

**Precipitation Enhancement Feasibility Study
in Aid of Cotton Irrigation**

**Report to DPIE Cotton Research and Development Corporation
Narrabri, NSW
Project Number : CSA1C**

Lynda E. Chambers and Alexis B. Long

July, 1992

9/1991 to 8/1992.



Contents

1	INTRODUCTION	1
2	THE GAUGE SYSTEM	1
3	RAINDAYS	2
4	RAINFALL	3
4.1	RAINFALL CORRELATIONS	4
5	SELECTION OF CONTROL AND TARGET GROUP STATIONS	8
6	STATISTICAL SIMULATION OF THE CLOUD SEEDING EXPERIMENT USING HISTORICAL PRECIPITATION DATA	10
7	CLOUD CONDITIONS	13
8	CONCLUSION	17
	APPENDIX 1 - Gauge Locations	19
	APPENDIX 2 - Explanation of Terms and Definitions	23
	APPENDIX 3 - Regression Estimates	25
	APPENDIX 4 - The Double Ratio	26
	APPENDIX 5 - Program Used in Simulations	27
	APPENDIX 6 - Summary of Steps Taken in Simulations	31

1 INTRODUCTION

The following report summarises the work from 1st July 1991 investigating the statistical feasibility of cloud seeding in the Copeton Dam region of NSW. The analysis covers a number of issues concerned with the number of days for which it is expected rain is to fall, the amount of rain expected to fall, cloud and wind conditions on rainy days, and the likelihood of detecting prescribed cloud seeding increases in rainfall.

The report is divided into a number of sections. The first section deals with the initial selection of the gauge sites. The second section looks at the analysis of raindays (a definition is given at the beginning of the raindays section). We then look at the rainfall amounts in and around the Copeton Dam region as well as the correlations of rainfall amounts between stations. Following sections discuss the breakdown of raindays according to wind direction and speed, the selection of control and target areas and their relationships to each other, simulated seeding experiments based on historical data, the detection of seeding increases in rainfall, and the determination of cloud conditions on rainy days.

The report below indicates that there is a good potential for cloud seeding in the Copeton Dam region. Should cloud seeding go ahead there are a number of individual gauges and groups of gauges which, in their present positions, could act as very good predictors of the rainfall in the Copeton Dam region. This is essential for a good cloud seeding experiment. The odds of detecting a 30 percent increase in precipitation are about 4:1 in a 5-6 year experiment and for a 20 percent increase about 2:1.

2 THE GAUGE SYSTEM

The first step in the analysis of rainfall in the Copeton region and its surrounds was to look at the present gauge system. Most of the gauges are privately operated by farmers who send their data returns to the Bureau of Meteorology. A list of gauges, with their latitudes and longitudes, was supplied by the Bureau of Meteorology. These gauge sites were then plotted and a circle of 100 km radius, from the Copeton Dam, was drawn to include the stations of interest. As it turns out this size circle included enough gauges for the simulation of the seeding effect. The list contained 549 gauges within 100km of the Copeton Dam. However, after removing the gauges with no records for the last 20 years the number of gauges remaining is 263.

Gauge stations were then eliminated if there was a high number of years with no data (at least seven missing years over a 30 year period) or if the stations were discontinued over three years ago. Stations which were missing the last year or two were marked and later found (through the Bureau of Meteorology) to be due mainly to slow returns of data, although one station was found to have closed in this period. All these gauges were saved. There were then 156 stations at this stage.

A list was then made of all the stations recording the number of incomplete months in five year periods from 1971 to 1991. When this number was particularly high, especially

within 1986-1991, these stations were excluded from the study. Since these stations had shown themselves to be unreliable in recent years and, hence, would be unlikely to be of any use if cloud seeding is to go ahead. These excluded stations typically had twenty or more missing months in each five year period. This reduced the number of stations in the study to 132. The latitudes, longitudes and names of these stations are given in Appendix 1. Figure 1 is a map of stations used throughout most of the analysis. Figure 2 is a map of some of the stations which were found to be less reliable later in the studies and were consequently discarded.

Because new gauges in regions where it commonly snows would need to be heated, in a prospective seeding experiment, all stations were categorised according to whether records showed it snowed (on average) :

- less than half a day per year
- half to one day per year
- at least one day per year.

