。数十年来, 国内外云物理工作者进行大量有设计的外场试验研究, 试图从巨大的降水变异背景上检出人工播云效果。根据外场作业和科研需要, 人们设计出两大类试验方案——随机化试验和非随机化试验。

随机化试验从原则上可以做到完全适用于随机抽样原则, 可以根据随机抽样理论, 定量地检验效果并指明其可信度。但其试验周期相当长, 如福建省古田试验长达 12 年, 才在 $\alpha < 0.05$ 上以 90% 检出率检出 20%—30% 的催化效果^[1], 以色列进行长达 15 年试验得出 15% 增雨效果。可见随机化试验方案不适于抗旱蓄水外场作业, 仅适用于研究。

人工降雨抗旱救灾, 水库蓄水外场作业中, 均采用非随机化作业方案, 非随机化试验通常有序列试验、区域对比试验和历史回归试验。这些方案假设自然降水在时间、空间上是平稳的, 但是自然降水在时间、空间上分布存在巨大差异, 使得这些假设条件在实际上难以满足, 因而上述三种方案不能客观、定量地估计、评价人工影响的效果。

1992 年 10 月 6 日收到, 1993 年 3 月 5 日收到修改稿。

本文在分析各种非随机化试验方案的基础上,提出一种新的试验方案——区域控制模拟试验,该方案采用剔除——逼近方法从时空分布上存在巨大差异的长序列历史雨量的资料中寻找与作业期降水特征(降水分布概型、均值、方差、相关性等)相似型。然后,采用统计数值模拟方法建立影响区自然降水预报方程,对效果进行统计评价。此外本文还采用统计数值模拟方法分析自然降水变异,区域雨量相关对模拟结果的影响,同时还探讨了试验周期。

二、区域控制模拟试验方案

该方案是在影响区上风方设计两个或两个以上对比区(如图1,其中对比区1为主对比区)。三个区雨量资料分别记为:

历史资料	影响区	Y'_{ns}	
	对比区1	X'_{1ns}	(n 组样本)
	对比区2	X'_{2ns}	
作业期资料	影响区	Y_s	
	对比区1	X_{1s}	(J 组样本)
	对比区2	X_{2s}	

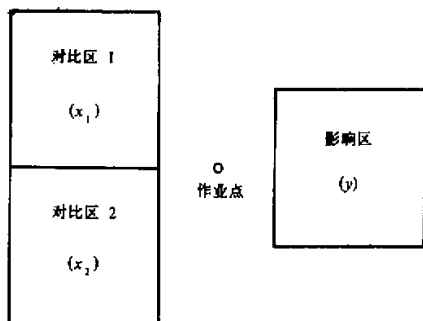


图1 区域控制模拟试验试验区设置图

区域控制模拟试验效果统计分析步骤如下:

1. 分析作业期两个对比区及历史资料中三区雨量特征:均值、相关性 & 雨量分布的概型。

2. 根据上述特征值对对比区1和对比区2的历史资料进行剔除,使保留下的对比区X1和对比区X2的历史雨量资料特征值与作业期两个对比区的雨量特征值相拟合,设剔除后历史资料样本数为 K 组,则

影响区	Y_{ns}	
对比区1	X_{1ns}	(K 组样本)
对比区2	X_{2ns}	

3. 对经上述处理后的三个区历史样本雨量资料采用统计数值模拟方法,建立影响区与主对比区1的自然雨量相关系数 $R_{(y-x1)}$,回归系数 $a_{(y-x1)}$ 、 $b_{(y-x1)}$ 与两个对比区这些参量的多元回归方程,即

$$\begin{cases} R_{(y-x1)} = R_0 + R_1 \cdot R_{(x1-x2)}, \\ a_{(y-x1)} = a_0 + a_1 \cdot a_{(x1-x2)} + a_2 \cdot a_{(x2-x1)}, \\ b_{(y-x1)} = b_0 + b_1 \cdot b_{(x1-x2)} + b_2 \cdot b_{(x2-x1)}, \end{cases} \quad (1)$$

式中 R_0 、 R_1 、 a_0 、 a_1 、 a_2 、 b_0 、 b_1 、 b_2 分别为 $R_{(y-x)}$ 、 $a_{(y-x)}$ 、 $b_{(y-x)}$ 的多元回归方程的回归系数。

