# COMMENTS ON SILVERMAN'S PAPER PUBLISHED IN THE WMA 2009 *JOURNAL OF WEATHER MODIFICATION* ENTITLED " AN INDEPENDENT STATISTICAL EVALUATION OF THE VAIL OPERATIONAL CLOUD SEEDING PROGRAM"

Don A. Griffith*, David P. Yorty, and Mark E. Solak

North American Weather Consultants, Sandy, Utah

## 1. INTRODUCTION

This is an interesting paper that was published by Silverman. The length of the Vail seeding program lends itself to detailed analysis. North American Weather Consultants (NAWC) does have a few comments as well as concerns regarding this paper. These comments and concerns are addressed in the following.

A quote from Silverman states "Silverman (2007) showed that it is imperative to use as a control or controls, to the extent that available data permits, the streamflow station or stations that yield the most precise results." He showed the control or combination of controls that had the highest correlation with the target and, especially, the lowest standard deviation of the residuals (differences between the observed and predicted values) will yield the most precise evaluation results. NAWC readily agrees with that statement, however, examination of the correlations obtained in this study as provided in Table 4 in The Vail paper (Silverman 2009a) suggests this ideal was not obtained. The $r^2$ values obtained were significantly lower than those previously obtained by Silverman in the analyses he has performed on long-term Sierra Nevada programs (Silverman, 2007, 2008, 2009b). Values from Table 2 from Silverman (2009) indicate the correlation coefficients in his Vail analysis ranged from 0.775 to 0.918 or $r^2$ values of 0.60 to 0.84; values considerably lower than he established for the three California programs where the $r^2$ values ranged from 0.96 to 0.98.

## 2. DISCUSSION

A general concern regarding Silverman's Vail analysis is the high estimates of seeding effects in some of the sub-basin target areas. For example, the GBH and GPT estimated average increases are +28.8 and +18.5%, respectively. These are high numbers especially considering that these are estimates of increases in annual streamflow values and that the cloud seeding program is only conducted during some of the winter months. The 28.8% value is considerably higher than the Weather Modification Association Statement of Capabilities (WMA 2005) for winter seeding programs that contains an expected range of 5-15% increases. There are several potential explanations for these high estimates. The two target stations (GBH and GPT) happen to have the two lowest correlations with the selected control station (FRR). The $r^2$ values are only 0.60 for the GBH site and 0.66 for the GPT site. The historical not seeded periods for these two stations are also short (13 not seeded seasons for the GBH site and only 9 not seeded seasons for the GPT site). As Silverman mentions in his Vail paper, "A potential control is a streamflow station that has not been seeded, is highly correlated with the target, and has a long enough record of full natural flow data during the historical and operational period to support a meaningful evaluation." NAWC questions whether these criteria have been satisfied in Silverman's analysis, especially for the two target stations that have the highest indicated seeding increases. The other factor of concern is the very small size of these sub-target basins. The sizes of these basins were not reported by Silverman but are 4.5 square miles for the GBH site and 5.3 square miles for the GPT site according to USGS records. These are very small drainages especially when placed in the context of the typical sizes of winter operational cloud seeding target areas that might range from several hundred to several thousand square miles. The data from Silverman's Tables 1 and 3 may be combined to provide an estimate of the average increases in streamflow for all of the sub-basin target areas for an average water year. The results are provided in Table 1.

The total calculated average increase (column 5) from Table 1 when divided by the average annual runoff (column 3) provides an estimated average increase for all the individual target basins combined in an average water year. The result is an estimated 8.4% for the combined watersheds. The total size of the combined watersheds is 149.2 square miles; still a small area in terms of an intended winter cloud seeding program target percentage increases

**Corresponding author*: Mr. Don Griffith, North American Weather Consultants, Inc., 8180 South Highland Dr., Suite B-2, Sandy, UT 84093 ; e-mail: dgriffith@nawcinc.com

in annual streamflow. A couple of explanations for these results could be the short historical periods and low correlations for these two sub-basins (as discussed in the above) and/or possible channeling of seeding material during seeded storms favoring these areas. When the results are merged over the entire area (albeit still a rather small area), the indicated average increase of 8.4% becomes much more reasonable. NAWC concludes that very small sub-areas in a large winter time cloud seeding area may show indications of rather high seeding increases but when results are averaged over larger target areas, the results are likely to fall within the 5-15% increase range as contained in the WMA Statement of Capabilities. Stated another way, it is highly unlikely that cloud seeding over a more typical sized winter cloud seeding target area could produce an average seasonal increase of 28%.