These were then plotted on a map of the area, Figure 3, to give an overall view of the regions with no, low or high snow falls.

A more detailed discussion of the adequacy of the existing Bureau of Meteorology gauges in the Copeton Dam region is given in section 5 once the final selection of gauges for the experiment is complete.

3 RAINDAYS

A rainday is defined as being a day on which at least 1mm of rain had fallen. Using the stations selected above the number of raindays in each month of each year for each station was recorded. Boxplots (see Appendix 2 for a complete description of how the boxplots were constructed and what it all means) were then used to give a quick idea of the distribution of the number of raindays for each station. Nevertheless, means, standard deviations, minimums, maximums and quantiles for the number of raindays were also recorded. (These terms are defined in Appendix 2) These boxplots are too numerous to shown in this report.

The expected number of raindays in each month and in a year were calculated for the groups of stations which are found in section 4.1. For the Copeton Dam region the expected number of raindays in each month are :

April	May	June	July	August	September	October
5.42	6.17	5.80	6.73	6.75	6.01	8.07

The expected number of raindays in a year in the Copeton region is 45. Areas immediately surrounding the Copeton Dam region had a similar number of raindays while areas to the north-west had fewer raindays, high twenties to low forties, and areas to the east and south-east had more raindays in a year, mid forties to mid sixties.

