Statistical Evaluation of Weather Modification Programs

Compare Seeded Cloud to Unseeded Clouds

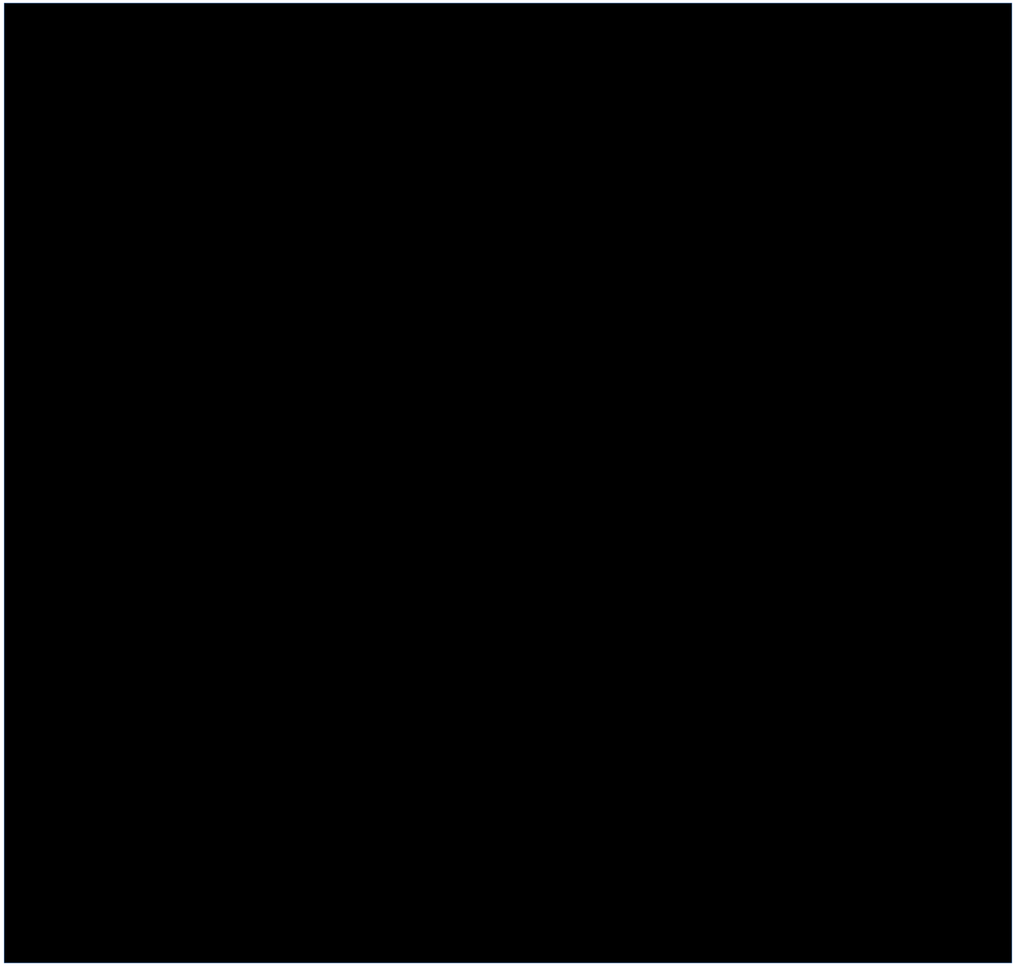
- Skeptics questioned whether the two classes of clouds were the same.
- Were the seeded clouds “better” than unseeded clouds?
- Randomization
  - Randomly seed and not seed clouds and compare the two populations.
- Blind Experiments
  - The people carrying out the experiments could not tell whether they were seeding or not seeding.

Initial Results – Statistical Studies

- The results were not nearly so spectacular as people were hoping.
- Many experiments showed no effects on precipitation.
  - There are a lot of ways to go about seeding a cloud.
  - Many of these seeding techniques may not be effective.
  - How can we tell how to optimize the seeding techniques?