4. 将作业期对比区 1 和对比区 2 雨量特征值 R 、 a 、 b 代入(1)得出作业期影响区和主对比区相关系数和回归系数的估计值 $\hat{R}_{(y-x)}$ 、 $\hat{a}_{(y-x)}$ 和 $\hat{b}_{(y-x)}$ 。从而模拟出作业期影响区自然降雨量和主对比区自然降雨量一元线性回归方程:

$$\hat{y}_s = \hat{a}_{(y-x)} + \hat{b}_{(y-x)} \cdot x_{1s}. \quad (2)$$

5. 将作业期主对比区自然降雨量代入(2), 求出作业期影响区自然降雨量的估计值 $\hat{y}_s$ 。

6. 催化作业效果

$$\Delta R = \frac{y_s - \hat{y}_s}{\hat{y}_s},$$

效果的显著性采用 t 检验法检验。

影响区域控制模拟方案效果评价准确性因子很多, 下面我们利用古田人工降雨随机试验^[1]资料采用本方案进行模拟统计。通过分析模拟结果与随机试验结果的差异, 来探讨影响区域控制模拟方案统计结果准确性及其因子的作用规律。为此, 在古田试验区原对比区 (称对比区 1) 正南另设一个面积相同的对比区 2。该区同样位于作业点上风方, 不受催化剂影响。

三、各种评价方案的比较

人工影响天气试验可分为三种不同的类型:

类型 1 重点旨在增加降雨量和减少冰雹, 属于纯业务作业项目。

类型 2 是人工影响天气的探索性的纯研究性项目。

类型 3 是业务和研究相结合的项目, 有时主要是研究性的。

由于各种类型目的不同, 它们设计方案, 效果评价方法也各不相同, 第一类采用非随机化方案; 第 2、3 类一般采用随机化方案。尽管各类试验方案及评价方法差异极大, 但实际上是通过各种方法, 或作不同假设, 采用统计方法估计出作业期影响区的自然降水, 并与催化后实测雨量相比较得出催化效果, 同时检验效果的显著性。本节将讨论各种方案的物理基础, 分析方法的特点及局限性。

1. 随机试验方案

随机试验把适于催化的机会分成二组, 按照随机化规则决定一组催化, 一组不催化, 留作对比。原则上讲, 随机试验可以做到完全适合于随机抽样原则, 因而可以根据随机抽样理论的一套成熟的运算方法, 定量地检验效果并指明其可靠程度。

该方案实际上认为根据随机化原则确定的催化与不催化二组机会中, 除了作业这一因素外, 其它因子对降水贡献没有系统性的误差。通过比较, 如果在这二组机会中, 观测到雨量有明显的差异, 就可以归因于作业。但是由于自然降水的巨大变异性, 使得这一假设并非在任何情况下都能满足。研究表明^[2,3]: 降水自然起伏影响随机试验效果统

计结果。功效分析表明：这种影响随样本数的增大而减小，即检出催化效果的几率随着样本容量和增雨效果的增大而增大。因此，在试验设计时应应对功效进行分析，确定试验周期，如古田人工降雨试验，要在检出率为 90% 条件下检出 20%—30% 增雨效果 ($\alpha < 0.05$) 需样本数为 250 个。此外，功效分析结果还表明：不同试验方案，功效差异极大，单区随机试验最低，随机交叉回归试验最高。

表 1 是利用古田人工降雨试验区 1975—1988 年 4—6 月份随机试验非催化单元资料，对随机试验各种方案进行功效模拟试验，计算所需的试验样本数。

表 1 不同随机试验方案所需样本数的统计值

方 案	80.0%		90.0%	
检 出 率 增 雨 效 果	20.0%	30.0%	20.0%	30.0%
单区随机试验	> 320	> 320	> 320	> 320
区域随机交叉试验	90	60	100	70
区域随机回归试验	200	150	250	200
区域随机交叉回归试验	70	50	90	60

上述分析表明：(1) 随机试验不同方案的功效差异极大；(2) 认为“随机试验中催化不催化二组样本中，除了作业这一因素外，其它因子对降水贡献没有系统的误差”是有一定条件的。引因，除进行功效分析外，在确定适宜作业机会时，还应尽可能周详地考虑降水宏、微观条件，以制定严格作业标准。

