Evaluation of the Wyoming Weather Modification Pilot Project (WWMPP) Using Two Approaches: Traditional Statistics and Ensemble Modeling

Roy M. Rasmussen, Sarah A. Tessendorf, Lulin Xue, Courtney Weeks, Kyoko Ikeda, Scott Landolt, and Dan Breed

National Center for Atmospheric Research, Boulder, Colorado

Terry Deshler

University of Wyoming, Laramie, Wyoming

Barry Lawrence

Wyoming Water Development Office, Cheyenne, Wyoming

ABSTRACT

The Wyoming Weather Modification Pilot Project randomized cloud seeding experiment was a crossover statistical experiment conducted over two mountain ranges in eastern Wyoming and lasted for 6 years (2008–13). The goal of the experiment was to determine if cloud seeding of orographic barriers could increase snowfall and snowpack. The experimental design included triply redundant snow gauges deployed in a target–control configuration, covariate snow gauges to account for precipitation variability, and ground-based seeding with silver iodide (AgI). The outcomes of this experiment are evaluated with the statistical–physical experiment design and with ensemble modeling. The root regression ratio (RRR) applied to 118 experimental units provided insufficient statistical evidence (p value of 0.28) to reject the null hypothesis that there was no effect from ground-based cloud seeding. Ensemble modeling estimates of the impact of ground-based seeding provide an alternate evaluation of the 6-yr experiment. The results of the model ensemble approach with and without seeding estimated a mean enhancement of precipitation of 5%, with an inner-quartile range of 3%–7%. Estimating the impact on annual precipitation over these mountain ranges requires results from another study that indicated that approximately 30% of the annual precipitation results from clouds identified as seedable within the seeding experiment. Thus the seeding impact is on the order of 1.5% of the annual precipitation, compared to 1% for the statistical–physical experiment, which was not sufficient to reject the null hypothesis. These results provide an estimate of the impact of ground-based cloud seeding in the Sierra Madre and Medicine Bow Mountains in Wyoming that accounts for uncertainties in both initial conditions and model physics.

1. Introduction

The main source of water in the western United States is snowmelt from major mountain ranges. This resource is renewed annually through snowfall that melts in the spring. The timing of the snowmelt corresponds well with the spring growing season, allowing agriculture to exist in this usually semiarid region. As the reality of climate change becomes more evident, water managers are increasingly concerned about the future amount and timing of snowmelt. The recent report by the Bureau of Reclamation (2012), "Colorado River Basin Water Supply and Demand Study," presents an in-depth analysis of the future of water resources in the Colorado River basin, United States, and concludes that future demand will likely outstrip the available water supply when accounting for climate change and expected increases in population. Thus, conservation and other measures are needed to avoid a future water crisis in this region. The report listed 13 measures to help avert this crisis including the possibility of cloud seeding of winter orographic clouds to increase snowpack.

ə Denotes content that is immediately available upon publication as open access.

Corresponding author: Roy M. Rasmussen, rasmus@ucar.edu

While orographic cloud seeding using silver iodide (AgI) is commonly performed over many western U.S. mountain ranges today, estimating the likely impact on snowpack is difficult because the seeded snowfall signal is small compared to the large and highly variable natural precipitation. It is impossible to know what the natural precipitation would have been during a storm that was seeded. Therefore, a randomized seeding program, lasting multiple years and using multiple correlated target areas, was recommended for the Wyoming Weather Modification Pilot Project (WWMPP; Breed et al. 2014, hereafter B14) to provide the data needed to perform an evaluation of the likely amount of extra precipitation produced as a result of cloud seeding over two mountain ranges in Wyoming.

The WWMPP following the design of B14 was a 6-yr, randomized cloud-seeding experiment in operation from 2008 to 2013. The statistical results of the WWMPP, presented here, fail to reject the null hypothesis that there is no seeding effect. While the statistical analysis indicates a small, positive seeding effect, it is not statistically significant (*p* value of 0.28), largely because of the relatively small effect compared to the large natural variability. The statistics could possibly be improved by conducting a longer program, but the costs preclude such an option. The WWMPP generated a wealth of data and provides opportunities for further investigations of cloud seeding based on the quality of the data collected and the length of the experiment. Here, results of ensemble modeling of the WWMPP are presented in addition to the statistical analysis of the physical experiment in section 2.

The ensemble modeling approach uses large numbers of ensemble members to evaluate the likely ground-based seeding effect on precipitation during the WWMPP. A modeling approach is possible because of recent advances in computational and modeling capabilities allowing thousands of simulations to be performed (section 3). The use of multiple simulations of a single case enables characterization of the seeding signal within highly variable natural precipitation for even small seeding effects. By using an ensemble of model runs that are initialized and configured to represent the range of recognized uncertainties in the model setup, an estimate of the seeding effect is possible using the ensemble mean and running the model with and without the impact of seeding. However, before trusting the ensemble mean, the model has to be shown to capture the essential physics of cloud and precipitation formation. This can be accomplished through comparisons with observations. This is done using the 6-yr quality-controlled snow gauge dataset from the WWMPP. Discussion of results and final comments and conclusions are presented in section 4.

2. Statistical results of cloud seeding from the WWMPP

The WWMPP was undertaken to determine whether seeding Wyoming's orographic clouds with AgI enhances the natural precipitation process resulting in a measurable increase in snowpack. The project design and background are described in detail in B14, including the design of the statistical evaluation. The evaluation was primarily focused on a statistical, randomized crossover design in which two similar ranges were observed concurrently. It also included a physical evaluation based on observations and modeling studies. This approach follows National Research Council recommendations that evaluations of cloud seeding include statistical, physical, and modeling components (National Research Council 2003).

For two mountain ranges whose precipitation is well correlated, a randomized crossover statistical design allows robust results to be computed with fewer experimental units (EUs), as long as the two ranges are well correlated, compared to two ranges evaluated independently using a single-target design (Gabriel 1999). In the WWMPP, an EU was chosen to be a 4-h period of time during which AgI ground generators on one of the two mountain ranges are active. The seeded range was randomly selected. Based on historical SNOTEL data, the daily snowfall between the Medicine Bow Mountains and Sierra Madre has a temporal correlation of ~0.5, sufficiently high to use these two ranges for the statistical evaluation using a randomized crossover statistical design.

The experimental design (B14) specifies that seeding of one range was to take place during 4-h EUs when conditions in the two ranges meet certain criteria: 1) 700-hPa temperature less than −8°C (from sounding or model), 2) supercooled liquid water (SLW) present in both ranges (from radiometers or model), and 3) 700-hPa wind direction between 260° and 315° (from sounding or model). The conditions these criteria characterize are considered optimal for AgI seeding effects as described more fully in B14.

Snowfall measurements using project gauges during EUs from the 6-yr project showed an average correlation of 0.58 between the two mountain ranges. This is slightly better than that estimated from SNOTEL data prior to the experiment, further supporting the use of the crossover design for this project.

The statistic specified for the evaluation (B14) is the root regression ratio (RRR; Gabriel 1999). This statistic requires data collected from covariate gauge sites located in areas upstream and crosswind of the two ranges. This allows for better accounting of the natural precipitation (Fig. 1; Gabriel 1999).

B14 describes the design, instrumentation, approach to selecting EUs, and proposed evaluation for the
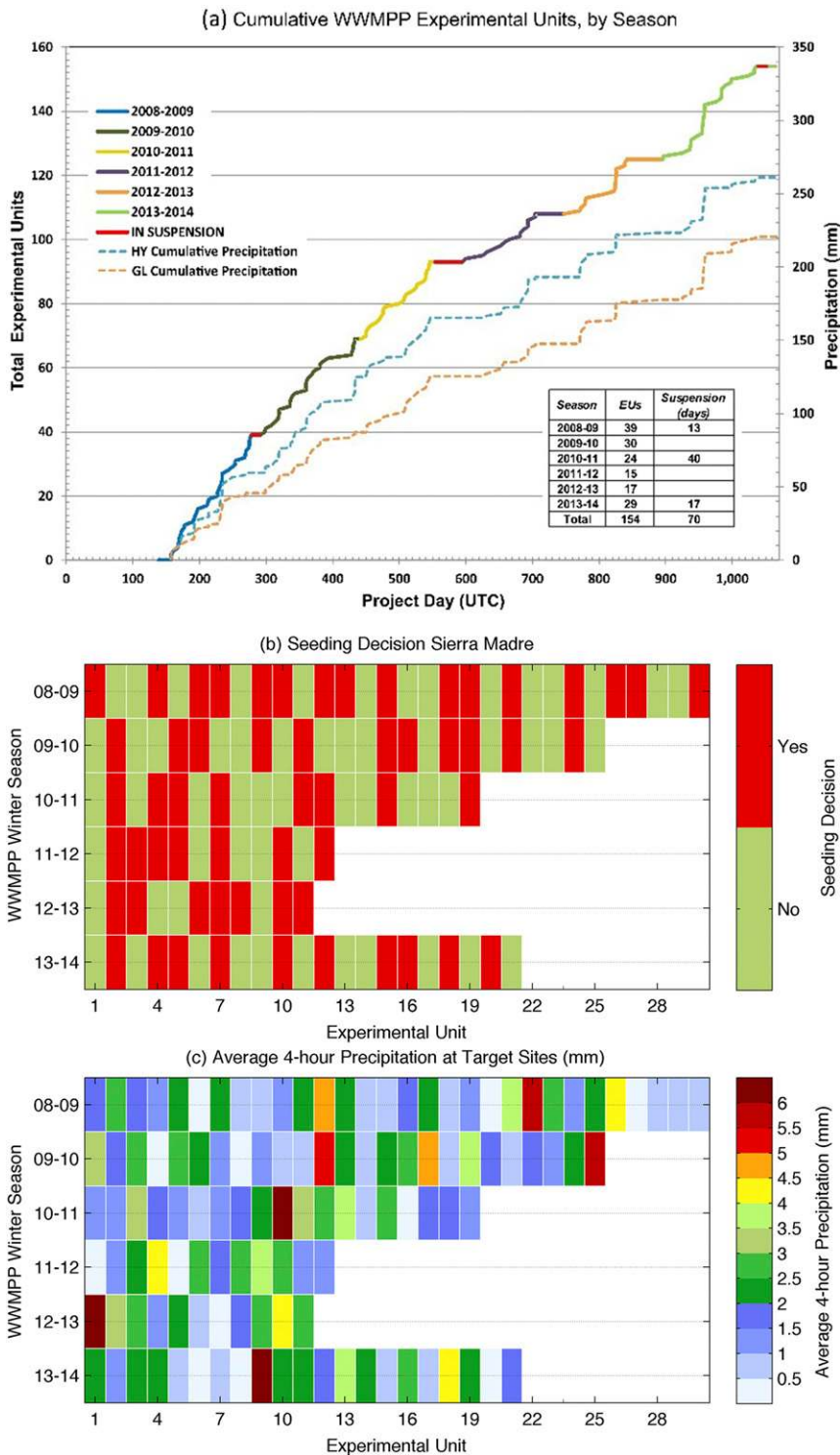
FIG. 1. Summary of EUs for all 6 years of the WWMPP. (a) History of accumulation of EUs over the 6 years as a function of project day. Red periods show when the project was suspended because of snowpack conditions exceeding 120% of normal. Figure courtesy of Bruce Boe. (b) Checkerboard plot of the seeding decision for the Sierra Madre by year. Red indicates a seeded case and green is an unseeded case. (c) Checkerboard plot of the 4-h precipitation averaged at both target sites (mm).

TABLE 1. The number of WWMPP EUs by year and month. In parentheses, the EUs eliminated because of QC on the precipitation data are shown in regular font and those EUs eliminated because of generator performance issues in boldface font. In situations where an EU was excluded both by precipitation data QC and by generator performance issues, it is considered eliminated by precipitation data. This occurred in two months for two cases each (indicated by a plus sign). Months in which suspension criteria were met are indicated by an asterisk.

