On the Use of Ratio Statistics for the Evaluation of Operational Cloud Seeding Programs

Bernard A. Silverman, PhD
Consulting Meteorologist
Centennial, CO, USA

Abstract. The purpose of this study is to describe and demonstrate the capability and merits of using ratio statistics in evaluating the effectiveness of operational (non-randomized) cloud seeding programs. The application of the ratio statistics methodology is illustrated by an independent statistical evaluation of the Kings River operational cloud seeding program over its entire period of operations from water years 1955 to 2004. The effect of seeding in terms of confidence limits was emphasized because they provide information on the strength of the seeding effect whereas null hypothesis significance tests indicate only whether there is any seeding effect at all. The effect of seeding on the Kings River-Pine Flat Dam streamflow station, the primary seeding target in the Kings River Basin, was evaluated using the control that gives the most precise evaluation results possible with the available data. The results of this evaluation study indicate that (i) for the data involved in this study, ratio statistics was found to be a more precise and more reliable evaluation methodology than the traditional historical regression methodology, (ii) evidence for positive, statistically significant and cost effective seeding effects was found at the target site in the Kings River Basin with an estimated increase in streamflow due to seeding of +5.1% with 90% confidence that the true effect of seeding is somewhere between +1.5 and +8.8%, and (iii) it was found that there was a marked improvement in seeding effectiveness that started around 1978, the physical cause(s) of which is worthy of further study.

1. INTRODUCTION

There are operational (non-randomized) cloud seeding programs being conducted in almost all of the major watersheds of the Sierra Nevada Mountains of California aimed at increasing precipitation in order to enhance streamflow for increased hydroelectric power generation, for downstream agriculture and/or reservoir recreation. All of these programs have been, for the most part, conducted continuously since their inception, the earliest one starting in the Upper San Joaquin River Basin in water year 1951. All the operational programs expect to increase precipitation according to the same seeding conceptual model, i.e., by seeding for microphysical effects to improve the precipitation efficiency of the clouds. Some operational programs try to accomplish this by conducting seeding operations on both summer and winter storms to increase rainfall and to augment snowpack, respectively, whereas some conduct seeding operations on only winter storms to augment snowpack. Both ground-based and aircraft seeding is being applied on some of the operational programs using seeding systems such as silver iodide ground generators, airborne silver iodide generators, airborne silver iodide flares, and/or airborne hygroscopic flares whereas some of the operational programs only use ground-based silver iodide generators or aircraft seeding.

The purpose of this study is to describe and demonstrate the capability and merits of using ratio statistics in evaluating the effectiveness of operational (non-randomized) cloud seeding programs. The application of the ratio statistics methodology is illustrated by an independent statistical evaluation of the Kings River operational cloud seeding program over its since it started in 1955 until 2004. Henderson (2003a) gives a concise description of the Kings River Basin, the meteorology that affects it, and the history of the operational seeding program. With the exception water years 1981-1987 when seeding was totally suspended while the Pine Flat Power Plant was being constructed and several partial year, weather-related suspensions thereafter, the seeding program has been operated each year since its inception. Seeding has been conducted with both ground generators and with aircraft 6-7 months each year in an effort to increase rainfall and snowpack. Henderson (1966) evaluated the first 10 years of the Kings River program using the historical regression method and reported annual increases ranging from –6.1% to +15.0% with an average increase in streamflow of +6.1% that was significant at the 0.005 level. Subsequently, Henderson (2003a) evaluated the first 47 years of the Kings River operational seeding program, including the non-seeded period 1981-1987, using the historical regression method and reported annual increases ranging from –9.5% to +25.6% with an average increase of +5.5% having a statistical probability of 99.9%.

2. EVALUATION PROCEDURES

The evaluation of the Kings River operational cloud seeding program is based on unregulated "natural" or full natural flow (FNF) streamflow data. The evaluation was conducted by two different statistical methods. The first evaluation of the seeding program was conducted using the historical regression method, as was used in the previous evaluations. This was done to put the results of the various evaluations that were conducted previously into perspective. The second evaluation of the seeding program was conducted using ratio statistics (Gabriel, 1999). Gabriel reported that ratio statistics produces results that approximate those from re-randomization for sample sizes of 100 or more; however, experience by this author has shown that comparable results are obtained for even smaller sample sizes. Ratio statistics were emphasized since it is easier to compute than re-randomization and it readily provides additional, useful information about confidence limits.

The main emphasis in the presentation of the results is on confidence limits because they indicate a range within which the true effect lies whereas null hypothesis significance tests infer only whether there is any effect at all (Gabriel, 2002; Nicholls, 2001). Use of confidence limits provides information on the strength of the seeding effect to allow informed judgments to be made about its cost effectiveness and societal significance. It should be noted, however, that statements of significance are implicit to confidence limits statements. Saying, for example, that there is 90% confidence that the true effect of seeding lies between a Single Ratio (SR) of SR_{90LO} and SR_{90HI} is tantamount to saying that the confidence limits result is significant at a 2-sided level of significance of 0.10 (1-sided level of significance of 0.05). Thus, 90% confidence limits that include an SR value of 1 indicates that the experimental result is not significant at a 2-sided level of significance of 0.10 (1-sided level of significance of 0.05).

