May 2007

Statistical Analysis of NAWC's Winter Cloud Seeding Program for Central and Southern Utah

By

Mason Statistical Consulting Group
David J. Mason, PhD.
Mohammed Chaara, MStat., MBA.

Mason Statistical Consulting Group (MSCG) of Salt Lake City, Utah analyzed 32 winter seasons of operational cloud seeding in the period of 1974-2006. North American Weather Consultants provided these data for the purpose of evaluating their program of augmenting precipitation in Utah's mountains by cloud seeding. The resulting runoff would be augmented, providing additional useful fresh water. The seasons in the study run from December through March and the precipitation data were derived primarily from high elevation SCS/NRCS precipitation gage sites. The SCS/ NRCS adjusted data were used, these adjustments were made to estimate monthly precipitation data from data that were made less frequently than monthly. Historical data for not seeded seasons between 1957 and 1973 with data from 1984 were selected in order to perform target/control analyses. The locations of the target sites and control sites are provided in **Figure 1**. The names of these sites are provided in **Figure 2**. The steps in this analysis included the following:

- **<u>Data preparation:</u>** An outlier analysis was done to come up with a sample for study. Both the seeded and historical data sets had outliers that needed to be excluded from the study. **Figures 6** and **12** display data plots for the historical and seeded values with outliers. MSCG used different techniques to identify outliers in both data sets. For the historical period, Seasons in 1969 and 1972 were outliers and thus, were excluded (check **Figures 7–10** for identifying the outliers). For the seeded data, MSCG excluded seasons 1980, 1986, 1989 and 2002. These seasons were found to be outliers according to **Figures 13-16.** The analysis resulted in 28 seeded and 16 unseeded seasons for further study, **Figures 11** and **17** display the normality check for the studied data sets.

- **<u>Target and control modeling:</u>** Various model-fitting techniques were considered during this analysis including data transformations like logs and first differences. MSCG found that a simple linear regression best describes the relationship between target and control areas and provided a strong correlation between the two areas, yielding a high R^2 value. **Figures 3** and **4** are plots of the control area versus target area data and the fitted line for both seeded and historical periods. The regression models of Y on X found were:

$$Y(n) = 1.49\ X(n) - 1.3327 \text{ (inches)} \quad \text{Adjusted } R^2 = 0.9673$$

$$Y(s) = 1.52\ X(s) - 0.0553 \text{ (inches)} \quad \text{Adjusted } R^2 = 0.9523$$

Where Y denotes the data from the target area, X denotes the control data and s or n denotes the time; seeded or non-seeded. From these regressions, we conclude that the slopes of the lines are the same and the intercepts are not same. Therefore, the difference between the intercepts is a good estimator of the effect of the cloud seeding program. A 90% significance level for the difference of the intercepts is

(0.218 inches, 2.437 inches)

This interval can be interpreted as percentage of change from not seeded to seeded seasonal precipitation:

(2.13% , 23.82%)

This supports the claim that the seeding program leads to 10% or more increase in precipitation. Check **Figure 5** for a graphic explanation.

- **<u>Hypothesis Testing:</u>** The null hypothesis that there was no difference between predicted and observed seasonal seeded precipitation was tested using a two sample non-parametric t-test, Mann-Whitney. It should be noted that these tests were developed for use on randomized data sets and therefore their use on non-randomized data sets (such as these) carry less significance than if the data had been randomized. The null hypothesis was rejected at the 5% significance level for a one-tailed t-test. This means that with 95% confidence the average precipitations observed and the calculated, predicted by the simple linear regression are different.

It is concluded from this analysis that there is a difference in the seeded versus non-seeded precipitation seasons. This difference falls in a range of 0.218 to 2.437 inches of increase in average December through March precipitation in the target area. The analysis led to a p-value of 0.0465 for the Mann-Whitney test for difference, this is significant at the 5% level. It is noted that these data were from a non -randomized data set.