2. 非随机化试验方案

1. 序列试验

序列试验以历史平均值作为试验期自然降水量的估算值。它假设试验区的自然降水量在历史上是平稳的时间序列。

效果检验可采用以下二种方法：

方法 1：雨量资料服从正态分布，当历史资料样本数 $n > 30$ 时，取统计量

$$u = \frac{R - \bar{R}}{S_R},$$

其中 $\bar{R}$, S_R 为历史降水资料的均值及标准差， R 为催化后降水量。当历史资料年份较长 ($n > 30$ 时)， u 近似服从正态分布。

方法 2：雨量资料分布未知情况下，采用秩和检验法。

降水时空分布巨大差异，使得该方法物理意义上存在局限。用历史平均雨量代替作业期自然雨量必然引起假效果。这种假效果随所选取的历史资料年限不同有极大差异。特别是干旱年份自然降水量明显偏低，使得假设明显地缺少物理基础。从统计学上看，自然降水标准差 S_R 很大，要得到统计上显著 ($\alpha = 0.05$) 的增雨结果，所需的增雨值要远大于人工催化可能的增雨效果，可见这种方法灵敏度低。

上述分析表明, 序列试验不宜用来检验人工降水的效果。

2. 区域对比试验

以同期对比区自然降雨量作为影响区自然雨量的估算值, 假设试验期的雨量的空间分布统计上是均匀的。从物理意义上来说这种假设是不合理的。显然由于降水在地理区域上差异极大, 以及对比区选择上不可避免存在主观性, 使得统计检验失去意义。这表明这种方案难以科学、客观评价人工降雨效果。

3. 区域回归试验

利用对比区自然降雨量作为预报因子, 对试验期目标区雨量进行统计推断。具体统计方法是: 借助于一个或一个以上的对比区(控制区), 根据历史资料建立目标区与对比区历史雨量回归方程, 然后利用历史回归方程从试验期对比区的雨量来估计目标区的自然雨量。这种方案显然没有假设降水在时、空分布上是平稳的, 比序列试验和区域对比试验前进了一步, 但是仍假定试验期影响区和对比区雨量相关关系与历史上同期的区域雨量相关性相同。然而, 试验期的天气形势与建立回归方程的历史时期的形势可能很不相同, 气候相似的依据不充分, 因而作业期 2 区雨量相关关系与历史上同期 2 区雨量相关关系相同假设难以满足, 这就影响了该方案物理基础。此外, 历史资料长短以及对对比区选择等都会影响统计检验的结果; 导致统计结果不科学、不客观。

4. 区域控制模拟试验

该方案承认降水时空分布上存在巨大不平稳性, 通过在长序列历史资料中寻找与作业期降水相似型的资料, 对作业期影响区自然降水进行估计。

具体方法是采用逐步剔除-逼近方法, 从时空分布上存在巨大差异的长序列历史资料中寻找与作业期降水特征(降水分布概型、均值、方差、相关性等)相似型。然后采用统计数值模拟方法对效果进行评价。逐步剔除-逼近的步骤如下:

1. 设历史样本 G 容量为 $N(g_1, g_2, \dots, g_i, \dots, g_n)$, 分别按顺序剔去一个个个体, 得到 n 个样本容量为 $(N-1)$ 的新的历史样本, 记作 $G_1', G_2', \dots, G_i', \dots, G_n'$ (其中 G_i' 为从 G 中剔去第 i 个个体而得到的新历史样本)。

2. 分别统计 n 个新的历史样本 $G_1', G_2', \dots, G_i', \dots, G_n'$ 的降水特征值(分布概型、均值、方差及相关性)。并与作业期降水特征值进行比较, 找出差别最小的样本 $G_i'(g_1, g_2, \dots, g_{i-1}, g_{i+1}, \dots, g_{n-1})$ 。至此, 完成第一步剔除-逼近。

3. 对新的历史样本 G_i' , 重复 1、2 步骤完成了第二步剔除-逼近, 得出与作业期降水特征更接近的新的历史样本 G_i' 。

4. 继续上述剔除-逼近步骤, 直至得出新的历史样本的降水特征与作业期降水样本的特征值之间的相对差异(ξ)小于规定值。设经过 k 次剔除-逼近则得出新的历史样本为 $G^k(g_1^k, g_2^k, \dots, g_{n-k}^k)$, 其样本容量为 $(n-k)$ 。

相对差异 ξ 值是根据对效果检验准确度要求而确定的。 ξ 值越小, 剔除-逼近后得到新的历史样本与作业期降水特征愈相近, 效果统计结果准确度亦愈高。

为保证统计分析有一定样本数, 规定经逐步剔除-逼近否的样本数下限为 30。如小于 30, 就认为无法从所给的历史样本中找到与作业期降水特征相似的样本, 即无法通过本方法评价效果。