Month	2008/09	2009/10	2010/11	2011/12	2012/13	2013/14	Total
Nov	0 (0, **0**)	2 (0, **1**)	4 (0, **0**)	1 (0, **1**)	0 (0, **0**)	1 (0, **0**)	8 (0, **2**)
Dec	12 (0, **2**)	9 (0, **0**)	6 (2, **0**)	3 (1, **0**)	5 (0, **0**)	6 (1, **0**)	41 (4, **2**)
Jan	8 (1, **0**)	7 (0, **1**)	5 (0, **2**)	3 (1, **0**)	6 (0, **0**)	11 (1, **0**)	40 (3, **3**)
Feb	9 (1, **0**)	6 (1, **0**)	9 (0, **1**)	6 (0, **0**)	6 (6, **0**$^+$)	7 (5, **0**$^+$)	43 (13, **1**$^+$)
Mar	9 (3, **0**)	0 (0, **0**)	0* (0, **0**)	2 (0, **0**)	0 (0, **0**)	4 (1, **0**)	15 (4, **1**)
Apr	1* (1, **0**)	6 (2, **0**)	0* (0, **0**)	0 (0, **0**)	0 (0, **0**)	0 (0, **0**)	7 (3, **0**)
Total	39 (6, **3**)	30 (3, **2**)	24 (2, **3**)	15 (2, **1**)	17 (6, **0**$^+$)	29 (8, **0**$^+$)	154 (27, **9**)

WWMPP. The WWMPP was conducted for six winters (2008–13) and collected 154 EUs, each consisting of 4-h blocks of precipitation data at target and covariate sites. This number was reduced to 118 through quality-control procedures. Figure 1 provides information about the EUs over the entire project, and Table 1 gives the distribution of EUs by month and year, with the eliminated cases shown in parentheses for each month.

In this study, the primary and associated analyses of the WWMPP are presented based on the quality-controlled 4-h snow gauge precipitation data collected during the 118 acceptable EUs. In Fig. 1a, the number of EUs is plotted as a function of project day and color coded according to the year the EU occurred. Precipitation at each of the target sites is also shown and visually demonstrates a strong correlation between the two. A color-coded plot of the seeding decisions by case and year for the Sierra Madre is shown in Fig. 1b. Since seeding occurred either in the Sierra Madre or in the Medicine Bow Mountains (but not both), a similar plot (not shown) would have the exact opposite shading for seeding in the Medicine Bow Mountains. Note that the seed decisions are randomly distributed as expected from the experimental design.

a. Data collection and quality control

The topography of the Sierra Madre and Medicine Bow Mountains is shown in Fig. 2, along with instrumentation used for the WWMPP. Seeding was conducted using 16 ground-based AgI generators (B14) located at the red triangles in Fig. 2. Eight generators are located upstream (to the west) of each range. Precipitation was measured in both ranges at the eight gauge sites (black squares with plus signs within), four sites for each range. The target site for the Sierra Madre was identified as the Highway 47 (HY) gauge location and the target site for the Medicine Bow Mountains was located at the Glacier Lakes Ecosystem Experiments Site (GL). The other gauge sites served as either covariates (upstream or crosswind sites), or as consistency checks for the target gauges [Battle Pass (BP) and Towner Lake (TL)]. Each gauge site was equipped with three snow gauges to allow for appropriate quality-control procedures to be applied. The Battle Pass and Towner Lake sites were initially considered as part of an extended target site but were later eliminated because of the failure of the Vaisala VRG snow gauges located at those sites. A description of the snow gauge deployment and type of gauges used is presented in B14.

Measuring snow is challenging because of a number of factors including wind-induced undercatch, snow accumulation (capping) on the gauge leading to random snow dumps, and temperature-dependent weighing sensors (Rasmussen et al. 2012). As a result, a key element of the design (B14) included quality-control protocol for the snow gauge data to ensure the best possible data was utilized for the project evaluation. The appendix presents the snow gauge quality-control methods including an innovative methodology to eliminate the effects of diurnal temperature variability on Electronic Technologies, Inc. (ETI) gauges.

To reduce the introduction of bias while performing the gauge analysis, a primary analyst was assigned the quality-control task at the beginning of the project. All information on seeding days, times, and ranges was withheld from the analyst. Three passes through the precipitation data were conducted. The first pass considered the daily raw gauge data during all six winters. This quality-control process categorized each day and time period as one of good, bad, or questionable data quality and provided an initial dataset for the statistician to use to calculate preliminary statistics.

The second pass involved a joint analysis by the primary analyst and a second analyst for only the gauge data from days containing EUs, however, still working without knowledge of which range was seeded and the time of seeding. This second analysis provided a quality-control evaluation that modified only about 2% of the
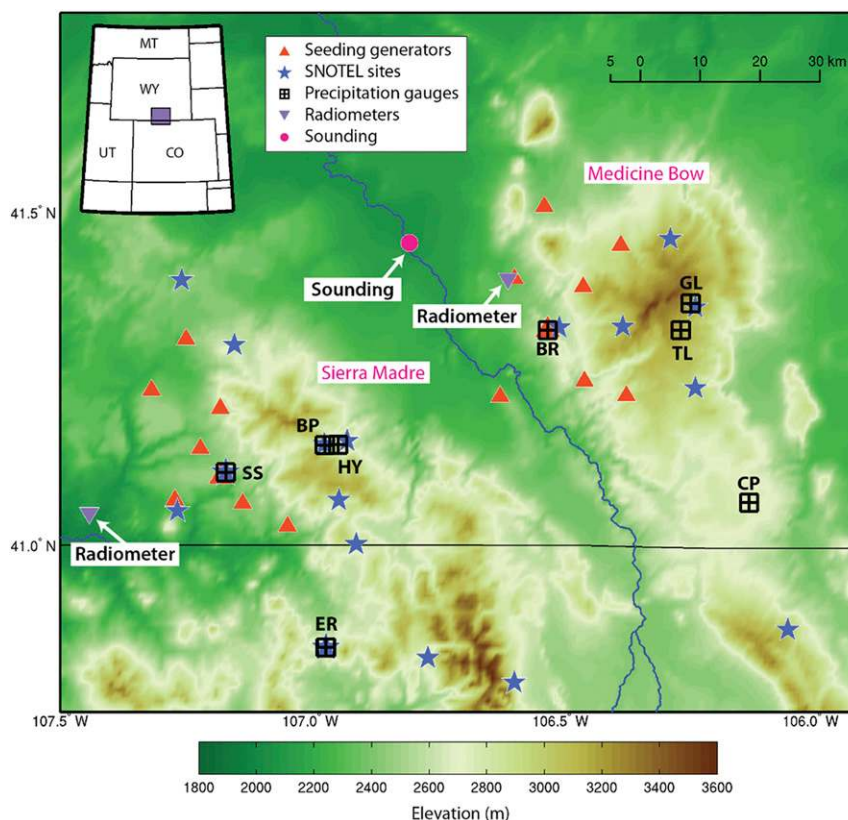
FIG. 2. Depiction of topography and instrumentation deployment in the Sierra Madre and Medicine Bow Mountains for the WWMPP. HY was the target site for the Sierra Madre, and GL was the target site for the Medicine Bow Mountains. The covariate sites were SS and ER for the Sierra Madre and BR and CP for the Medicine Bow Mountains. Alternative target sites were BP for the Sierra Madre and TL for the Medicine Bow Mountains.

cases that were accepted during the first pass by the primary analyst.

The third and final pass used the temperature-corrected data resulting from applying the corrections given in the appendix. This final analysis was conducted by the primary and secondary analysts as well as the analyst that conducted the gauge quality control (QC) of the ETI gauge data. These three analysts were chosen based on their more than 50 years of combined experience analyzing snow gauge data (Rasmussen et al. 2012, 1999; Ikeda et al. 2010) from projects associated with the Federal Aviation Administration (FAA), National Oceanic and Atmosphere Administration (NOAA), and the World Meteorological Organization (WMO).

About 10% of the categorizations made in the second pass through the data were modified based on the third analysis. If bad or questionable data existed, or data were missing at any of the eight gauge sites during an EU, that EU was eliminated from the dataset. This quality-control procedure resulted in the elimination of 27 EUs (~17%) and are indicated by the numbers in the parentheses of Table 1. An additional nine cases were eliminated as a result of AgI generators failing to perform as specified in B14. These are indicated by the boldface numbers in parentheses in Table 1. The final number of EUs was thus reduced to 118. This dataset was used for the primary statistical analysis.

b. Primary statistical analysis

The primary statistical analysis follows the design described in B14 and consists of a statistical evaluation using the RRR statistic. In the following section, the environmental characteristics of the 4-h EUs is examined first to assess the consistency of the EUs, followed by the primary statistical analysis results.

Each EU was called by the meteorologists working for the seeding contractor, Weather Modification, Inc. (WMI), based on the pre-established criteria (see B14 for details):

- 700-hPa temperatures less than or equal to $-8°C$.
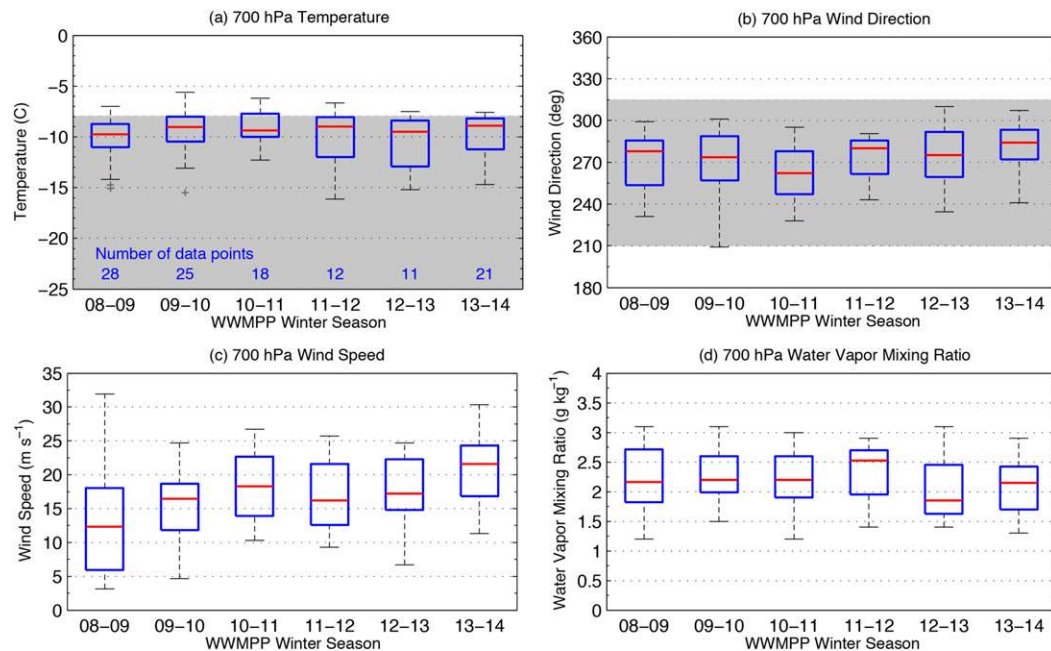- 700-hPa wind direction between $210°$ and $315°$.

FIG. 3. Box-and-whisker plots of 700-hPa (a) temperature, (b) wind direction, (c) wind speed, and (d) water vapor mixing ratio from the Saratoga sounding for the EUs with available sounding data (115 of the 118 total EUs) during the WWMPP by winter season. The red line indicates median values, the blue boxes represent the 25% and 75% inner quartiles of the data, and the whiskers illustrate data beyond the inner-quartile range that are not considered outliers and defined as anything beyond 1.5 times the interquartile range. The number of data points included in each season (one per EU with data) are listed in (a) in blue. The gray shading indicates seeding criteria for temperature and wind direction.

- The simultaneous occurrence of SLW in both ranges as determined by the microwave radiometers.

These criteria were established by the project design to produce a consistent set of EUs that would have a high likelihood of exhibiting a seeding effect, while also meeting the requirements for a crossover design. In the following, we examine how well these conditions were met throughout the years of the project.

Figure 3 uses the Saratoga soundings launched just prior to or during most EUs to create box-and-whisker plots of temperature, wind direction, wind speed, and water vapor mixing ratio (an indication of the moisture available in the atmosphere). Figure 3a shows that the median 700-hPa temperature of the EUs ranged between $-9°$ and $-10°$C, with most of the values colder than $-8°$C, consistent with case-calling criteria. Observed wind direction at 700 hPa ranged from $210°$ to $310°$ (Fig. 3b), also consistent with the case-calling criteria.

While not in the EU case-calling criteria, wind speed and water vapor mixing ratio are shown in Figs. 3c and 3d. Median wind speeds at 700 hPa ranged from 12 to $22 \, \mathrm{m \, s^{-1}}$ with the highest wind speeds occurring in 2013/14 and the lowest in 2008/09. A plot of the ranges of wind direction and wind speed (Fig. 4) shows the cases to be relatively consistent from year to year. Atmospheric

moisture (as indicated by the 700-hPa water vapor mixing ratio; Fig. 3d) per EU was very consistent from season to season.