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 hypotheses/analyses, some of which are suggested by the results of previous analyses. After applying the Bonferroni method (see, e.g., Gabriel, 2000), it is unlikely that any test statistic could be rejected at the resulting level of significance. On the other hand, with such a large number of tests, a few are likely to yield significant results purely 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 and not as measures of statistical significance.
From a rigorous statistical standpoint, the suggested effects that are indicated must be confirmed through new, *a priori,* randomized experiments specifically designed to establish their validity.

3. SELECTION OF THE TARGET AND CONTROL

The Kings River operational cloud seeding program was designed to increase the annual flow of the Kings River into Pine Flat Reservoir. Consequently, Kings River-Pine Flat Dam (CDEC Station ID KGF), representing a drainage area of 1545 square miles, was selected as the primary target for evaluation. Since KGF was used by Henderson (1966; 2003a) as the target in his evaluation of seeding effects, it afforded the opportunity to compare the results produced here with those of previous evaluations. The FNF data for KGF was obtained from the CDEC (California Data Exchange Center) web site online at http://cdec.water.ca.gov.

The criteria for the choice of a control was that it should yield the highest correlation (ρ) with the target (KGF) and, therefore, the smallest standard deviation of the residuals (differences between the observed and predicted values), s_o, since this should yield the most precise evaluation results (see Appendix A). There are four (4) potential control stations, Merced River at Pohono Bridge (USGS site #11266500, hereafter referred to as MDP), Merced River at Happy Isles Bridge near Yosemite (USGS site # 11264500, hereafter referred to as MHI), Success Dam in the Tule River Basin (hereafter referred to as SCC), and Cottonwood Creek (hereafter referred to as CCR). Suitably long records of FNF data were available for each of these control stations. The potential control stations were investigated by themselves and in physically reasonable combinations. The resulting correlation coefficients (ρ) and standard deviations of the residuals (s_o) are given in Table 1.

Based on the results given in Table 1, the combination of MDP and CCR (hereafter referred to as MDP+CCR) was chosen as the control site for the evaluation of KGF. It is interesting to note that Henderson (2003a) used MDP+CCR as the control in his evaluation of the Kings River operational cloud seeding program. Henderson (1966) used the combination of MDP and the Kern River measured at Kernville in his earlier, 10-year evaluation but could not use the Kern River control in his 47-year evaluation because operational cloud seeding of the Kern River Basin started in 1978.

Table 1. Comparison of regression analysis results using the potential controls against the target (KGF). Also shown is the source and period of record (water years) of each dataset.

Control	Data Source	Record Length	Corr.Coef. ρ .	Std. Dev. s_o
MDP	1	1917-2004	0.959	251,702
MHI	1	1916-2004	0.957	254,608
SCC	2	1931-2004	0.912	371,916
CCR	3	1935-2004	0.924	348,111
MDP+CCR	**-**	**1935-2004**	**0.988**	**141,686**
MHI+CCR	-	1935-2004	0.987	149,420
SCC+CCR	-	1935-2004	0.947	294,476

1. Obtained from the United States Geological Survey (USGS) web site online at http://waterdata.usgs.gov/ca/nwis/nwis
2. Obtained from the California Data Exchange Center via CDEC web site online at http://cdec.water.ca.gov
3. Obtained from the Los Angeles Power and Water Department (Paul Scantlin, Personal Communication)

4. EVALUATION BY THE HISTORICAL REGRESSION METHOD

Gabriel and Petrondas (1983) point out that the validity of the statistical inferences based on the historical regression method depends on two assumptions: (1) all differences between the historical and operational periods are due to either the effect of seeding or to random year-to-year variation and there are no other effects confounded with these differences, and (2) the variability of streamflow between the two periods behaves like variability of two independent random samples of years. These assumptions are highly suspect for streamflow data that are subject to some year-to-year dependence, cycles and/or trends, especially since there are no exact models of this behavior available. In addition, Brownlee (1960) has shown that this analysis method requires homoscedasticity (constancy of variance) between the data in the historical and operational periods and this is not the case here. A test of the equality of variances between the historical and operational streamflows indicates that they are statistically different. However, many statistical methods are quite robust to departures from the assumptions under which they were derived and that may be the case for the historical regression evaluation of the Kings River Basin operational cloud seeding program.

In accordance with this evaluation method, a regression equation was derived for the 20-year historical period (1935-1954) prior to the start of operational seeding that predicts the streamflow at the target station (KGF) as a function of the streamflow at the control station (MDP+CCR). The following predictor equation for water year streamflow at KGF was obtained:

$$\text{KGF (AF)} = 2.619476 * \text{MDP (AF)} + 31.98236*\text{CCR (AF)} - 102144.0$$

Dennis (1980) warned that the most serious difficulty with the historical regression method has to do with the lack of stability of the target-control relationship. The regression line obtained for the historical period was, therefore, checked to see how well it fit the data during the operational period. A statistical comparison of the regression line for the historical period with that derived for the operational period indicated there was no statistical difference between the slopes of the two lines and, in fact, from a statistical standpoint the two lines can be regarded as a single coincident line. The regression equation was then used to predict the streamflow at the target station that would have occurred in the absence of seeding for each water year in the operational period. The difference between the actually observed streamflow and the predicted streamflow at the target station for each water year is assumed to be the result of seeding.