Of potential interest to this discussion is another type of NAWC analysis that has been applied to longer-term winter cloud seeding programs. This is an engineering analysis technique commonly called “double mass” plots. Using this technique, two variables can be plotted in a cumulative fashion that will demonstrate how the two variables may be correlated over time. Each successive season’s data are added to the accumulated values for the combined prior seasons of data. If the two variables are well correlated, a straight line can be drawn through the individual points. If there is a change in the relationship between the two variables with time, a “break” in the straight line will appear.

NAWC applied the double mass technique to one of the target basins found in Silverman’s analysis (Upper Gore Creek, GUP) and one of the control stations (Fryingpan River near Ruedi, FRR). The Upper Gore Creek and Fryingpan sites were selected since they had long historical not seeded data (1948-1976) and since Silverman had concentrated his analyses on those using the Fryingpan site as his primary control site. Annual data were plotted using the double-mass technique for the period of 1948-2005. This plot is provided as Figure 1.

There are a couple of interesting features in Figure 1. First, there is a break upward in the plot indicating more streamflow at the target station than at the control station, which appears to be coincident with the start of the seeding program in 1977. This provides strong, independent support that the indicated increases in Silverman’s analysis are real and are related to the cloud seeding activities. There are a couple of more subtle differences that may be important. The slope of the line decreases during the years from approximately 1981 through 1989. The slope of the line increases beginning approximately with 2001 continuing through 2005. If it is assumed that the breaks we are seeing on this plot are due to seeding effects, then why did the effectiveness of seeding appear to decrease during the 1981-1989 period and why did the apparent effectiveness increase beginning in 2001?

The double mass approach seems to be more sensitive in suggesting differences in seeding effectiveness than Silverman’s technique of plotting the “time evolution of seeding effect.” Compare Silverman’s Figure number 3 from his paper, reproduced here as Figure 2, with the above Figure 1. A couple of clarifications are necessary regarding Figure 2.

**Table 1.** Target Basin Characteristics and Calculated Increases in Annual Streamflow

| Gaging Station Name | Gaging Station Symbol | Drainage Area ($mi^2$) | Ave. Annual Runoff (ac. ft.) | Ave. % Increase | Calculated Ave. Increase (ac.ft.) |
|---|---|---|---|---|---|
| Piney R. | PNY | 86.2 | 54,234 | +6.3 | 3,416 |
| Booth Cr. | GBO | 6.0 | 8,091 | +9.3 | 752 |
| Middle Cr. | MID | 5.9 | 3,944 | +7.9 | 312 |
| Pitkin Cr. | GPT | 5.3 | 7,395 | +18.5 | 1368 |
| Bighorn Cr. | GBH | 4.5 | 5,482 | +28.8 | 1579 |
| Upper Gore Cr. | GUP | 14.4 | 20,523 | +11.1 | 2278 |
| Black Gore Cr. | GBL | 12.6 | 12,052 | +4.6 | 554 |
| Turkey Cr. | TMW | 14.3 | 10,312 | -2.0 | (21) |
| **Totals** | | **149.2** | **122,033** | **+8.4** | **10,238** |