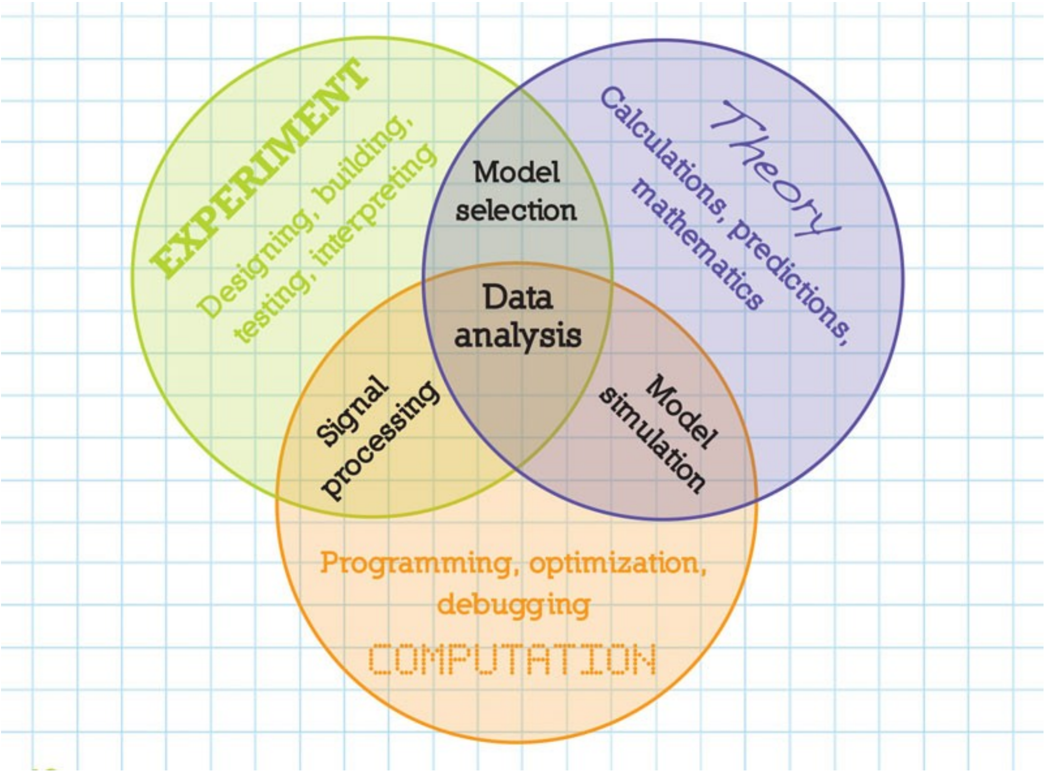
Black Box Experiments

- Most of the early seeding experiments were of the “Black Box” type.
- There was an input (seeding) and an output (precipitation)
- There was no knowledge of what went on in-between.

