The Israel 4 Cloud Seeding Experiment: Primary Results

YOAV BENJAMINI,[a] AMIR GIVATI,[b] PAVEL KHAIN,[c] YOAV LEVI,[c] DANIEL ROSENFELD,[d] URI SHAMIR,[e] AYAL SIEGEL,[f] ASSAF ZIPORI,[c] BARUCH ZIV,[g] AND DAVID M. STEINBERG[a]

[a] *Department of Statistics and Operations Research, Tel Aviv University, Tel Aviv, Israel*
[b] *Environmental Studies Department, Porter School of the Environmental and Earth Sciences, Tel Aviv University, Tel Aviv, Israel*
[c] *Israel Meteorological Service, Bet Dagan, Israel*
[d] *Institute of Earth Sciences, Hebrew University, Jerusalem, Israel*
[e] *Faculty of Civil and Environmental Engineering, Technion-Israel Institute of Technology, Haifa, Israel*
[f] *The Israel Water Authority, Tel Aviv, Israel*
[g] *Department of Natural and Life Sciences, Open University, Ra'anana, Israel*

(Manuscript received 22 May 2022, in final form 3 November 2022)

ABSTRACT: After 38 years of operational cloud seeding for rain enhancement in northern Israel, the Israel 4 experiment was conducted to reassess its effect on rainfall and provide a basis to evaluate its utility. Operational seeding started after two randomized experiments, the second ending in 1976, found a large and statistically significant effect of cloud seeding on rainfall. Observational studies in later years raised doubts as to the magnitude of the effect, possibly because of changing climatological conditions. A carefully designed randomized experiment was conducted from 2013 to 2020. A unique feature of the design was the use of forecast rainfall on target, rather than rainfall in an unaffected area, as a control variate to attenuate variability. The Israel 4 experiment was stopped a year earlier than planned, because the result was disappointing: a 1.8% increase, p value = 0.4, and 95% confidence interval of (−11%, 16%). These results led to a decision by the Israel Water Authority to stop operational seeding.

SIGNIFICANCE STATEMENT: The recent cloud seeding experiment in northern Israel did not show a significant rainfall increase—unlike the sequence of seeding experiments conducted in Israel in the previous century.

KEYWORDS: Cloud seeding; Experimental design; Orographic effects; Weather modification

1. Introduction

Cloud seeding has been studied for more than 70 years as a means for increasing rainfall or snow precipitation. If successful, seeding is a cost-effective procedure and experiments continue to examine its efficacy. French et al. (2018) presented the scientific basis for the potential of cloud seeding in orographic regions, using both radar measurements and airborne cloud probes to obtain data illustrating the effects of glaciogenic seeding. Dong et al. (2020, 2021) also described microphysical signatures of seeding on cloud structure. Models that carry out detailed simulations of cloud microphysics have also illustrated seeding effects (Xue et al. 2013a,b; Chu et al. 2014; Xue et al. 2016; Geresdi et al. 2017, 2020). Tessendorf et al. (2019) describe the use of further tools, combining computer models with radar and Doppler measurements to study cloud processes. A consistent finding in the microphysical studies is that seeding effectiveness is strongest when natural rainfall efficiency is low.

Despite the scientific basis for the potential of cloud seeding in orographic regions (e.g., French et al. 2018), most experiments failed to demonstrate beneficial effect (WMO 2010). For example, the Wyoming Weather Modification Pilot Project (Breed et al. 2014), a 6-yr study, estimated a small positive seeding effect, but far from strong enough to discredit the null hypothesis that seeding has no effect (Rasmussen et al. 2018). The Snowy Precipitation Enhancement Research Program in Australia reported similar results (Manton and Warren 2011). The review by Rauber et al. (2019) described additional experiments, with little support for rainfall enhancement by seeding.

One notable exception was the sequence of experiments conducted in Israel in the 1960s and 1970s (Gagin 1981). Those experiments concluded that seeding was effective in the north of Israel and served as the basis for conducting operational seeding there into the twenty-first century.

Over the years of operational seeding, unseeded periods resulting from technical difficulties provided more recent estimates of the seeding effect. These analyses showed decreased seeding effects: 6%–11% through 1990 (Nirel and Rosenfeld 1995), and no enhancement at all by 2002 (Sharon et al. 2008). Indirect hydrological analysis questioned the size of the effect (Benjamini and Harpaz 1986), and further doubts were raised by Alpert et al. (2008) and Levin et al. (2010). These studies, together with the possible effect of changing atmospheric conditions, including the increased concentration of aerosols due to pollution (Givati and Rosenfeld 2004, 2005), prompted debate at the Israel Water Authority (IWA) (then the Israel Water Commission) on the potential benefit of conducting a new experiment.

Another important motivation for reevaluation of cloud seeding was the introduction to Israel, starting in 2006, of seawater desalination plants, so that additional water resources

Corresponding author: David M. Steinberg, dms@tauex.tau.ac.il

DOI: 10.1175/JAMC-D-22-0077.1

from cloud seeding should be weighed and compared with desalination. Ultimately, the Water Authority appointed a steering committee, which planned, oversaw, and supervised a new experiment, Israel 4, to evaluate the effect of seeding from airplanes and ground generators on precipitation in the Lake Kinneret watershed basin. The experiment started in December 2013, and the IWA ended the experiment in spring 2020. We report here on the design and the primary results and conclusions of the Israel 4 cloud seeding experiment.

2. Cloud seeding in Israel

Although initial seeding efforts date back to 1949, the first large seeding experiment (Israel 1) was launched only in February 1961 (Gagin and Neumann 1974, 1981). It ran through April 1967 and included 364 experimental days. There were two regions, one in northern Israel and the other in central Israel, with a neutral zone in between them. The northern region was relatively close to the Mediterranean coast and did not extend eastward to Lake Kinneret (the Sea of Galilee), which is an important water source for Israel. The design was a crossover, with each day randomized to seeding in exactly one of the two regions. The seeding effect, combined over the two regions, was estimated by a root double ratio statistic as 15% enhancement (Gabriel 1970; Gagin and Neumann 1974, 1981). The hypothesis that seeding has no effect was rejected by a statistical test with p value $= 0.009$. The maximal effect, a 22% enhancement, was estimated for part of the northern region extending from near the coast to about 10 km from Lake Kinneret.

The Israel 2 experiment focused on the northern region, extending over a target area that covered the drainage basin for Lake Kinneret. The experiment included 388 days over the years 1969–75. Israel 2 also had several target regions, with two primary regions just east (and hence downwind on typical rain days) of the seeding line, one covering the more northern areas, the other farther to the south, and additional target areas farther to the east, surrounding the lake. The area to the west of the seeding line was used as a control area to help neutralize some of the large natural variation in daily rainfall amounts via a double ratio statistic (Gabriel and Feder 1969). The final analysis estimated a 13% (p value $= 0.028$) increase in rainfall from seeding (Gagin and Neumann 1981). More detailed analysis of the results pointed to an especially strong benefit of seeding when the cloud-top temperatures were between $-20°$ and $-15°C$ (Gabriel and Rosenfeld 1990).