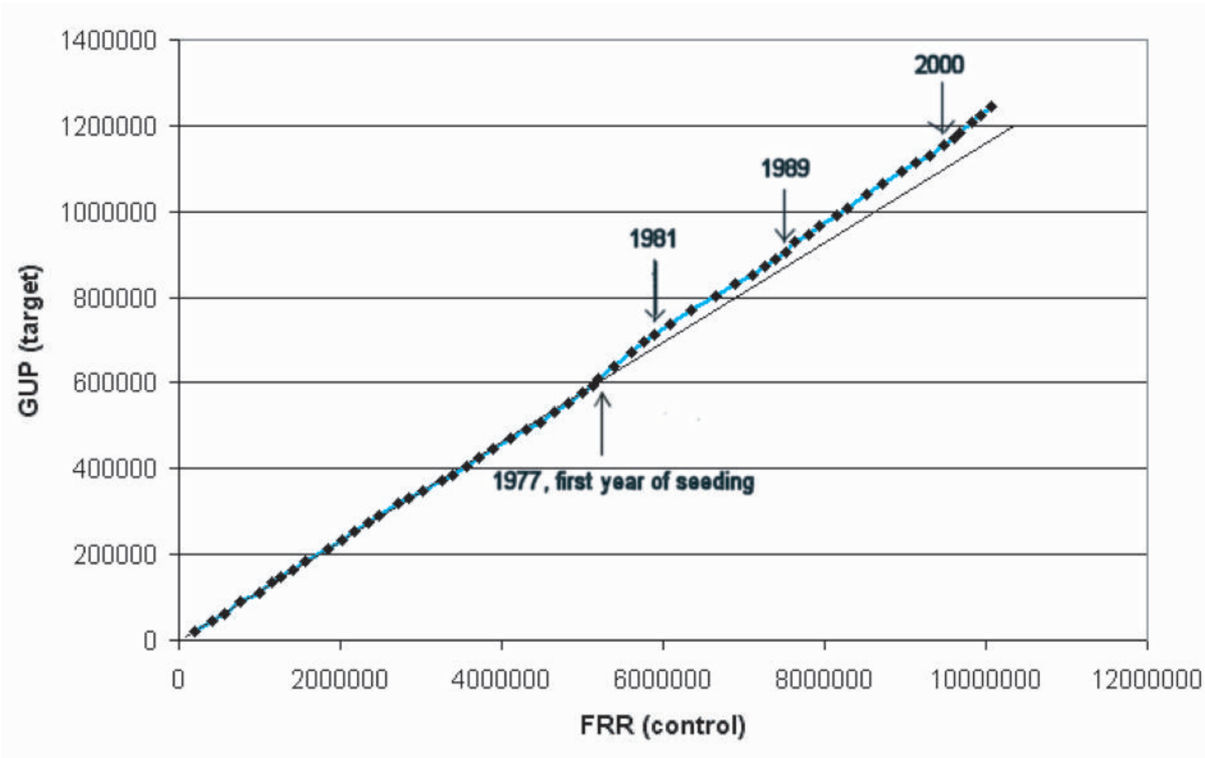


*Figure 1. Double Mass Plot of Annual Streamflow, Upper Gore Creek versus Fryingpan River*

This display begins in 1985 but the seeding program began in 1977. For this reason it is easier to look for the increase in seeding effectiveness during the 2001-2005 period rather than the decrease in the 1981-1989 period. Additionally, the reader needs to look at trace number 5 in Figure 2, which is the one for Upper Gore Creek. This trace does not appear to indicate the increase in effectiveness depicted in Figure 1.