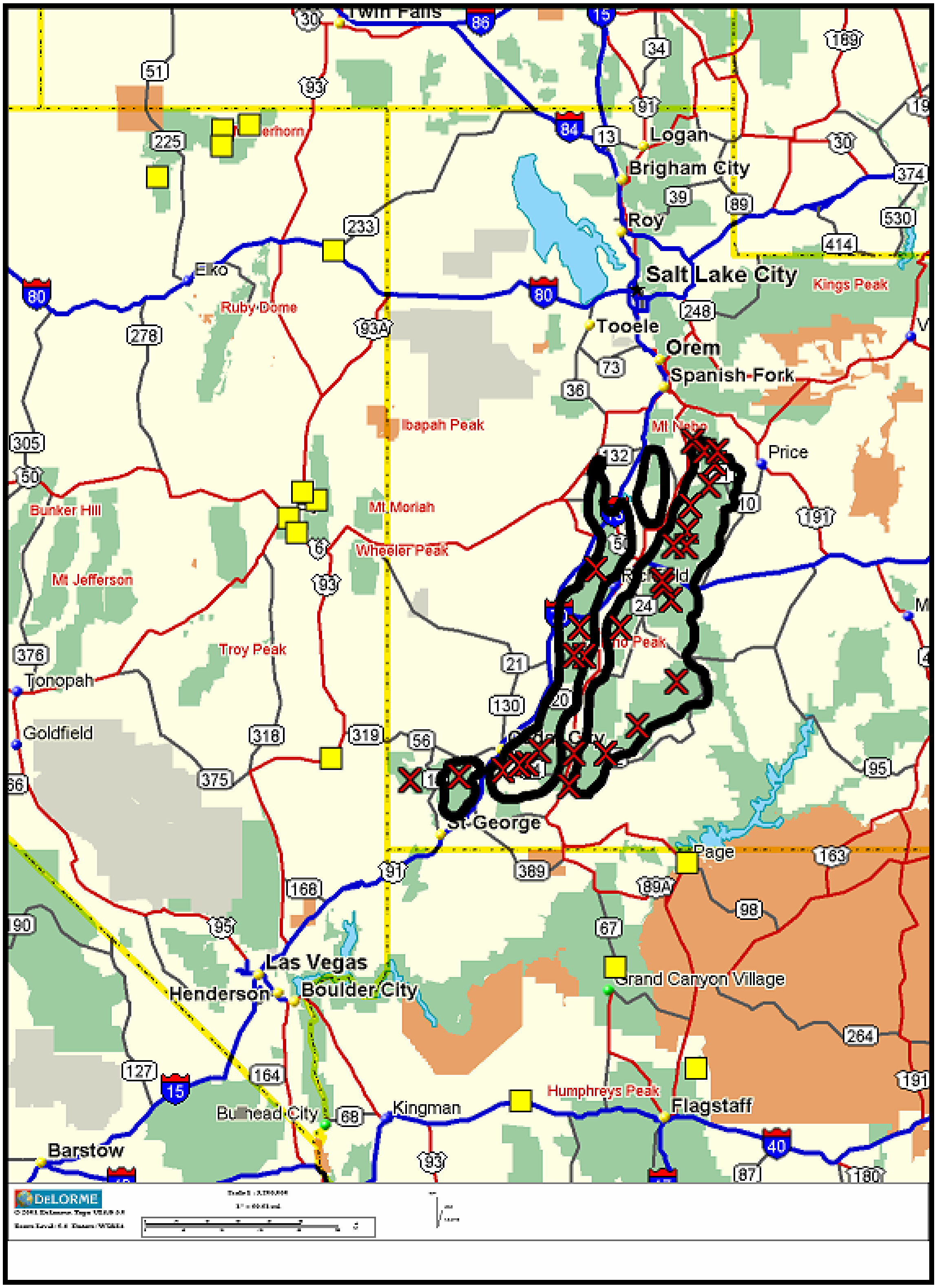


Figure 1: Control and Target Stations

Southern/Central Utah Target and Control Sites

Site Name	Lat(N)	Long(W)	Elev (Ft)
Control Sites			
Bear Creek Tel, Nv	41E50'	115E27'	8040
Berry Creek Tel, Nv	39E21'	114E39'	9100
Bright Angel R.S., Az	36E13'	112E04'	8400
Caliente, NV	37E37'	114E31'	4440
Ely, NV	39E17'	114E51'	6250
Jacks Peak Tel, NV	41E32'	116E01'	8420
McGill, Nv	39E24'	114E46'	6340
Oasis, Nv	41E02'	114E30'	5830
Page AZ	36E56'	112E27'	4270
Pole Creek RS, Tel Nv	41E52'	115E15'	8330
Seligman, Az	35E19'	112E53'	5250
Seventy-Six Ck Tel Nv	41E42'	115E28'	7100
Ward Mountain, Tel #2 Nv	39E08'	114E49'	9200
Wupatki NM, Az	35E31'	111E22'	4908
Target Sites			
Alton	37E26'	112E29'	7040
Beaver Dams	39E08'	111E33'	8000
Big Flat	38E18'	112E21'	10290
Black Fl. UM Ck.	38E41'	111E36'	9400
Box Creek	38E30'	112E02'	9300
Bryce Can. NP Hdq.	37E39'	112E10'	7915
Buck Flat	39E08'	111E27'	9800
Castle Valley	37E40'	112E44'	9580
Dills Camp	39E02'	111E28'	9200
Electric Lake UP&L	39E36'	111E13'	8380
Fairview 8N	39E45'	111E25'	6750
Farnsworth Lake	38E46'	111E40'	9600
Gooseberry R.S.	38E48'	111E41'	7920
Hatch	37E39'	112E26'	6910
Kimberly Mine	38E29'	112E23'	9300
Kolob	37E32'	113E03'	9250
Little Grassy Ck.	37E29'	113E51'	6100
Long Flat	37E30'	113E25'	8000
Mammoth-Cottonwood	39E41'	111E19'	8800
Merchant Valley	38E18'	112E26'	8750
Midway Valley	37E34'	112E50'	9800
Pickle Keg Spring	39E02'	111E35'	9600
Pine Creek	38E53'	112E15'	8800
Red Pine Ridge	39E27'	111E16'	9200
Scofield-Skyland Mine	39E41'	111E12'	8710
Seeley Ck. R.S.	39E19'	111E26'	10000
Webster Flat	37E35'	112E54'	9200
Widtsoe-Esc. # 3	37E50'	111E53'	9500