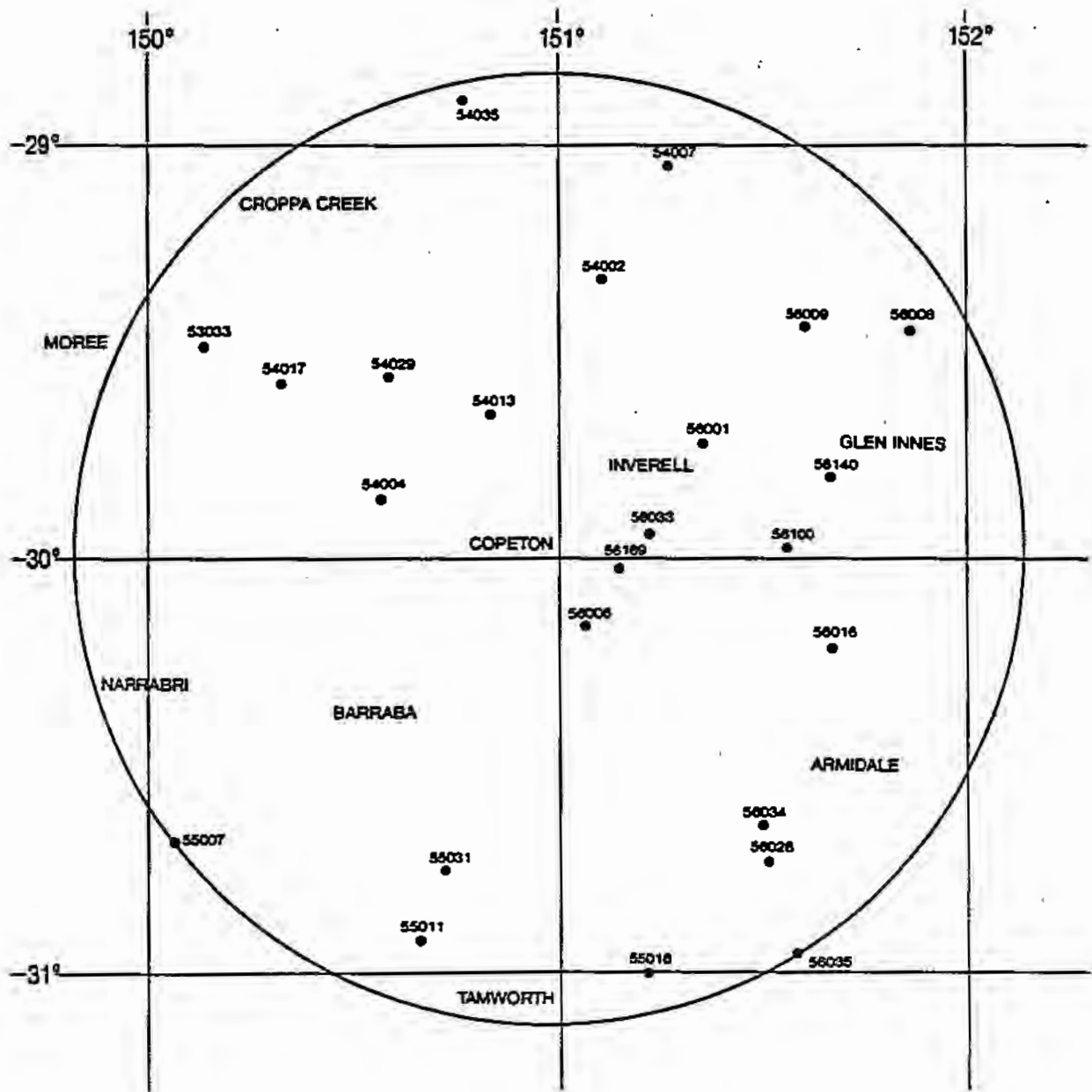


FIGURE 2: THE LOCATIONS OF SOME OF THE GAUGE STATIONS ELIMINATED DURING THE ANALYSIS

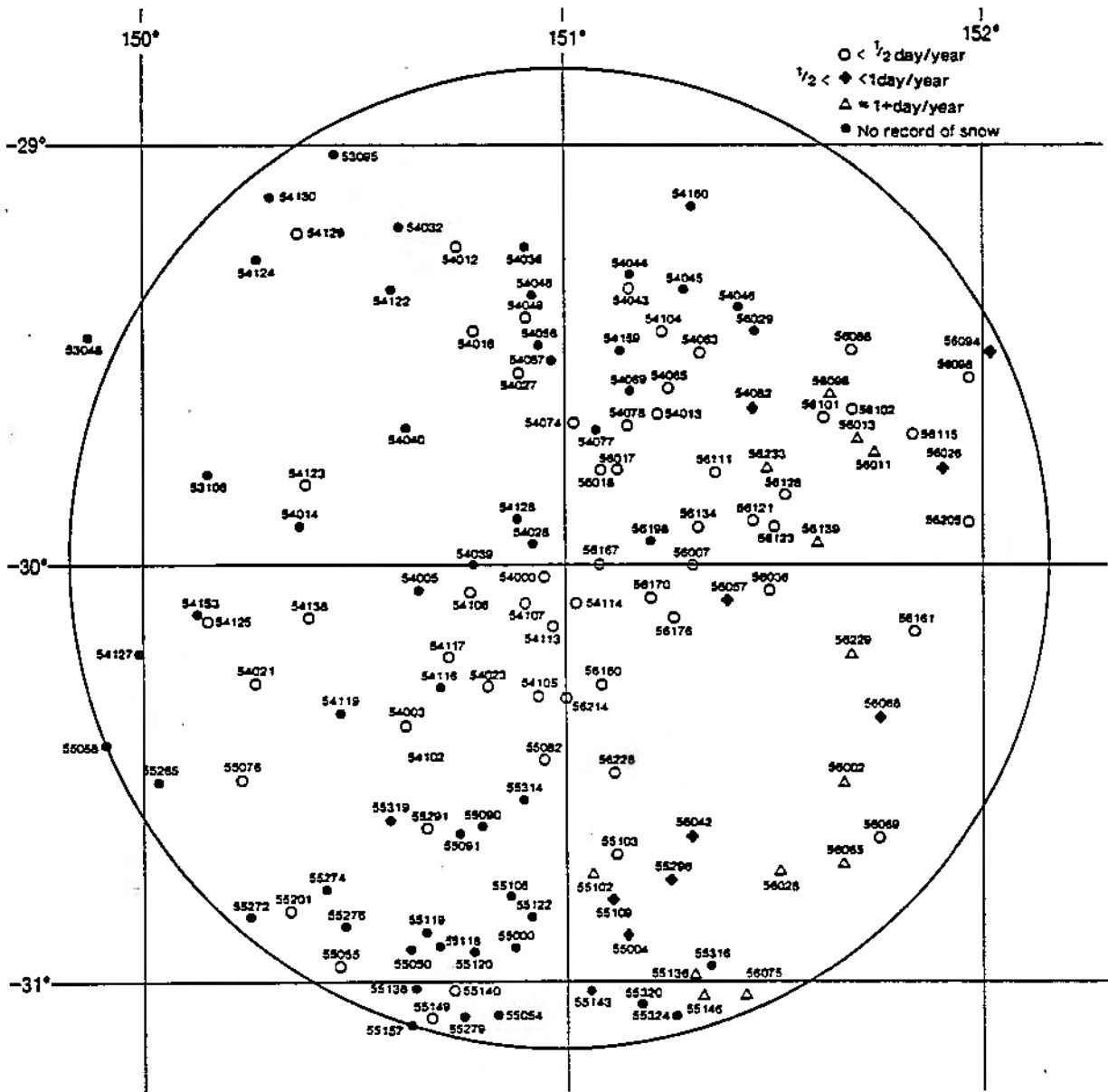


FIGURE 3: THE DISTRIBUTION OF SNOWFALL AMOUNTS IN AND AROUND THE COPETON DAM REGION

4 RAINFALL

To obtain a quick overview of the amount of rainfall in and around the Copeton Dam region monthly and yearly summaries were made. Boxplots (see Definition in Appendix 2) were made and means, standard deviations and quantiles were recorded for both the months and the year.

Some stations had periods when the gauges were not checked daily so it is not certain on which day(s) the rain fell. The period for which the gauge was not checked is known as the accumulated period and the total rainfall for that period as the accumulated rainfall. For accumulated rainfall days we averaged the total rainfall over the accumulated period. The days with dew or frost were coded as zero rainfall and days when it snowed were treated as rainfall but notes made of when snow occurred. Notes were also made of stations which had many accumulated rainfall periods and, when this number was exceedingly high, with most months containing accumulated rainfall data, these stations were not used. When there are large numbers of accumulated periods the actual pattern of rainfall is distorted. Instead of knowing when it rained and how much rain fell on a particular day we only know how much rain fell over a given period often several days long. This results in a loss of information. There were four stations with high numbers of accumulated periods. Almost all of the other stations had at least one period of accumulated rainfall. There were 128 stations remaining at this stage.