In passing it is worth noting the utility of the double mass plot in selecting target and control sites when using the historical target/control evaluation technique. The stability of the site's relationships can be readily assessed using this technique. Silverman considered several stream gaging stations as potential control sites. One station that he considered was West Divide Creek (WDC). Two other control sites considered by Silverman were the South Fork of the White River (WSF) at Buford and the North Fork of the White River (WNF) at Buford. NAWC prepared double mass plots for WSF versus WNF, Figure 3 and WDC versus WSF, Figure 4. Figure 3 indicates a stable relationship between the South Fork and the North Fork of the White River sites. Figure 4, however indicates a break in the record for West Divide Creek versus the South Fork of the White River beginning in about 1966. The plot in Figure 4 becomes quite variable after 1966. Since Figure 3 indicates stability in the relationship between the South and North Forks of the White River, it is concluded that there is considerable variability in the West Divide Creek streamflow records after 1966 for unknown reasons. As a consequence, West Divide Creek would be a poor choice for a control station. Fortunately, although Silverman initially considered West Divide Creek as a control site, he based most of his analyses on using the Fryingpan site near Ruedi (FRR) as his primary control gage. This appears to be a good choice based upon a similar double mass plot that NAWC prepared for FRR versus WSF (South Fork of the White River at Buford) provided in Figure 5. Figure 5 indicates a stable relationship between these sites. Another interesting insight can be gathered from Figure 5. One of the assumptions in selecting control sites is that they are not affected by the seeding program to be evaluated or by other cloud seeding programs in the area for that matter. Figure 5 substantiates this assumption.

## 3. GENERAL COMMENTS

Finally, some general comments on Silverman's Vail paper are as follows. The Vail seeding program is not a randomized experiment and Silverman's analyses are *a posteriori*. Quoting from Hess (1974), "The weather scientist recognizes the large

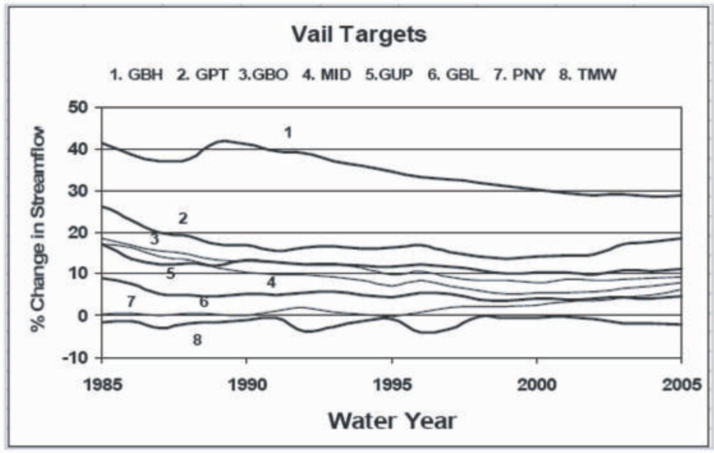


*Figure 2. Time Evolution of the Seeding Effect (% change in streamflow)*

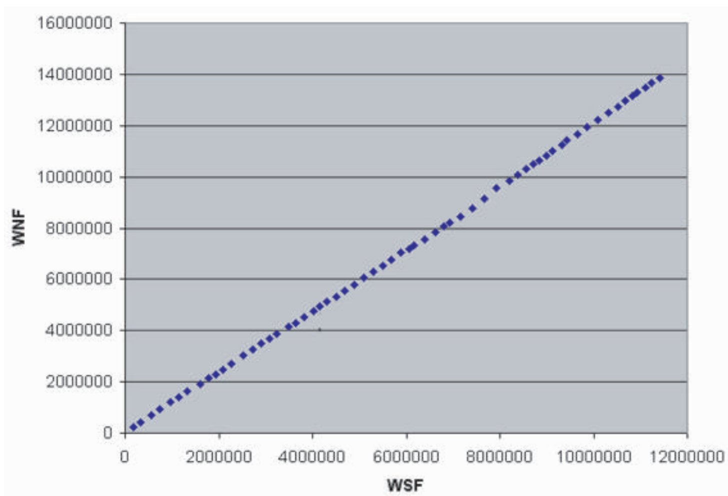


*Figure 3. Double Mass Plot of Annual Streamflow, South Fork of White River versus North Fork of White River*