Figure 2: Target and Control Sites

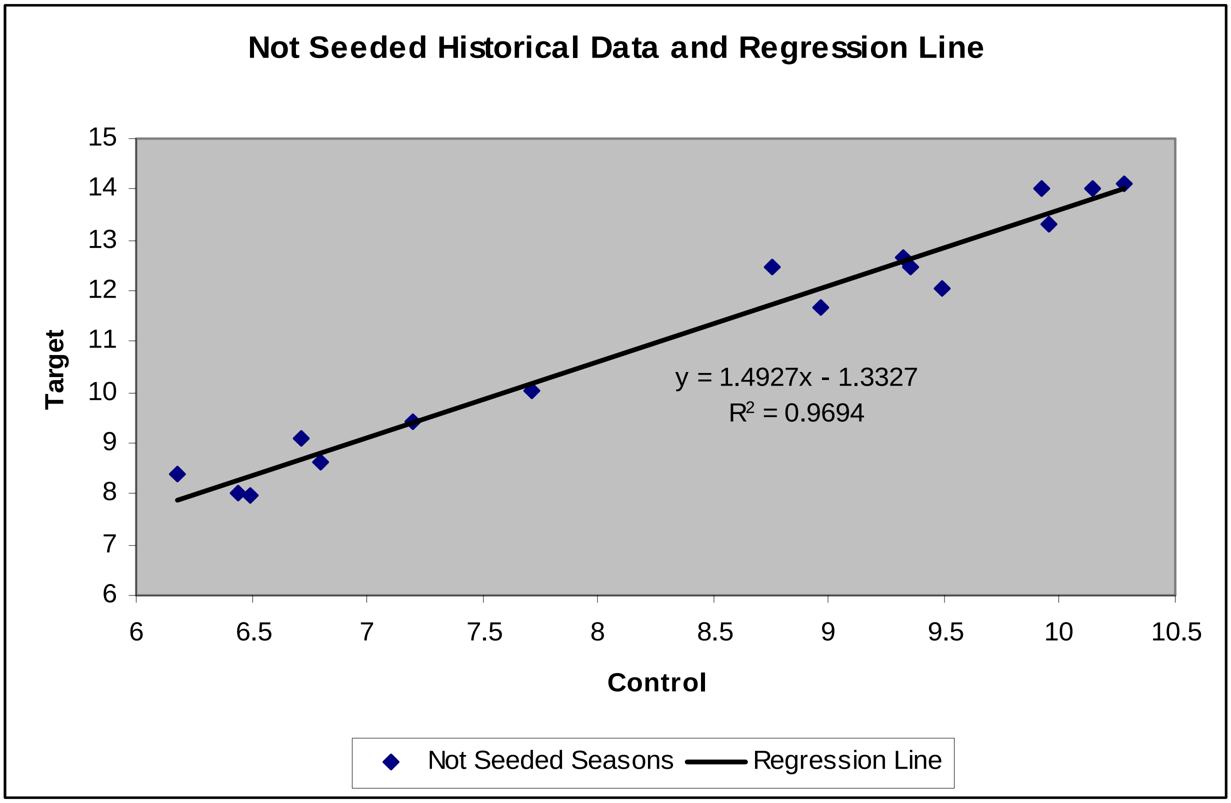


Figure 3: Regression Plot for Not Seeded Seasons

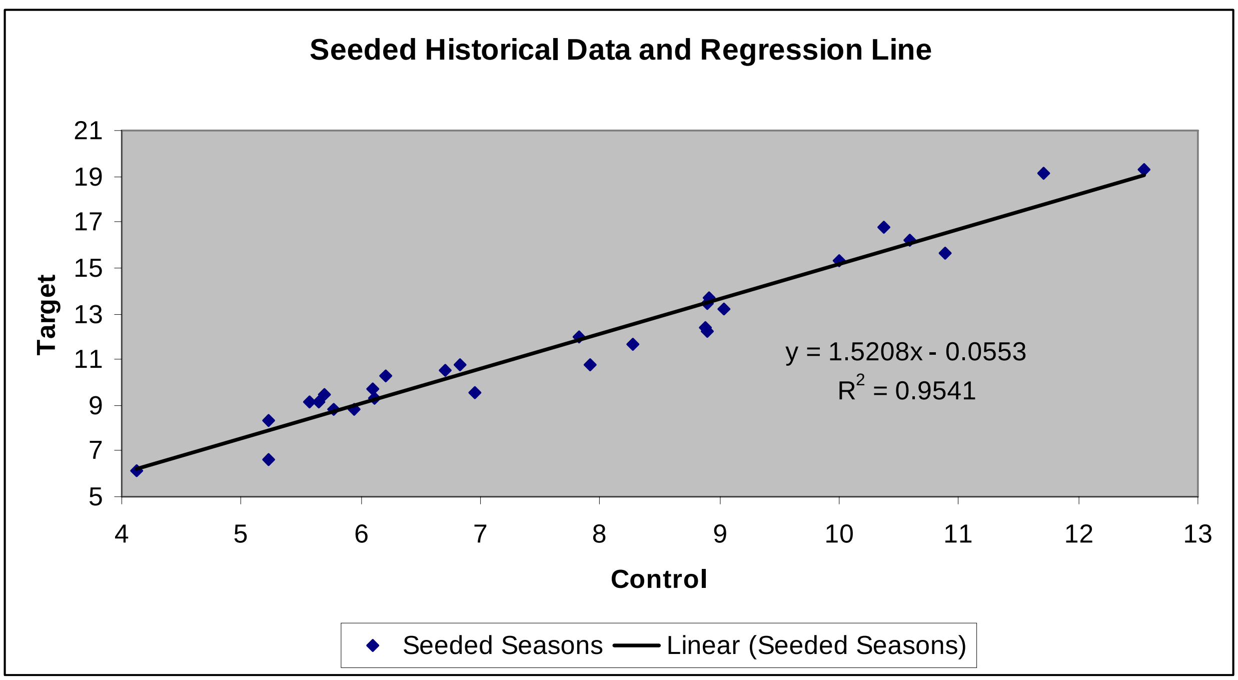


Figure 4: Regression Plot for Seeded Seasons

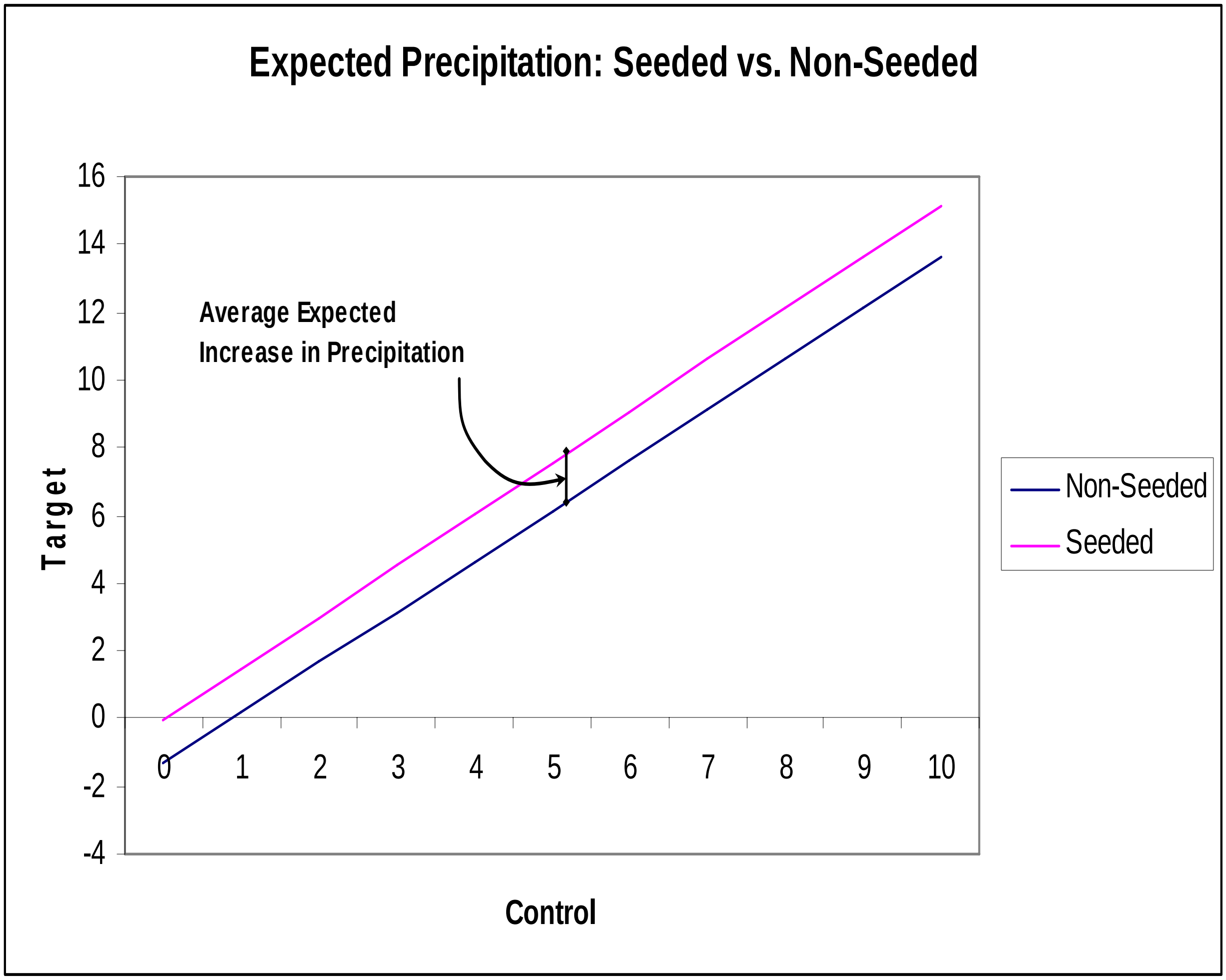
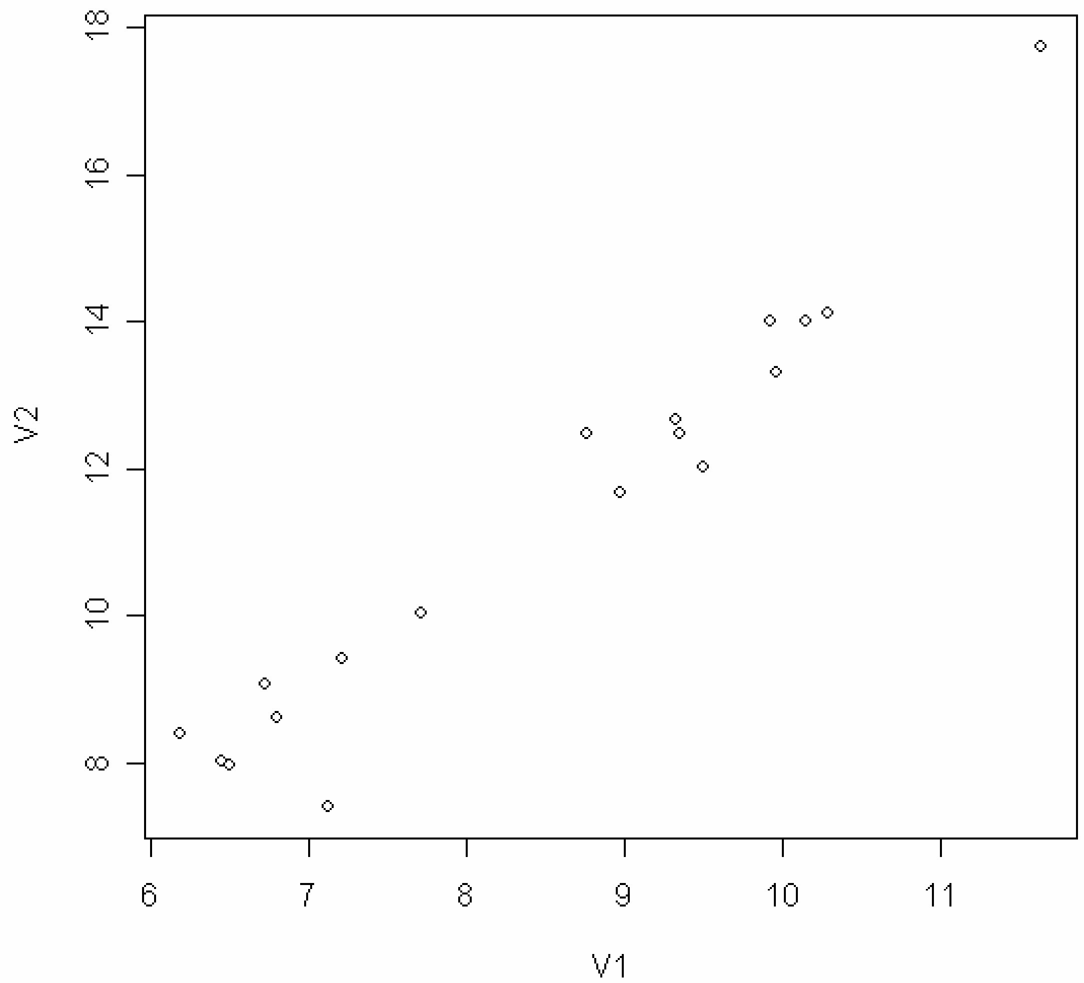


Figure 5: Expected Increase in Precipitation



<u>Figure 6:</u> Historical data plot before excluding outliers.