A third experiment, Israel 3, run during 1975–95, examined the effects of cloud seeding on rainfall in southern Israel. Israel 3 did not show a beneficial effect of seeding there (Rosenfeld 1998).

The positive conclusions from Israel 1 and Israel 2 for seeding in the north were consistent with the scientific understanding that seeding was most likely to have a beneficial effect on orographic clouds (French et al. 2018), which are typical of northern Israel inland from the Mediterranean Sea (Freud et al. 2015). They served as the basis for ongoing operational seeding, which was carried out from 1975 to 2013.

Between 2007 and 2012, two studies were carried out to provide better scientific support for the statistical seeding experiment. Zipori et al. (2012) examined the chemical signature of rainwater from three sampling stations in the target area, and one station west of the seeding flight path, which served as a control. They found statistically significant silver enrichment (as compared with aluminum, an indicator for dust) in the target area relative to the control station, supporting the hypothesis that the seeding material (silver iodide) was present in clouds in the target area. Furthermore, using satellite data, they also showed that the maximal silver enrichment was seen when the rain was collected from mixed-phase clouds, consistent with scientific understanding of the role of silver iodide in precipitation-forming processes. Freud et al. (2015) performed 27 research flights above the coastal region and the target area. They showed that clouds above the target area often contain high concentrations of supercooled water, which makes them a good target for glaciogenic seeding. In addition, they showed that there are two main natural seeding processes in the area: 1) hygroscopic seeding of sea salt aerosols by strong winds; and 2) a "seeder–feeder" mechanism, as high-level clouds "seed" lower-level clouds with hydrometers.

3. The Israel 4 experimental plan

a. *Goal and target area*

The broad goal of the Israel 4 cloud seeding experiment was to assess the effect of operational seeding, conducted according to best agreed upon practices, from both planes and ground generators, on the precipitation in the Lake Kinneret watershed basin. This led to the specific, and directly measurable, goal of assessing the effect of seeding on rainfall in particular regions of the watershed. Figure 1 shows a map of the lake and the surrounding basin. The primary target area from the experiment was composed of the areas N6, N7, N8, and N9 in the figure. Area N3 was defined as a secondary target area. Several reasons led to the decision to focus on the Lake Kinneret basin: the orographic conditions prevailing in this watershed, and the fact that the lake is a major water resource for Israel, and the earlier sequence of experiments, which found positive seeding effects only in northern Israel. The division into regions followed the classification from the earlier seeding experiments, to facilitate comparison of results with those experiments.

The experimental null hypothesis was that there would be no positive effect of seeding on rainfall in the primary target area and would be assessed with a one-sided hypothesis test, using 0.05 as the level needed to declare a statistically significant effect of the seeding. The effects on subregions were further defined as secondary hypotheses, and their conclusions could be used for assessment of economic value.

The coastal strip in Israel (C2 in Fig. 1) is west of the seeding routes and was defined in the initial planning stages as a control region, as in the Israel 2 experiment. However, the control region was not subsequently used to analyze the experimental data; see the section on the statistical analysis method for details.

b. *Study period*

The Israel 4 experiment began in December 2013, after final approval from the IWA. The experimental seeding was carried out each year through 30 April, and in all subsequent years the

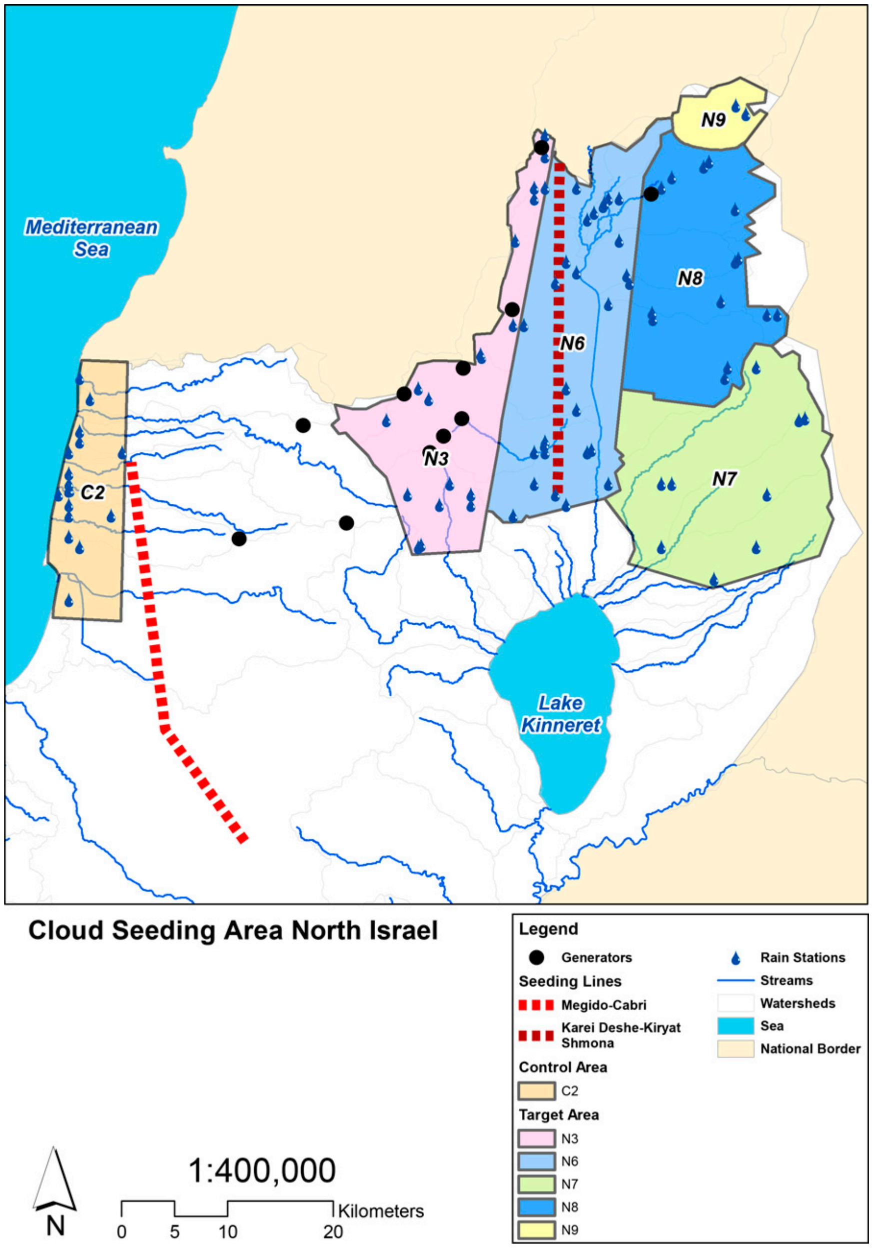

FIG. 1. Map of the study region, showing Lake Kinneret, the primary target region (N6, N7, N8, and N9), the secondary target region (N3), the control region on the coast (C2), the rain gauge and ground generator locations, and the flight paths. There were two flight paths near the coast, one matching the entire red line and the other beginning farther north on this line.

seeding period began on 1 November. The experiment was limited to this period because there is little rainfall in Israel outside those months.

