THE SEARCH FOR A PERFECT WEATHER MODIFICATION EXPERIMENT: AN ESSAY

Arnett S. Dennis

Institute of Atmospheric Sciences
South Dakota School of Mines and Technology
Rapid City, South Dakota 57701

1. INTRODUCTION

Tom Henderson (your editor) wrote to me on 14 November 1977 as follows: "At nearly every meeting I attend, there is always one person who strongly points out that "There has never been one program which has conclusively demonstrated that weather modification technology can either increase rainfall or decrease hail damage' . . . . In the general literature, there seems to be plenty of information which strongly suggests positive results . . . . I think it is time to start saying something about these results, at least in the context of the total weight of the evidence."

Mr. Henderson's letter sets out starkly the contrast between those persons who are awaiting a single conclusive demonstration of the value of weather modification and the rest of us, who take a more pragmatic approach.

Discussions with certain scientists awaiting a definitive experiment show that what they have in mind is an ideal, which in this essay is called the "perfect experiment". The perfect experiment features randomization with a double blind and objective means for selecting test cases. Response variables and predictors are well defined. There are no changes in experimental design after the experiment starts; there are no missing data; and there are no minor annoyances such as the moving of rain gauges to make way for construction projects.

There is an analogy between the search for "perfect experiments" and the demands of some trial lawyers that their clients receive "fair trials," in the sense that the trials must be technically perfect. The demand for "fair trials" in the legal system has led to such incongruities as dismissal of charges against an accused person, who appeared guilty on the basis of the preliminary evidence, because of a typographical error in a summons served upon him. "Our deification of the notion of a 'fair trial' has so far submerged the value of an <u>accurate</u> trial that the latter has no real legal significance" (Bailey, 1978).

Persons demanding the perfect experiment in weather modification sometimes appear ready to write off an entire experiment if, for example, one of the rain gauges used in the evaluation is moved during the course of the experiment. Brownlee (1960) raised the question of whether or not armed guards should be stationed at rain gauges used in the evaluation of cloud seeding experiments to guard against "salting" by "interested parties".

Just as the preoccupation with "fair trials" has hampered the search for justice for defendants as well as society (Bailey, 1978), the preoccupation with perfect experiments has hampered progress in development of weather modification technology. The uncritical acceptance of "results" from certain experiments which conformed (more or less) to the requirement for randomization, and the steadfast rejection of data from other experiments and from operational projects as "inadmissible evidence" have combined to create false pictures of the state of the art in weather modification, and have also undermined the design of experiments on which future progress depends.

This essay explores the development of the perfect experiment concept, gives examples of its deleterious effects, and sounds a warning against possible further damage.

2. HISTORICAL DEVELOPMENT

> It is frustrating to contemplate the possibility that the above money, $795,000, if devoted to one or two deliberate, well-planned and executed experiments, might have settled the question.
> Brownlee, 1960, p. 453.

The idea that a few experiments could lead to final answers about the value of weather modification technology has been around since 1946. The argument to justify the idea is beguilingly simple. One simply mounts an experiment in which a fraction of the suitable test cases are reserved by random selection as control cases. It is assumed that all differences between seed and no-seed samples are due to the seeding. Through replication and application of standard statistical techniques to the resultant data set, one is able to estimate the effect of the seeding treatment upon any selected response variable.

In 1951, representatives of several federal agencies met to establish a prestigious Advisory Group which, in turn, met to design a number of experiments (Petterssen, 1957). Several experiments were started around 1953. They included Project SCUD on the east coast and the U.S. Weather Bureau's ACN Project in the Pacific Northwest. However, the Advisory Group soon found that the matter was not as simple as had at first appeared; initial analysis of the data showed the random variations to be large. The Advisory Group decided in 1954 to discontinue the experiments because "It appeared that several more years were necessary to determine the effects with any degree of certainty" (Petterssen. 1957, p. 4).