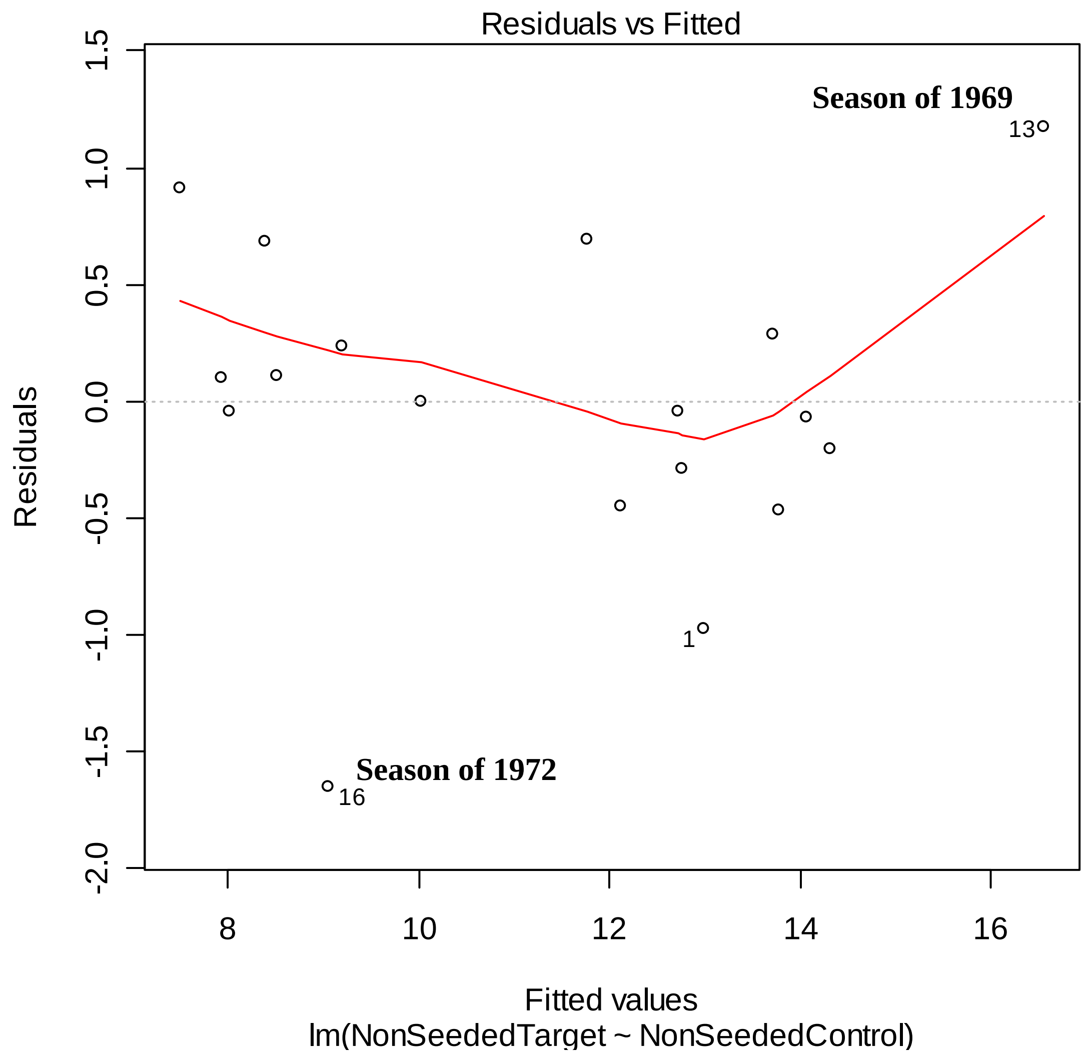


Figure 7: Residual technique for identifying outliers; Historical data set.

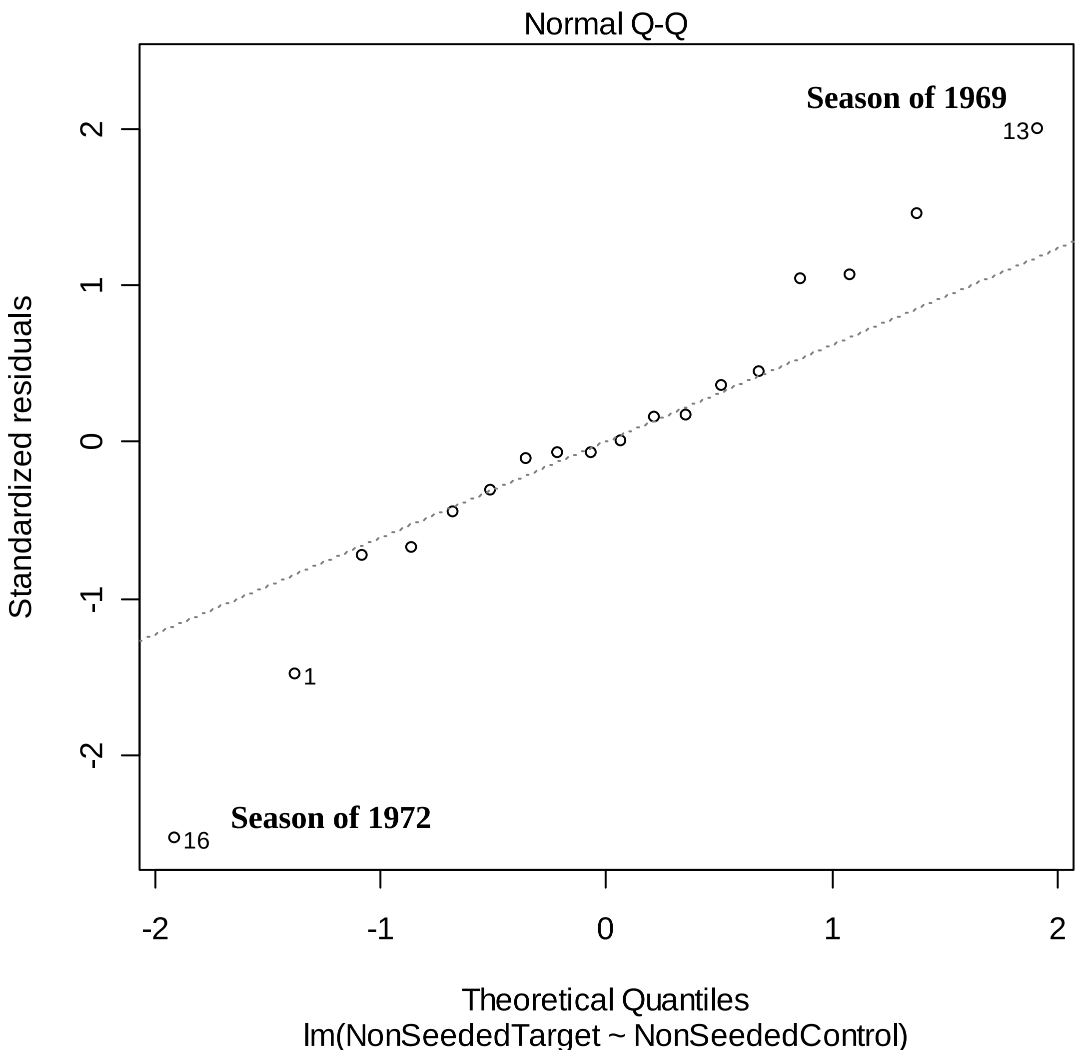


Figure 8: Quartiles technique for identifying outliers; Historical data set.