Simulations based on rainfall in prior years showed that about 300 experimental units would be needed to achieve statistical power of 80% to detect a 20% enhancement in rainfall. With an expectation of 35–40 units per year, the steering committee decided on an 8-yr time horizon for Israel 4. An interim analysis was planned after the first four years with the option of early termination in case of clear-cut evidence that seeding is beneficial for rainfall or that seeding had no benefit. The experiment continued for seven years, ending after the winter of 2019/20.

c. *Rainfall measurements*

Rainfall was measured by the Israel Meteorological Service (IMS) at a network of rain gauges across the target area. Measurements were made daily and covered the period from

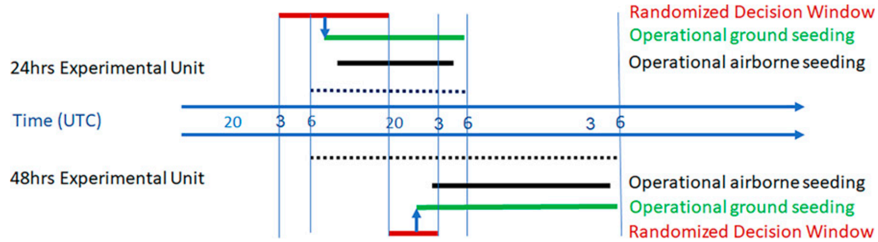

FIG. 2. Time line showing 24- and 48-h experimental units, the decision windows that define them, and typical corresponding seeding times. The figure shows two rain days, each from 0600 to 0600 UTC the following day. The remaining information relates to experimental units declared during the first of the two rain days, with a typical 24-h unit above the time line and a 48-h unit below the time line. The red outside lines show the decision windows. The small lines with arrows extending from the decision windows show hypothetical decision times. Airborne seeding began 2 h after the decision was made, because of operational delay. Ground seeding was initiated at the time of decision, unless the decision was made between 0300 and 0530, in which case it was delayed to 0530. If not ended earlier, ground seeding was stopped at 530 and airborne seeding was stopped at 500 prior to the end of the experimental unit.

0600 UTC the previous day to 0600 UTC the day of the measurement. These 24-h measurement units are referred to as "days" throughout this work and are matched to the dates on which the 24-h unit began.

The rain gauge locations are shown on Fig. 1. The experimental plan included 36 gauges in the primary target area, 16 in the secondary target area, and 15 in the control region.

d. Seeding

Seeding was carried out on allocated days (as described below), both from the air and from ground generators in the primary target area. The seeding was directed by the Rain Enhancement Branch (a division of EMS Mekorot, the national water company of Israel) financed by the IWA. Seeding followed best current practice, so as to mimic operational seeding. This meant that the branch staff would activate the seeding plane and the generators only when local weather conditions were considered appropriate (see section 3e).

Airborne seeding was conducted by burning silver iodide (AgI) solution in acetone (4.5%) at a rate of approximately 40 L in 3.5 h. The steering committee agreed that seeding should be carried out by all available means following best current practices and gave analysts at the branch and airborne operators the liberty to choose the route to best match current conditions. Three primary flight paths were set at the planning stage, as shown in Fig. 1. Two of the routes were about 10 km from the Mediterranean coast and differed only in that one was slightly longer than the other; the third route was in the east of Israel, much closer to the primary target area, for use on days when the wind direction would carry the AgI to the north of the target region if one of the western flight routes was used. Seeding was not conducted west of these primary flight routes.

Because measurements were at the resolution of rain days, it was essential to match the window when seeding could occur to the 24-h period beginning at 0600 UTC. There was usually a delay of about 2 h between the decision to seed and the arrival of the plane to the start of the flight route, from the need to contact the pilot and to get the plane to the seeding line. Furthermore, a transport time of about 30 min was required for AgI dispersed on the western seeding routes to reach the target area (Zipori et al. 2012). Consequently, the time window for airborne seeding for a given day was from 0500 (just prior to the start of the rain day) to 0530 on the following calendar day (just prior to the end of the rain day). See Fig. 2 for a timeline; section 3f explains the decision windows and the division in Fig. 2 into 24- and 48-h experimental units.

Airborne seeding was accompanied by seeding from 16 ground generators (each with two chimneys) located in the target area (see Fig. 1 for locations). The generators dispersed AgI at a concentration of 1%, and at 3.8 L h^{-1} for both chimneys. Simulation results on the dispersion of inert particles indicated that the particles from the ground generators reach the −10°C isotherm, and so reach the clouds, which are usually about 1000 m above sea level. On days allocated to seeding, ground generators were operated from 0530 until 0530 on the following calendar day.

e. Conditions for seeding

Days were included in the experiment only when they were classified as appropriate for seeding from satisfying a number of meteorological and operational conditions. The meteorological conditions reflected scientific understanding of rain generation in orographic clouds, for example that cloud-top temperatures must be sufficiently cold for AgI seeding to be active. The conditions also required weather patterns, and in particular wind directions, that would support transport of the silver iodide from the airborne seeding routes to the target area. Sufficient wind speed is also needed for transport and to achieve some orographic lifting of the seeded material into the clouds. Two of the conditions exploited model-based forecasts:

1) At least three hourly forecast images during the day for which (i) the direction of the wind along the seeding route at 850-hPa was between 210° and 290° azimuth for at least

80% of the pixels and (ii) the temperature at the cloud top was less than −8°C for at least 5% of the pixels in the target area.
2) The forecast rain integrated over the target area for that day was at least 2 mm.

The requirement of a forecast of minimal integrated rain in the target area was related to our goal of evaluating the effect of operational seeding which, even if scientifically effective on low-rainfall days, was not considered to be cost effective. Therefore, days with low forecast rainfall would not have been seeded under standard operating procedures.

Initially forecasts were made using the Weather Research and Forecasting (WRF) Model (https://www.mmm.ucar.edu/weather-research-and-forecasting-model) both at times 0000 and 1200 UTC at a resolution of 1.3 km (for the next 36 h) and at a resolution of 4 km (for 36–72 h) were used. Later the steering committee decided to switch to use of the Consortium for Small-Scale Modeling (COSMO) forecast model (Baldauf et al. 2011; see also http://www.cosmo-model.org/content/model/cosmo/general/dynamics.htm).