Evaluations of the multi-year and annual (water year-by-water year) seeding effectiveness were conducted. Since there was no seeding during the period 1981-1987, the multi-year evaluation was based on 43 years of seeding operations. The non-seeded period was purposely included in the evaluation of annual seeding effectiveness as a check on the validity of this application of the historical regression evaluation method.

4.1 Multi-Year Evaluation of Seeding Effects

It was found that the average difference between actual and predicted water year streamflow at KGF over the 43 years of seeding operations was +78,338 AF. This represents an increase in streamflow of 5.1%, with 90% confidence that the increase was between +1.0% and +9.3%. A t-test of the null hypothesis that the average difference between actual and predicted KGF streamflow in the seeded years minus that in the historical years was equal to zero yielded a 2-sided P-value of 0.044. Thus, there is sufficient evidence to reject the null hypothesis at the 2-sided level of significance of 0.10. As expected, these results after 43 years of seeding operations are consistent with the results obtained by Henderson (2003a) after the first 40 years of seeding operations.

4.2 Evaluation of Annual Seeding Effects

The series of water year streamflow differences, represented as "percentage changes due to seeding", was plotted as a function of water year (Fig. 1). It can be seen from Fig. 1 that the differences during the operational period were negative in about 28% of the water years and positive in about 72% of the water years with the percentage change due to seeding ranging from –16.9% to +40.3%.

Henderson (2003a) also reported percentage changes due to seeding ranging from –9.5% to +25.6% in his historical regression evaluation of 47 years of seeding operations in the Kings River Basin. Although he found that negative seeding effects occurred in 12 of those years, he emphasized that the average percentage increase in flow for the total period was +5.5%.

NAWC (1978) obtained results similar to those in Fig. 1 in their examination of the first 25 years of cloud seeding operations in the Upper San Joaquin River Basin. NAWC interpreted their finding of highly divergent seeding effects (similar to those in Fig. 1) as the result of some year-to-year climatic variability not explained through the use of the historical regression equation prior to and following the start of seeding in 1951. NAWC concluded that the effects of seeding are superimposed on this residual natural variability during the operational period.

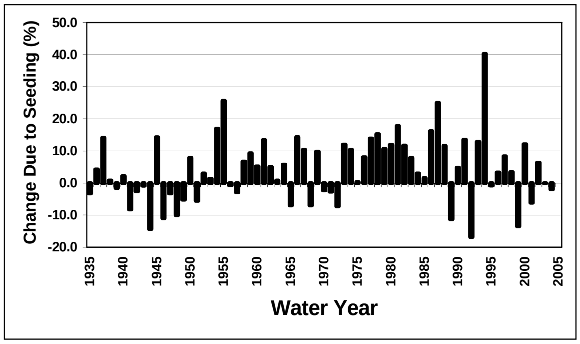

Fig. 1. Water year differences between the actual streamflow at the KGF target site and that predicted by the historical regression method, represented as percentage changes due to seeding

The natural variability of the streamflow data is illustrated in Fig. 2 in which the water year streamflow of the controls, MDP and CCR, is plotted as a function of water year. The year-to-year variability, short-term cycles, as well as a long-term trend of streamflow with time are quite evident. The annual seed/no seed differences produced by the historical regression equation are apparently confounded with and dominated by effects due to natural variability. Because of the bias introduced by these non-random effects on annual streamflow, the attempt to predict seeding effects on an annual basis is not valid and yields misleading results. Consider, for example, the evaluation results during the period 1981-1987 when seeding operations were suspended during the construction of the Pine Flat power plant. Despite the lack of any seeding, substantial increases in streamflow due to seeding are indicated in Fig. 1.

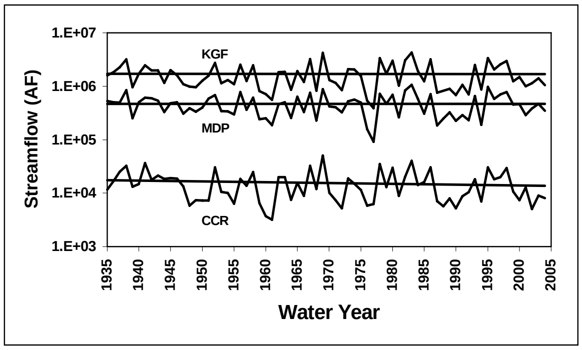

Fig. 2. Streamflow at the Kings River-Pine Flat Dam (KGF) target and the controls, Merced River at Pohono Bridge (MDP) and Cottonwood Creek (CCR), as a function of water year. Note that the steamflow is plotted on a logarithmic scale.

Because of the lack of an exact model that describes the behavior of these non-random effects, the variability of streamflow cannot be taken into account properly in predicting annual streamflows. If one were to actually believe in the evaluation of annual effects, one would have to conclude that seeding frequently decreases as well as increases the streamflow and by fairly sizable amounts. It is unlikely that seeding acts in this manner to produce these kinds of highly divergent effects.