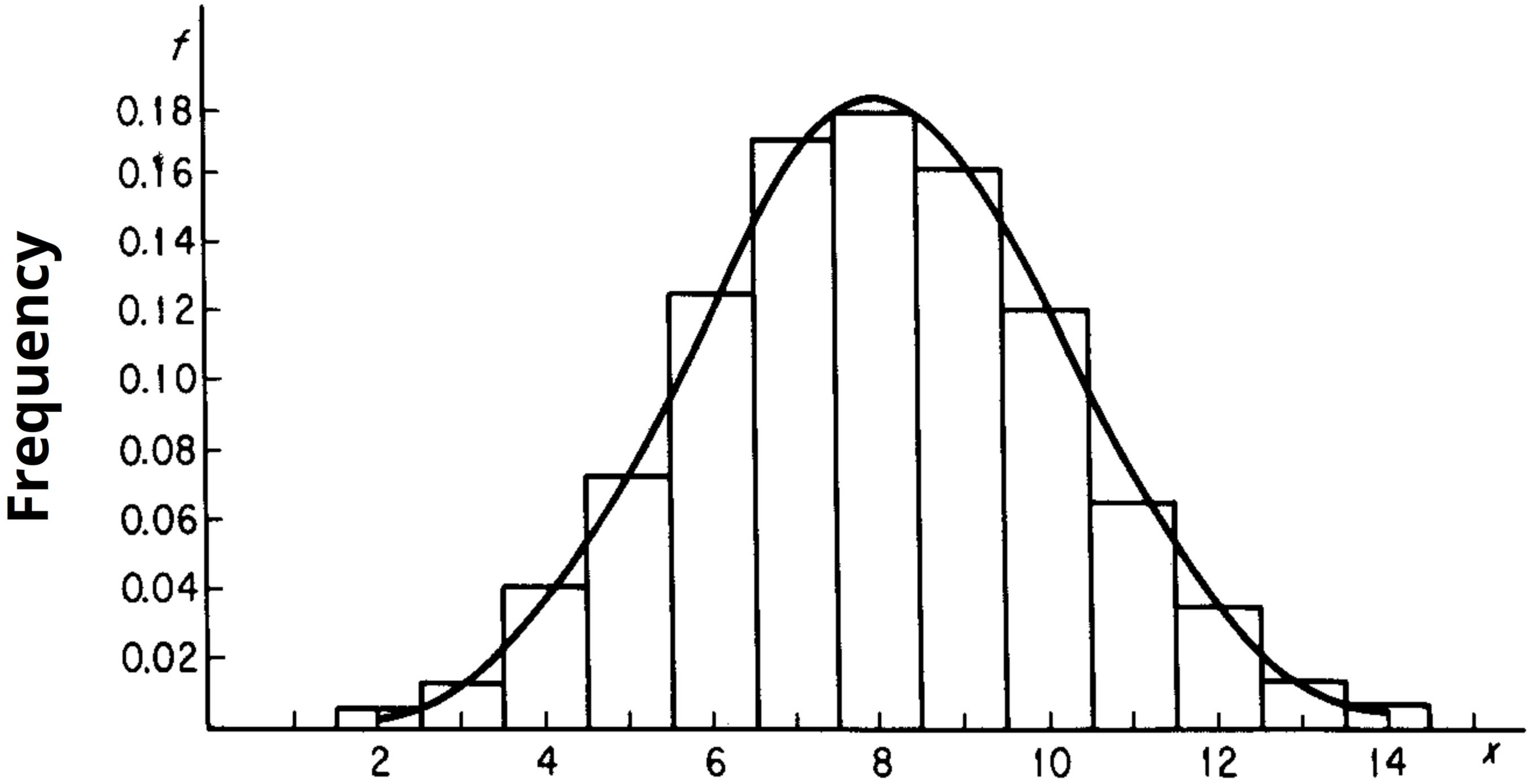


Physical Experiments

- Rather than just looking at the input and output, the intermediate steps were examined.
- If there was a step that was missing, the reason(s) could then be investigated.