When the forecast conditions were met, the day was monitored for the development of more specific conditions that would justify initiating seeding. The following conditions were required:

3) existence of low clouds in satellite images up to 150 km upwind from the western seeding route and/or the development of low clouds in the seeding area itself along with cloud peak temperatures less than −5°C upwind from the seeding line, or less than −8°C in the target region,
4) condition of wind direction as detected in the radar software ("EDGE"), up to a distance of 150 km from the western seeding line and/or in the seeding area, between 220° and 280° with a minimal speed of 15 kt (1 kt ≈ 0.51 m s^{-1}),
5) availability of data from at least one radar station,
6) detection of radar echoes up to 150 km upwind from the western seeding route and/or the development of radar echoes in the seeding area,
7) availability of one of the planes designated for seeding,
8) condition of open air space with no safety issues restricting flight, and
9) availability of VHF radio communication.

Days meeting all the above conditions were defined as "appropriate for seeding" and were included in the experiment. The decision could be made at any time during the day.

f. Experimental units

As noted, "rain days" were 24-h periods beginning at 0600 UTC. Experimental units were either single days (24-h units) or two consecutive days (48-h units). The guidelines for forming these units were based on the desire to avoid declaring a day as a 24-h unit when the weather conditions that justified seeding first occurred near the end of the 24-h period; typical examples are depicted in Fig. 2. The decision interval for each rain day was defined as 0300 UTC just prior to the start of the rain day until 0300 on the next day to compensate for the delay from the time a decision was made to the initiation of airborne seeding. For example, a decision to include the current rain day in the experiment was made at 0430 or 0800 the same calendar day, or 0100 on the next calendar day. When a positive inclusion decision was made early in the rain day, that 24-h period served as a 24-h experimental unit. However, when a positive decision was made late in the rain day, the steering committee preferred to merge the current day with the following day, generating a 48-h (2 day) experimental unit. Preliminary research (Hall 2015), using past rainfall data, showed a small advantage to setting the decision boundary at 2000 UTC over other cutoff times in terms of reducing the variance associated with estimating the effect of the seeding. Thus 48-h units were generated when the decision time was between 2000 and 0300 UTC. Although these units reported rainfall from 48-h time windows, the rain began late in the first 24 h, so that the potential duration of rainfall was comparable to that of the 24-h experimental units.

g. Randomization

Experimental units were randomly allocated to be either seeded or unseeded by the steering committee's statistical team. Separate sets of closed randomization envelopes were provided in advance by the statisticians for the 24- and 48-h units. When a positive inclusion decision was made, the staff of the Rain Enhancement Branch opened the next envelope for the unit type. Thus, the decision to include a day in the experiment was made without knowledge of whether or not that unit would be assigned to seeding. The randomization enforced some balance in both lists. Specifically, for the 24-h units, 7 of each 14 successive units were allocated to seeding; for the 48-h units, 6 of each 12 successive units were allocated to seeding. The operational staff was not aware of this balancing and so could not anticipate the allocation of the final envelope(s) for each type of unit within each of these blocks.

h. Statistical analysis

The statistical analysis used the double ratio (DR) statistic (Gabriel and Feder 1969),

$$\mathrm{DR} = \frac{R_s/R_u}{C_s/C_u},$$

where R_s and R_u are the average rainfall over the target area on seeded and unseeded experimental units, respectively, and C_s and C_u are matching averages for an appropriate control variate. In computing the numerator, first each unit was summarized by averaging the results from all target area stations with rainfall data, and then the units were averaged.

The DR gives a direct estimate of the seeding effect and has been widely used to analyze cloud seeding experiments (Rangno and Hobbs 1987; Super and Heimbach 1983), including past experiments in Israel. Values greater than 1 indicate a positive effect. The control ratio helps to neutralize some of the large variability that is inherent in rainfall data, thus providing a more precise estimate of the seeding effect.

A novel aspect of Israel 4 was the choice of the control variate as forecast rain in the target area from a numerical weather prediction model that did not include any effects of cloud seeding. The common choice in cloud seeding experiments (including the

Israel 2 and 3 experiments) has been to use rainfall in an unseeded area as the control variate. A control variate must be unaffected by seeding and, to effectively reduce variance, it should be highly correlated with rainfall in the target area. In the planning stages of Israel 4, we found that rainfall in the target area had higher correlations with forecasts for the target area from the WRF Model than with observed rainfall in the coastal region. For example, in the N8 region, the correlation with the forecasts was 0.81 as compared with 0.69 with rain in the coastal region. In the N3 region, the matching correlations were 0.89 (forecasts) and 0.78 (coast). The forecast rainfall was computed in an exactly analogous manner to the actual rainfall. Forecasts were computed for each rain gauge location (as the average of the forecasts at the four model grid points that surround the station location) and then those forecasts were averaged over gauges to provide regional forecasts.

The final analysis used forecasts from the COSMO model. We found that COSMO forecasts were more highly correlated with rainfall on unseeded days in Israel 4 than were the WRF forecasts. The model was applied retroactively to the prevailing weather conditions for each included day from the beginning of the experiment. For COSMO, the nearest model grid point to each station was used to generate the rainfall forecast rather than the mean of the surrounding grid points, as in WRF.

The use of forecast rainfall rather than rainfall in the coast also made it possible to link the control variate average directly to the available rainfall data. If a rain gauge failed to report rainfall for an experimental unit, the grid point matching that station was also removed from the forecast rainfall for the unit.

A randomization test was used to assess the null hypothesis of no positive seeding effect. In this test, we generated 10 000 new random allocations to seeded/unseeded units, applying the same block balancing rules used in the experiment. We then computed for each one of these phantom allocations the DR from the observed data. If seeding had an effect, the actual DR from the experiment should be larger than most of the phantom DRs; the fraction of them that exceed the actual DR is the p value for testing the hypothesis. The analysis was carried out using the "intention to treat" principle (Gupta 2011), by which a unit enters the analysis as seeded/unseeded according to its random allocation assignment, even if that does not match what happened in practice. For example, if a unit was allocated to be seeded, but no seeding actually took place, the unit would still enter the analysis as "seeded." Analysis by intention to treat is necessary to justify the randomization test that we used and helps in reducing unintentional biases.

The standard error of the DR as an estimate of the true seeding effect and 95% confidence intervals were computed by bootstrap analysis (Efron and Tibshirani 1994, chapter 13).

The statistical analysis was carried out using the *R* environment, version 4.1.

i. Early termination

If seeding has a positive effect, an added cost of the experiment is the loss in rainfall due to seeding only on half of the appropriate days. The steering committee included an option for early termination of Israel 4 after 4 years and a return to operational seeding if the data collected during that time provided clear-cut evidence of a positive effect. To preserve an overall 5% level of significance for the experiment, the steering committee required evidence exceeding an unadjusted significance threshold of 0.001 for the test after year 4 and then applied a 0.049 threshold for the final test. The use of adjusted thresholds for experiments with an interim analysis has been thoroughly studied in the context of clinical trials in medicine (Demets and Lan 1994; O'Brien and Fleming 1979).

TABLE 1. Units allocated to seeded/unseeded conditions by rainy season.