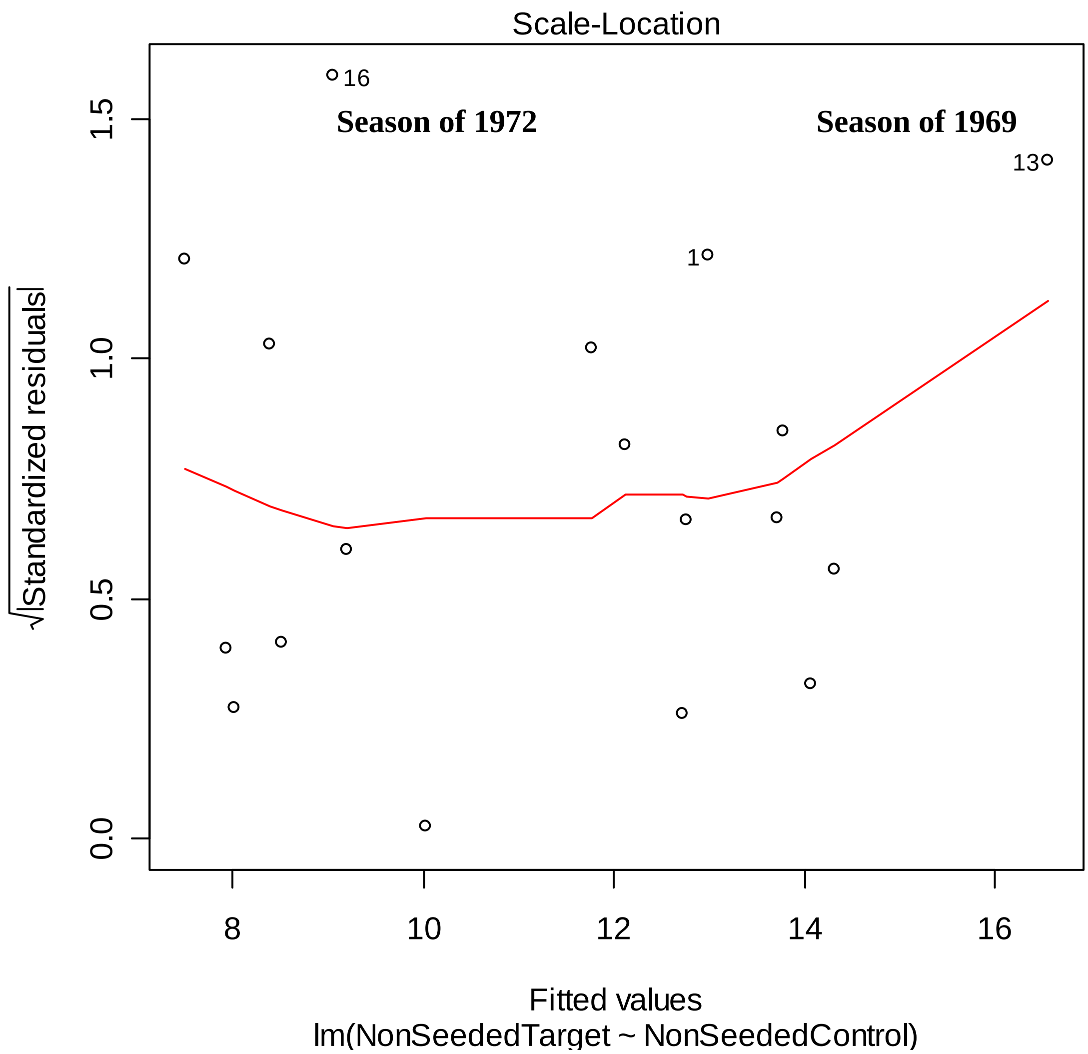


Figure 9: Scaled-Values technique for identifying outliers; Historical data set.

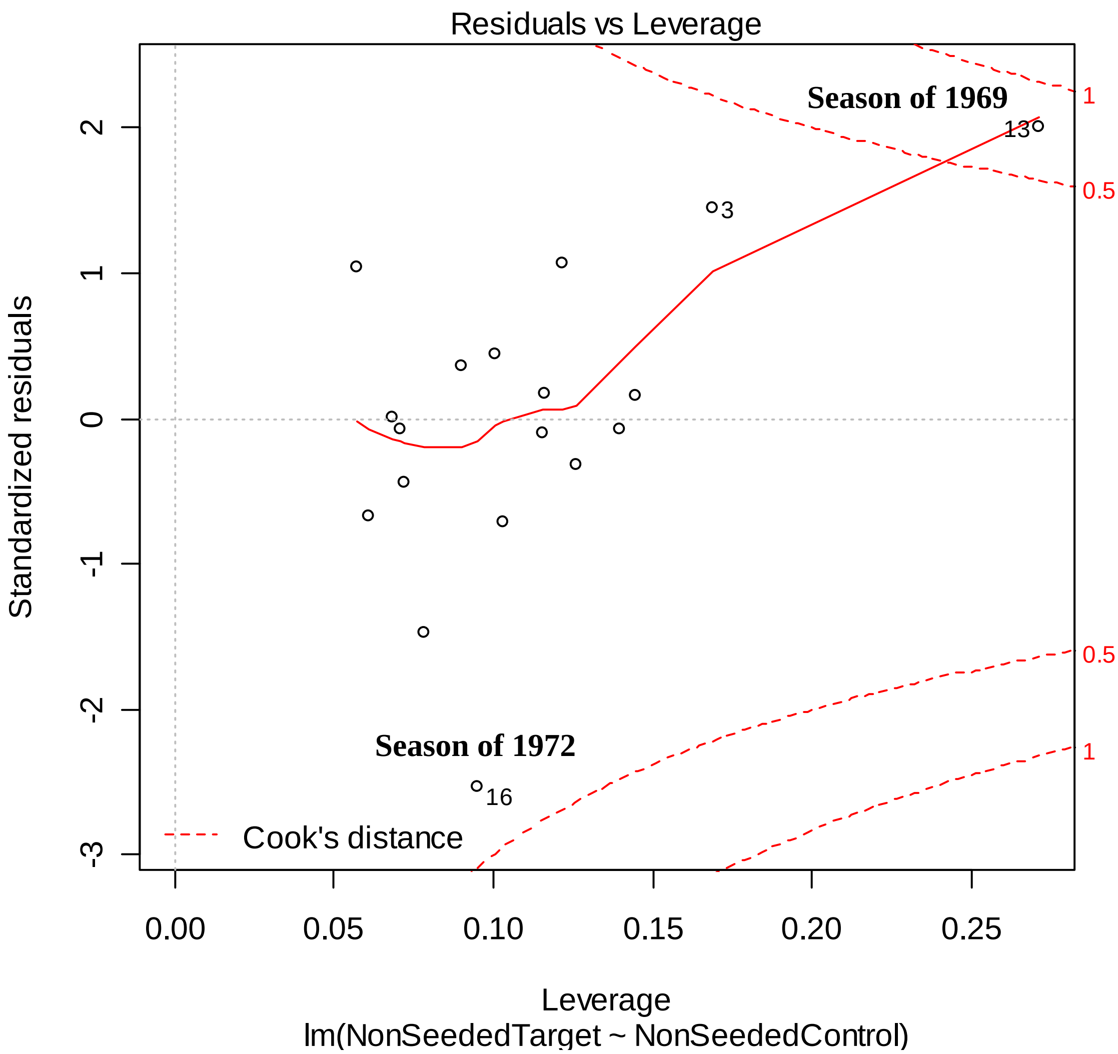


Figure 10: Leverage technique for identifying outliers; Historical data set.