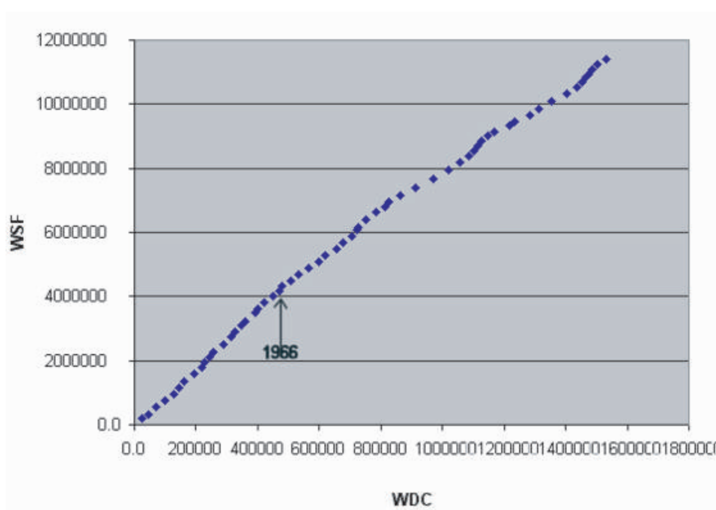


*Figure 4. Double Mass Plot of Annual Streamflow, West Divide Creek versus South Fork of White River*

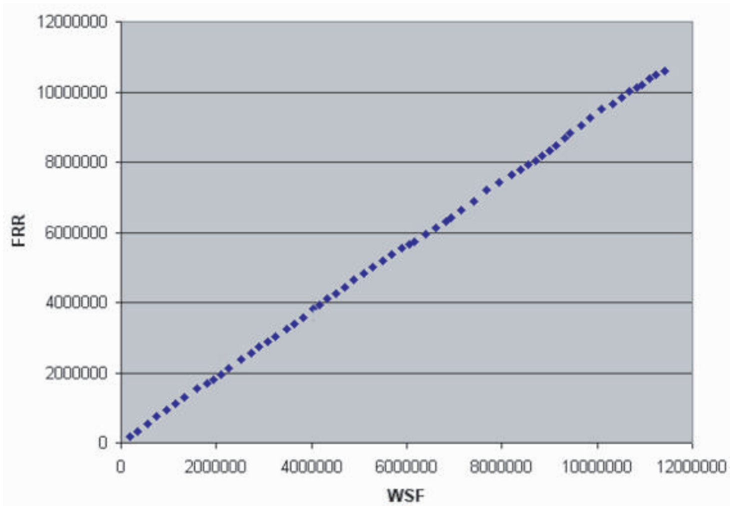


*Figure 5. Double Mass Plot of Annual Streamflow, South Fork of White River versus Fryingpan River*

natural variability of rainfall and cloud characteristics in space and time, and sees the need for appropriate statistical methods to cope with the problems of uncertainties, for, as expressed by F. Mosteller and J.W. Tukey in 1968, 'One hallmark of the statistically conscious investigator is his firm belief that however the survey, experiment or observational program actually turned out, it could have turned out somewhat differently.' Statistical methods were designed by R.A. Fisher (1960) to handle these problems in connection with the design and analysis of comparative experiments in biological and agricultural research, where large and only partly controllable variability is present. The basic ideas involve (1) *replication*, from which a quantative estimate can be made of the variability of the response to treatment, and (2) *randomization,* a process of allocating treatments to the experimental material by tossing a coin (or equivalent procedure), which may make it possible for the experimenter to attribute whatever effects he observes to the treatment and the treatment only. **Together, these two principles enable one to make a valid assessment of the uncertainty of the result in terms of a probability statement or by setting confidence limits** (emphasis added)."

A reference that explains the ratio statistics methodology as applied by Silverman in his Vail paper is "Ratio Statistics for **Randomized** (emphasis added) Experiments in Precipitation Stimulation (Gabriel, 1999).