Cloud-top temperature was not a criterion for case calling, yet may be a source of variability in cloud-seeding efficacy. Therefore, the cloud-top temperature from infrared satellite measurements was also examined for each of the EUs (not shown). The results indicate that the seasonal temperature medians ranged between $-22°$ and $-26°$C. Cirrus often obscured lower clouds, so this analysis may not accurately reflect the cloud-top temperature of the seeded clouds. Other quantities not considered are aerosol content and updraft velocity. In this analysis, these quantities are assumed, for a given flow regime, to be comparable across both mountain ranges.

Radiometer data collected at sites upwind of each range (one instrument per range) indicate that SLW was present in the EUs as required by the case-calling criteria (Fig. 5). It is not possible to directly compare the measurements at the Sierra Madre site (Savery, Wyoming) with that of the Medicine Bow (Cedar Creek) site given the complex geometry of the viewing angles, siting relative to each range, and other factors. These physical differences may be more significant contributors to the observed differences between the sites than the SLW differences themselves. Nonetheless,
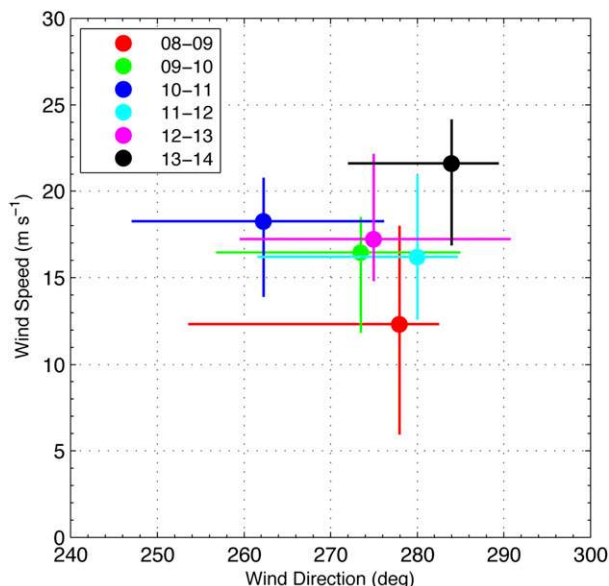
FIG. 4. Inner-quartile range of 700-hPa wind speed and wind direction of the 118 EUs for the six winter seasons of the project. Center points represent the median of each distribution.

the results show that SLW above 0.02 mm was present in the vast majority of EUs in both ranges.

Finally, the range of median precipitation from year to year for the 4-h EUs (Fig. 6) was between 1.5 and 2.0 mm, with a very small range in the monthly values. There is a large difference (a factor of 3–5) in the number of precipitation events occurring in December– February compared with November, March, and April. Yet, there is almost no difference in the precipitation amount, probably reflecting the colder temperatures and low water content during the winter months.

c. Primary statistical evaluation

The environmental analysis of the EUs shows that the case-calling criteria were generally met, allowing us to proceed to the primary statistical analysis using quality-controlled snow gauge data from the EUs to perform the primary statistical evaluation, as described in B14.

The requirements for confidence in a statistical test of precipitation enhancement for the WWMPP was set at a (one tailed) statistical significance level of 0.05 to minimize the potential for false-positive results and a statistical power of 0.8 to control for false-negative results. The RRR test statistic was used to compare precipitation during seeded and unseeded events (Gabriel 1999; List et al. 1999). The equation for the RRR statistic is given below. A detailed description of the RRR statistic and its derivation is given in B14.

A key aspect of the current program design is the use of covariate gauge sites. Data at these sites are used to
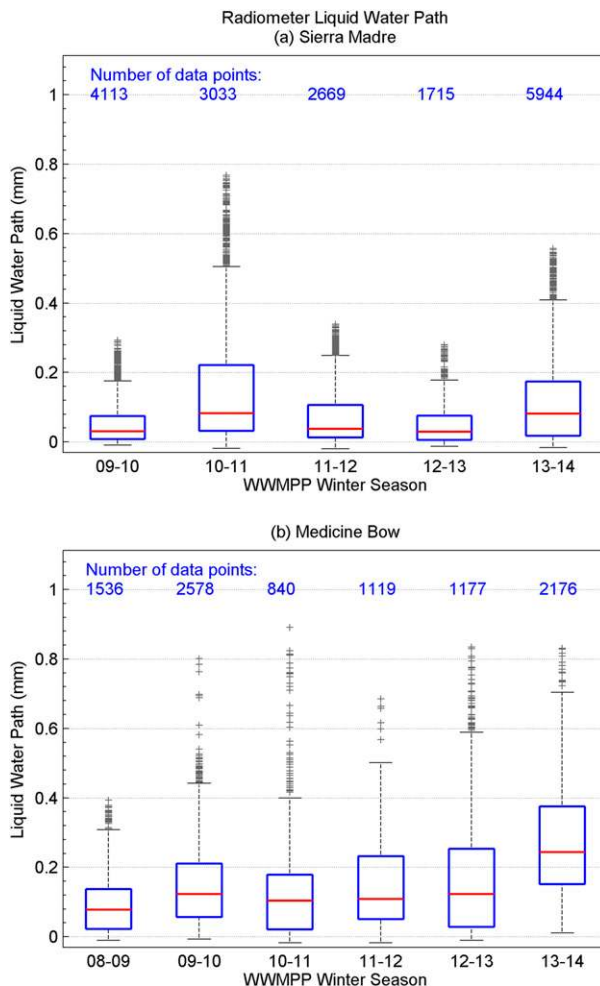

FIG. 5. Box-and-whisker plots (as in Fig. 3) of the radiometer liquid water path (LWP) during EUs when data were available for (a) the Sierra Madre radiometer and (b) the Medicine Bow radiometer by winter season. The number of data points included in each season (up to 240 per EU using 1-min data) are listed atop each panel in blue.

create a statistical model that predicts differences in measured natural precipitation between the targets. This model does not use information about seeding actions and is used to estimate differences in precipitation at the target sites due to natural variation. Accounting for this difference reduces the impact of natural variability in the precipitation and thus makes differences due to seeding easier to detect. This adjustment can be accomplished by estimating the coefficients b_1 and b_2, which minimize the value of the squared error:

$$[y_i - x_i - b_1(z_{1i} - 1) - b_2(z_{2i} - 1)]^2, \qquad (1)$$

where y_i and x_i denote relative precipitation values measured at the two target gauge sites (Y and X) and
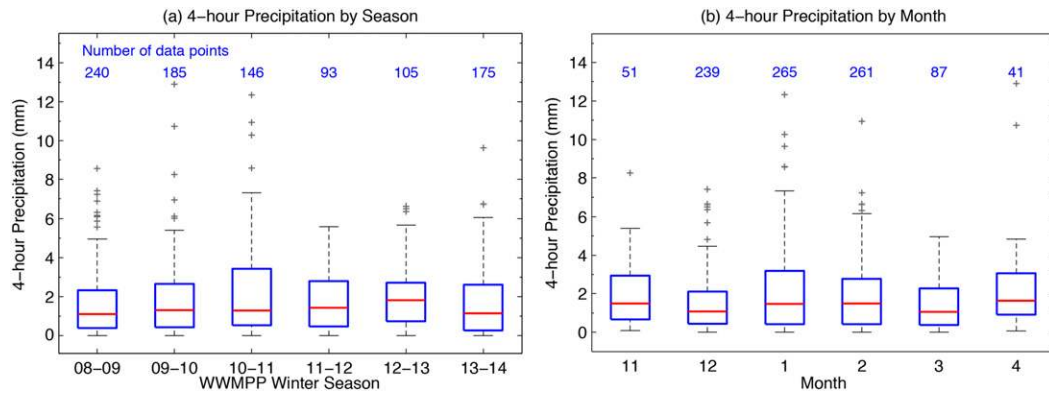
FIG. 6. Box-and-whisker plots of the 4-h precipitation from each of the eight WWMPP snow gauge sites for the 118 EUs during the WWMPP by (a) winter season and (b) month over all six winter seasons.

z_{1i} and z_{2i} denote relative precipitation values measured at the two covariate gauge sites (Z_1 and Z_2). This results in a multiple regression equation that relates the differences in relative precipitation at the two targets to the relative precipitation at the two covariates (Gabriel 1999). If the precipitation values from the covariate gauge sites are not correlated with the precipitation values at the target gauge sites, the coefficients b_1 and b_2, which represent weights of the covariate gauges, will approach zero.

This information is used to scale the root double ratio (RDR) equation:

$$\text{RDR} = \sqrt{\frac{\sum_{i=1}^{n} \theta_i y_i}{\sum_{i=1}^{n} (1-\theta_i)y_i} \frac{\sum_{i=1}^{n} (1-\theta_i)x_i}{\sum_{i=1}^{n} \theta_i x_i}}. \quad (2)$$

The precipitation values at the two targets for EU event i are denoted by x_i and y_i. For this ratio test statistic, $\theta_i = 0$ when target X is seeded and $\theta_i = 1$ when target Y is seeded.

Equation (2) is scaled by the square root of product of the covariate single ratios to provide the RRR equation,

$$\text{RRR} = \frac{\text{RDR}}{\sqrt{\text{SR}_{\text{covariate1}}^{\beta_1} \text{SR}_{\text{covariate2}}^{\beta_2}}}. \quad (3)$$

The true, but unknown, population parameters β_1 and β_2 are estimated by b_1 and b_2, which minimize Eq. (1). In the denominator, $\text{SR}_{\text{covariate1}}$ and $\text{SR}_{\text{covariate2}}$ are the single ratios for covariate sites 1 and 2. The entire denominator in Eq. (3) adjusts the RDR statistic, the estimated seeding effect without covariates for a double-barrier design (B14). The adjustment depends on information from the covariates, which are weighted by the strength of the relationship between the precipitation differences at the covariate and target gauge sites. For the analysis of the WWMPP, the coefficients were determined at the end of the experiment with the precipitation data that passed the quality-control steps (118 experimental units). The test statistic—RRR—is used to perform the null hypothesis testing described next.

d. Statistical hypothesis

An RRR value of 1.0 indicates that there is no difference between seeded and unseeded populations. Thus, for this study, evaluating the null hypothesis for the ratio test statistic would be H0: RRR = 1.0. Since we are interested in whether seeding leads to an increase in precipitation, the alternative hypothesis is HA: RRR > 1.0. Thus, if the null hypothesis is rejected, the alternate hypothesis (RRR > 1.0) is accepted indicating that there is likely a seeding effect.

The evaluation of the p value, the probability that the null hypothesis is correct, was done using a rerandomization procedure as recommended in Brillinger et al. (1978). The rerandomization procedure, also known as a permutation test, is a nonparametric test, meaning that it does not make any distribution assumptions. In theory, the value of the test statistic is calculated under all possible permutations of the seeding decision—that is, every possible ordering of the seeding decision while maintaining the same number of cases. The resulting p value is the proportion of permutations that give a ratio test statistic that exceeds that found by the experiment. In practice, it is unrealistic to calculate all possible values of the test statistic, which is 118 factorial, a number larger than 10^{190}. Therefore, the seeding decision is instead permuted a large number of times, in our case, 100 000. For statistical significance, the p value should be less than the chosen significance level (<0.05), meaning there would be less than a 5% probability that the null hypothesis is rejected by chance.

TABLE 2. Final summary statistics for the RSE.

Winter seasons	No. of EUs	Correlation between targets	Ratio statistics	
			RRR	P
2008–14	118	0.584	1.033	0.283

The target-to-target (HY to GL) correlation was 0.58 (Table 2), with some variability from year to year. This compares well with the estimated correlation prior to the randomized seeding experiment (RSE) of 0.5 from SNOTEL data and validates the randomized crossover design statistical approach used in this study.

The RRR for the final dataset was 1.03 ($p = 0.28$; Table 2). The probability distribution of RRR values arising from the 100 000 rerandomization cases is shown in Fig. 7, overlain in red with the RRR from the WWMPP, 1.03. Note that 28% of the rerandomization cases fall to the right of RRR = 1.03, indicating that the probability that the RRR value could have occurred by chance is 28%, much higher than the p value of 5% needed to reject the null hypothesis and accept the alternate hypothesis that seeding is effective as specified in the design of the experiment. Thus, the statistical analysis fails to show that cloud seeding is effective. This does not mean that there is no effect of cloud seeding as the RRR value was 1.03, only that with the number of EU cases collected (118) one cannot distinguish between a small 3% seeding effect and background weather noise. If there is actually a 3% effect, this small signal could be detected by collecting more randomly determined cases, but the number needed would likely be over 1000 based on examination of the data.