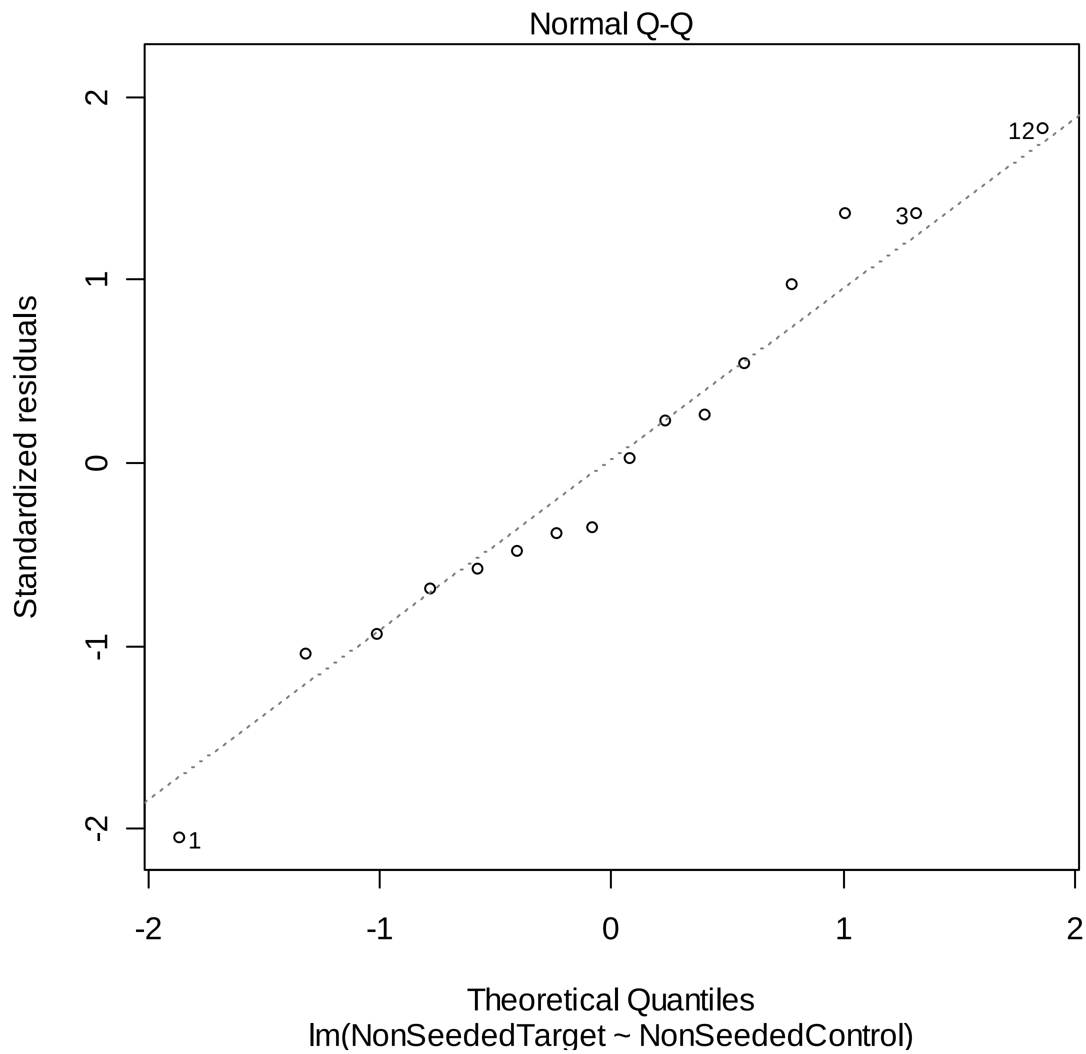
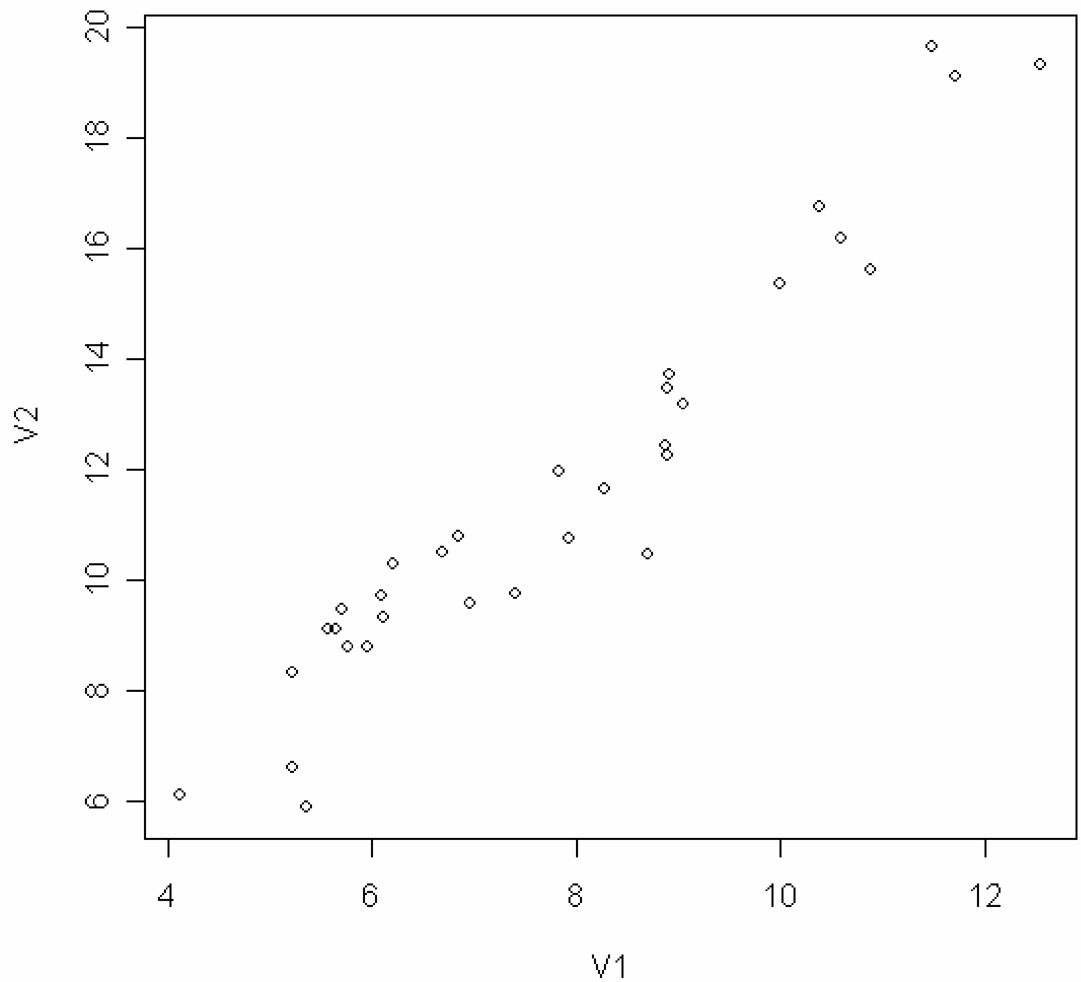


Figure 11: Q-Q plot for historical data after excluding outliers.



<u>Figure 12:</u> Seeded data plot before excluding outliers.

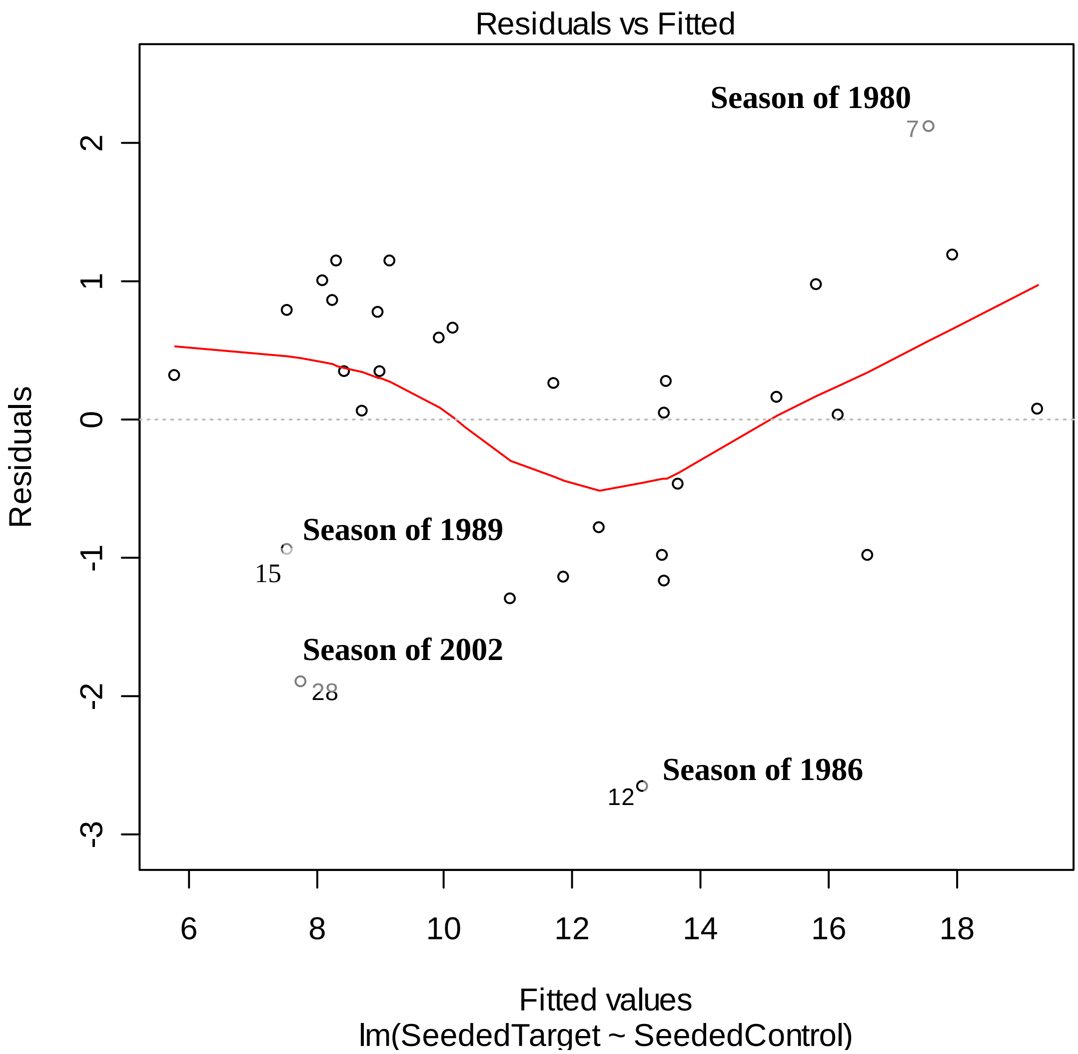


Figure 13: Residual technique for identifying outliers; Seeded data set.