5. EVALUATION BY RATIO STATISTICS

The second evaluation was based on the use of ratio statistics. For single targets, as is the case here, 3 ratio statistics are relevant, i.e., the Single Ratio (SR), the Double Ratio (DR) and the Regression Ratio (RR). Gabriel (1999) has shown that, when a suitable control area is available, using ratios based on regression is always preferable to using single or double ratios; therefore, the use of the Regression Ratio (RR) is emphasized in this analysis.

The regression ratio (RR) is given by the relationship, $RR = SR / SR_{PRED}$ where the single ratio (SR) is the ratio of the average target streamflow during the operational period (TS_O) to the average streamflow for the seeding target during the historical period (TS_H), i.e., $SR=TS_O/TS_H$, and SR_{PRED} is the SR as predicted by the target-control regression relationship. By dividing the SR by SR_{PRED} , SR is adjusted for effects due to natural differences in streamflow between the target and control and, by taking advantage of the high correlation between the target and control streamflows over the entire period of analysis (including both the historical and operational periods), the variance of the regression ratio is reduced with respect to the variance of the single ratio for the target station only. This enables the detection of smaller effects due to seeding with greater probability.

The RR results were then adjusted for biases that can occur when operational data are compared to historical records in an *a posteriori* evaluation of non-randomized seeding programs. Brier and Enger (1952) have shown that the variation between sequences of years is different from the variation between random samples of years. Gabriel and Petrondas (1983) have shown that reliable conclusions cannot be drawn from comparisons of operational data with historical records, and have demonstrated the problems encountered in trying to do so. The calculated P-values are likely to be lower than the true level of significance and the calculated confidence limits are likely to be more precise than they really are. Because of these problems, standard statistical methods are prone to indicate effects when there might not be any. Gabriel and Petrondas (1983) suggest that the computed P-values from analyses of this type should be augmented by a factor whose magnitude is proportional to the number of years involved in the evaluation in accordance with a rough guideline that fit their findings. A comparison with re-randomization analysis results indicated that an adjustment to the RR results based on doubling the computed P-value was appropriate, this factor being somewhat smaller than that indicated by the rough guideline of Gabriel and Petrondas (1983). At various stages in the evaluation the results using the regression ratio that were adjusted for bias in this way were compared to those from re-randomization analysis to verify their accuracy for the sample sizes involved. The water year streamflow was evaluated by re-randomization analysis using 30,000 permutations (3 runs of 10,000 permutations each to establish the stability of the results). It was found that the results were indeed approximately equal. The ratio statistics results of the 43-year evaluation of the Kings River operational cloud seeding program were statistically comparable to those from the re-randomization analysis. Although randomization is the only sure way of safeguarding against bias and its influence on the evaluation results, the factor of two adjustment is deemed adequate to provide reasonably unbiased results because (1) the historical and operational periods are quite long so that the potential effect on average streamflows due to year-to-year variability and short-term cycles is mitigated, and (2) the regression ratio takes into account the effect of the long-term trend in natural streamflow through the regression between the target and control, and (3) the ratio statistics methodology is very robust to departures from its inherent assumptions and approximate results from re-randomization analyses extremely well. The bias-adjusted regression ratio values are designated as RR_A.

As was the case for the evaluation by the historical regression method, KGF and MDP+CCR were used as the target and control stations, respectively, for the regression ratio evaluation. The results of the evaluation for the water year streamflow are given in Table 2. In addition to the 90% confidence limits, the level of confidence that the seeding effect was positive (≥ 0%) is shown. The level of confidence that the annual average seeding effect is ≥ 1% is also shown. The level of confidence that the annual average seeding effect is ≥ 1% is taken as a conservative indicator of the cost-effectiveness of the seeding operations. Henderson (2003b) has shown that the benefit-to-cost ratio of increasing streamflow by 1% through the snowpack augmentation programs in the Sierra Nevada mountain watersheds is about 6.8:1, so a threshold of cost effectiveness of 1% is very conservative indeed.

It can be seen from Table 2 that the results of the evaluation suggest a positive effect of seeding. It is estimated there is an increase in streamflow due to seeding at KGF of +5.1% with 90% confidence that the true effect of seeding is somewhere between +1.5% and +8.8%. This result was indeed statistically comparable to the 90% confidence limits estimated by re-randomization analysis that, using the method of Fletcher and Steffens (1996), was +0.9% and +8.9%. The probability that the seeding effect is ≥ 0% and ≥ 1% is equal to or greater than 99.1% and 97.0%, respectively.

It should be noted that the estimate of the effect of seeding on KGF by the historical regression method given in the previous section is the same as that obtained using the bias-adjusted regression ratio, the historical regression estimate being only slightly larger than the regression ratio estimate prior to round-off. This suggests that the historical regression method, as applied to these data, was robust to departures from the assumptions under which it was derived. It remains to be determined whether the historical regression method does, in general, yield reasonably precise estimates of a multi-year effect of seeding. This matter is addressed again in the next section.

Table 2. Results of the KGF evaluation of water year streamflow. Results are given for the proportional effect of seeding, $\delta(\%) = 100*(RR_A-1)$, where RR_A is the bias-adjusted regression ratio, P_0 is the probability (%) that $\delta \geq 0\%$, and P_1 is the probability (%) that $\delta \geq 1\%$ (the threshold of cost effectiveness).