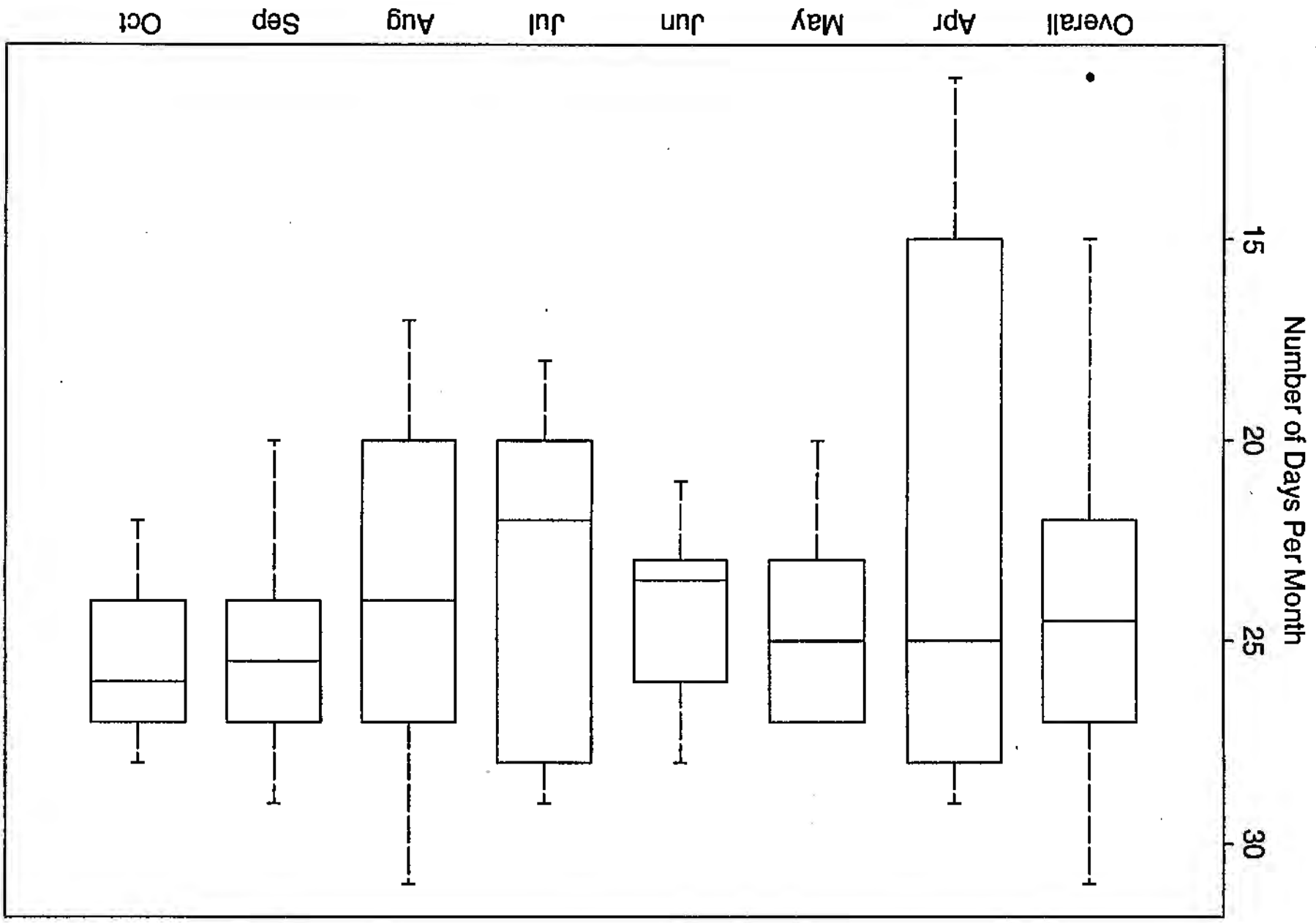
Once an overall idea of the number of raindays per month, and the amount of rain each month, at each station were obtained the daily rainfall data were investigated. The data were grouped according to the amounts of rainfall per day. The rainfall categories used were :

0	mm
(0,.5]	mm
(.5,1]	mm
(1, 2]	mm
(2, 5]	mm
(5,10]	mm
over 10	mm

The analysis of daily rainfall was carried out as follows. For each of the categories mentioned above the mean, standard deviation, minimum, maximum and quantiles were found for each month and for an overall average of the months. This was carried out for all stations. Most of the stations seemed to show similar patterns, eg peaks in rainfall at the same times, but this is investigated further through correlation analysis.

Figure 4 shows boxplots of the number of days in the 0mm rainfall category (ie which had no rainfall) for station 54153. This is an example of what some of the boxplots looked like for the various stations. The "overall" boxplot is a boxplot using all the data for April to October, whereas the other boxplots use only the data of their respective months. The April boxplot is a little unusual in that there is a greater variation in the number of days without rain resulting in a longer box. The boxplots were constructed using 30 years of data from 1962 to 1991. (see Appendix 2)

Figure 4 : Boxplots of the Number of Days Without Rain, per Month, for Station 54153



4.1 RAINFALL CORRELATIONS

Correlations, for daily rainfall totals, were then calculated for groups of nearby stations to see the similarity, or lack of it, between neighbouring stations. To do this accumulated raindays were averaged (as before), days with dew or frost counted as zero rainfall and snow treated as rainfall. All correlations were greater than .4000 but with most (72%) being greater than .7000.

Scatterplots of the station groups were also used to identify any unusual points that could have an influence on the correlation coefficient or any unusual patterns in the data. For example, we would not expect a station to have a large rainfall if another nearby station had no rain on the same day. This type of problem shows up quite clearly in a scatterplot. Figure 5 is an example of a typical scatterplot. The rainfall for each day for one station is plotted against the same day's rainfall at another station. If the stations have similar rainfall patterns then the points in a scatterplot should lie close to a straight line. The scatterplot of stations 54077 and 54078 shows an example of an event we would not expect for stations which are fairly close together. There were a few occasions when station 54077 showed little or no rainfall when station 54078 had quite high rainfall. It is not possible to say why this occurred due to a lack of information in the data set supplied by the Bureau of Meteorology. It could be due to a large isolated rainfall only affecting station 54078, but I do not feel that this is the reason. There may have been an error in the recording of either the amount of rain that fell, or the day on which it fell, or some sort of data error once the data were recorded on the computer. However, this type of problem did not occur all that often.