Based upon the above, the statement in the Abstract of Silverman's Vail paper, "Evidence for statistically significant (underline added) seeding effects ranging from +6.3 to +28.8% was found for 5 of the 8 seeding targets" is highly questionable. Silverman makes a contradictory statement later in this same paper as follows stating, "It is emphasized that this study is an *a posteriori* evaluation of a non-randomized seeding operation. In addition, this evaluation is an exploratory study that involves consideration of a multiplicity of analyses, some of which are suggested by the results of previous analyses. With such a large number of tests, a few are likely to yield significance strictly by chance. In view of these considerations, the results of the evaluations in this study must be viewed with caution. **It is emphasized that the results should be interpreted as measures of the strength of the suggested seeding effect. From a rigorous statistical viewpoint, the suggested effects that are indicated must be confirmed through new, *a priori*, randomized experiments specifically designed to establish their validity.**" (emphasis added). This self-stated contradiction imposes limitations on the interpretation and statements regarding statistical significance and confidence intervals in Silverman's Vail analysis. Silverman's analysis provides **"estimates"** of the results of cloud seeding but does not provide **"statistical proof"** of the significance of these estimates as strongly implied in this paper. Similar contradictions are found in Silverman's other three recently published papers in the WMA regarding analyses of long-term non-randomized winter cloud seeding programs (Silverman, 2007, 2008, 2009b). Furthermore, the titles of two of the four papers by Silverman use the term "Statistical Evaluation", a term, which NAWC considers misleading, based upon the above discussion.

NAWC believes that the standard historical target/control analysis technique, when applied correctly, is entirely adequate in providing estimates of potential increases due to cloud seeding from long-term, non-randomized programs. More sophisticated techniques, such as those proposed by Silverman (which is actually an adaptation of the historical target/control technique) add little to this standard approach since the data are non-randomized and statements on statistical significance and confidence intervals are therefore not valid. There are several considerations that go into the development of good historical target/control evaluations. Dennis (1980) listed several questions regarding the use of the historical target/control regression technique. These concerns and the approach or approaches that NAWC uses to address each are summarized in the following:

1. "Reliability of the results unless the underlying data sets conform to the normal distribution which, for precipitation data, requires an appropriate data transformation." Quoting from Dennis (1980) "Rainfall observations say for one hour or day at a point, tend to be highly skewed, with most observations near zero and a long tail extending to large amounts." NAWC utilizes longer-term data, either three or four-month cumulative precipitation values or April 1st snow water contents that are also cumulative values. Further, NAWC deals with averages of multiple sites (not a point measurement) for the control and target average values. These data are not highly skewed.

2. "Unconscious bias in the selection of data in post-hoc evaluations." As described in Griffith, et al, 2009, NAWC normally establishes target and control stations for use in its evaluations early in the lifetime of its operational programs. These target and control sites and the resultant regression equations are typically maintained throughout the lifetime of the seeding program. Changes are typically made only if observations are discontinued at one or more target or control sites. As a consequence, NAWC evaluations would be considered *a priori* evaluations.

3. "Difficulty in eliminating residual uncertainties." Dennis (1980) in discussing this concern states, " A number of possible biases are dealt with rather simply. Agreements before a program begins as to which rain gages are to be included in calculating target and control rainfall, for example, go far toward eliminating both unconscious bias and any temptation to select data to demonstrate a desired result (Court, 1960). As discussed in the above, NAWC typically follows this recommendation in evaluating its operational programs usually following the first season of operation.

4. "Representativeness of the target-control relationship and its stability in time." Quoting from Dennis (1980) "The most serious difficulty with the historical regression method has to do with the stability in time of the target-control relationship." ... "Neyman and Scott (1961) have hypothesized that the lack of stability in the target-control relationship is related to the occurrence of specific storm types, some of which favor the target area and some of which favor the control area." ... "The best that can be done appears to follow the criteria noted above for the selection of control areas and to be alert to any obvious changes in weather patterns that could distort the target-control relationship. One must not go to extremes in this regard; obviously, if one looked long enough, one could always find *something* that was different between the historical period and the operational period (Gabriel, 1979)." NAWC has generally attempted to address this concern by careful selection of target and control sites (as described in Griffith, et al, 2009) as recommended by Dennis (e.g. as close to target areas as possible, in areas typically upwind to avoid contamination, and selecting target and control sites at similar elevations). In fact, we often attempt to geographically bracket the target area with control sites in order to address the concern of storms favoring one area over another during specific storms or perhaps extending through an entire season. NAWC interprets the discussion contained in Dennis (1980) to be directed at short time intervals, e.g., individual storm periods. NAWC's evaluations utilize seasonal data over periods of as many seasons as possible for which quality records are available that would be less subject to this concern since storm tracks change from storm to storm and any large departures between two areas are frequently dampened over longer time periods.