The withdrawal of the U.S. Government from weather modification experiments in 1954 left the field largely to commercial operators, a fact reflected in the Final Report of the Advisory Committee on Weather Control (Advisory Committee, 1957). However, randomized experiments continued in some foreign countries, notably Australia, India, and Mexico, and some experiments started in the United States in 1960 under the newly authorized National Science Foundation program.

It should have been obvious that, if enough randomized experiments were conducted, some of them would come up with "significant results," even in the absence of cloud seeding effects. Therefore, it is not surprising that the literature began to be filled with such terms as "uncontrolled background variation", "bad draws", and "unrepresentative draws". All of these, of course, are covered under the more generally accepted term of Type I statistical errors. Particularly troubling was a salt seeding experiment in India, which provided evidence of rainfall increases of up to 40% over substantial areas with an acceptable level of statistical significance (Biswas et al., 1967). It was troubling because physical meteorologists were at a loss to describe the mechanism by which the apparent rainfall increases could have been produced.

Some scientists had already been urging physical studies in association with randomized field experiments to replace the "black box" concept of cloud seeding. By 1970 it was generally agreed that no field experiment would be accepted as having produced significant positive results unless some physical process had been postulated to demonstrate the physical plausibility of the results indicated by statistical analysis. At first, qualitative or conceptual models of the processes were accepted but, as time went on, more and more persons argued that the models should be quantitative and detailed.

The term "proof of concept experiment" has surfaced in the last few years to designate tests of well defined physical hypotheses. The results of a proof of concept experiment would theoretically be transferable to other geographic regions after due allowances for climatological differences. In a sense, the proof of concept experiment is the currently accepted version of the perfect experiment.

3. SOME QUASI-PERFECT EXPERIMENTS

No perfect experiment has been completed to date, in part because we keep raising the standards used to judge field experiments. While awaiting the perfect experiment, some persons have tried to make do with a small sample of quasi-perfect ones.

The recent thinking of a conservative scientist with regard to acceptance or rejection of results of field experiments is illustrated in the remarks by Mr. J. Warner at the WMO/IAMAP Scientific Conference on Weather Modification at Tashkent in 1973 (Warner, 1974). Mr. Warner stated that, in his opinion, there had been two experiments in the world up to that time which could serve as the basis for concluding that precipitation could be increased by cloud seeding, these being the Climax experiment and the Israeli experiment. He thought that the Tasmanian experiment of 1964-70 might also qualify when the analysis was completed.

Unfortunately, the credibility of an experiment varies with its location, the scientific prestige of the participants and sponsors, the amount of sophisticated equipment used, and the skill with which the results are presented, as well as upon its scientific merit. Mr. Warner's attention to the Israeli experiment may be traceable to long and useful associations between Australian and Israeli scientists. Indeed, the Israeli program perhaps owes its very existence to a visit which Dr. E. G. Bowen of Australia paid to Israel about 1958.

In making a list of quasi-perfect experiments, I would include the Necaxa

Project in Mexico. Although I have never visited the project, I have been privileged to know the principal scientist involved, Mr. E. Perez Siliceo, since the mid-1950's. Necaxa has proven the most durable randomized project in the world. There is considerable confusion concerning this point, because one of the first papers about it in English made use of some results for non-randomized seeding prior to 1956 (Perez Siliceo et al., 1963). The project was operated with strict randomization and no important change in methods of operation for 13 years beginning in 1956. Ground generators were operated to release AgI to seed orographic cumulus clouds. Admittedly, the Necaxa Project did not include the physical studies now considered a part of the perfect experiment, but the reader should note that, when the randomized experiment began in 1956, the present emphasis on physical studies did not exist.

The Necaxa Project yielded evidence of overall changes in rainfall distributions significant at the 0.001 level (Perez Siliceo, 1970). Material which Mr. Perez Siliceo has shown me in private sessions indicates that under some stratifications significance levels ranged around 10^{-4} to 10^{-5}. By the standards of his time, Mr. Perez Siliceo and his associates in the Mexican Light and Power Company conducted a perfect experiment and obtained significant evidence of rainfall increases in the predesignated target area.