While the WWMPP statistical experiment was not able to show a cloud-seeding signal above the observed weather noise, it did create a unique dataset of carefully controlled double-barrier seeded and unseeded cases. These data are valuable for further analysis such as an ensemble modeling approach to evaluate the impact of cloud seeding presented below. Such an approach is effective because the direct seeding effect can be estimated by running a model with and without seeding, alleviating the need for large numbers of cases to reduce the background weather noise as encountered in the above statistical approach. However, large numbers of simulations (ensemble approach) are still needed to reduce the uncertainty of model forcing, model physics, and transport dynamics in estimating the seeding effect. While the model can be used to perform these types of calculations, it needs to be grounded in the reality of surface snowfall observations as well as upper-air soundings and radiometer liquid water to make the results believable. The unique set of observations collected by the WWMPP
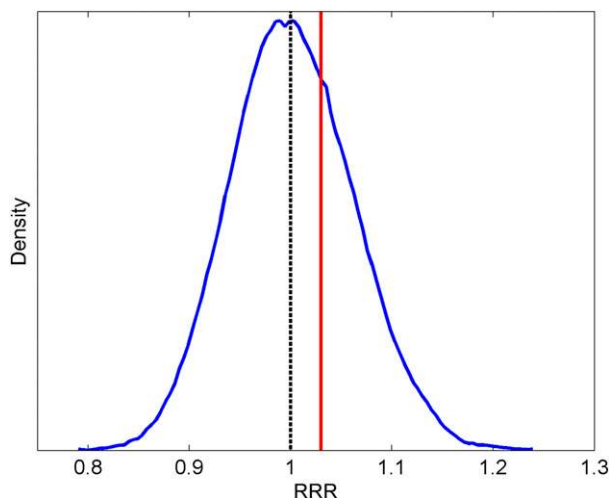

FIG. 7. Probability distribution of the RRR calculated by randomly assigning seeding to each EU with replacement 100 000 times, using the quality-controlled gauge dataset from the WWMPP (blue line). The red line gives the RRR value calculated from the actual seeding cases.

provide these observations for the current study. The following section describes the combined model–observations ensemble approach to evaluate cloud seeding during the WWMPP RSE.

3. Evaluation of the WWMPP using an ensemble modeling approach

a. Description of ensemble modeling approach

The traditional statistical approach to cloud-seeding evaluation involves a randomized seeding program in which a target is randomly selected to either be seeded or not seeded and a covariate measurement to account for natural storm variability. This approach was applied to the quality-controlled snow gauge data from the WWMPP and resulted in a value of RRR = 1.03 that was not statistically significant (Table 2). While a statistical approach can provide robust results in cases in which the seeding effect is relatively large, the number of cases needed becomes quite high when the seeding effect is small. This was the case for the statistical analysis of the WWMPP program.

As part of the WWMPP, a cloud-seeding module was developed by Xue et al. (2013a,b) and implemented into the Weather Research and Forecasting (WRF) Model through the Thompson (Thompson et al. 2008; Thompson and Eidhammer 2014) microphysical scheme. This allowed simulations of each EU with and without seeding. The seeding effect could then be estimated as the net increase in precipitation realized by differencing the model runs with and without the

seeding module activated. However, because of variations in precipitation due to model initial condition and model uncertainty, one deterministic pair of simulations for each of the EUs provides only one possible outcome of the model-estimated seeding effect. To really characterize the possible seeding effect from such a model analysis requires accounting for variations in initial conditions and model uncertainty through an ensemble of model simulations.

An ensemble modeling approach applied to the EUs of the RSE involves performing many simulations of each EU using different initializations and model physics configurations. Results from the ensemble modeling is provided by the ensemble mean and the spread of the ensemble members.

The ensemble approach has become standard practice in weather forecasting since the 1990s based on Lorenz's (1963) groundbreaking results showing that small perturbations in the initial condition of the atmosphere can lead to significant changes in the forecast, and thus a single deterministic forecast is not likely representative of the actual weather. Ensemble forecasting requires the use of probabilistic methods to analyze the results of running the ensemble members forced with appropriate perturbations to cover the most likely possible inputs.

For the ensemble spread to accurately estimate confidence and realistic variability (Slingo and Palmer 2011), the ensemble needs to appropriately sample the likely uncertainty in 1) the initial conditions driving the model and 2) the uncertainty in the model formulation.

In the following, we use these guidelines in creating an ensemble that estimates the impact of cloud seeding for the RSE. Three requirements were established to optimize the ensemble analysis of seeding effect:

1) The model physics need to accurately estimate precipitation over orographic barriers.
2) The seeding effect needs to be well characterized.
3) Sufficient numbers of ensemble members need to be created to reliably estimate the likely seeding effect and its variation.

Regarding the first requirement, Ikeda et al. (2010) and Rasmussen et al. (2011, 2014) have shown that the WRF model configured with the Thompson microphysics (Thompson et al. 2008) is able to estimate orographic precipitation in Colorado and Wyoming within 5%–10%.

Regarding the second requirement, Xue et al. (2013a,b) developed a cloud-seeding module to estimate the impacts of orographic cloud seeding using AgI for both ground and airborne seeding. While the scheme has been verified for AgI dispersion (Boe et al. 2014; Xue et al. 2014), the seeding effect, while reasonable, has not been directly tested. The various seeded ensemble members investigated here attempt to capture this uncertainty, recognizing that more research is needed to have full confidence in the scheme. The recent Seeded and Natural Orographic Wintertime Clouds: The Idaho Experiment (SNOWIE) project (French et al. 2018; Tessendorf et al. 2018) was largely motivated to reduce this uncertainty, and data from this field program is currently being analyzed and used to evaluate and improve the seeding algorithm.

Details to satisfy the third requirement are included in section 3b below (and Table 4). Briefly, for each EU there are two boundary conditions, two planetary boundary layer (PBL) schemes, and six microphysical initializations (see Table 4) leading to eight model configurations for each run. Three reanalyses were used for model initialization, leading to 24 control ensemble members for each EU for the natural/unseeded members. For the seeded members, three possible seeding configurations were used to model the likely impact of AgI on the cloud, leading to 72 ensemble members to bracket the range of seeding effects for each EU. This results in the creation of 2832 (24 control configurations × 118 EUs) unseeded simulations and 8496 seeded simulations (3 seeding configurations × 2832 control simulations) to estimate the impact of seeding for the 118 EUs.

b. Ensemble formulation

Based on the above discussion, the WRF Model was used to conduct model simulations for each EU by running a 24-h simulation for the outer 2.7-km-grid-spacing domain (Fig. 8a) and a 12-h simulation centered on the time of the EU for the inner 900-m domain (Fig. 8b). The inner domain is initialized 12 h into the outer domain simulation. Table 3 provides further information on the model configuration used.

Since each EU is only 4 h long, the probability of a single high-resolution simulation capturing the snow gauge accumulation of precipitation for that specific time period is quite low (Rossa et al. 2008). However, by capturing the likely variability in initial conditions and model physics using an ensemble approach, the model simulation of snow gauge accumulation is significantly improved. Part of this improvement is because biases in the time-dependent forcing of precipitation (i.e., moisture flux, synoptic or mesoscale forcing) are reduced by the averaging process used to produce an ensemble mean, and the forcing due to orographic uplift is reinforced (Rasmussen et al. 2011).

The following modeling approach and analysis method (summarized in Table 4) was used to address these requirements:
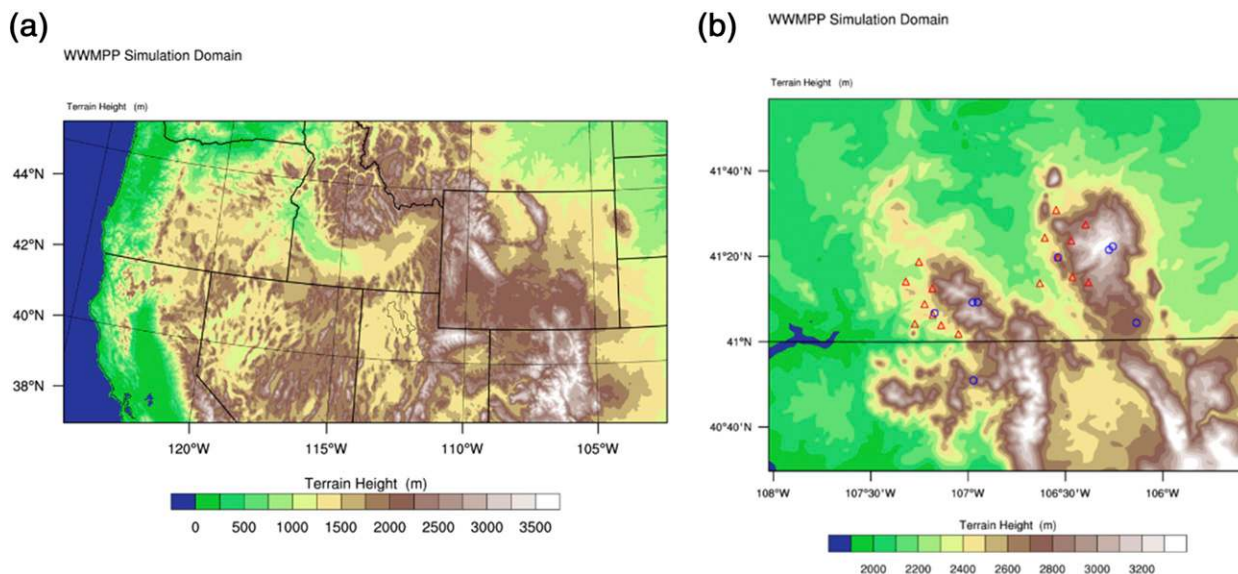
FIG. 8. Outer and inner domains used for model simulation of each EU. (a) The outer domain has 27-km grid spacing and is run for 24 h per EU. (b) The inner domain has 0.9-km grid spacing and is run for 12 h per EU.

1) Uncertainty in model forcing

ECMWF interim reanalysis (ERA-Interim), Climate Forecast System Reanalysis (CFSR), and North American Regional Reanalysis (NARR) products are used as initial and boundary conditions for all the conditions in Table 4. These reanalyses represent three of the most commonly used estimates of the four-dimensional (4D) historical atmospheric conditions currently available.

2) Uncertainty in boundary-condition treatment

Boundary conditions with and without hydrometeors are used. Traditional WRF boundary condition treatment does not include hydrometeors at the boundary between model domains; this capability was added to the model for this experiment; boundary condition column in Table 4.

3) Uncertainty in model physics for natural cloud and precipitation processes

(i) Different PBL schemes are used to examine any spread in the results because of the treatment of turbulence in the boundary layer (second column under control simulation in Table 4).

(ii) Different cloud condensation nuclei (CCN) concentrations are listed in the CCN column in Table 4. The standard climatology (CCN in Table 4) was based on 7 years of GOCART climatology following Thompson and Eidhammer (2014). The range of possible CCN is represented by 1/5 and 5 times the climatology for a clean and polluted background.

(iii) Different ice nuclei (IN) parameterizations are listed in the IN column in Table 4. The Cooper (1986) ice nucleation parameterization (IN1) and Meyers et al. (1992) ice nucleation parameterization (IN2) are used to represent different ice initiation processes.

(iv) The ID column in Table 4 is used to reference subsets of the PBL and microphysical schemes listed in the three previous columns in Table 4 and both the seeding module configurations described next.

TABLE 3. Model configuration.

Domain (m)	2700	900
Horizontal grids	720×360	231×180
Vertical levels	81 terrain-following eta levels	
Time step (s)	12	5
PBL scheme	Mellor–Yamada–Nakanishi–Niino (MYNN) 2.5 order	MYNN/Yonsei University (YSU)
Microphysics scheme	Thompson and Eidhammer (2014) scheme	Thompson and Eidhammer (2014) with the seeding parameterization

TABLE 4. Matrix of ensemble simulations. The top row indicates whether the simulation is a control or seed simulation. The leftmost column indicates the ensemble case for the control runs (C1–C8). The second column from the left indicates the type of boundary condition used (HYD BC = hydrometeors in the boundary; NO HYD BC = no hydrometeor). The third column indicates the PBL scheme used (MYNN or YSU). The fourth column indicates the CCN concentration used and the fifth column indicated the ice nucleation scheme used. For example, ensemble C1 uses the MYNN PBL scheme, a CCN concentration of $100 \, cm^{-3}$ and the IN1 scheme for the control simulation. The sixth column (ID) indicates the ensemble case for the seeded simulations. The seventh column indicates the AgI nucleation methodology used for the ensemble member, while the eighth column indicates the AgI scavenging approach used. AN1 nucleation is based on Meyers et al. (1992) and AS1 scavenging is based on Xue et al. (2013a). For example, ensemble seed simulation S1 uses Meyers et al. (1992) AgI ice nucleation (AN1) and Xue et al. (2013a) scavenging (AS1) applied to all eight control simulations C1–C8.