Evaluation Statistic	Result
$\delta(\%)$	+5.1
90% Confidence Limits	
Lower Bound (LB) (%)	+1.5
Upper Bound (UB) (%)	+8.8
P_0	99.1
P_1	97.0

6. TIME EVOLUTION OF SEEDING EFFECTIVENESS

The strength and consistency of the seeding effect over time was investigated in the next step of the evaluation. Towards this end, a progressive statistical evaluation using ratio statistics, called the cumulative year statistical evaluation, was conducted. Using this method the seeding is evaluated as a function of the cumulative number of years of seeding, i.e., initially the first 10 seeded years, then the first 11 seeded years, then the first 12 seeded years, ... , and finally all seeded years. The seeding effect calculated for each seeded water year is the value that would have been obtained if the evaluation were done for all seeded years up to and including that water year. A significant change in trend in the plot of the cumulative year evaluation results as a function of water year (number of seeded years up to and including that water year) should be indicative of a significant change seeding effectiveness and, in turn, a significant change in some aspect of the meteorology and/or seeding procedure that affected the seeding effectiveness. In interpreting the plot of the cumulative year evaluation results in the next section, it should be recognized that, for a given change in trend, the indicated change in seeding effectiveness increases as the number of seeded years, N, included in the evaluation increases. Thus, a small change in trend at large N represents a larger change in seeding effectiveness than it does for smaller values of N.

Figure 3 shows how the water year seeding effect and its 90% confidence limits evolved over time for KGF, the target selected for evaluation in the Kings River Basin. Apparent in Fig. 3 is a noteworthy change in trend that is indicative of a change in seeding effectiveness. It can be seen that there is a 33% increase in the cumulative year seeding effect during the period 1978-1980. For the cumulative year seeding effect to increase by this much, there had to be a marked improvement in seeding effectiveness starting around 1978. The physical cause(s) of this improvement in seeding effectiveness is worthy of further study. It is tentatively postulated that this improvement in seeding effectiveness may be due to (a) a change in meteorological (seedability) conditions, and/or (b) a change in seeding operational procedures (seeding strategy and/or seeding agent).

In the previous section it was found that the seeding effect estimated by the historical regression method only slightly overestimates the seeding effect estimated by the regression ratio method. This raised the question as to whether the historical regression method does, in general, yield reasonably precise estimates of a multi-year effect of seeding. A clue to the answer to this question can be found by comparing the time evolution of the seeding effect estimated by each method (see Fig. 4).

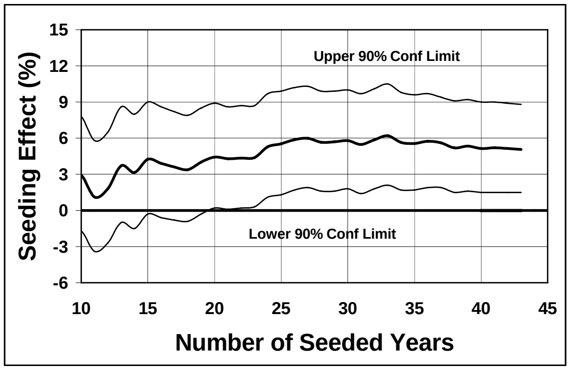

Fig. 3. The proportional seeding effect for KGF, δ(%) = 100*(RR_A-1) where RR_A is the bias-adjusted Regression Ratio, as a function of the cumulative number of seeded years. Also shown are the 90% confidence limits. The seeding effect calculated for each seeded water year is the value that would have been obtained if the evaluation were done for all seeded years up to and including that water year.

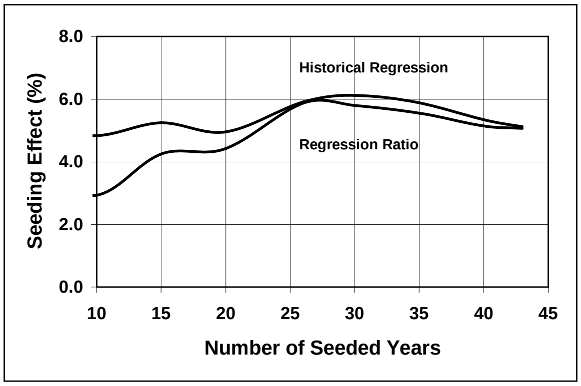

Fig. 4. Cumulative year effect of seeding estimated by the historical regression method and the regression ratio method.

It can be seen from Fig. 4 that the cumulative effect of seeding estimated by the historical regression method always overestimates the cumulative effect of seeding by the regression ratio method. After 10 years the estimated seeding effect by the historical regression method is 65% greater than that by the regression ratio method, after 25 years of seeding it decreases to less than 1% greater, and then it increases to no more than 5.8% greater before falling to 1% greater after 43 years of seeding. Assuming that the regression relationship derived from the historical period is representative of the operational period, as is the case here, the historical regression method may, indeed, yield reasonably precise estimates of a multi-year effect of seeding provided that the natural variability is averaged over a sufficiently long period of years. For the KGF evaluation, it appears that at least 25 years is needed to average out the effects of natural variability on the estimated seeding effect.