Because the averaging of accumulated rainfall could have an effect on the resulting correlations two groups of stations were used to test for such an effect. For these groups the accumulated days were removed and the correlations recalculated. However, the correlations obtained after the removal of the accumulated data points were very similar to the original correlations only slightly higher. This means that there is little difference in using accumulated data over data with the accumulated rainfalls removed. The results are given below, for station to station correlations :

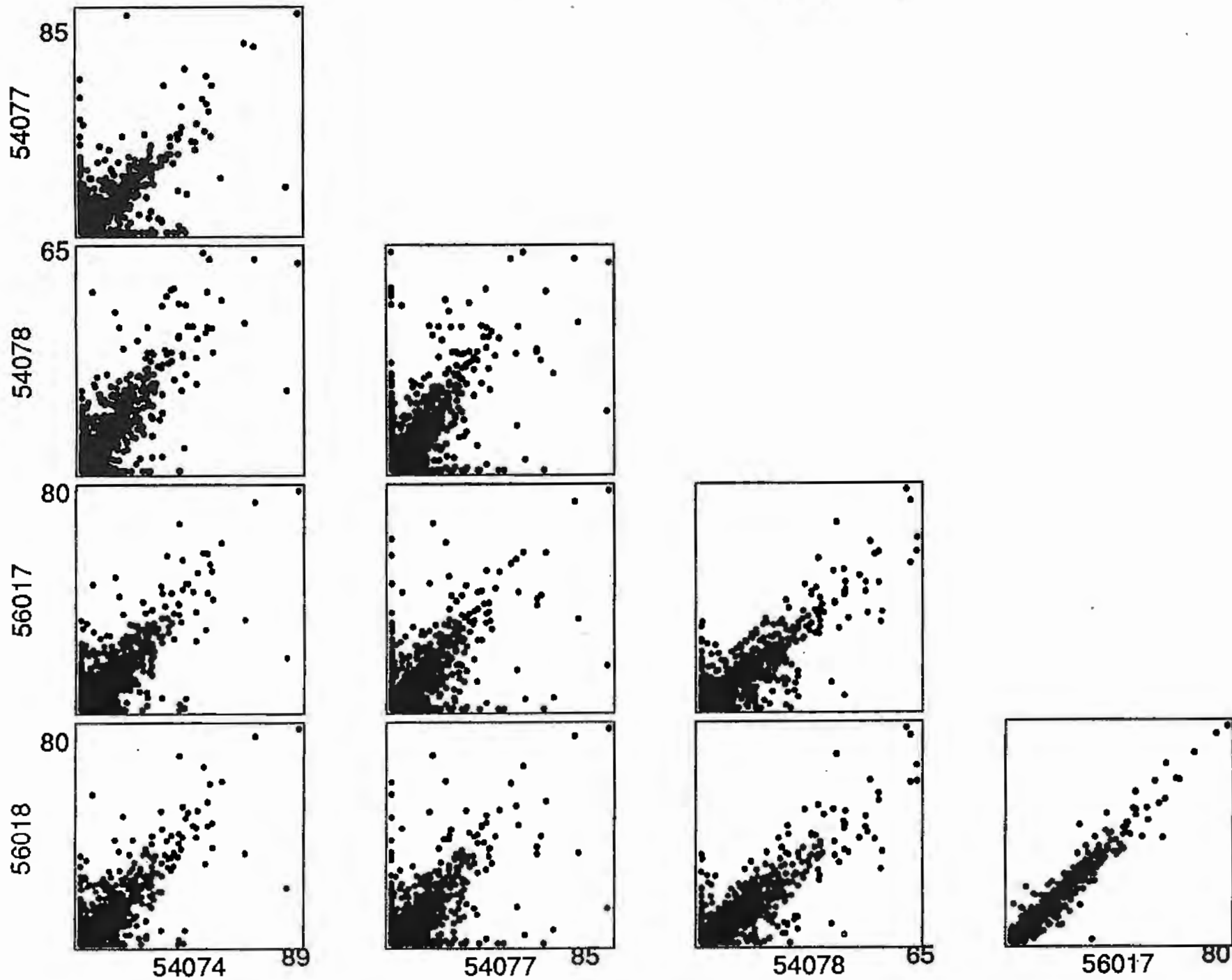
WITH ACCUMULATED RAINFALLS

	53048	53095	53108
53048			
53095	.7177		
53108	.6321	.5336	

WITHOUT ACCUMULATED RAINFALLS

	53048	53095	53108
53048			
53095	.7180		
53108	.6360	.5353	

Figure 5 : Scatterplots of daily rainfall (mm)



WITH ACCUMULATED RAINFALLS

	56057	56176	56170	56007	56134	56121	56036
56057							
56176	.7720						
56170	.7673	.8406					
56007	.6609	.6592	.6841				
56134	.7995	.7684	.7971	.6864			
56121	.6787	.6606	.6828	.6348	.7523		
56036	.7047	.6814	.6855	.5849	.6938	.6615	

WITHOUT ACCUMULATED RAINFALLS

	56057	56176	56170	56007	56134	56121	56036
56057							
56176	.7733						
56170	.7685	.8413					
56007	.6605	.6611	.6832				
56134	.8017	.7699	.7995	.6909			
56121	.6825	.6614	.6835	.6345	.7537		
56036	.7121	.6839	.6888	.5857	.6977	.6661	

Stations with station-to-station correlations greater than 0.8 were combined to form a new group to reduce the amount of analysis needed since little information would be lost. The combined stations had their daily rainfall totals averaged to give an area average. The stations in the Copeton Dam Group (see also Figure 1) with inter-station correlations greater than 0.8 were :

54000
 54005
 54028
 54106
 54107
 54128