	Control simulations					Seed simulations	
ID	Boundary condition	PBL scheme	CCN	IN	ID	AgI nucleation	AgI scavenging
C1	HYD BC	MYNN	CCN	IN1	S1	AN1	AS1
C2		YSU	CCN	IN1			
C3		MYNN	0.2 CCN	IN1			
C4		MYNN	20 CCN	IN1	S2	0.2AN1	0.2AS1
C5		MYNN	CCN	IN2			
C6		MYNN	0.2 CCN	IN2			
C7		MYNN	20 CCN	IN2	S3	5AN1	5AS1
C8	NO HYD BC	MYNN	CCN	IN1			

4) Uncertainty in seeding processes
   (i) The different PBL schemes will impact the seeding effect through AgI dispersion.
   (ii) In the AgI nucleation column in Table 4 three schemes are listed. AgI nucleation option 1 (AN1) follows the DeMott (1995) and Meyers et al. (1995) parameterization. To capture the likely uncertainties due to activation rate, the default rate was decreased by a factor of 5 (0.2AN1) and increased by a factor of 5 (5AN1).
   (iii) The AgI scavenging column in Table 4 includes in AgI scavenging option 1 (AS1) AgI self-coagulation, scavenging of AgI by hydrometeors, and AgI dry deposition (Xue et al. 2013a). This removal rate was decreased (increased) by a factor of 5, that is, 0.2AS1 (5AS1), to capture the likely uncertainty associated with AgI removal processes.
5) Uncertainty in spatial and temporal distribution of precipitation
   Precipitation in 3×3 grids surrounding the gauge sites was analyzed to address spatial uncertainty of the precipitation.

c. Verification of the ensemble modeling approach

The WRF Model results are compared to the RSE gauge data to establish confidence that the model is able to provide a reasonable estimate of precipitation over the Sierra Madre and Medicine Bow Mountains using the configuration presented in section 3a.

The distribution of differences between paired errors of snow gauge observations and model-estimated precipitation at the gauge sites acting as either control or target (remember that the target and control location varies from EU to EU because of the random assignment of seeding in the crossover design) are shown in Figs. 9 and 10. The differences are nearly equally distributed about zero for every model configuration. The results are given as box-and-whisker plots and mean differences. The same paired error analysis was conducted for each of the model configurations for the covariate gauge sites and further supports these conclusions (not shown). This indicates that there was no single model configuration in the ensemble biasing the statistical results. Moreover, these results show that the ERA-Interim model initialization had the least spread of error compared to the NARR and CFSR.

Figure 11 provides for the two target snow gauges (HY and GL) and the four covariate gauges [Sandstone (SS), Barrett Ridge (BR), Elk River (ER), and Chimney Park (CP)] the distribution of the 4-h precipitation accumulation differences of the ensemble mean to the observations for each EU. For each gauge, the model ensembles are separated by the three reanalyses used. The data pairs are for the specific EU ensemble simulation (24 members for each reanalysis) with the corresponding EU gauge accumulation. The results show that the model ensemble mean is within 0.5 mm of the 4-h observations 50% of the time (boxplots). However, there are some outliers that extend to ~3-mm difference when considering 5%–95% of the data (whiskers). Overall, the ensemble mean provides a reasonable estimate of the observed precipitation in that differences between the model ensemble mean and the observations are near zero for all snow gauge sites, and the distribution of the differences is nearly symmetric about zero.

A direct comparison of the actual values of precipitation to the model ensemble results from each EU is
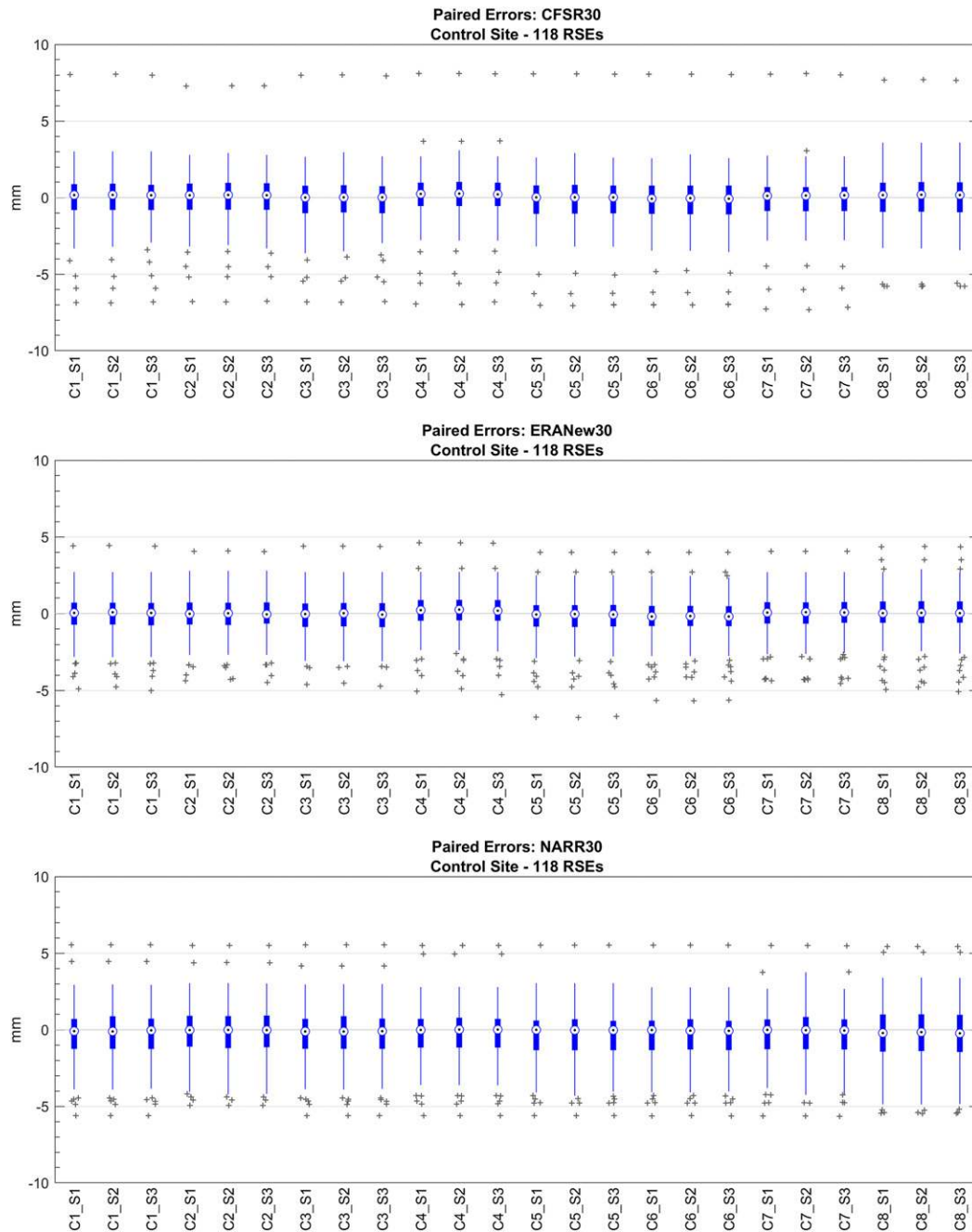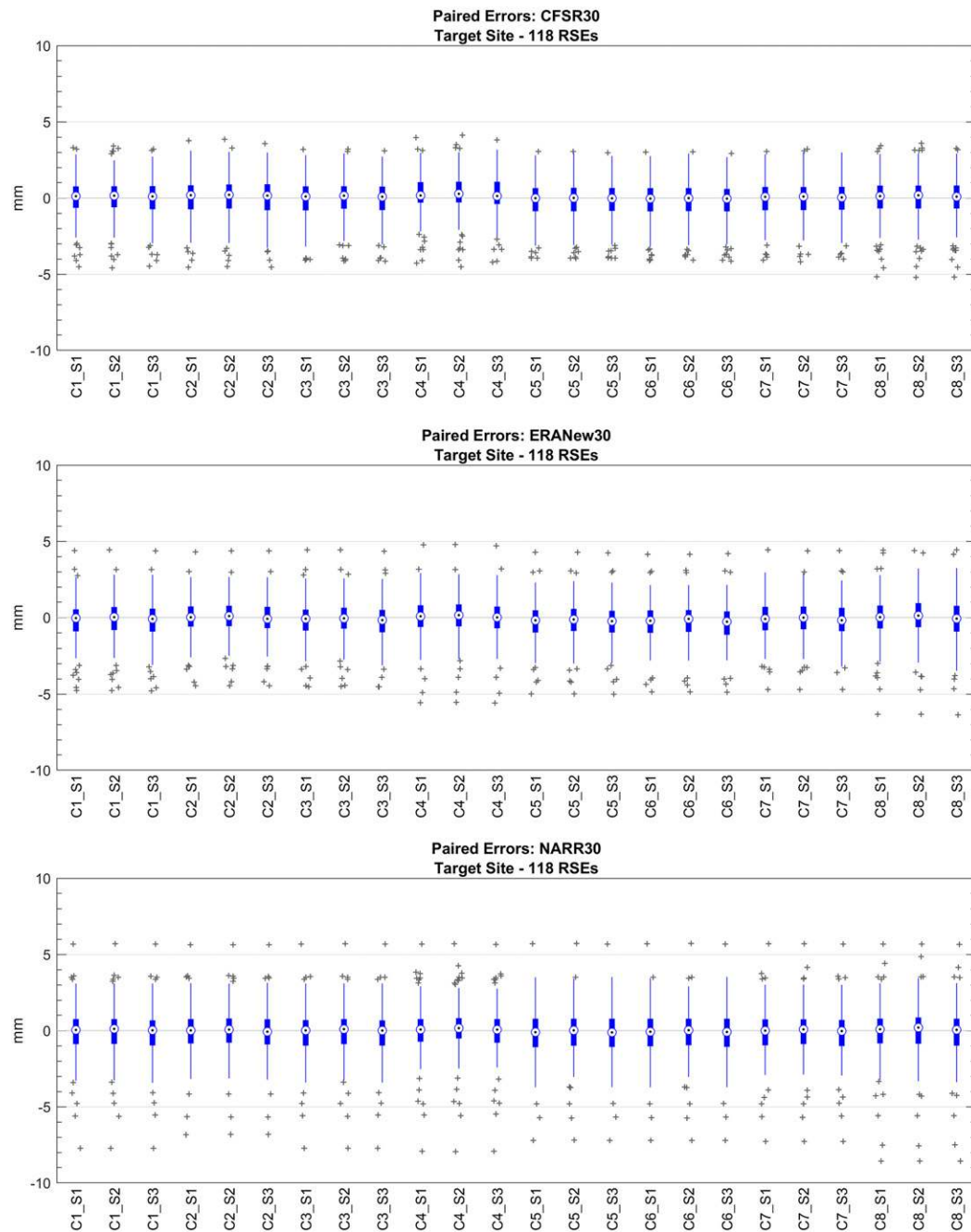
FIG. 9. Paired errors, observations − ensemble mean, for the control site for all 118 RSE cases (control in this case can be either the HY or GL sites, depending on which range was not targeted for seeding). Error is observed 4-h gauge-measured precipitation minus the model 4-h precipitation. (a) All 24 ensemble members per RSE case initialized with CFSR, (b) all 24 ensemble members per RSE case initialized with ERA-Interim, and (c) all 24 ensemble members per RSE case initialized with NARR.

shown in Fig. 12a for GL and in Fig. 12b for HY. The uncertainty in the observed 4-h precipitation at these sites is less than 0.2 mm, the width of the circles plotted. While the observed precipitation amount is usually contained within the spread of the ensemble, as

discussed above, in a number of cases it is not. While the ensemble mean in most cases improves the precipitation estimate over randomly choosing an ensemble member, there are a number of cases where it does not. Thus, while the present ensemble-mean approach improves

FIG. 10. Paired errors, observations − ensemble mean, for the target sites for all 118 RSE cases (target in this case can be either the HY or GL sites, depending on which range was the target for seeding) for all 118 RSE cases. (a) All 24 ensemble members per RSE case initialized with CFSR, (b) all 24 ensemble members per RSE case initialized with ERA-Interim, and (c) all 24 ensemble members per RSE case initialized with NARR.

the seeding estimate over choosing just one deterministic simulation, there are still some uncertainties in the precipitation estimation. These uncertainties may be reduced by increasing the ensemble size even further but may also be due to the inherent difficulty of simulating precipitation at a point in space for a very short period of

only 4 h. It should be noted, however, that the cases with large uncertainty are relatively rare, as shown in Fig. 11.