上述逐步剔除-逼近运算, 可根据要求先确定 ξ 值, 全部由计算机完成。一旦 ξ 值

确定, 对同一组作业期和历史资料剔除一逼近的结果是唯一的, 不存在人为的随意性。

四、不同雨量相关对区域控制模拟方案
统计结果的影响

区域控制模拟试验也是通过对地面雨量采用统计学方法来评价效果, 因此, 影响区与对比区域雨量相关将直接影响到效果的评价。下面将讨论不同雨量相关对本方案统计结果的影响。统计时 R 分别取 0.9551、0.9013、0.8826、0.7954、0.5761、-0.1690 六组。表 2 列出 $\theta = 20.0\%$ 时, 这六种 R 情况下, 效果统计量 E 的各个特征量, 图 2 是 $\theta = 0.0\%$ 时不同区域雨量相关下, 增雨效果统计量频率分布图。图 3 是标准离差, 平均相对误差($\bar{\eta}$), 相对误差 $\leq 20\%$ 频率(Q)及变异系数(C 、 V)随区域雨量相关系数变化图。

分析表明:

1. 区域控制模拟方案统计结果与区域雨量相关的关系极为密切。增雨效果统计量的统计特征量、标准离差、平均相对误差均随相关系数的提高单调下降, 表明本方案统计出的结果的误差随 R 增大减小, 越来越接近真值, 且变异系数也随 R 增大而减小。当 R 小时($R < 0.70$)变化缓慢; 随着 R 增大, 变化加速。

表 2 不同雨量相关时效果统计量 $E(\%)$ 的统计特征量($\theta = 20.00\%$)

R 特征量	-0.1690	0.5761	0.7954	0.8826	0.9013	0.9551
均值(%)	21.23	20.90	19.42	19.53	19.75	19.97
标准离差(%)	17.82	10.68	5.73	4.04	3.54	1.95
$Q(\%)$	23.0	32.5	62.1	87.5	91.1	97.2
$\bar{\eta}(\%)$	6.15	4.50	2.90	2.35	1.25	0.15
C 、 $V(\%)$	83.39	51.10	27.65	21.20	15.95	9.76
α	<0.05	<0.05	<0.05	<0.05	<0.05	<0.05

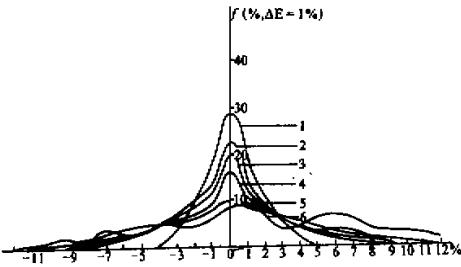


图 2 $\theta = 0.0\%$ 不同雨量相关系数效果统计值频率分布

1. $R = 0.9551$, 2. $R = 0.9013$, 3. $R = 0.8825$, 4. $R = 0.7954$, 5. $R = 0.5761$, 6. $R = 0.1690$

2. 降水自然变异的影响，也是随着 R 的增大而减小。从图 2 可以看出，当 $R \leq 0.7$ 时， $\theta = 0.0\%$ ，效果统计值 E 极为分散，而且出现波动，随着 R 的增大频率分布曲线的波动消失并向效果统计值 $E = 0.0\%$ 集中，即降水自然变异的影响越来越小。
3. 相对误差 $\leq 20\%$ 的频率随 R 增大而增大，当 $R > 0.70\%$ 时提高极快，若以 $\theta = 20.0\%$ ，相对误差 $\leq 20\%$ 的检出率为 80.0%、90.0%为条件，则区域雨量相关应分别为 0.85 和 0.90。

五、不同增雨效果对区域控制模拟方案统计结果的影响

由于降水时空差异，使得本方案统计出的增雨效果统计值与实际增雨值存在误差，这种误差不仅与样本数、区域雨量相关有关，而且还与实际增雨值有关。本节分析历史样本为 300，作业样本为 120 实际增雨值不同时本方案增雨效果统计量的变化规律。表 3 列出增雨效果统计量 E 的特征量。图 4 是 E 的变异系数及相对误差 $\leq 20\%$ 频率随增雨值 θ 变化图。