	24-h units		48-h units	
Rainy season	Seeded	Unseeded	Seeded	Unseeded
2013/14	6	5	1	4
2014/15	18	18	1	0
2015/16	5	7	3	1
2016/17	13	15	3	3
2017/18	10	8	4	2
2018/19	22	23	0	1
2019/20	16	15	1	1
Total	90	91	13	12

4. Results of Israel 4

a. Experimental days

The Israel 4 experiment began in November 2013 and continued for seven years, through the winter of 2019/20. Israel 4 included 206 experimental units and 231 rain days. Table 1 shows the number of 24- and 48-h units allocated as seeded/unseeded each year. The number of units varied considerably from one year to another, primarily from weather variation, with only a small number of rainy days in some of the years. However, logistical problems also occurred. The low number of experimental units in the winter of 2015/16 resulted in part from a technical problem that grounded the seeding plane for about 1 month in the middle of the rainy season. That entire period was excluded from Israel 4.

There were some minor deviations from the planned randomization scheme and in implementing the randomized allocation. Duplicate envelopes were provided for some of the randomization numbers and resulted in erroneously using four random unit assignments twice. In addition, one of the 48-h units was randomized using a 24-h unit envelope (the envelope was labeled 24–48 and intended for use with the 48th unit that was 24 h long). Implementation exactly matched the random allocation for each experimental unit. The only minor deviation was one unit allocated to seeding in which the generators were engaged but there was no airborne seeding. On that day there was almost no rain and conditions never were sufficiently conducive to justify seeding from the air.

The basis of the randomization test is the random assignment to the unit numbers of whether or not to seed. Therefore, in carrying out the test, we repeated these deviations. Each phantom randomization resulted in a new set of assignments (seed/no seed) to

TABLE 2. Average (by experimental unit) rainfall (mm) and average forecast rainfall in the primary target area (boldface type) and in each of its subareas (N6, N7, and N8) and in the secondary target area (N3).

	Avg rainfall		Avg forecast rainfall			
Region	Seeded	Unseeded	Seeded	Unseeded	Double ratio	95% confidence interval
Primary target	**15.77**	**13.27**	**11.90**	**10.21**	**1.018**	**0.89**, **1.16**
N6	14.39	12.01	11.17	9.73	1.046	0.90, 1.23
N7	16.73	14.00	11.09	9.11	0.980	0.83, 1.16
N8	20.01	17.25	15.98	13.50	0.980	0.87, 1.10
Secondary target N3	19.40	15.95	14.24	11.65	0.995	0.87, 1.14

the unit numbers, and each experimental unit was then matched to the same unit number actually used to randomize that unit.

b. Rainfall

Complete rainfall data were collected for most of the rain gauges in the target area. A small number of stations reported no data. Notably, that was the case for the sole station in N9, which was thus dropped from our analysis. Three gauges in the primary target area, one in the secondary target area and one in the control area provided partial data, failing to report on some days, with substantial loss of data from three gauges (Yassur—79 units, Ortal—41 units, and Kfar Blum Manual—33 units). Yassur is in the control area and was not used in our final analysis. For the other gauges, the extent of missing data was similar for both seeded and unseeded units, mitigating concern that differential missing data biased the results.

In the primary target area, 79.5% of the total rainfall during the period of Israel 4 occurred on the days included in the experiment. The inclusion percentages at the different target area rain gauges varied from 74.9% to 83.2%. The rainy days not included resulted either from unsuitable meteorological conditions (e.g., winds from the east) or technical problems (e.g., the period when the seeding plane needed repair).

Table 2 shows average (by experimental unit) rainfall and average forecast rainfall in the primary target area, and in each of the subareas that compose it, for the entire seven years of the experiment. It also shows the DR for each of these areas with a 95% confidence interval. The DR for the full primary target area was 1.018 with standard error (SE) of 0.07; that is, an estimated 1.8% increase in rainfall contributed by seeding. The p value for the one-sided significance test was 0.40. Thus, the Israel 4 experiment did not provide evidence that would support the rejection of the null hypothesis that there is no positive effect of seeding on rainfall in the primary target area. Similar DRs were found for each of the three subareas. The DR for the secondary target area N3 was 0.99 [95% confidence interval of (90.87, 1.10)].

Somewhat higher amounts of rainfall were observed on 48-h experimental units than on 24-h units. On the single day units, the median rainfall in the primary target area was 10.0 mm, with an interquartile range of 4.6–18.7 mm. The matching numbers for the 48-h units were 16.7 and 5.6–23.9 mm.

Figure 3 plots daily average rainfall in the primary target area on the experimental days against the forecast average rainfall by COSMO. There is a strong linear relationship between the forecast and the actual rainfall on the unseeded days with a correlation coefficient of 0.85. Figure 3 shows a regression line through the origin relating the actual to the forecast rainfall on the unseeded days (dashed line). The regression line has a slope of 1.12, indicating that the actual rainfall was, typically, about 12% higher than the forecast rainfall. The underprediction is also evident in Table 2, where the forecast rainfall is consistently less than the actual rainfall, but the ratio of observed to forecast rainfall is 1.17 on the unseeded days, slightly higher than the regression slope.

Earlier cloud seeding experiments in Israel used rainfall on the coastal strip as the control variate, rather than forecast rainfall in the target area. The corresponding correlation coefficient between rainfall in the primary target area and on the coastal strip during Israel 4 was 0.65. As expected, correlations were higher for those subareas that are closer to the coast (0.69 in N6, 0.53 in N7, and 0.57 in N8). These correlations are substantially lower than those between actual and model forecast rainfall (0.80 in N6, 0.85 in N7, and 0.87 in N8).

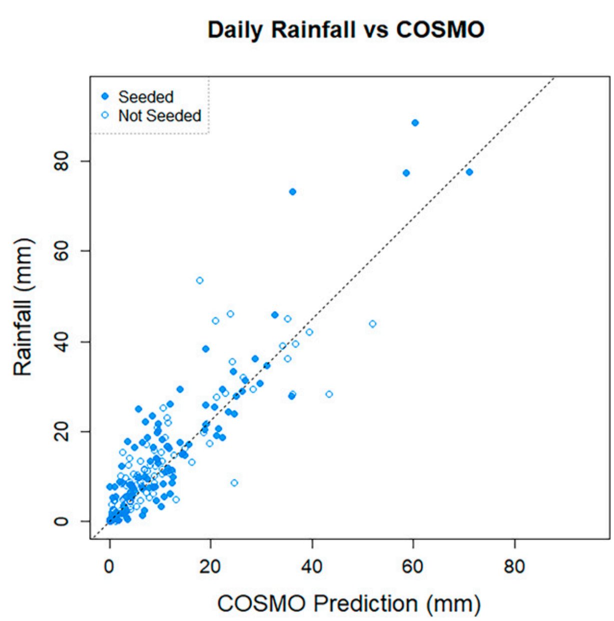

FIG. 3. Average rainfall in the primary target area (mm) on the experimental days against the forecast average rainfall by COSMO (r = 0.85). The dashed line is a regression line through the origin relating actual to forecast rainfall on the unseeded days.

c. *Early termination*

The experimental protocol included a specific option for early termination after 4 years if Israel 4 showed a clear positive effect of seeding. The first 4 years of the experiment included only 119 rain days. The DR was slightly below 1 and there was wide uncertainty because of the small sample size. Consequently, the steering committee recommended continuation of the experiment under the same protocol.