Figure 13 shows, for each EU at the seeded target gauge site, the difference between the mean of the 72 ensemble members characterizing the precipitation estimated with the active seeding module and the mean
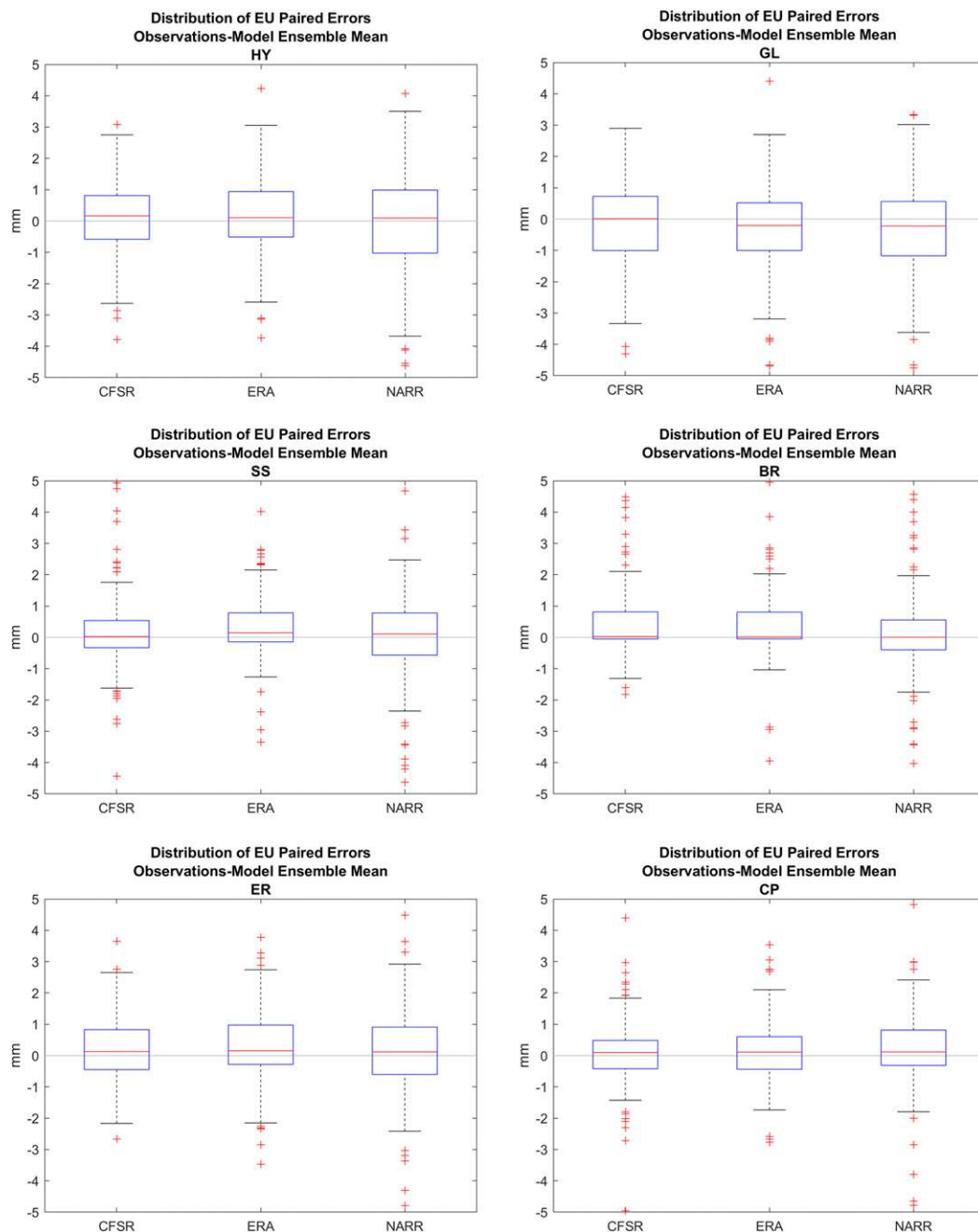
FIG. 11. Distribution of EU paired errors by subtracting 4-h observed precipitation accumulation at each gauge from the ensemble mean of the 24 ensemble members' accumulation at each gauge. Each box-and-whisker plot represents the paired differences of the 118 EU cases for a particular EU snow gauge location (see Fig. 1 for details). HY is the target in the Sierra Madre; GL is the target in the Medicine Bow Mountains; SS is the upwind covariate site in the Sierra Madre; BR is the upwind covariate in the Medicine Bow Mountains; ER is the southern covariate for the Sierra Madre; and CP is the southern covariate for the Medicine Bow Mountains. The box represents 25%–75% of the differences, while the whiskers represent 5%–95% of the data. Outliers are shown as plus signs. The mean difference is represented by the horizontal red line. Results are shown in each plot for the three different analyses (CFSR, ERA-Interim, and NARR).

GLEES 4-Hour Precipitation
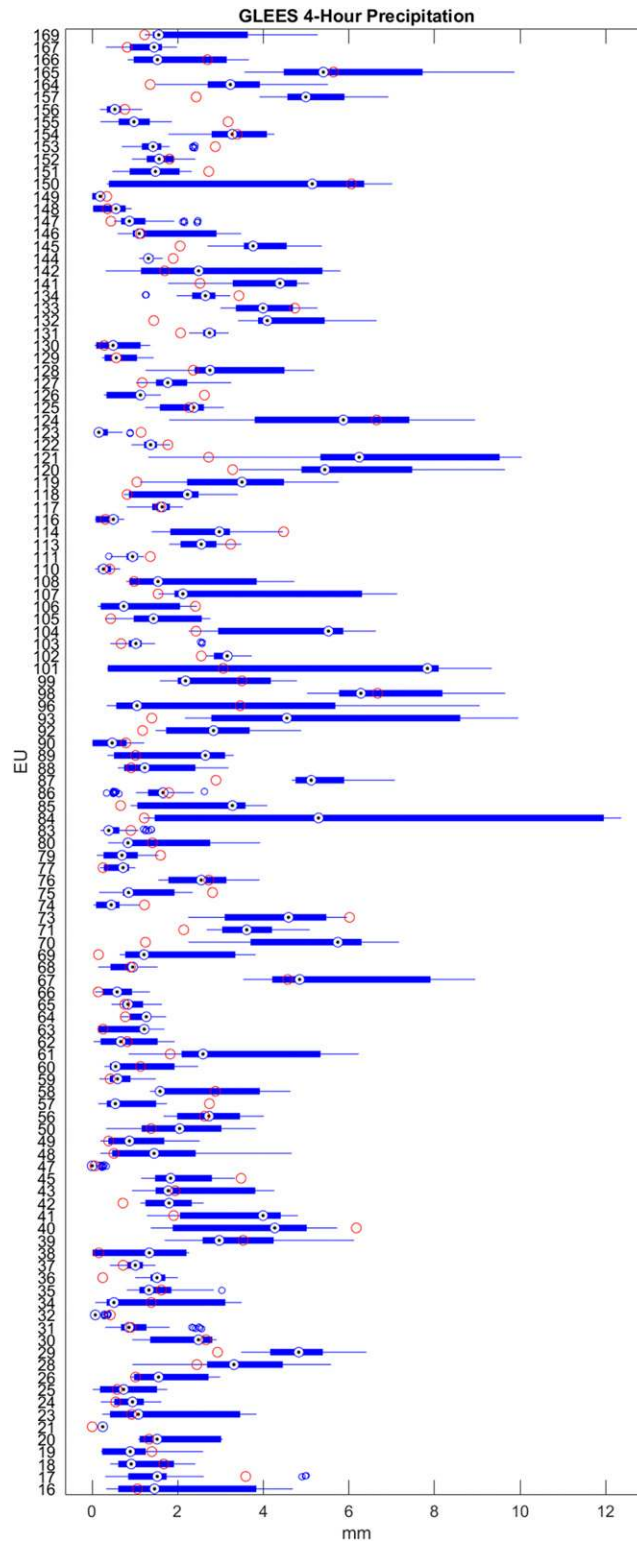
FIG. 12. (a) Comparison of the ensemble-modeled 4-h precipitation for each EU (box-and-whisker plot, median indicated by the open circle with a dot at the center, 25% and 75% of range indicated by the thick blue horizontal bar, 5% and 95% of the range by the light blue horizontal bar) to the observed 4-h precipitation (open red circle) at the GL site (for both target and control conditions). The width of the red circle representing the observations (0.2 mm) is greater than the uncertainty of the data point (typically 0.1 mm). (b) As in (a), but for the HY site.
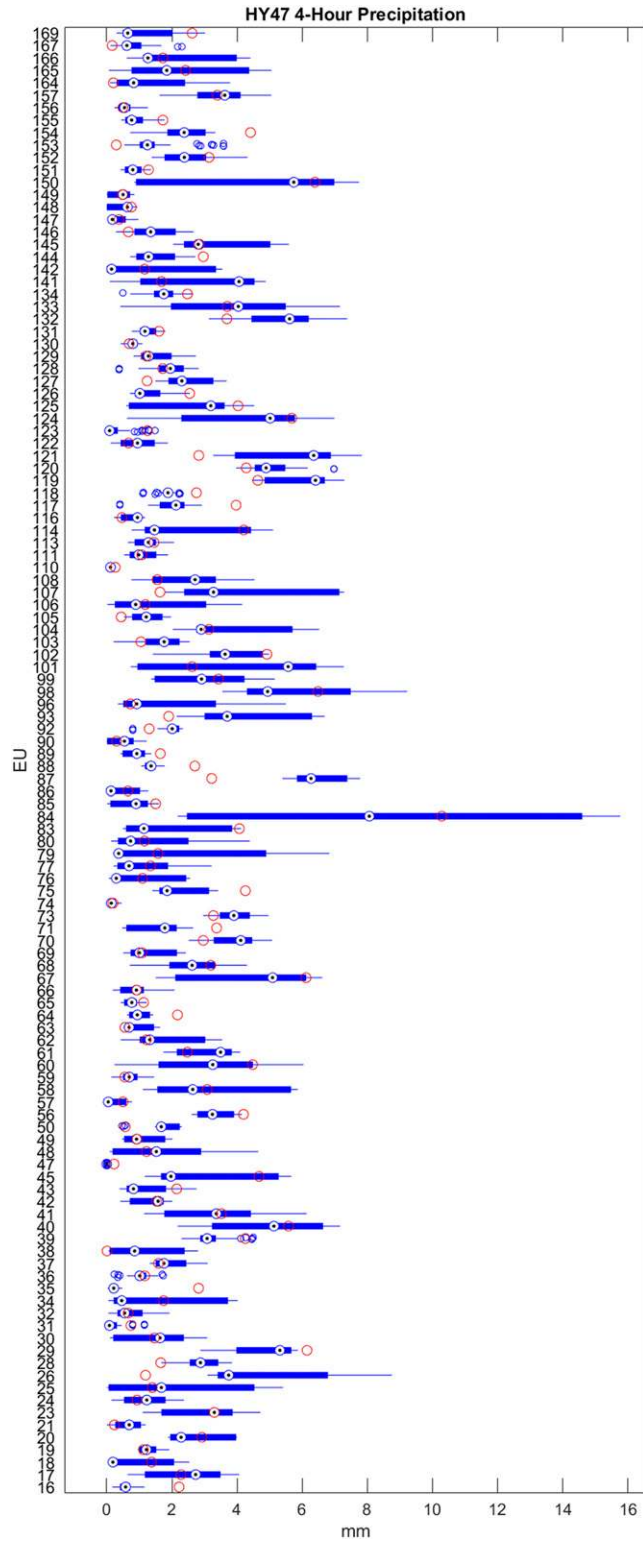
FIG. 12. (*Continued*)
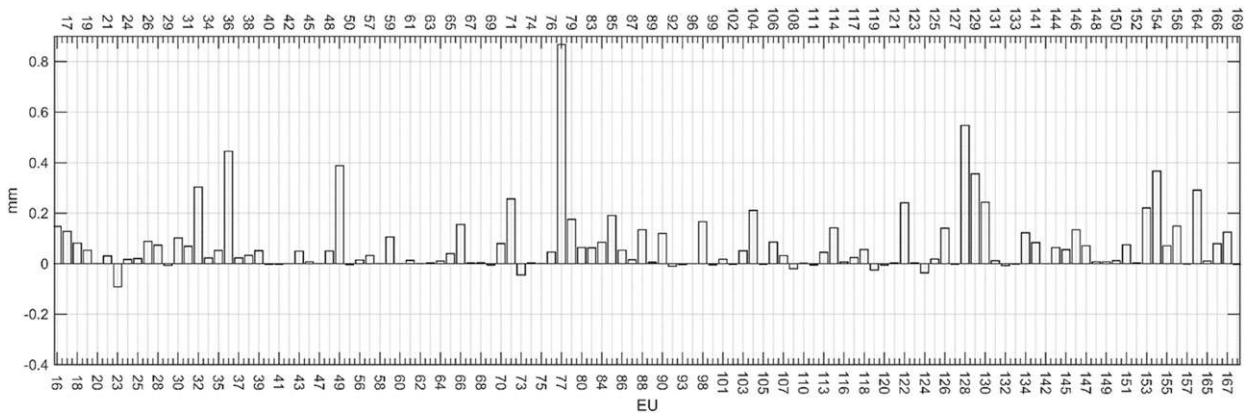
FIG. 13. ERA-Interim model-estimated seeding effect (seeding simulation 4-h precipitation accumulation minus control simulation precipitation) for the targeted gauge site (mm), either GL or HY, in each EU from the model ensemble mean.

