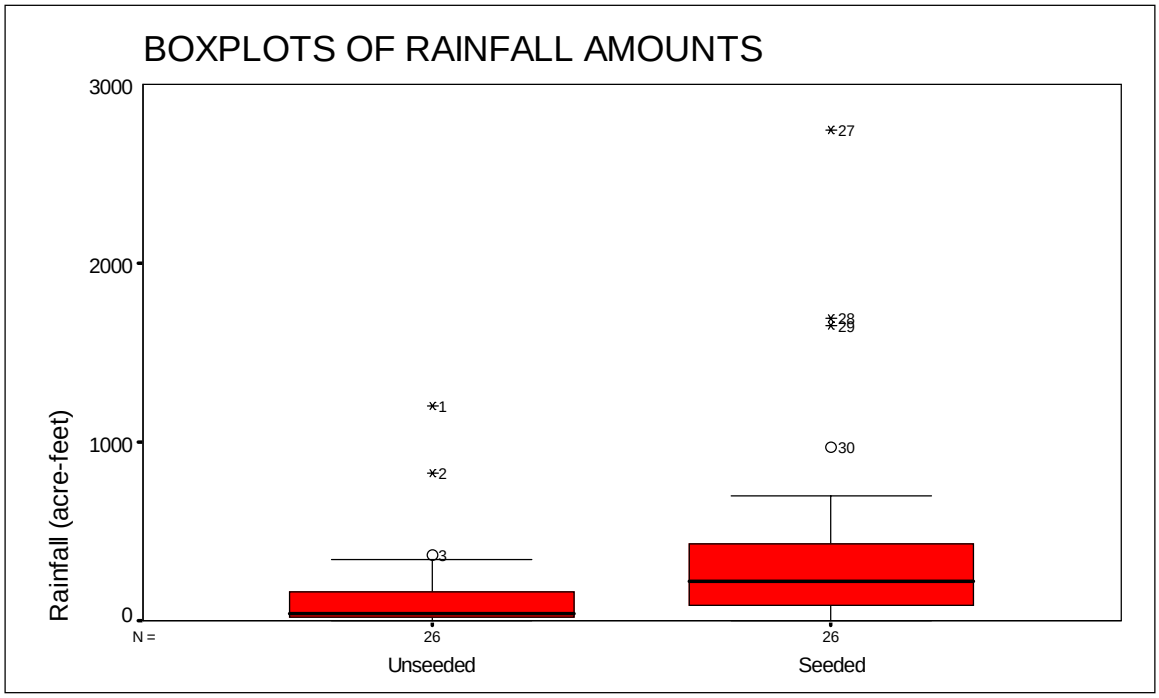
CLOUD SEEDING EXPERIMENT

4. Displaying Seeded and Unseeded Rainfalls

We will examine the data to answer the questions about the effects of cloud seeding on rainfall. First we will use SPSS graphical tools to display the data in our study. The bolded hyperlinks displayed below lead directly to the corresponding topics.

4.1 The side-by-side boxplots of rainfall amounts on the original scale.
4.2 The side-by-side boxplots of rainfall amounts on the scale of the natural logarithm of acre-feet.

4.1 First we will obtain the boxplots of the rainfall amounts on the acre-feet scale for seeded and unseeded days. SPSS (see Section 10) produces the following boxplots:



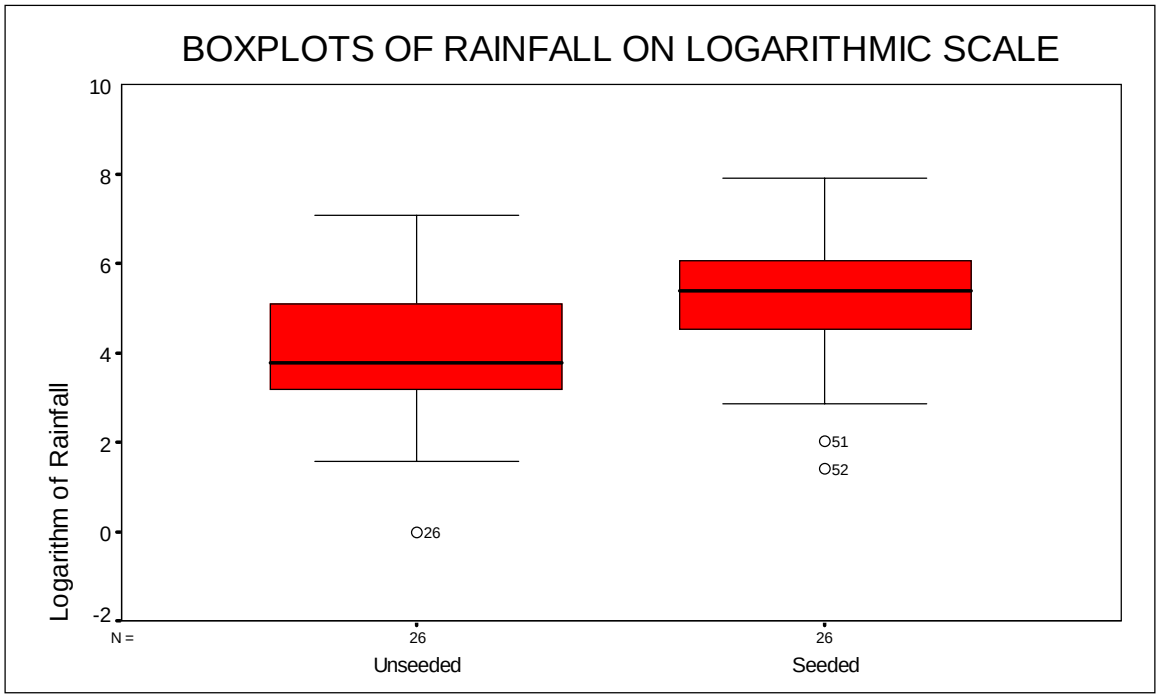
The above drawing displays the side-by-side boxplots of rainfall amount on the acre-feet scale. First we describe each of the two boxplots and then we compare them.

The boxplot for unseeded days is highly skewed to the right. It indicates that most of the time the rainfall volume from unseeded clouds is very small or zero. However, there are outliers and extreme observations in the distribution. The third observation of 372.4 is an outlier, and the first and second observations equal to 1202.6 and 830.1 are considered as extreme observations. The spread of rainfall amount is relatively small.

The boxplot for seeded days is slightly skewed to the right. There is one outlier and three extreme observations in the distribution. The spread of rainfall amount is relatively large.

The side-by-side boxplots indicate that the rainfall tended to be larger on the seeded days. Both distributions are quite skewed, and more variability occurred in the seeded group than in the control group.

4.2 Now we will use SPSS to display and compare the distributions of the natural logarithm of rainfall for seeded and unseeded days.



From just a quick glance, we see that the logarithm transformation has helped compress the numbers. In general, the logarithm transformation tends to pull in the long tail of the distribution on the right, but stretch it out on the left. In other words, small numbers spread out more, while large numbers are squeezed more closely together.

The positions of the quartiles and whiskers indicate that on the logarithmic scale, both distributions are approximately symmetric, and have approximately the same spread. The boxplots confirm the conclusion we have reached in 4.1 that the rainfall tended to be larger on the seeded days.

In general, if the graphical displays of two samples show them both to be skewed and if the sample with the larger average also has a larger spread, the log transformation is likely to be a good choice.