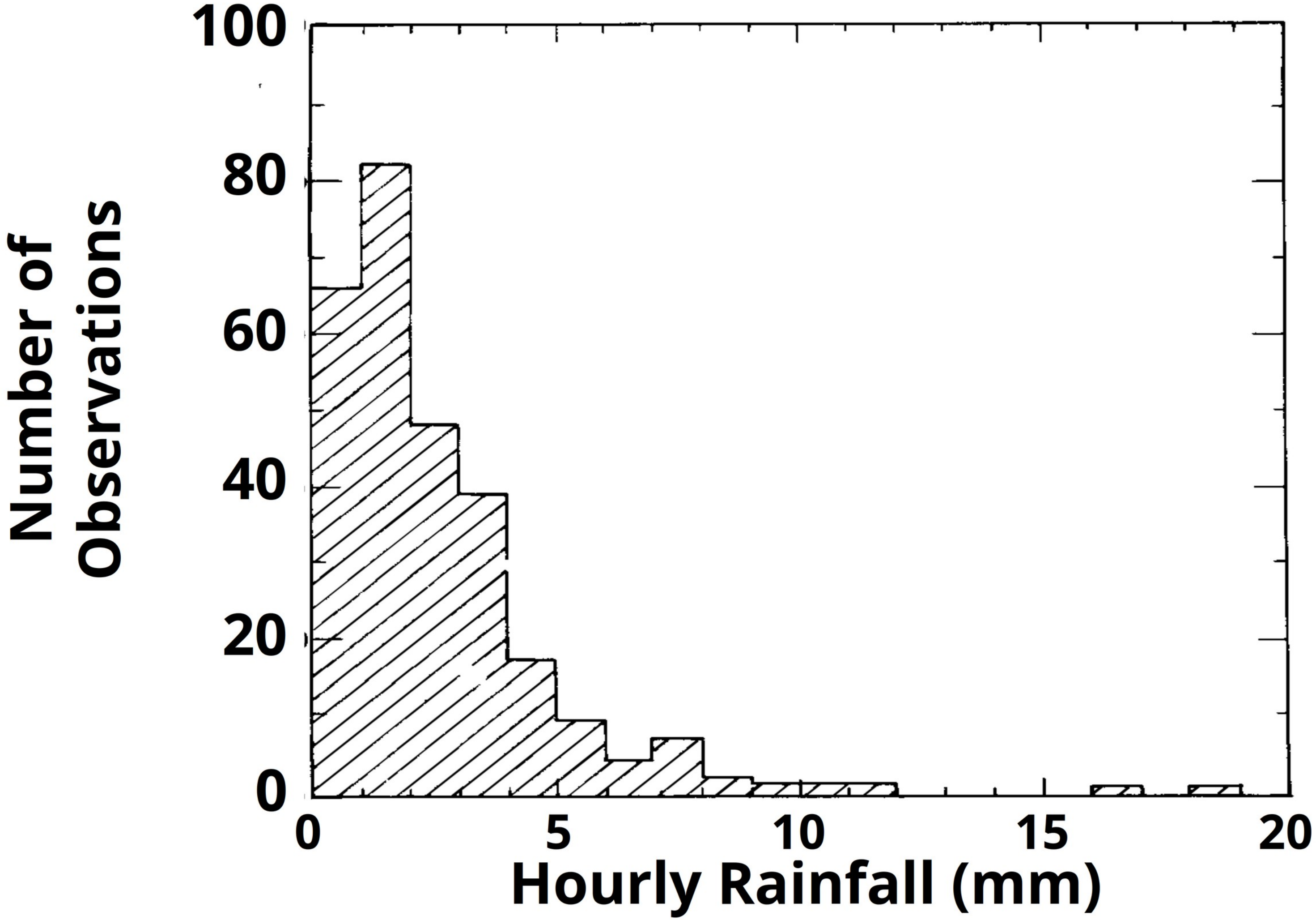


Statistical Analysis - Normal Distribution



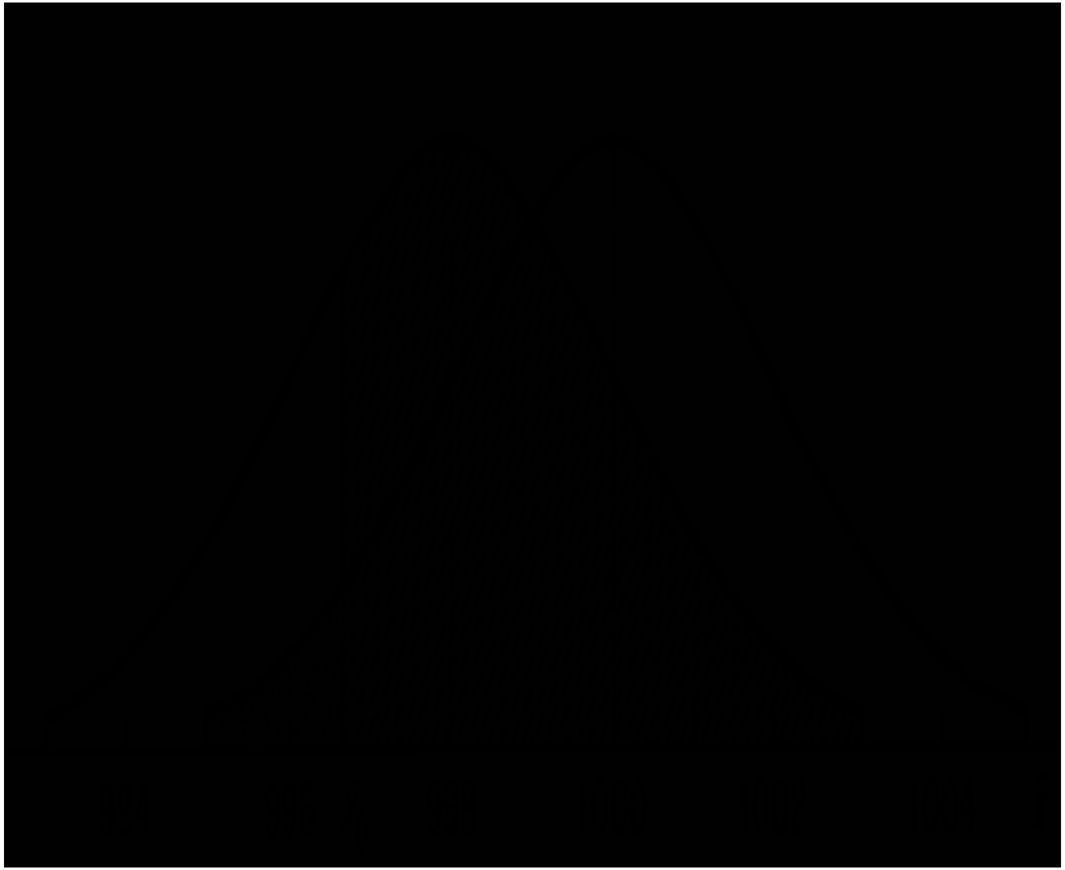
Normal Curve Approximation to the Binomial