of the 24 ensemble members representing the natural precipitation expected for that target and that EU. While there are a variety of seeding effects, the seeded precipitation exceeded 0.4 mm for only 3 EUs and exceeded 0.2 mm on only 13 cases, about 10% of the EUs. Smaller amounts are much more frequent. In a few cases, the seeding effect is slightly negative.

The frequency plot of the model-estimated seeding (Fig. 14) shows that the seeding effect resembles a decreasing exponential distribution above 0.05 mm. This result is similar to the near-exponential shape of the distribution of radiometer-measured and model-simulated integrated liquid water (not shown), consistent with the seeding effect being proportional to the amount of supercooled liquid water present.

d. Evaluation of the WWMPP using the ensemble modeling approach

The model ensemble-mean seeding effect for the Sierra Madre and Medicine Bow target gauge sites, individually and combined, is given in Fig. 15. The results show a mean model-estimated seeding effect of ~5% for the two ranges. The 25th percentile is 3% and the 75th percentile is 7%. The actual statistical result from the physical experiment is comparable to these results, falling at the 25th percentile level (3%). Interpreting these percentage increases requires some caution. They represent the possible increases in precipitation resulting from seeding of the EUs generated by the WWMPP. They do not represent the percentage increases in annual snowfall from winter orographic seeding. Answering that question requires understanding the fraction of precipitation that may occur from the cases considered as EUs (Ritzman et al. 2015). Ritzman et al.'s conclusions were that the cases considered by the WWMPP represent about 30% of the annual precipitation expected over the two ranges from

snowfall. Taking these results at face value, along with the results of Fig. 15, suggests that the expected impact from seeding these two ranges, applying the criteria of the WWMPP, is on the order of a 1.5% increase in the total annual precipitation.

4. Summary and conclusions

The first part of this paper describes the statistical evaluation of the WWMPP randomized seeding experiment (RSE), including a detailed description of the quality control processes used to screen the snow gauge data. As part of the quality-control effort, a new and robust method to correct ETI snow gauge data was developed and discussed in the appendix.
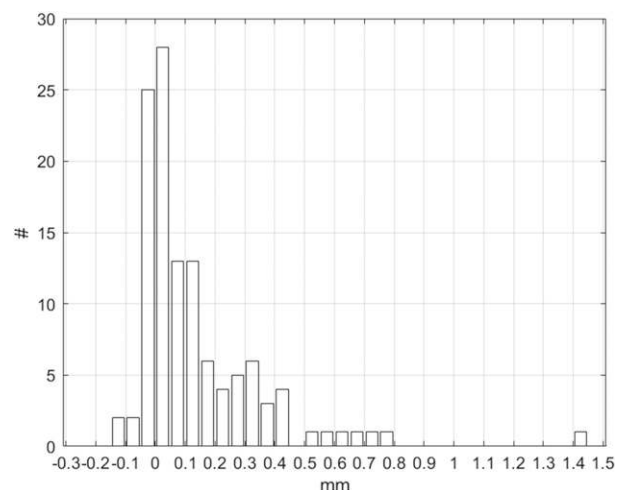

FIG. 14. Number of ERA-Interim model-estimated seeding effects for the target gauge sites, either GL or HY, estimated from the model ensemble mean as a function of the estimated effect (mm).
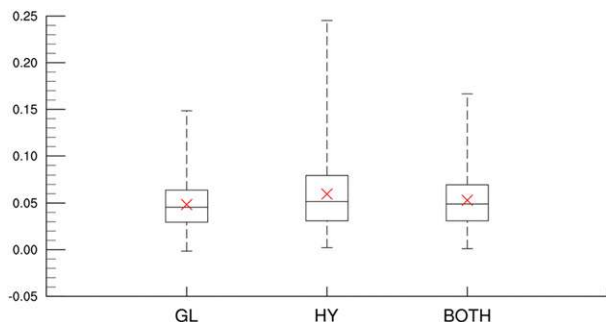
FIG. 15. Distributions of the fractional increase in precipitation at the Medicine Bow Mountains target (GL), the Sierra Madre target (HY), and both (average of the two) based on the difference between seeded and unseeded model ensemble simulations (8946 EU cases simulated). The horizontal bar is the median and the red × is the mean value. The top and bottom of the boxes represent the quartile ranges.

There were 118 EUs that met quality-control and seeding criteria. The statistical results using these 118 EUs collected over six years revealed that there was insufficient statistical evidence (p value of 0.28) to reject the null hypothesis that there was no effect of ground seeding. While it may be possible to statistically reject the null hypothesis by collecting on the order of 1000 cases for a 5% signal, the time and expense required makes this prohibitive.

Instead of collecting additional cases at large expense, an ensemble modeling approach to estimate the impact of ground-based seeding was conducted and presented in the second part of this paper. The modeling approach has the advantage that both the conditions with and without seeding can be simulated, allowing the difference of the model simulations to estimate the seeding effect. An ensemble approach allows one to better account for uncertainty in initial conditions, model biases, and random errors in the model simulations. A prerequisite to using a model, however, is that the simulations must reasonably represent reality. The WWMPP RSE comprehensive snow gauge, sounding, and radiometer data were used to perform this function. The comparison of the model ensemble simulations and the snow gauge data per EU showed reasonable agreement, with the model mean and median errors being close to zero, and the distribution of errors in the paired comparison symmetric about zero error in most cases.

It should be noted, however, that while in the majority of cases the ensemble mean improved the estimate of precipitation over a single deterministic simulation, in some cases the observed 4-h precipitation was significantly different than the ensemble mean. Since the seeding effect is calculated from the difference of a seeded and unseeded simulation, the estimated seeding

effect may still be reasonable. As these cases are randomly distributed (Fig. 11), any bias in the results contributed by these relatively few cases is estimated to be small.

The model ensemble seeding effect was estimated with the 24 model ensemble members for each of three reanalysis forcing datasets, with a total of 8946 simulations to simulate each of the 118 EUs. This approach minimized potential biases due to time variation of forcings of precipitation for each EU and emphasized the orographically forced component of precipitation.

The results of the model ensemble in which the paired seeding effect was estimated for each ensemble member produced a mean precipitation enhancement of 5%, with an inner-quartile range of 3%–7%. These results provide an estimate of the impact of ground-based cloud seeding in the Sierra Madre and Medicine Bow Mountains in Wyoming for the RSE EUs that accounts for key uncertainties in both initial conditions and model physics. As indicated by Ritzman et al. (2015), the RSE cases account for only 30% of the total precipitation, thus the overall seeding impact on annual precipitation over this region is ~1.5%.

a. Next steps

This study estimated the additional precipitation likely produced by ground-based seeding with AgI over the Sierra Madre and Medicine Bow Mountains in Wyoming using a seeding module within the WRF Model. The next step is to use additional observations to further verify that the model estimate of the seeding effect is reasonable. This would require specific field studies for both ground and airborne seeding. The recent SNOWIE field effort (Tessendorf et al. 2018; French et al. 2018) has provided an excellent dataset for such an evaluation.

b. Application of the ensemble approach to other mountain ranges

It is tempting to apply the ensemble approach to other mountain ranges to estimate potential seeding effects. However, verifying that the model is estimating snowfall and supercooled liquid water correctly through observational comparisons is still necessary. It is also important to verify that the model-generated wind and thermodynamic structure of storms are reasonable through comparison with appropriate data from the mountain range of interest. If such a comparison is performed and the comparison reasonable and the model-generated seeding effect is also verified, then the ensemble modeling approach provides a powerful tool to estimate the likely seeding effect of AgI seeding in orographic clouds for both ground and airborne seeding. While this approach
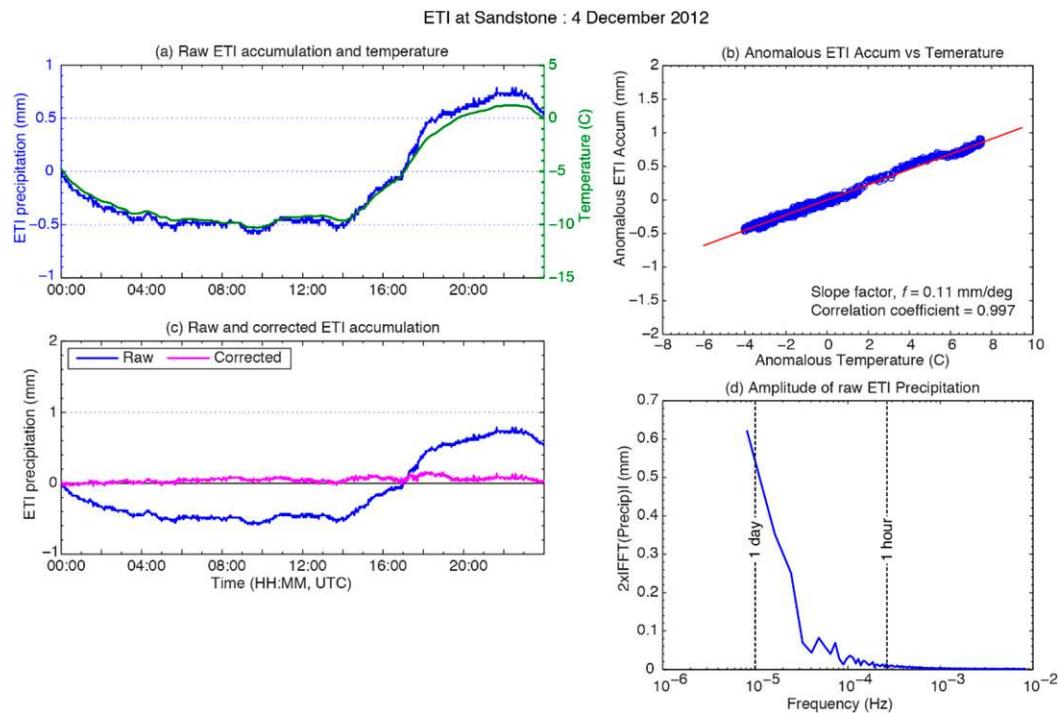
ETI at Sandstone : 4 December 2012

FIG. A1. (a) Time history of raw ETI precipitation accumulation (blue line) and temperature measurements (green line) on a clear day (relatively humidity <80%) at SS. The raw precipitation data decrease with a temperature decrease between 0000 and 0500 UTC [1700 and 2200 mountain standard time (MST)] and increase with a temperature increase from 1400 UTC (0700 MST). (b) Scatterplot of anomalous ETI precipitation accumulation vs anomalous temperature on the same day. Anomalous measurements are the departures from the 24-h average values. Red line is the least squares fit of the data. For this dataset, the slope of the best fit f is 0.11 mm $°C^{-1}$ and the correlation coefficient is 0.997. (c) Time history of raw and corrected ETI precipitation accumulation. (d) FFT of the raw ETI precipitation data, indicating frequency and wave amplitude of the raw ETI precipitation.

requires a significant investment of time and money, the statistical approach to evaluate cloud-seeding programs using snow gauge observations from either ground or airborne seeding can often require over 1000 cases to detect the signal from the noise, depending on the strength of the a priori unknown seeding signal.