4. AN EXAMPLE OF HOW RELIANCE UPON A QUASI-PERFECT EXPERIMENT LEADS TO MISAPPREHENSIONS CONCERNING THE STATE OF THE ART

> Someone ought to have the temerity to ask whether the result was correct, not simply whether the rituals were acceptable.
> Bailey, 1978.

As noted in the Introduction, the uncritical acceptance of results from certain experiments which conformed to the requirement for randomization has led to erroneous ideas concerning the state of the art in several fields of weather modification technology. The special case of hail suppression will be considered here as an example of this effect and as justification for the statement.

The quasi-perfect experiment to be considered in this case is the National Hail Research Experiment (NHRE). NHRE was based upon a physical concept for hail suppression picked up by American scientists in a number of visits to the Soviet Union in the mid-1960's. The experiment was carried out with randomization and with objectively defined test cases and response variables. The associated physical studies were most impressive and extensive.

The analysis of the NHRE data following three seasons of field experimentation showed that the data on hail mass were consistent with any interpretation between approximately a 60% suppression of hail mass and a 300% to 500% increase (e.g., Crow et al., 1977). At this point the reader may wonder how an experiment with such an inconclusive result could be considered perfect. However, just as the concept of a "fair trial" has little to do with the requirements for an accurate trial (Bailey, 1978), so there was nothing in the definition of a perfect experiment up to 1970 that required it to give an accurate result in as short a time as three years.

The statisticians and other scientists responsible for interpreting the NHRE data and communicating their findings to the scientific community have emphasized repeatedly the associated uncertainties. Nevertheless, because it was the National Hail Research Experiment, the "results" have been accorded great respect both here and abroad.

Due to the publicity accorded the NHRE findings, the results of other randomized experiments and operational programs of hail suppression, which in total involved treating hundreds of times as many hail cells as NHRE and which indicate a modest suppression effect by silver iodide seeding, are widely ignored or rejected. Last summer a man wrote to me from Germany asking how an operational program in South Dakota could have suppressed hail. He wrote, "Claims of success in hail prevention are in contradiction with the results of the NHRE which was suspended because of lack of clearcut results."*

Changnon (1977) has reviewed the field of hail suppression and reports that "The results of five of six recent suppression projects show suppression levels of 20-50%, but the results are largely not significant at the 5% level. This difference between average beliefs of experts** and the results of recent projects suggests the need for an extensive investigation of the data and results of these recent projects." Changnon (1977, p. 26) further suggests that the review be conducted by an "august body." While endorsing his suggestion, I doubt that the research establishment will abandon its search for the perfect hail suppression experiment in order to conduct such a survey.

5. AN EXAMPLE OF A DELETERIOUS EFFECT UPON AN EXPERIMENTAL DESIGN

As noted above, proof of concept experiments should yield results which are transferable to other geographic locations. However, the insistence that there have been only two or three credible experiments to date leaves one with a totally inadequate kit of models to use in the many varied places where weather modification may be applied. Furthermore, the existence of a physical model does not automatically imply that it is the correct model, and is no guarantee of transferability.

Attempts to transfer those models which have proven applicable to the simple situations under which quasi-perfect experiments are possible can lead to serious errors. For example, the Climax model used microphysical processes exclusively, particularly growth of snowflake by deposition of water vapor, to explain apparent increases in orographic snowfall due to AgI seeding (Grant and Kahan,

---

*Private communication from George Breuer. The reader with experience in statistical inference may note that Mr. Breuer's statement contains a flaw in its logic, but his point of view is widely accepted.

**The average belief of the experts corresponded to "no knowledge of a hail suppression capability."

1974, p. 290). In setting up the Colorado River Basin Pilot Project (CRBPP) on the San Juan mountains in southwestern Colorado, the U.S. Bureau of Reclamation was forced to base its design on the Climax results in order to gain scientific backing for its new project. At a Skywater workshop in May, 1968, the Bureau of Reclamation officials were pressed by their own statistical consultants to say what evidence existed to justify launching the CRBPP. The only answer deemed acceptable was the Climax results as presented by Grant and Mielke (1967). The result was an experimental design that emphasized stratification in terms of cloud top temperature, a la Climax.