The steering committee met annually to review the results of Israel 4. One of the questions that was examined each year was the probability of reaching a result that would show a positive effect with a small p value, if the experiment continued for the fully planned eight years. At the end of year 7, the estimate of the seeding effect from the combined data was very low (DR = 1.018). The assessment at the end of year 7 indicated that the estimated effect would remain small and would not achieve a small p value if the experiment was continued to its eighth year. As a result, the steering committee recommended termination of Israel 4, and that was the decision reached by the Israel Water Authority. The details of the assessment are given in the appendix.

The steering committee realized that the decision to close the experiment after seven years might affect the analyses that we have reported thus far. Early termination of an experiment because of positive results is known to lead to biased estimates of the effect being assessed (Whitehead 1986). The bias arises because, had the effect been small or modest, the experiment would have continued. Ergo, early termination is more likely to occur when initial results are overoptimistic. Procedures have been developed to adjust effect estimates and confidence intervals following early termination (Todd et al. 1996; Weinstein et al. 2013; Woodroofe 1992).

These problems were examined in the context of Israel 4 by Engel (2020), who found that, if early termination were to follow a small observed effect, standard procedures have a very small bias. Because this was the case for the Israel 4 results, the standard analyses reported here give valid statistical inferences.

5. Summary

The team at the Rain Enhancement Branch, under two different leaders, took great care to conduct the experiment in strict adherence to its protocol. All deviations from the experimental design were approved by the steering committee and are reported in this paper (section 4a). The two statisticians on the steering committee closely monitored the experiment throughout its duration.

The primary goal of the experiment was to assess whether regular yearly seeding operations, conducted using best currently known practices, increases the amount of rainfall in the primary target region, the Kinneret basin. The final analysis of the experimental data estimated a 1.8% increase in rainfall (95% confidence interval, from −11% to +16%) resulting from seeding. This effect was much too small, relative to the large variability in rainfall, to prove the existence of a positive seeding effect in the target area. In assessing the estimated increase, note that it takes time to fly up and down the seeding line and that there was often a delay in seeding when switching from one aircraft to the other. Thus, even with the policy of best practices, the actual clouds affected by seeding are only a small fraction of the potential clouds.

Our primary goal led us to carry out the Israel 4 experiment on precisely those days when best practices would have triggered the use of operational seeding. As noted in our literature review, a number of recent studies of the microphysical effects of glaciogenic seeding have found that seeding efficiency is greatest when natural precipitation efficiency is low. Thus, an experiment devoted strictly to the scientific goal of assessing the maximal effect of seeding might have removed days with high forecast rainfall. Our goal, on the contrary, led us to remove days with low forecast rainfall, as the additional water that could be generated was not considered to be sufficient to make seeding cost-effective on such days.

A unique feature in the design of the Israel 4 experiment was the use of model forecast rainfall as a control variate. Previous rain enhancement experiments, including Israel 1 and Israel 2, have used rainfall in a control region. The estimated effect for Israel 4 is based on the ratio of average rainfall on seeded versus unseeded days, divided by the matching comparison of the average forecast rainfall on those days. The forecasts were found to be biased negatively (low) by about 12%. Although this may diminish their value as forecasts, it does not limit their value as a control variate. The importance of a control variate is to reflect the natural difference in rainfall between those days allocated to seeded and unseeded conditions. The bias in the forecasts affects both the seeded and unseeded units in exactly the same way, thus it does not affect their ratio, which serves as our control measure. What is important is to adopt a control variate that has a high correlation with actual rainfall in the target area. Our preliminary research indicated that the model forecasts would be much more effective in reducing variation, a conclusion borne out in the actual experimental results, as seen in the higher correlations of actual rainfall in the target area with the forecast rainfall than with the rainfall in the control area along the Mediterranean coast.

We have intentionally limited this article to the central analysis for both the primary and secondary target areas. It is, of course, natural to consider using the experimental data to investigate additional questions, for example to compare the seeding effect under different meteorological conditions or cloud characterizations or to assess its hydrological impact. The experimental protocol allowed for such analyses to be carried out, but only after the primary goal had been examined. Had we included those analyses in the protocol, the additional hypotheses would have necessitated multiple testing adjustments, such as the Bonferroni procedure. Calculations at the design stage showed that such adjustments would greatly reduce the power to prove rain enhancement for the primary hypothesis. Such post hoc analyses may be reported in separate papers, subject to the provision in the protocol that they employ false discovery rate adjustments (Benjamini and Hochberg 1995) that enable statistical procedures to account for multiple testing.

The experiment was stopped with the conclusion that the effect of operational seeding according to current best practices is small. Given the inherent variability in daily rainfall, the results of the Israel 4 experiment are consistent with the hypothesis that seeding has no positive effect at all. The Israel Water Authority accepted the final result of Israel 4 delivered by the steering committee and used it to weigh the cost of conducting operational seeding against the value of the expected additional rainfall that would result. That analysis led to the decision to end the experiment and also to stop operational seeding.

The results of the experiment pertain to the effect of seeding following current best practices for the Lake Kinneret basin. They do not imply that cloud seeding cannot increase rainfall in Israel, or elsewhere, under certain meteorological conditions. Unlike Israel Experiments 1 and 2, conducted almost five decades ago, the Israel 4 experiment failed to provide evidence that seeding causes increased rainfall in the target area.

Acknowledgments. We are grateful to Nati Glick, who headed the Branch for Rainfall Enhancement for part of Israel 4, to Rotem Rozenblum for her assistance with the data analysis, and to Tony Yovel for his assistance in preparing the map of the study region. The work of Ms. Rozenblum was supported by research grant 4500963137 from the Israel Water Authority to D. M. Steinberg and Y. Benjamini.

Data availability statement. The data and code are available online (github.com/dmsteinberg/Israel-4).

APPENDIX

Assessment of the Results after Year Seven

As reported, at the end of year 7, the DR statistic estimated the effect of seeding on rainfall in the primary target area as 1.018. Simulations showed that, to a very close approximation, the logarithm of the DR (equal to 0.018) has a normal distribution whose mean is the logarithm of the true seeding effect. The standard error of the DR, based on our bootstrap analysis, is 0.07. The same value holds for the log(DR) statistic. {A Taylor series expansion relating the DR to the log(DR) shows that $\mathrm{SE}[\log(\mathrm{DR})] \approx (1/\mathrm{DR})\times\mathrm{SE}(\mathrm{DR})$. The DR for the Israel 4 experiment was very close to 1—hence the near equality between the standard deviation (SD) of the log(DR) and that of the DR itself.