7. SUMMARY AND CONCLUSIONS

An independent statistical evaluation of the Kings River operational cloud seeding program over its entire period of operations through water year 2004 was conducted. Two statistical approaches were used in the evaluation of the seeding program, both based on streamflow. The first evaluation of the seeding program was based on the historical regression method, as was used in previous evaluations of the seeding operations. An evaluation of the seeding program was then conducted using ratio statistics and, in particular, the bias-adjusted regression ratio. The following is a summary of the main findings of this evaluation study:

(1) The use of ratio statistics and, in particular, the bias-adjusted regression ratio is a more precise and more reliable evaluation methodology for this seeding project than the traditional historical regression methodology. Although randomization is the only sure way of safeguarding against bias and its influence on the evaluation results, the adjustment to the regression ratio is deemed adequate to provide reasonably unbiased results. The ratio statistics methodology is very robust to departures from its inherent assumptions and it was shown that ratio statistics results approximate those from re-randomization analyses extremely well.

(2) The estimation of seeding effects on an annual basis using the historical regression method is not statistically valid for the Kings River streamflow data. The annual seed/no seed differences produced by the historical regression equation are confounded with and dominated by effects due to natural variability. Consequently, the evaluation yielded misleading results that indicate negative effects of seeding in 28% of the operational years and positive effects in 72% of the operational years. It also indicates positive seeding effects in years when seeding operations were suspended. The historical regression method did, however, yield reasonably precise estimates of the effect of seeding on KGF. This suggests that the historical regression method, as applied to these data, was robust to departures from the assumptions under which it was derived. Assuming that the regression relationship derived from the historical period is representative of the operational period, the historical regression method may yield reasonably precise estimates of a multi-year effect of seeding provided that

the natural variability is averaged over a sufficiently long period of years.

(3) The most precise evaluation results will be obtained by choosing a control that yields the highest correlation with the target and, therefore, the lowest standard deviation of an observation about its true value. Even seemingly small differences in the values of the correlation coefficient and/or the standard deviation between prospective controls could make meaningful differences in the evaluation results they produce. The importance and value of the choosing the most precise control in applying the regression ratio method is discussed in Appendix A.

(4) The Kings River-Pine Flat Dam (KGF) streamflow station in the Kings River Basin was evaluated for the effects of seeding using the ratio statistics method. Evidence for positive, statistically significant and cost effective seeding effects was found. The evaluation of 43 years of seeding on water year streamflow through 2004 resulted in an estimated increase of +5.1% with 90% confidence that the true effect of seeding lies somewhere between +1.5% and +8.8%. The probability that the effect of seeding is equal to or greater than 0% and 1% (the threshold of cost effective operations) is 99.1% and 97.0%, respectively.

(5) An analysis of the time evolution of seeding effectiveness revealed that there was a marked improvement in seeding effectiveness that started around 1978, the physical cause(s) of which is worthy of further study.

It is reiterated 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 hypotheses/analyses, some of which are suggested by the results of previous analyses. In view of these considerations, the results should be interpreted as measures of the strength of the suggested seeding effect and not as measures of statistical significance. Nevertheless, the estimated effects of seeding should be of considerable value to the Kings River Conservation District (KRCD) in determining the past, present and future value of their cloud seeding operations according to risk criteria used in their business operations. As Boe et. al.(2004) state, "… if a potential sponsor of a cloud seeding program, following careful deliberation, decided they had an 80% likelihood of obtaining a 10% increase in precipitation that would yield a benefit/cost ratio of 10:1, they would probably choose to support the program." According to Henderson (2003b) an increase in streamflow of only 1.5% in the Sierra Nevada Mountain watersheds would yield a benefit/cost ratio of 10:1.

8. REMARKS

The results of these analyses provide incentive to do a more in-depth study of past seeding operations with the ultimate aim of optimizing the cost effectiveness of future seeding operations. The strength and consistency of the increases in streamflow due to seeding suggested by this study are larger than one would expect based on 1) the results of previous physical studies of the seedability of orographic clouds over the Sierra Nevada Mountains of California, and 2) the problems encountered in the transport and diffusion of the seeding material to the right place and at the right time in the clouds (see, e.g., Marwitz, 1987; Reynolds and Dennis, 1986, Reynolds, 1988, Deshler et. al., 1990, Reynolds, 1996). New studies are, therefore, needed to clarify and extend the results, and to resolve the uncertainties and inconsistencies in the statistical and physical evidence obtained thus far. Progress in physical understanding comes from noting the unexpected and following it up as well as from confirming the expected. Mindful that the results from *a posteriori* analyses might evince a physically interesting result that in fact might only reflect chance, strong statistical support for a result alerts the physical scientist even though there is no ready theory to explain the results or the findings run counter to the postulated seeding conceptual model or the findings appear to be inconsistent with the findings of previous physical studies. Physical understanding is clarified and advanced through follow-up statistical and physical studies and experiments prompted by such findings. From a rigorous statistical standpoint, the suggested effects that are indicated must be confirmed through new, *a priori,* randomized experiments specifically designed to establish their validity. However, the continuity of the operational cloud seeding program should not be disturbed so the randomized experiments should be piggybacked on the operational program in a manner similar in concept to that suggested by Gabriel and Changnon (1982).