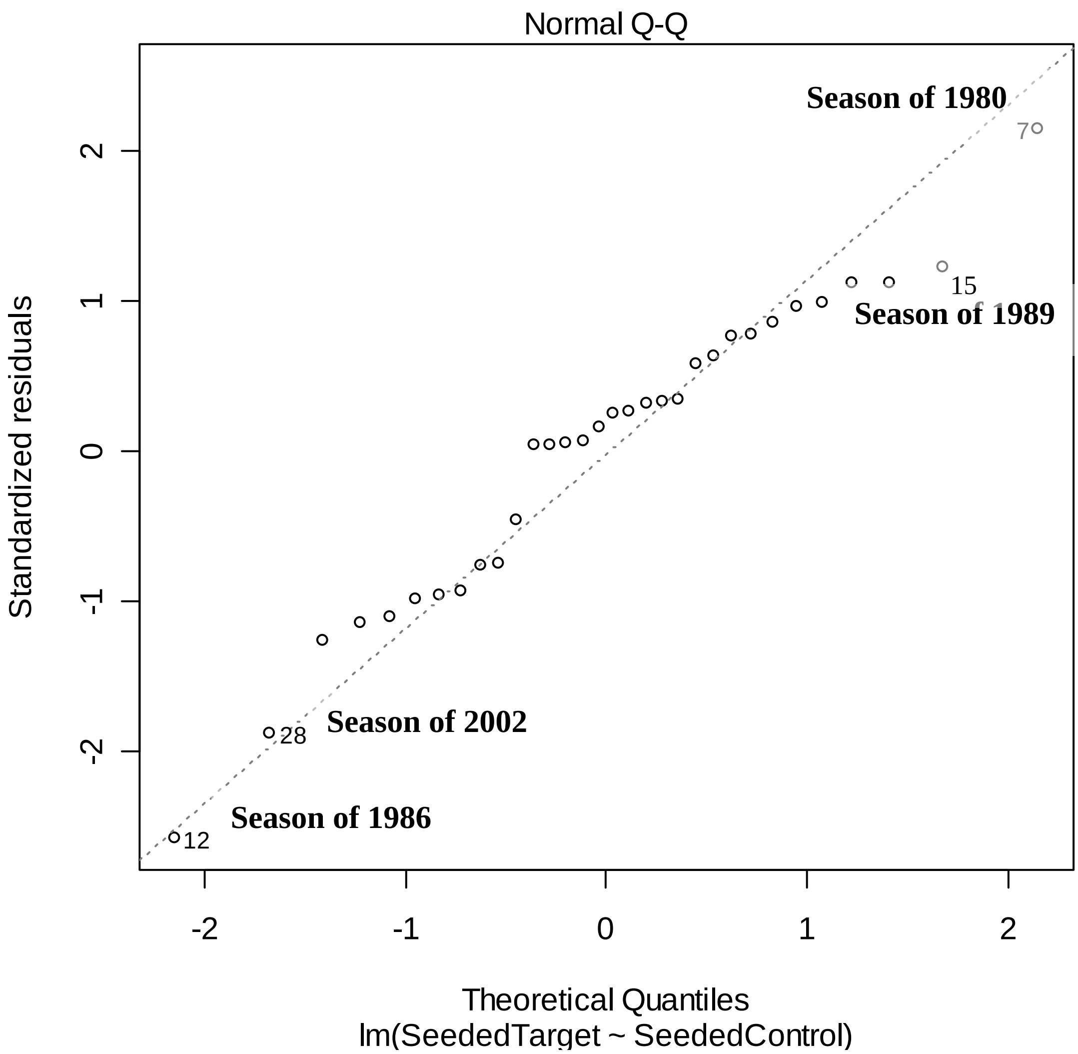


Figure 14: Quartiles technique for identifying outliers; Seeded data set.