The Copeton Dam group average was then compared to all the other area averages and to various other stations which were not highly correlated with other stations elsewhere in search of individual stations or groups of stations which would be good predictors of rainfall in the Copeton Dam region. The comparison yielded the correlations shown below.

STATION GROUP OR
INDIVIDUAL STATION

CORRELATION
WITH COPETON
REGION

NUMBER OF *
OBSERVATIONS

54039	.6443	4343
56167	.7666	4434
54113	.6583	4067
54125,54138	.8424	2935
54040	.8011	4006
54123	.7413	4373
53108	.5651	2049
54014	.6257	3762
54125,54021,54138,54119	.8509	2935
54127	.5530	3729
55058	.6162	4189
55076	.4958	4160
55265	.6793	3944
54117,54003,54102,54023,54105, 56180,56228,56314	.9052	2048
55090,55091,55291,55314	.8140	1100
55274,55276,55055	.7650	4434
55272	.5429	4066
55201	.5206	4434
55105,55122,55000,55120,55119,55118	.7652	4374
55050	.6215	4189
55138,55140,55157,55279,55054, 55120,55050	.7627	4158
55143,55136,55146,55316,56075	.7595	1223
55296,55102,55103,55109,55004	.7923	3057
56002,56065,56068,56229	.7879	1925
56111,56128,56123,56139	.8487	4249
56170,56176	.8763	3945
56057	.8154	4374
56069	.7030	4220
56161	.6121	3333
56007	.7237	4404
56134	.8322	4434
56121	.7270	4434
56036	.7746	4373
56205,56026,56115,56011,56098,56094	.7392	4098
56013,56102,56101,56096,56088,56011	.8067	4131
54044,54043,54045,54046	.8086	4404
54104,54063,54082,54065,54073, 54078,54069	.8520	3456

STATION GROUP OR INDIVIDUAL STATION	CORRELATION WITH COPETON REGION	NUMBER OF * OBSERVATIONS
54078,54074,56018,56017	.8991	4220
53095	.7159	3792
53048	.7693	4434
54032	.6922	4434
54122	.6842	4159
54130	.6871	4067
54129	.7394	4343
54124	.7419	4434
54102,54036,54048,54049,54027,54057	.8642	3882
54016	.7201	4281
54153	.6851	1284
54116	.7613	4403
56214	.8414	4434
55082	.6423	4434
54056	.7258	3762
55149	.4708	4434
56042	.7497	4068
54077	.7021	4434
56029	.6612	4404

* Number of Observations is the number of days used to calculate the correlation.