The Climax target area extends above 10,000 ft MSL and the air masses that reach it have already dropped most of their moisture on mountain ranges further west. Storms in the San Juan Mountains have more abundant moisture. If more attention had been paid to the results of operational projects and some randomized experiments elsewhere in the western United States, the CRBPP design would have allowed for the following reasonably well established facts, which had already been published in the refereed literature:

a. The success of cloud seeding to increase precipitation in many orographic situations is dependent upon the degree of atmospheric stability or instability, rather than upon the cloud top temperature. Very favorable results are obtained in certain unstable situations. (e.g., Elliott, 1962; Dennis and Kriege, 1966.)

b. Much of the solid precipitation growth in orographic situations with embedded convective cells is by accretion rather than by deposition. (e.g., Dennis and Kriege, 1966.)

c. Seeding results in orographic situations are strongly influenced by dynamic effects as well as by microphysical effects. (Elliott, 1966.)

It is sufficient to note that the CRBPP was not notably successful in terms of its original statistical design, although *post hoc* analysis of the data is turning up useful suggestions for future experiments, as well as for operational programs in snowpack augmentation.

6. IMPLICATIONS FOR FUTURE EXPERIMENTS

The desire to conduct perfect or quasi-perfect experiments may well lead to additional errors in planning for future weather modification research. In this essay I have emphasized the dichotomy which has existed in the past between perfect experiments and accurate experiments. This dichotomy may be disappearing. Most scientists planning a field experiment in weather modification now look for a statistical design using data stratification and enough covariates or predictors to ensure that their experiment is not only unbiased but has a high probability of reaching an accurate conclusion.

The perceived danger at this point is more subtle. It is the distinct possibility that the next generation of experiments will be unbiased and accurate, but meaningless. This would occur if, for example, an experiment were designed solely on the basis of making the hypotheses amenable to easy testing or to exploitation in a numerical cloud model, or both. Basing an experiment upon a simple hypothesis that we can test or model, rather than upon the more complex but controlling factors, is like looking for a lost coin under a lamp post, where the light is good, rather than where the coin was dropped.

The evidence from the Necaxa Project and elsewhere on the importance of mesocale dynamic effects, including (at Necaxa) rainfall suppression effects in control areas upwind of the generator locations, must not be ignored. It suggests that some of the present experiments, with their great emphasis on microphysical observations within individual seeded clouds, are testing only a small subset of the important hypotheses on how cloud seeding affects the weather. Although the mesoscale dynamic effects have not yet been captured in a numerical model. they may dominate any experiment dealing with clouds in convectively unstable situations. In this connection, we must be especially wary in selecting predictors for experiments on convective clouds. Choosing target-control or randomized crossover designs in preference to the target-only design involves making certain assumptions about what seeding does or does not, and those assumptions really require a restatement of the hypotheses being tested.

Another point that should be considered afresh in the design of any weather modification experiment is whether randomization is necessary at all. Randomized experiments have increased our knowledge in some areas. However, the conclusions of experiments on orographic precipitation (e.g., Vardiman and Moore, 1977) are not that much different from those published 20 years ago on the basis of Thom's target-control analysis of operational projects (Advisory Committee, 1957). Since 1957, we have come to accept the reality of inadvertent weather modification by industrial complexes and large cities, and have done so without the imposition of any randomization on industrial work schedules. Planning for possible future Stormfury experiments includes the design of statistical techniques for extracting maximum information from data sets where all suitable storms would be seeded, although it is still possible that randomization will be imposed on Stormfury seedings. In time to come, atmospheric scientists may wish to engage in weather modification experiments where randomization would not be feasible at all (Dennis and Gagin, 1977, p. 80).