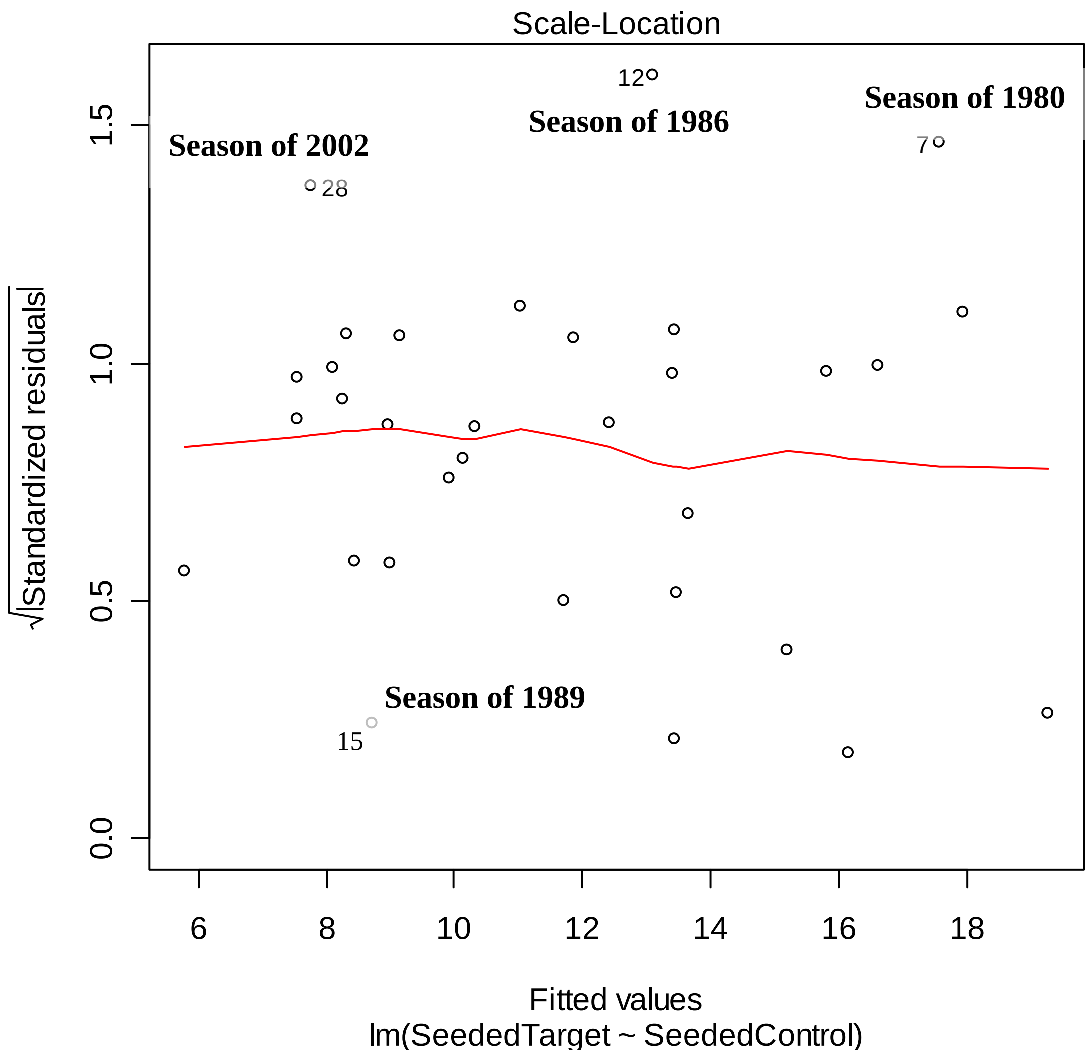
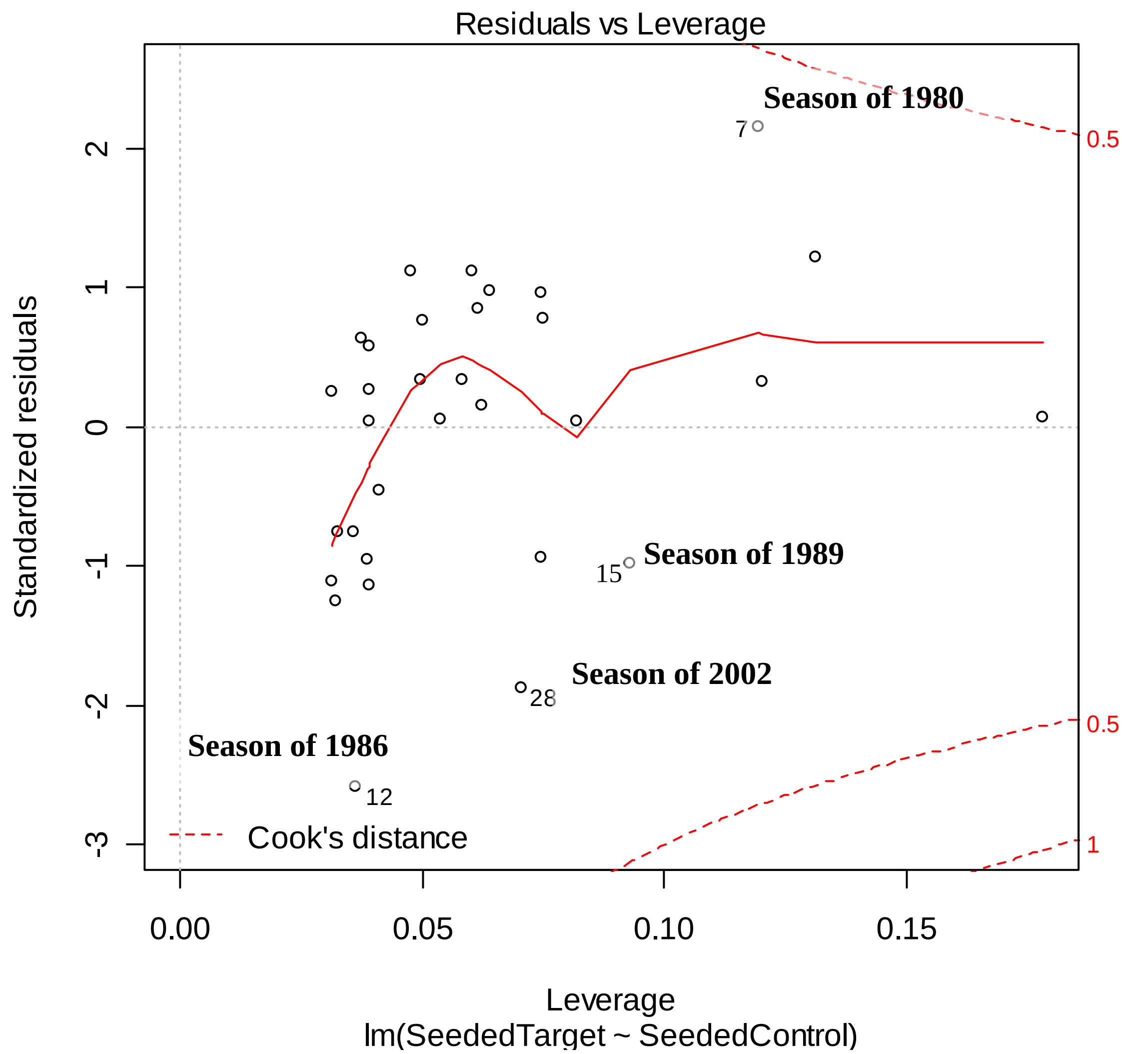


Figure 15: Scaled-Values technique for identifying outliers; seeded data set.



<u>Figure 16:</u> Leverage technique for identifying outliers; seeded data set.

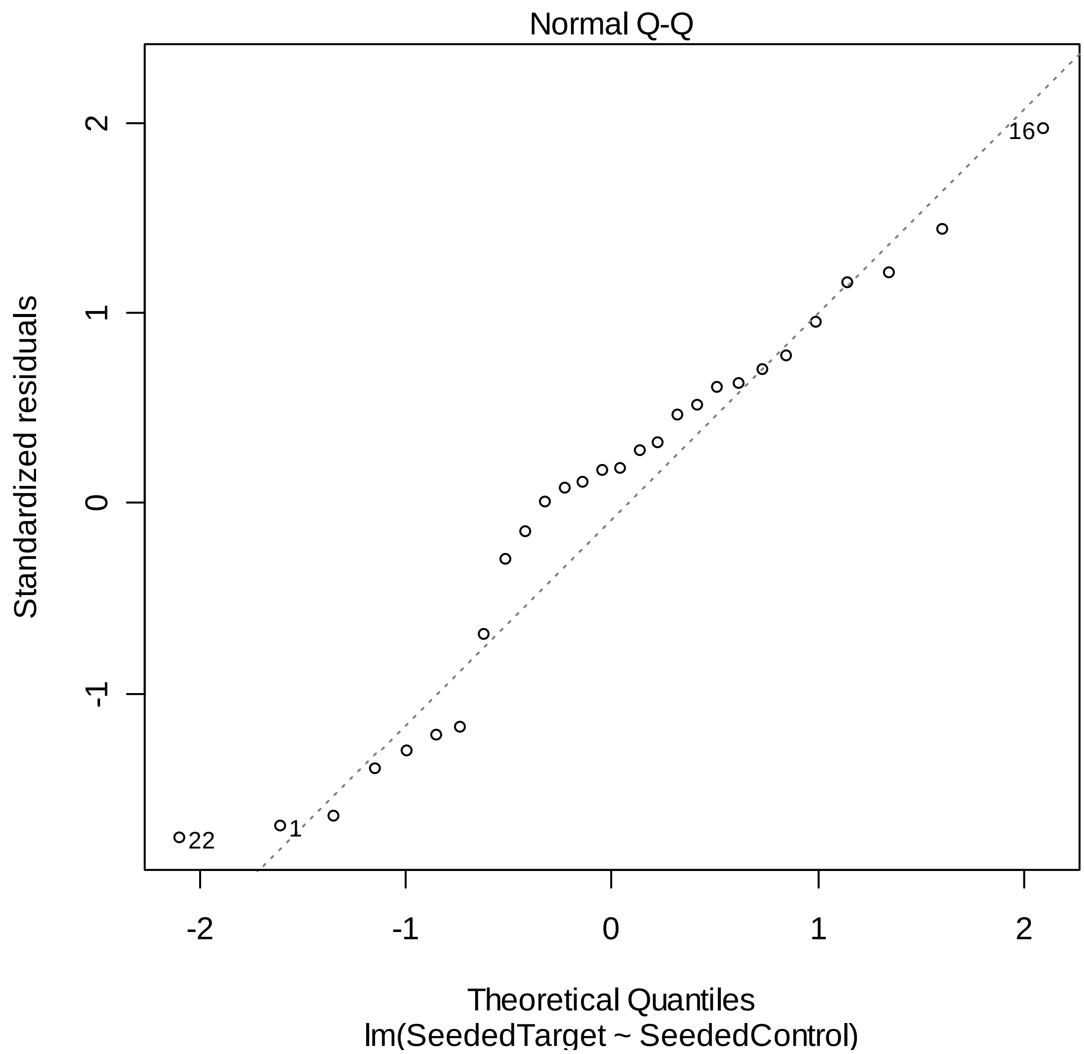


Figure 17: Q-Q plot for seeded data after excluding outliers.