Example of Actual Rainfall Distribution



Hypothesis Testing

- Generally, we have two sets of measurements:
  - The unseeded set
  - The seeded set
- Are there any differences between these two sets of measurements that are not the result of random chance?

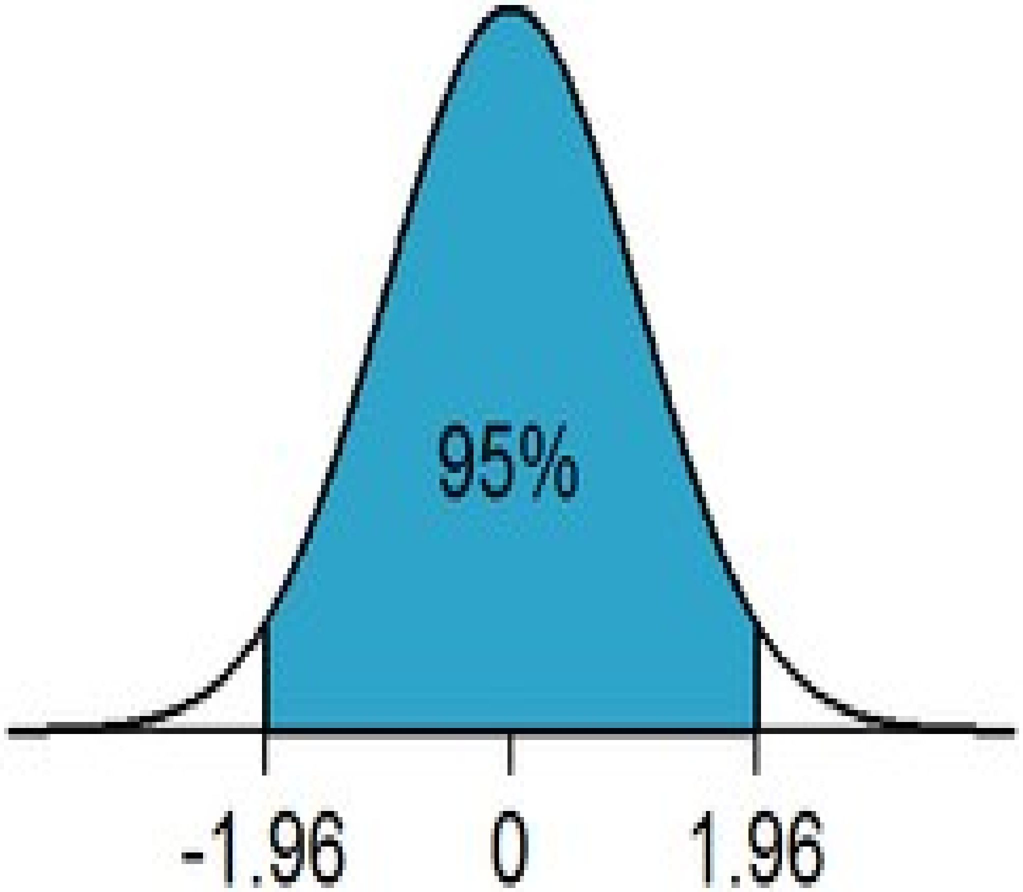


Formulation of the “Null Hypothesis”