This essay is not intended to justify laxity in the design of future experiments. Quite the contrary. For too long we thought that adherence to the randomization ritual would eventually solve our problems. An unbiased experiment must be a minimum standard, rather than an ultimate objective. Regardless of whether randomization is used or not, the design of future experiments should take sufficient account of all available data from previous experiments, operations, and numerical models of atmospheric processes to be a meaningful exercise, and have enough built-in controls to provide accurate as well as unbiased results.

7. ACKNOWLEDGMENTS

Preparation of this paper was supported by the State of South Dakota. I thank Dr. Harold Orville for reviewing the manuscript.

8. REFERENCES

Advisory Committee on Weather Control (H. T. Orville, Chairman), 1957: Final Report of the Advisory Committee on Weather Control. U.S. Government Printing Office, Washington, D.C., Vol.1, xii+32pp., Vol. 2, 422 pp. [L. C. Card 58-600.06].

Bailey, F. L., 1978: Watch out for trial lawyers. Newsweek, Vol. XCI, No. 1, p.7.

Biswas, K.R., and R. K. Kapoor, K.K. Kanuga, and Bh. V. Ramana Murty, 1967: Cloud seeding experiment using common salt. J. Appl. Meteor., 6, 914-923.

Brownlee, K. A., 1960: Statistical evaluation of cloud seeding operations. J. Amer. Stat. Assoc., 55, 446-453.

Changnon, S. A., 1977: On the status of hail suppression. Bull. Amer. Meteor. Soc., 58, 20-28.

Crow, E. L., A. B. Long, J.E. Dye, P. W. Mielke, Jr., and C. W. Ulbrich, 1977: Primary statistical evaluation of the National Hail Research Experiment randomized seeding experiment 1972-74. Preprints 6th Conf. Planned and Inadvertent Wea. Modif. (Champaign-Urbana), Amer. Meteor, Soc., Boston, Massachusetts, 150-153.

Dennis, A. S., and D. F. Kriege, 1966: Results of ten years of cloud seeding in Santa Clara County, California. J. Appl. Meteor., 5, 684-691.

Dennis, A. S., and A. Gagin, 1977: Recommendations for future research in weather modification. Weather Modification Program Office, NOAA/ERL, Boulder, Colorado. 112pp.

Elliott, R. D., 1962: Note on cloud seeding evaluation with hourly precipitation data. J. Appl. Meteor., 1, 578-580.

Elliott, R. D., 1966: Effects of seeding on energy of systems. J. Appl. Meteor., 5, 663-668.

Grant, L. O., and P. W. Mielke, 1967: A randomized cloud seeding experiment at Climax, Colorado, 1960-65. Proc. 5th Berkeley Symposium on Mathematical Statistics and Probability. Vol. V: Weather Modification Experiments. (L. LeCam and J. Neyman, Editors). Univ. of California Press, Berkeley, 115-131.

Grant, L. O., and A. M. Kahan, 1974: Weather modification for augmenting orographic precipitation. Weather and Climate Modification (W. N. Hess, Editor). John Wiley and Sons, New York, 282-317.

Perez Siliceo, E., A. Ahumada A., and P. A. Mosiño, 1963: Twelve years of cloud seeding in the Necaxa watershed, Mexico. J. Appl. Meteor., 2, 311-323.

Perez Siliceo, E., 1970: 19 years of cloud seeding operations in the Necaxa, Puebla and Lerma, Mexico, watersheds from the periods 1949-1951 plus 1953-1968. Preprints 2nd Nat'l Conf. Wea. Modif. (Santa Barbara), Amer. Meteor. Soc., Boston, Massachusetts, 87-90.

Petterssen, S., 1957: Reports on experiments with artificial cloud nucleation. Meteor. Monogr., 2, 11, 1-4.

Vardiman, L., and J. A. Moore, 1977: Generalized criteria for seeding winter orographic clouds. Skywater Monogr. No. 1, U.S. Department of the Interior, Denver, Colorado.

Warner, J., 1974: Rain enhancement - A review. Proc. WMO/IAMAP Scientific Conf. Wea. Modif. (Tashkent), WMO-No. 399, World Meteor. Org., Geneva, Switzerland, 43-50.