Appendix A
Choosing the Best Available Control for the Evaluation

In conducting a target-control statistical evaluation of a cloud seeding program using the regression ratio method, sometimes there are a number of potential controls from which to choose. A potential control is a streamflow station that has not been seeded, is highly correlated with the target, and has a long enough period of full natural flow data during the historical and operational period to support a meaningful evaluation. It is important to select the control that yields the most precise evaluation results. The usual practice is to choose the control that is most highly correlated with the target. In doing so, the control with the smallest standard deviation of the residuals (differences between the observed and predicted values), s_o, is chosen. In regression ratio analysis, the control that maximizes ρ and minimizes s_o will usually be the most efficient control. Even seemingly small differences in the values of ρ and/or s_o between prospective controls could make meaningful differences in the evaluation results they produce.

Take, for example, what happens when the choice of control is limited to the best single control rather than a combination of controls. In the case of evaluation of the Kings River-Pine Flat Dam (KGF), one would use MDP (ρ = 0.959 and s_o = 251,702) instead of the combination MDP+CCR (ρ = 0.988 and s_o = 141,686) as used in obtaining the results given in Table 2. Although the correlation coefficient for MDP alone is quite high and the differences in ρ and s_o between MDP and MDP+CCR are relatively small, the evaluation results they produce are quite different. The estimated increase in streamflow using MDP+CCR as the control is +5.1% with 90% confidence that the true effect of seeding is somewhere between +1.5 % and +8.8%. On the other hand, the estimated increase in streamflow using just MDP as the control is –0.2% with 90% confidence that the true effect of seeding is somewhere between -6.5 % and +6.4% Whereas the estimate of the seeding effect using MDP+CCR as the control is statistically significant, the estimate of the seeding effect using just MDP as the control is not.

Figs. A1a and A1b show the differences in the time evolution of the bias-adjusted regression ratio (RR_A) and the standard error of the estimate $SEE(RR_A)$, respectively, in the evaluation of KGF using MDP+CCR and MDP alone as the control. It can be seen in Figs. A1a and A1b that the values of RR_A are consistently higher and the values of the $SEE(RR_A)$ are consistently lower in the evaluation using MDP+CCR than that using MDP alone. The marked increase in the cumulative effect of seeding that started around 1978 is common to both evaluations. However, it can be seen that the dramatic decrease in seeding effectiveness that started around 1992 that is readily apparent in the evaluation using MDP as the control is no longer apparent in the evaluation using MDP+CCR as the control. In addition, the short period fluctuations in the time evolution of seeding effectiveness that appear in the evaluation using MDP+CCR as the control appear to be exaggerated in the evaluation using MDP as the control.

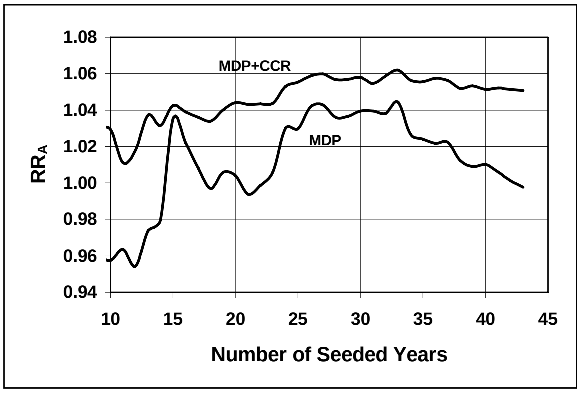

Fig. A1a. Time evolution of the bias-adjusted regression ratio (RR_A) in the evaluation of Kings River-Pine Flat Dam (KGF) using MDP and MDP+CCR as the control.

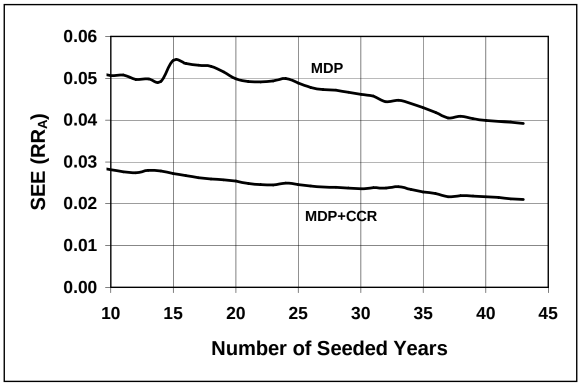

Fig. A1b. Time evolution of the standard error of the estimate, $SEE(RR_A)$, in the evaluation of Kings River-Pine Flat Dam (KGF) using MDP and MDP+CCR as the control.

Why does the addition of CCR make such a large difference in the evaluation results? The statistical reasons for this can be found in the test statistic used by the regression ratio method upon which the results are based. The test statistic $X = \mu_{1-P/2}$ (where μ is the parameter of the standardized normal distribution and P is the two-sided P-value) is directly proportional to. RR_A and inversely proportional to the standard error of the estimate, $SEE(RR_A)$, i.e.,

$$X = \mu_{1-P/2} \propto \ln(RR_A)/SEE(RR_A) = \ln(SR/SR_{PRED}) / \{SEE(SR)^* (1- \rho^2)^{1/2}\}$$

The P-value and other probabilistic results (like P_0 and P_1) of the evaluation are determined by the magnitude of X. As can be seen from Fig. A1a, the use of MDP+CCR improves the precision of the value of SR_{PRED} in the calculation of the regression ratio and results in a larger estimate of the value of RR_A. The use of MDP+CCR also increases the correlation coefficient (ρ) and, as can be seen in Fig. A1b, that results in a decrease in $SEE(RR_A)$, since SEE for the regression ratio, $SEE(RR_A)$, is equal to SEE(SR) times $(1- \rho^2)^{1/2}$. Both change in the appropriate direction to increase X. Since the P-value is inversely proportional to the absolute value of X in the standardized normal distribution, the P-value is decreased and the values of P_0 and P_1 increase. The regression ratio is uniquely capable of taking advantage of the relevant information in the target-control regression relationship to produce the most precise evaluation results possible with the available data.