APPENDIX

Precipitation Gauge Data Quality Control

This appendix describes a new methodology used as part of the precipitation gauge quality-control procedures to remove diurnal temperature variations from the ETI gauge dataset.

The load cell used to make the weight measurement of snow accumulation for the ETI gauges has a significant temperature dependence, leading to biased precipitation estimates under changing environmental conditions. For instance, a diurnal temperature change is obvious in the raw precipitation data, especially on clear days (Fig. A1). This section describes a methodology to remove much of the temperature effect on the raw ETI weight measurements. The methodology calculates a correction factor for each ETI gauge for every seeded and unseeded day between November 2008 and April 2014.

The main steps in the methodology are as follows:

1) Find a nonprecipitating 24-h period before and after a precipitation event containing a 4-h experimental unit (EU).
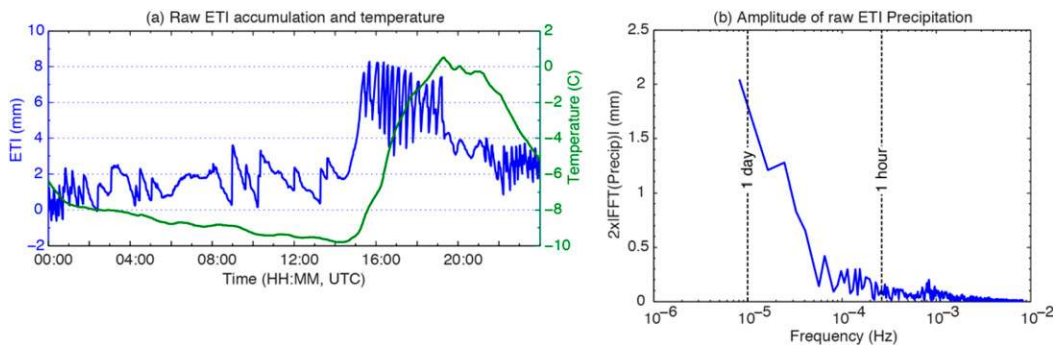
FIG. A2. (a) ETI accumulation and temperature for a clear day (22 Dec 2012) at TL. (b) FFT of the raw ETI precipitation data for 22 Dec 2012.

2) Subtract 24-h average ETI accumulation and temperature from the raw data to discern departures from the mean.

3) Fit a least squared line through the anomalous raw precipitation P' and temperature T' data to obtain a linear equation relating the precipitation perturbation to the temperature perturbation: $P' = fT' + b$. The slope f represents the change in precipitation accumulation with respect to a change in temperature.

4) Examine the quality of the clear-day data (described below). Repeat steps 1–3 until at least two good clear days are found.

5) Finally, correct the raw precipitation measurement for each EU (in between the two nonprecipitation days) by

$$P_f = P_r + \Delta P,$$

where P_f is the corrected precipitation accumulation and P_r is the raw precipitation accumulation—both being zeroed out at the beginning of the 4-h EU. The value of ΔP is the temperature-induced measurement response computed from

$$\Delta P = f \times (T_r - \overline{T}),$$

where f is the correction factor from 3), T_r is the temperature measurement, and $\overline{T}$ is the average over the 4-h evaluation period (f is the time-weighted average value from the two values found in step 3).

Figure A1b shows the raw and corrected ETI measurement on a clear day at Sandstone as an example. The 4-h accumulation was determined after a 9-min running-mean filter was applied to the datasets.

The quality of the clear-day ETI data chosen was checked based on the goodness of the least squared fit (residual of the fit and correlation coefficient) and by examining the fast Fourier transform of the raw precipitation data. The clear-day ETI data, from which the correction factor f is derived, was rejected if

1) The correlation coefficient was <0.95
2) Residual of the fit was >350 mm (i.e., large scatter in the data)
3) The amplitude of raw precipitation data oscillation for subhourly frequencies was >0.01 mm.

Figure A1d shows an example of fast Fourier transform (FFT) plot from good clear-day data. For this day, the amplitude of the precipitation measurement excursion induced by the diurnal temperature change is as large as 0.55 mm (see vertical line indicating wavelength of 1 day). As a contrast, measurement noise related to natural variability of the sensing transducer is much less than 0.01 mm (see vertical line indicating wavelength of 1 h). Figure A2b is the FFT plot from rejected clear-day data. The subhourly precipitation oscillation is much larger than the previous example, and the time series (Fig. A1a) clearly shows a problem in the data. Obviously, the correlation coefficient would be less than the threshold value if there were no temperature-induced precipitation measurement excursion. In this case, f is zero.

REFERENCES

Boe, B. A., J. A. Heimbach Jr., T. W. Krauss, L. Xue, X. Chu, and J. T. McPartland, 2014: The dispersion of silver iodide particles from ground-based generators over complex terrain. Part I: Observations with acoustic ice nucleus counters. *J. Appl. Meteor. Climatol.*, **53**, 1325–1341, https://doi.org/10.1175/JAMC-D-13-0240.1.

Breed, D., R. Rasmussen, C. Weeks, B. Boe, and T. Deshler, 2014: Evaluating winter orographic cloud seeding: Design of the Wyoming Weather Modification Pilot Project (WWMPP). *J. Appl. Meteor. Climatol.*, **53**, 282–299, https://doi.org/10.1175/JAMC-D-13-0128.1.

Brillinger, D. R., L. V. Jones, and J. W. Tukey, 1978: The Management of Weather Resources: Statistical task force to the weather modification board Rep., Vol. II, 94 pp.

Bureau of Reclamation, 2012: Colorado River basin water supply and demand study. U.S. Department of Interior Rep., 34 pp., https://www.usbr.gov/watersmart//bsp/docs/finalreport/ColoradoRiver/CRBS_Executive_Summary_FINAL.pdf.

Cooper, W. A., 1986: Ice initiation in natural clouds. *Precipitation Enhancement—A Scientific Challenge*, Meteor. Monogr., No. 43, Amer. Meteor. Soc., 29–32.

DeMott, P. J., 1995: Quantitative descriptions of ice formation mechanisms of silver iodide-type aerosols. *Atmos. Res.*, **38**, 63–99, https://doi.org/10.1016/0169-8095(94)00088-U.

French, J. R., and Coauthors, 2018: Precipitation formation from orographic cloud seeding. *Proc. Natl. Acad. Sci. USA*, **115**, 1168–1173, https://doi.org/10.1073/pnas.1716995115.

Gabriel, K. R., 1999: Ratio statistics for randomized experiments in precipitation stimulation. *J. Appl. Meteor.*, **38**, 290–301, https://doi.org/10.1175/1520-0450(1999)038<0290:RSFREI>2.0.CO;2.

Ikeda, K., and Coauthors, 2010: Simulation of seasonal snowfall over Colorado. *Atmos. Res.*, **97**, 462–477, https://doi.org/10.1016/j.atmosres.2010.04.010.

List, R., K. R. Gabriel, B. A. Silverman, Z. Levin, and T. Karacostas, 1999: The rain enhancement experiment in Puglia, Italy: Statistical evaluation. *J. Appl. Meteor.*, **38**, 281–289, https://doi.org/10.1175/1520-0450(1999)038<0281:TREEIP>2.0.CO;2.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141, https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

Meyers, M. P., P. J. DeMott, and W. R. Cotton, 1992: New primary ice-nucleation parameterizations in an explicit cloud model. *J. Appl. Meteor.*, **31**, 708–721, https://doi.org/10.1175/1520-0450(1992)031<0708:NPINPI>2.0.CO;2.

——, ——, and ——, 1995: A comparison of seeded and nonseeded orographic cloud simulations with an explicit cloud model. *J. Appl. Meteor.*, **34**, 834–846, https://doi.org/10.1175/1520-0450(1995)034<0834:ACOSAN>2.0.CO;2.

National Research Council, 2003: *Critical Issues in Weather Modification Research.* National Academies Press, 123 pp., https://www.nap.edu/read/10829.

Rasmussen, R. M., J. Vivekanandan, J. Cole, B. Myers, and C. Masters, 1999: The estimation of snowfall rate using visibility. *J. Appl. Meteor.*, **38**, 1542–1563, https://doi.org/10.1175/1520-0450(1999)038<1542:TEOSRU>2.0.CO;2.

——, and Coauthors, 2011: High-resolution coupled climate runoff simulations of seasonal snowfall over Colorado: A process study of current and warmer climate. *J. Climate*, **24**, 3015–3048, https://doi.org/10.1175/2010JCLI3985.1.

——, and Coauthors, 2012: How well are we measuring snow: The NOAA/FAA/NCAR winter precipitation test bed. *Bull. Amer. Meteor. Soc.*, **93**, 811–829, https://doi.org/10.1175/BAMS-D-11-00052.1.

——, and Coauthors, 2014: Climate change impacts on the water balance of the Colorado headwaters: High-resolution regional climate model simulations. *J. Hydrometeor.*, **15**, 1091–1116, https://doi.org/10.1175/JHM-D-13-0118.1.

Ritzman, J. M., T. Deshler, K. Ikeda, and R. Rasmussen, 2015: Estimating the fraction of winter orographic precipitation produced under conditions meeting the seeding criteria for the Wyoming Weather Modification Pilot Project. *J. Appl. Meteor. Climatol.*, **54**, 1202–1215, https://doi.org/10.1175/JAMC-D-14-0163.1.

Rossa, A., P. Nurmi, and E. Ebert, 2008: Overview of methods for the verification of quantitative precipitation forecasts. *Precipitation: Advances in Measurement, Estimation and Prediction*, S. Michaelides, Ed., Springer, 419–452, https://doi.org/10.1007/978-3-540-77655-0_16.

Slingo, J., and T. Palmer, 2011: Uncertainty in weather and climate prediction. *Philos. Trans. Roy. Soc.*, **369A**, 4751–4767, https://doi.org/10.1098/rsta.2011.0161.

Tessendorf, S. A., and Coauthors, 2018: A transformational approach to winter orographic weather modification research: The SNOWIE Project. *Bull. Amer. Meteor. Soc.*, https://doi.org/10.1175/BAMS-D-17-0152.1, in press.

Thompson, G., and T. Eidhammer, 2014: A study of aerosol impacts on clouds and precipitation development in a large winter cyclone. *J. Atmos. Sci.*, **71**, 3636–3658, https://doi.org/10.1175/JAS-D-13-0305.1.

——, P. R. Field, R. M. Rasmussen, and W. D. Hall, 2008: Explicit forecasts of winter precipitation using an improved bulk microphysics scheme. Part II: Implementation of a new snow parameterization. *Mon. Wea. Rev.*, **136**, 5095–5115, https://doi.org/10.1175/2008MWR2387.1.

Tukey, J. W., J. R. Brillinger, and L. V. Jones, 1978: Report of the Statistical Task Force to the Weather Modification Board. Vol. II. U.S. Government Printing Office, 94 pp.

Xue, L., and Coauthors, 2013a: Implementation of a silver iodide cloud-seeding parameterization in WRF. Part I: Model description and idealized 2D sensitivity tests. *J. Appl. Meteor. Climatol.*, **52**, 1433–1457, https://doi.org/10.1175/JAMC-D-12-0148.1.

——, S. Tessendorf, E. Nelson, R. Rasmussen, D. Breed, S. Parkinson, P. Holbrook, and D. Blestrud, 2013b: Implementation of a silver iodide cloud-seeding parameterization in WRF. Part II: 3D simulations of actual seeding events and sensitivity tests. *J. Appl. Meteor. Climatol.*, **52**, 1458–1476, https://doi.org/10.1175/JAMC-D-12-0149.1.

——, X. Chu, R. Rasmussen, D. Breed, B. Boe, and B. Geerts, 2014: The dispersion of silver iodide particles from ground-based generators over complex terrain. Part II: WRF large-eddy simulations versus observations. *J. Appl. Meteor. Climatol.*, **53**, 1342–1361, https://doi.org/10.1175/JAMC-D-13-0241.1.