- The null hypothesis is generally stated in a way that it will be rejected if there is a likely difference between the two samples.
- An example of the null hypothesis is “The two distributions are the same.”
- Assumption is made that the distribution is sampled Randomly

Statistical Significance Level

- As a rule, the acceptable confidence level should be established prior to the experiment.
- Set the confidence level at 5% or lower, means that we will reject the null hypothesis if there is only a 5% chance that the two data sets came from the same population.



Sample Size Required

- For large effects as compared to natural variability, number would be small
- For small effects relative to natural variability, number needed is large!
- Also depends on how “good” the samples are.
- Operational program seed everything; hence, n unseeded cases available for analysis.

Historical Regression Analysis

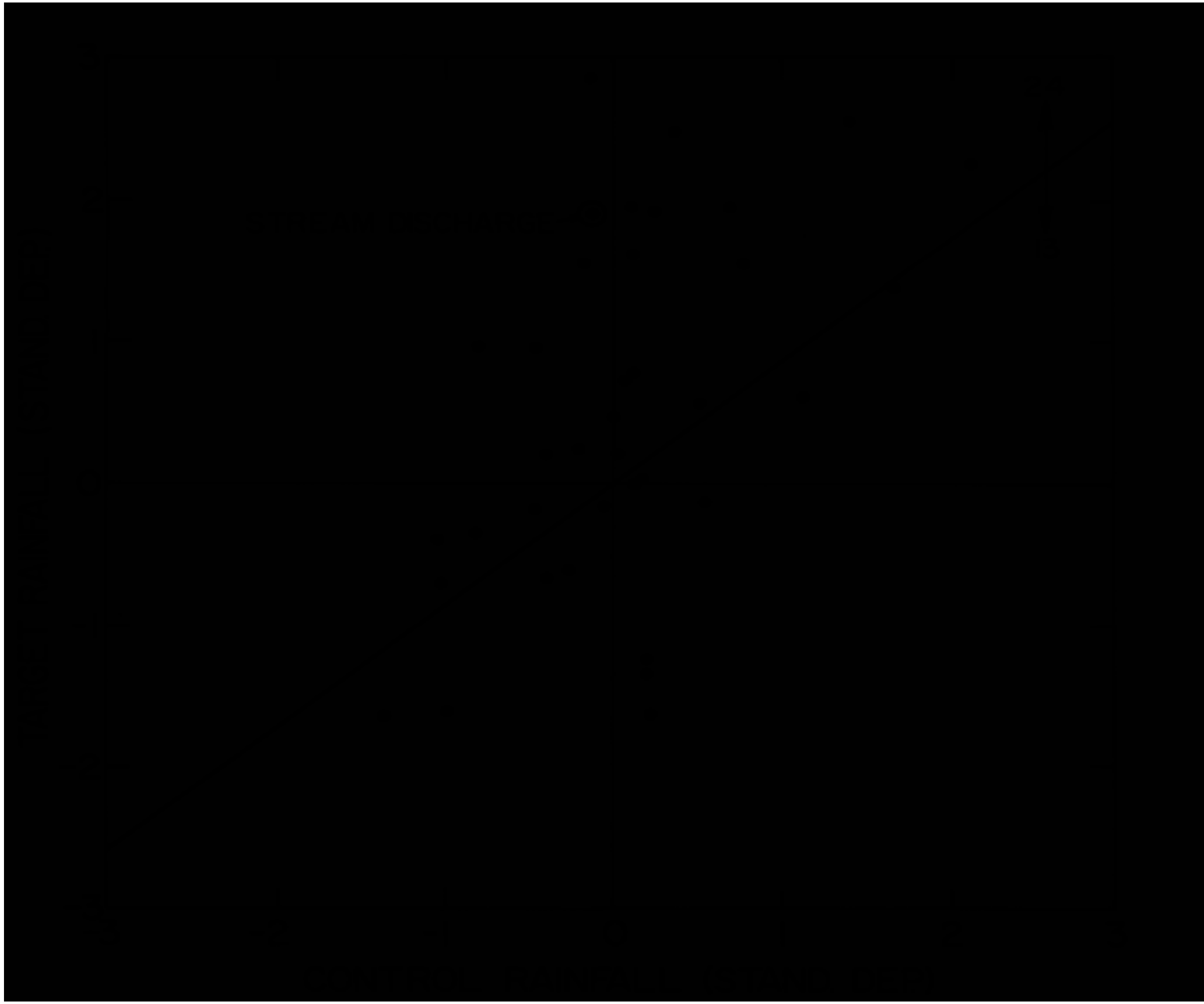
- Uses historical data as the unseeded data set.
- Uses some sort of covariate as a predictor of the precipitation.
  - Covariate can be anything that has a reasonably good correlation with the actual precipitation.
  - The covariate has often been a “control area” where there is no seeding taking place.