Denote by $\hat{\theta}_7$ the observed log(DR) at the end of 7 years and by $\hat{\theta}_8$ the log(DR) we would have observed after one more year of experimentation, had Israel 4 included an eighth year. Further, denote by $\hat{\gamma}_8$ the observed log(DR) for year 8 only. Although there is not a simple formula relating these quantities, numerical comparisons show that, to a good approximation, $\hat{\theta}_8 \approx (7/8)\hat{\theta}_7 + (1/8)\hat{\gamma}_8$. The weights correspond to the assumption that the amount of experiment day rainfall in year 8 would be close to the average of the preceding seven years. Further simulations showed that, even with just one year of data, the log(DR) statistic has an approximately normal distribution. The mean will be the logarithm of the true seeding effect. The standard deviation will be approximately

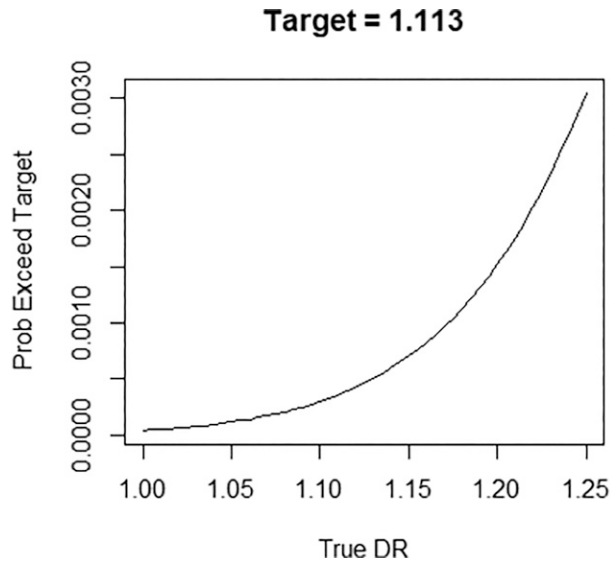

FIG. A1. The probability that continuing the Israel 4 experiment to year 8 would result in a p value of 0.05 for rejecting the hypothesis of no seeding effect.

$\sqrt{7}$ times as large as the SE of $\hat{\theta}_7$, that is, $0.07\sqrt{7} = 0.185$. Thus, conditional on the first seven years of data, we approximate the distribution of $\hat{\theta}_8$, conditional on the results of the first seven years, as $N(0.016 + 0.125\theta_{\mathrm{True}}, 0.023)$. The mean is the corresponding weighted average of $\hat{\theta}_7$ and the logarithm of the true seeding effect; the SD is 1/8 times the SD of $\hat{\gamma}_8$. The above distribution can be used to compute the probability that the 8-yr double ratio would exceed any given threshold, conditional on the 7-yr results.

The scientific goal for Israel 4 was to achieve a one-sided p value of 0.05 or less. We use this goal to derive the relevant threshold for the above calculation. The final p value, computed from the randomization test, can be approximated using the same normality approximations described above. The relevant threshold would use the marginal standard error of $\hat{\theta}_8$ and not the conditional standard deviation used above. Following standard statistical reasoning, we can approximate the SE for N years of data by $\sigma/\sqrt{N}$, where σ is the standard deviation for a single year. Thus, we can approximate the marginal SE of $\hat{\theta}_8$ by $0.07\sqrt{7/8} = 0.065$. The resulting threshold for a p value of 0.05 or less is 0.107 for $\hat{\theta}_8$, equivalent to a final DR of 1.113.

Figure A1 shows the probability of exceeding an 11.3% seeding effect for true effects ranging from 1 (no effect) to 1.25 (slightly larger than the upper limit of our 95% confidence interval for the seeding effect). Even for strong effects as high as 1.25, the probability of obtaining a result that achieved the 5% p value was only about 0.003.

REFERENCES

Alpert, P., N. Halfon, and Z. Levin, 2008: Does air pollution really suppress precipitation in Israel? *J. Appl. Meteor. Climatol.*, **47**, 933–943, https://doi.org/10.1175/2007JAMC1803.1.

Baldauf, M., A. Seifert, J. Förstner, D. Majewski, M. Raschendorfer, and T. Reinhardt, 2011: Operational convective-scale numerical weather prediction with the COSMO model: Description and sensitivities. *Mon. Wea. Rev.*, **139**, 3887–3905, https://doi.org/10.1175/MWR-D-10-05013.1.

Benjamini, Y., and Y. Harpaz, 1986: Observational rainfall–runoff analysis for estimating effects of cloud seeding on water resources in northern Israel. *J. Hydrol.*, **83**, 299–306, https://doi.org/10.1016/0022-1694(86)90158-7.

——, and Y. Hochberg, 1995: Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. Roy. Stat. Soc.*, **57B**, 289–300, https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.

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.

Chu, X., L. Xue, B. Geerts, R. Rasmussen, and D. Breed, 2014: A case study of radar observations and WRF LES simulations of the impact of ground-based glaciogenic seeding on orographic clouds and precipitation. Part I: Observations and model validations. *J. Appl. Meteor. Climatol.*, **53**, 2264–2286, https://doi.org/10.1175/JAMC-D-14-0017.1.

Demets, D. L., and K. K. G. Lan, 1994: Interim analysis: The alpha spending function approach. *Stat. Med.*, **13**, 1341–1352, https://doi.org/10.1002/sim.4780131308.

Dong, X., C. Zhao, Y. Yang, Y. Wang, Y. Sun, and R. Fan, 2020: Distinct change of supercooled liquid cloud properties by aerosols from an aircraft-based seeding experiment. *Earth Space Sci.*, **7**, e2020EA001196, https://doi.org/10.1029/2020EA001196.

——, and Coauthors, 2021: Increase of precipitation by cloud seeding observed from a case study in November 2020 over Shijiazhuang, China. *Atmos. Res.*, **262**, 105766, https://doi.org/10.1016/j.atmosres.2021.105766.

Efron, B., and R. J. Tibshirani, 1994: *An Introduction to the Bootstrap*. 1st ed. Chapman and Hall, 456 pp.

Engel, D., 2020: Estimating the effect of cloud seeding in an experiment with an early termination option. M.Sc. thesis, Dept. of Statistics and Operations Research, Tel Aviv University, 99 pp.

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.

Freud, E., H. Koussevitzky, T. Goren, and D. Rosenfeld, 2015: Cloud microphysical background for the Israel-4 cloud seeding experiment. *Atmos. Res.*, **158–159**, 122–138, https://doi.org/10.1016/j.atmosres.2015.02.007.

Gabriel, K. R., 1970: The Israeli rainmaking experiment 1961–67 final statistical tables and evaluation. Hebrew University Department of Statistics Tech. Rep., 47 pp.

——, and P. Feder, 1969: On the distribution of statistics suitable for evaluation of rainfall stimulation experiments. *Technometrics*, **11**, 149–160, https://doi.org/10.1080/00401706.1969.10490667.

——, and D. Rosenfeld, 1990: The second Israeli rainfall stimulation experiment: Analysis of precipitation on both targets. *J. Appl. Meteor. Climatol.*, **29**, 1055–1067, https://doi.org/10.1175/1520-0450(1990)029<1055:TSIRSE>2.0.CO;2.