The physical reasons why the addition of CCR makes such a large difference in the evaluation results are yet to be determined. Clearly CCR is capturing an element of variability in KGF that is not being captured entirely by MDP alone. Henderson (2003a) showed that there are several different types of weather patterns with their associated storm tracks and physical characteristics that affect the watersheds of the southern Sierra Nevada Mountains of California. It appears that MDP and CCR do a better job of sampling the storms that affect KGF than does MDP alone. The meteorology of storms that affect Kings River Basin should be investigated to determine why this is so.

Acknowledgments: This study was motivated by the preliminary results of the author's ongoing evaluation of the Southern California Edison (SCE) Upper San Joaquin River Basin Weather Modification Program. The SCE-sponsored evaluation of the San Joaquin weather modification program is the first operational (non-randomized) cloud seeding program to be evaluated using the ratio statistics methodology. Support for the preparation and publication of this paper is provided by the UAE International Prize for Excellence in Advancing the Science and Practice of Weather Modification. The author is grateful to the Department of Atmospheric Studies, Ministry of Presidential Affairs of the United Arab Emirates (UAE) for establishing the UAE Prize in coordination with the WMO. The author thanks the distinguished Evaluation Committee for the honor of being selected as a winner of the UAE Prize. The author also thanks Mr. Lyn Garver of the Kings River Conservation District for his cooperation in providing information concerning the operational seeding program that helped in the interpretation of the evaluation results.

9. REFERENCES

Boe, B., G. Bomar, W.R. Cotton, B.L. Marler, H.D. Orville (Chair), and J.A. Warburton, 2004: The Weather Modification Association's Response to the National Research Council's Report Titled "Critical Issues in Weather Modification Research", *J. Wea. Mod.*, **36**, 53-82.

Brier, G.W. and I. Enger, 1952: An analysis of the results of the 1951cloud seeding operations in central Arizona. *Bull. Amer. Meteor. Soc.*, **33**, 208-210.

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

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

Fletcher, L. and F.E. Steffens, 1996: The use of permutation techniques in evaluating the outcome of a randomized storm seeding experiment. *J. Appl. Meteor.*, **35**, 1546-1550.

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

Gabriel, K.R., 2000: Planning and evaluation of weather modification projects. *Seventh WMO Sci. Conf. on Wea. Mod.*, Chiang Mai, Thailand, Supplement, Vol. III, World Meteor. Org., 39-59.

Gabriel, K.R., 2002: Confidence regions and pooling – some statistics for weather experimentation. *J. Appl. Meteor.,* **41**,505-518.

Gabriel, K.R. and S.A. Changnon. 1982: Piggyback weather experimentation: Superimposing randomized treatment comparisons on commercial cloud seeding operations. *J. Wea. Mod.*, **14,** 7-10.

Gabriel, K.R. and D. Petrondas, 1983: On using historical comparisons in evaluating cloud seeding operations. *J. Climate Appl. Meteor.*, **22**, 626-631.

Henderson, T.J., 1966: A ten year non-randomized cloud seeding program on the Kings River in California . *J. Appl. Meteor.,* **5**, 697-702.

Henderson, T.J., 2003a: The Kings River weather resources management program. *J. Wea. Mod.*, **35**, 45-51.

Henderson, T.J., 2003b: New assessment of the economic impacts from six winter snowpack augmentation projects. *J. Wea. Mod.*, **35,** 41-44.

Marwitz, J. D., 1987: Cloud Physics Studies in SCPP from 1977-1987. Report No. AS 156, Department of Atmospheric Science, University of Wyoming, Laramie, WY, 101 pp.

NAWC, 1978: Report on cloud seeding operations in the Upper San Joaquin Basin 1974-75 and evaluation of twenty-five years of operations 1951-1975. Prepared for the Southern California Edison Company. NAWC Report No. 6-138, September 1978, 39 pp [available from Southern California Edison Company, P.O. Box 800, Rosemead, CA 91770]

Nicholls, N., 2001: The insignificance of significance testing. *Bull. Amer.Meteor. Soc.*, **82**, 981-986.

Reynolds, D.W., 1988: A report on winter snowpack augmentation. *Bull. Amer. Meteor. Soc.*, **69**, 1290-1300.

Reynolds, D.W., 1996: The effects of mountain lee waves on the transport of liquid propane-generated ice crystals. *J. Appl. Meteor.*, **35**, 1435-1456.

Reynolds, D.W. and A.S. Dennis, 1986: A review of the Sierra Cooperative Pilot Project. *Bull. Amer. Meteor. Soc.*, **67**, 513-523.