表 3 不同增雨值时效果统计量 E (%) 的统计特征量

$\theta(\%)$	0.0	10.0	20.0	30.0	50.0	70.0	100.0
特征量							
均值 (%)	0.024	10.026	20.028	30.031	50.036	70.040	100.047
标准离差 (%)	3.45	3.80	4.15	4.49	5.18	5.88	6.91
C_v (%)	143.75	38.10	20.72	14.95	10.35	8.40	6.91
$\bar{q}$ (%)		0.26	0.14	0.10	0.07	0.06	0.05
Q (%)		39.3	87.5	92.2	97.1	99.5	100.0
α	> 0.10	≈ 0.10	< 0.05	< 0.05	< 0.05	< 0.05	< 0.05

分析表 3、图 4 可以得出：

1. 区域控制模拟方法统计的效果统计量与增雨值关系密切，随着增雨值增大效果统计量越接近真值，平均相对误差越小，且变异系数也随之减小。
2. 效果统计量 E 的特征量随增雨值不是呈线性变化，在 $\theta = 0\% - 30\%$ 间变化迅速， $\theta > 30\%$ 变化缓慢。
3. 若以相对误差 $\leq 20\%$ ，检出率为 80%、90%作条件，则所需增雨值为 20%—30%。

六、不同试验方案统计结果的比较

通过分析各种试验方案统计出的增雨效果的准确性（以相对误差或相对差异来表示）来评比各方案的优劣。为此，我们用区域控制模拟方案和常用的非随机试验方案

(序列试验、区域对比试验、历史回归试验)对古田随机试验结果进行再评价,同时还对江西、河南、山东等省高炮、飞机抗旱作业效果进行分析、并作相互比较。

1. 古田人工降雨试验效果评价

表 4 列出采用不同方案对古田随机回归试验效果进行评价的结果,统计时,统计量取雨量本身。

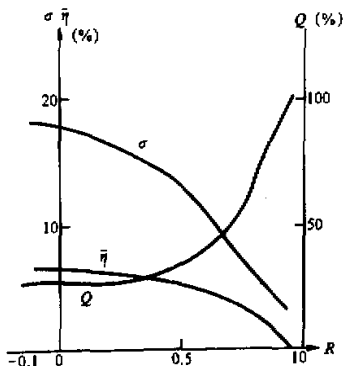


图 3 σ 、 $\bar{\eta}$ 、 Q 随区域雨量相关系数变化

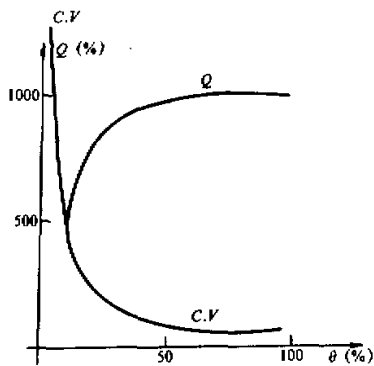


图 4 C 、 V 、 Q 随不同增雨值 θ (%)变化

表 4 古田人工降雨试验各种方案评价结果

项 目 \ 试验方案	随机回 归试验	区域控制 模拟试验	历史回 归试验	递减对 比试验	序列试验
绝对增雨(mm/3h)	1.55	1.50	1.26	1.12	3.21
相对增雨(%)	23.05	22.07	17.96	15.68	65.45
α	<0.01	<0.01	<0.05	<0.05	<0.05
J (%)		-4.25	-22.08	-31.97	179.61

注: J 为各方案统计的相对增雨与随机试验统计的相对增雨的相对差异。

比较表中各种方案统计结果可以看出区域控制模拟试验效果统计结果与随机回归试验结果最为一致,其相对差异仅-4.35%,历史回归试验次之为-22.08%,序列试验最差高达 179.61%。

2. 河南省(1988—1989年)飞机降雨效果评价

对河南省 1988—1989 年飞机人工降雨作业(作业次数 14 次)效果评价的结果列于表 5。

各种方案统计结果比较表明:

1. 不同方案产生假效果相差极大, $\theta=0.0\%$ 时(1988—1989 年非作业样本)区域

表5 河南省1988—1989年飞机人工降雨效果评价

样本分类	试验方案		区域控制 模拟试验	历史回 归试验	区域对 比试验	序列试验
	项 目					
1988—1989 非作业 样本(7个)	A—B1	相对增雨(%)	4.15	29.46	52.37	-11.03
		α	>0.10	<0.05	>0.05	>0.05
	A—B2	相对增雨(%)	2.53	34.37	71.87	-23.75
		α	>0.10	<0.05	<0.05	>0.05
1988—1989 作业样本 (14个)	A—B1	相对增雨(%)	25.05	17.55	-1.71	51.96
		α	<0.01	$\approx$ 0.05	>0.05	>0.05
	A—B2	相对增雨(%)	20.10	8.68	-27.12	33.93
		α	<0.01	>0.05	>0.05	>0.05

注：A、B1、B2分别为影响区，对比区1和对比区2。

控制模拟试验最小为2.53%—4.15%，且显著度很低($\alpha>0.1$)，即在检验水平 $\alpha=0.05$ 下认为假效果无统计意义。区域对比试验最大为52.37%，历史回归试验次之。

2. 历史回归试验、区域对比试验、序列试验方案统计结果产生的假效果很大，且明显受对比区的选择和历史资料长短的影响，因而无法在这些方案产生的假效果“噪声”中检出人工影响的效果。

3. 区域控制模拟方案，产生假效果很小且稳定，同时无统计意义。因此可以在降水自然变异背景上检出人工影响效果，对1988—1989年作业样本统计结果也表明，该方案统计结果稳定，而其它三种方案结果本身受对比区影响极大且不稳定。

3. 江西、山东省人工降雨效果评价

采用区域控制模拟试验方案对江西柘林水库地区1979—1982年高炮人工降雨效果及山东省1989—1991年飞机人工降雨效果进行分析，结果表明：区域控制模拟方案比其它非随机试验方案能客观准确地评价效果，产生的假效果极小；山东省飞机人工降雨试验统计结果产生假效果为0.68%($\alpha>0.20$)，无统计意义；江西省柘林水库地区高炮人工降雨试验统计结果产生的假效果为-4.27%($\alpha>0.20$)，无统计意义。这些起伏均远小于可能的人工增雨效果，即可以在自然降水变异的背景上检出人工增雨效果。

七、结 论

1. 统计数值模拟结果表明：区域控制模拟方案较其它非随机化试验（历史回归试验、区域对比试验、序列试验）能客观、定量评价非随机化人工降雨作业效果。

2. 历史样本长短、作业样本长短、区域雨量相关系数、增雨效果都直接影响到区域控制模拟试验统计结果。随着样本数增大、区域雨量相关提高、增雨效果提高，统计结果越接近实际增雨值、相对误差越小、检出率越大。

3. 采用区域控制模拟方案对古田人工降雨试验再评价表明，对122次催化作业相

对增雨 22.07%，与随机试验结果比较，相对差异仅-4.25%。如采取本方案试验周期可以缩短一半。

4. 采用区域控制模拟方案对山东、河南、江西飞机(高炮)人工降雨外场作业效果进行评价表明，产生的假效果远小于其它各种非随机试验方案，可以在降水变异“噪声”上检出作业效果。

参 考 文 献

- [1] 曾光平, 1991, 1975—1986 年古田水库人工降雨效果总分析, 大气科学, 15(4), 97—108.
- [2] 曾光平, 肖锋, 1983, 降水自然变异对人工降水效果检验的影响, 气象科学, 第 2 期, 79—87.
- [3] 曾光平, 1993, 人工降雨试验效果检验的统计模拟方法研究, 气象学报, 51, 第 2 期, 241—247.

Research on the Method of Evaluating the Efficiency of the Non-Randomized Artificial Precipitation Experiments

Zeng Guangping, Zheng Xingzhao, Fang Shizhen and Li Shunlai

(Meteorological Institute of Fujian, Fuzhou 350001)

Abstract

Based on various non-randomized artificial precipitation experiments, a new method of the evaluating the efficiency of non-randomized artificial precipitation, Regional Control Simulation Experiment Scheme is presented. Analyses of the statistical numerical simulation method and the reevaluated results of in the Gutian Reservoir area, as well as the evaluated results of artificial precipitation in Henan, Shandong, Jiangxi provinces by this method, indicate that the Regional Control Simulation Experiment Scheme can more objectively and more accurately evaluate the results of artificial precipitation than other non-randomized experiments. The influence of the regional rainfall correlation and seeding results on the statistical results of the Regional Control Simulation Experiment Scheme is also discussed by the statistical numerical simulation method.

Key words: non-randomized; artificial precipitation; regional control simulation.