Gagin, A., 1981: The Israeli rainfall enhancement experiments: A physical overview. *J. Wea. Modif.*, **13**, 108–120, https://doi.org/10.54782/jwm.v13i1.41.

——, and J. Neumann, 1974: Rain stimulation and cloud physics in Israel. *Weather and Climate Modification*, W. N. Hess, Ed., John Wiley and Sons, 454–494.

——, and ——, 1981: The second Israeli randomized cloud seeding experiment: Evaluation of the results. *J. Appl. Meteor. Climatol.*, **20**, 1301–1311, https://doi.org/10.1175/1520-0450(1981)020<1301:TSIRCS>2.0.CO;2.

Geresdi, I., L. Xue, and R. Rasmussen, 2017: Evaluation of orographic cloud seeding using a bin microphysics scheme: Two-dimensional approach. *J. Appl. Meteor. Climatol.*, **56**, 1443–1462, https://doi.org/10.1175/JAMC-D-16-0045.1.

——, ——, N. Sarkadi, and R. Rasmussen, 2020: Evaluation of orographic cloud seeding using a bin microphysics scheme: Three-dimensional simulation of real cases. *J. Appl. Meteor. Climatol.*, **59**, 1537–1555, https://doi.org/10.1175/JAMC-D-19-0278.1.

Givati, A., and D. Rosenfeld, 2004: Quantifying precipitation suppression due to air pollution. *J. Appl. Meteor. Climatol.*, **43**, 1038–1056, https://doi.org/10.1175/1520-0450(2004)043.

——, and ——, 2005: Separation between cloud-seeding and air-pollution effects. *J. Appl. Meteor. Climatol.*, **44**, 1298–1314, https://doi.org/10.1175/JAM2276.1.

Gupta, S. K., 2011: Intention-to-treat concept: A review. *Perspect. Clin. Res.*, **2**, 109–112, https://doi.org/10.4103/2229-3485.83221.

Hall, S., 2015: Cloud seeding effect estimation using WRF forecast data in lieu of control area measured rainfall. M.Sc. thesis, Dept. of Statistics and Operations Research, Tel Aviv University, 71 pp.

Levin, Z., N. Halfon, and P. Alpert, 2010: Reassessment of rain enhancement experiments and operations in Israel including synoptic considerations. *Atmos. Res.*, **97**, 513–525, https://doi.org/10.1016/j.atmosres.2010.06.011.

Manton, M. J., and L. Warren, 2011: A confirmatory snowfall enhancement project in the snowy mountains of Australia. Part II: Primary and associated analyses. *J. Appl. Meteor. Climatol.*, **50**, 1448–1458, https://doi.org/10.1175/2011JAMC2660.1.

Nirel, R., and D. Rosenfeld, 1995: Estimation of the effect of operational seeding on rain amounts in Israel. *J. Appl. Meteor. Climatol.*, **34**, 2220–2229, https://doi.org/10.1175/1520-0450(1995)034<2220:EOTEOO>2.0.CO;2.

O'Brien, P. C., and T. R. Fleming, 1979: A multiple testing procedure for clinical trials. *Biometrics*, **35**, 549–556, https://doi.org/10.2307/2530245.

Rangno, A. L., and P. V. Hobbs, 1987: A reevaluation of the climax cloud seeding experiments using NOAA published data. *J. Appl. Meteor. Climatol.*, **26**, 757–762, https://doi.org/10.1175/1520-0450(1987)026<0757:AROTCC>2.0.CO;2.

Rasmussen, R. M., and Coauthors, 2018: Evaluation of the Wyoming Weather Modification Pilot Project (WWMPP) using two approaches: Traditional statistics and ensemble modeling. *J. Appl. Meteor. Climatol.*, **57**, 2639–2660, https://doi.org/10.1175/JAMC-D-17-0335.1.

Rauber, R. M., and Coauthors, 2019: Wintertime orographic cloud seeding—A review. *J. Appl. Meteor. Climatol.*, **58**, 2117–2140, https://doi.org/10.1175/JAMC-D-18-0341.1.

Rosenfeld, D., 1998: The third Israeli randomized cloud seeding experiment in the south: Evolution of the result and review of all three experiment. Preprints, *Conf. on Cloud Physics and 14th Conf. on Planned and Inadvertent Weather Modification*, Everett, WA, Amer. Meteor. Soc., 565–568.

Sharon, D., A. Kessler, A. Cohen, and E. Doveh, 2008: The history and recent revision of Israel's cloud seeding program. *Isr. J. Earth Sci.*, **57**, 65–69, https://doi.org/10.1560/IJES.57.1.65.

Super, A. B., and J. A. Heimbach Jr., 1983: Evaluation of the Bridger Range winter cloud seeding experiment using control gages. *J. Appl. Meteor. Climatol.*, **22**, 1989–2011, https://doi.org/10.1175/1520-0450(1983)022<1989:EOTBRW>2.0.CO;2.

Tessendorf, S. A., and Coauthors, 2019: A transformational approach to winter orographic weather modification research: The SNOWIE project. *Bull. Amer. Meteor. Soc.*, **100**, 71–92, https://doi.org/10.1175/BAMS-D-17-0152.1.

Todd, S., J. Whitehead, and K. M. Facey, 1996: Point and interval estimation following a sequential clinical trial. *Biometrika*, **83**, 453–461, https://doi.org/10.1093/biomet/83.2.453.

Weinstein, A., W. Fithian, and Y. Benjamini, 2013: Selection adjusted confidence intervals with more power to determine the sign. *J. Amer. Stat. Assoc.*, **108**, 165–176, https://doi.org/10.1080/01621459.2012.737740.

Whitehead, J., 1986: On the bias of maximum likelihood estimation following a sequential test. *Biometrika*, **73**, 573–581, https://doi.org/10.1093/biomet/73.3.573.

WMO, 2010: WMO statement on weather modification. WMO Doc., 13 pp., https://www.skywaterventures.com/uploads/7/0/6/1/70616003/wmr_documents.final_27_april_1.final.pdf.

Woodroofe, M., 1992: Estimation after sequential testing: A simple approach for a truncated sequential probability ratio test. *Biometrika*, **79**, 347–353, https://doi.org/10.1093/biomet/79.2.347.

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. A. 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, and B. Geerts, 2016: A case study of radar observations and WRF LES simulations of the impact of ground-based glaciogenic seeding on orographic clouds and precipitation. Part II: AgI dispersion and seeding signals simulated by WRF. *J. Appl. Meteor. Climatol.*, **55**, 445–464, https://doi.org/10.1175/JAMC-D-15-0115.1.

Zipori, A., D. Rosenfeld, J. Shpund, D. M. Steinberg, and Y. Erel, 2012: Targeting and impacts of AgI cloud seeding based on rain chemical composition and cloud top phase characterization. *Atmos. Res.*, **114–115**, 119–130, https://doi.org/10.1016/j.atmosres.2012.05.023.