In addition to these concerns, the development of good target/control relationships needs to be concerned with the following:

1. Selecting target and control sites with quality data (no breaks in records, no station moves, continuity of data between stations which can be checked utilizing the double mass plotting technique) with adequate periods of not seeded historical data upon which the regression equations may be based.

2. Insuring that neither the cloud seeding program being evaluated nor other cloud seeding programs in the area do not impact the selected control stations either during the historical or seeded periods.

3. Determining which types of data to use. For example, each type of high elevation precipitation measurement technique has various disadvantages (e.g., precipitation gage catch reductions in high winds, drifting snow over snow pillows, snow melt at south facing sites as compared to north facing sites at similar elevations).

4. Achieving good relationships between the target and control sites as evidenced by high correlation coefficients.

NAWC carefully considers the above concerns in the development of evaluations of the operational programs conducted by NAWC.

## 4. SUMMARY

Silverman's analyses of the Vail Project and seeding projects in California have provided some interesting insights into prospective techniques for estimating the effectiveness of winter orographic cloud seeding efforts. In the case of the Vail analyses, areal averaging of the results for the small project target area provides an estimated 8.4 % increase in annual streamflow, falling within the generally-accepted 5% to 15% range of expected possible effects for that type of project. However, because the Vail and similar California analyses are *a posteriori* and are applied to non-randomized projects, the analyses and their indicated results carry the same caveats as similar analyses conducted by others over the years. Accordingly, the results must be viewed with caution and presented appropriately as indications, and certainly not proof, of seeding effects. Estimations of the effectiveness of non-randomized operational seeding projects are important but challenging. Such efforts must continue and will, no doubt, generate lively debate as they do.

**Acknowledgement.** Dr. Bernard Silverman provided copies of the streamflow data used in his analyses to North American Weather Consultants.

## REFERENCES

Court, A., 1960: Journal of the Irrigation and drainage Division. *Proc. Am. Soc. Civ. Eng.,* **86**, No. IRI, 121.

Dennis, A.S., 1980: *Weather Modification by Cloud Seeding*. Academic Press, New York, NY, 267pp.

Gabriel, K.R., 1979: *Commun. Statist. Theory Methods*. A8, 975-1015.

Gabriel, K.B., 1999: Ratio statistics for randomized experiments in precipitation stimulation. *J. Appl. Met.*, **38**, 290-301.

Fisher, R.A., 1960: The Design of Experiments, 7th Edition. Hafner, New York.

Hess, W.N., 1974: *Weather Modification by Cloud Seeding*. John Wiley & Sons, New York, 842pp.

Mosteller, F. and J.W. Tukey, 1968: *Data Analysis, Including Statistics, The Handbook of Social Psychology*, 2nd Edition, Addison-Wesley, Reading, Mass., Vol. 2, 819pp.

Neyman, J. and E.L. Scott, 1961*: J. Am. Stat. Assoc*., **56**, 580.

Silverman, B.A., 2007: On the use of ratio statistics for the evaluation of operational cloud seeding programs*. J. Wea. Mod*., **39**, 50-60.

Silverman, B.A., 2008: A Statistical Evaluation of the Kern River Operational Cloud Seeding Program. *J. Wea. Mod*., **40**, 7-16.

Silverman, B.A., 2009a: An independent statistical evaluation of the Vail Operational Cloud Seeding Program. *J. Wea. Mod*., **41**, 7-14.

Silverman. B.A., 2009b: An Evaluation of the San Joaquin Operational Cloud Seeding Program using Monte Carlo Permutation Statistics. *J. Wea. Mod*. , **41**, pp. 15-22.

Weather Modification Association, 2005: WMA Capabilities Statement on Weather Modification. WMA, Fresno, California.