Historical Regression Example

- Scatter diagram showing normalized monthly target and control rainfall amounts for operational cloud seeding projects in western U.S.

- There are 24 entries above the historical regession line and 13 below it.
  - Indicating a increase in precipitation.

H. C. S. Thom (1957b)



Problems with Historical Regression

- The samples have not been taken randomly.
- Time may be changing the relationship between the covariate and the actual response variable.
  - Typicality not accepted, if not randomized.
  - Therefore, very difficult to evaluate operational programs.

Single Area Experiments

- The single area type of experiment uses randomized seeding in one target area.
  - Storm units (perhaps days) are selected on a random basis and either seeded or not seeded.
  - At the end of the program, the response variable from the seeded cases is compared to the that from the unseeded cases.
  - This tends to be a rather inefficient means of conducting an experiment.

Target-Control Experiments

- The single area type of experiment uses randomized seeding in one target area.
  - Storm units (perhaps days) are selected on a random basis and either seeded or not seeded.
  - At the end of the program, the response variable from the seeded cases is compared to the that from the unseeded cases.
  - This tends to be a rather inefficient means of conducting an experiment.

Randomized Cross-over

- Two areas are selected and for each “storm”, one of the areas is selected at random to be seeded and the other is left as a control.
  - The seeded areas are compared to the unseeded areas.
  - This is the most efficient means of conducting a randomized experiment.

Use of Ratios for Evaluation

- What about using the ratio of rainfall from seeded (S) an unseeded (NS) events?
- Ratio = Rain(S)/Rain(NS)

Seeded Cases	Non-seeded Cases	Ratio (S/N)
0.01”	0.10”	0.1
.10”	.01”	10
Total .11”	**Total** .11”	**Mean Ratio** 5.05

- For typical precipitation data, we generally expect the mean ratios to be between 1.10 and 1.25.
- This does not mean that the precipitation has been increased by 10-25%.

Problems with Statistical Analyses

- One of the problem areas is uncertainty about the exact shape of the probability density distribution function.
- Even if the normal precipitation is well defined (BIG if), there may well be changes in the seeded distribution, other than just the mean.
- It takes a lot of data to define the distribution function well.

Other Types of Analyses

- Nonparametric statistical tests make no assumptions about the about the distribution function.
- The most common of these are the “rank” tests.
  - Rather than looking at the amount of precipitation from each storm, the events are ranked from the smallest to the largest.
  - The question that is addressed is, “Do the ranks of the seeded storms differ significantly from the ranks of the unseeded storms?”

Stratification of Data

- Perhaps seeding will be more effective under some conditions than others; in fact, there may actually be some decreases expected under some conditions.

- It might be advantageous to look at those cases separately.

  - One possible conceptual model is when the cloud top temperatures are warmer than -20°C and decreases in the precipitation at colder temperatures.

  - Use only those cases that fell into the “warm” category and analyze that subset of the total data set.