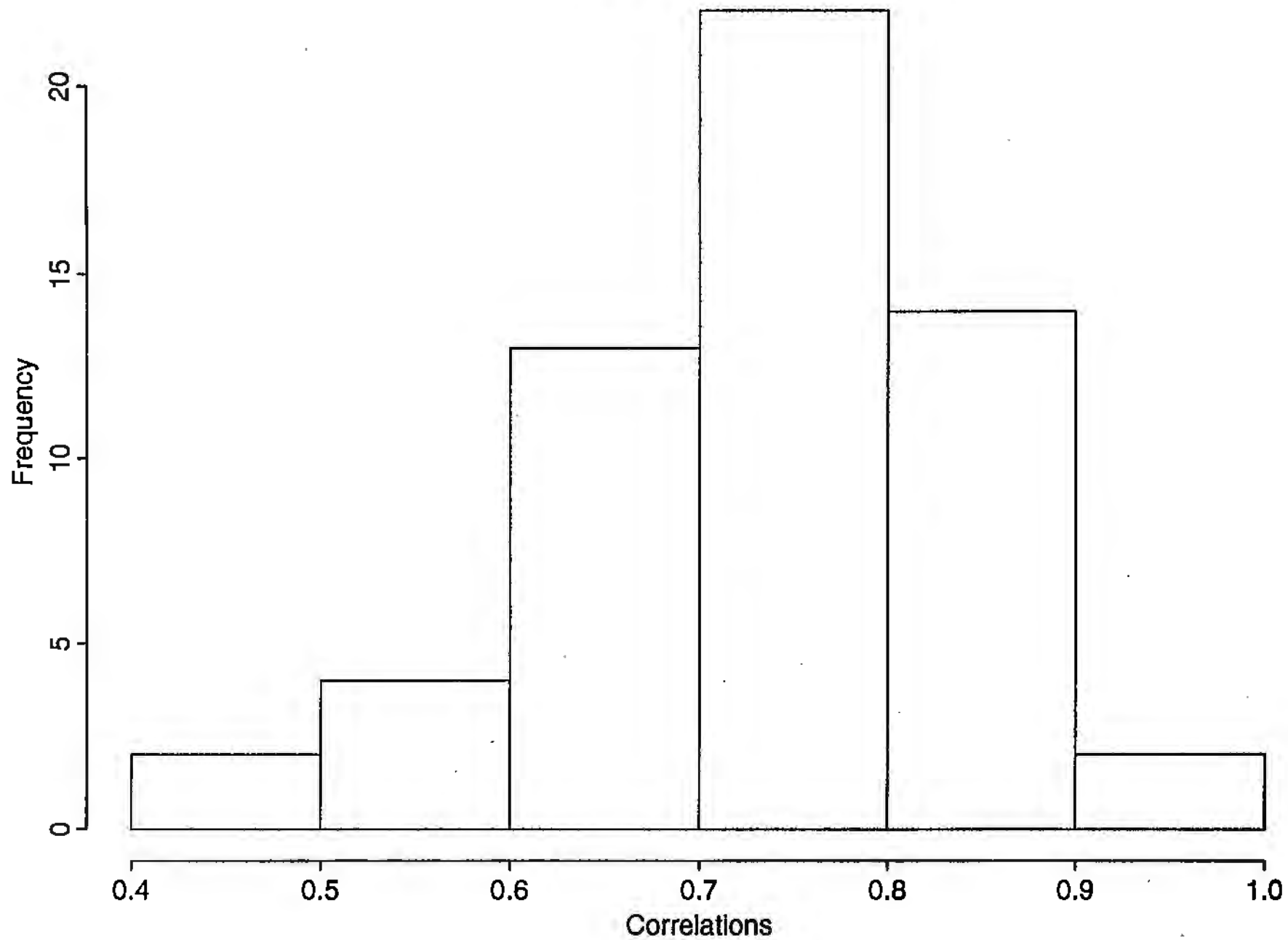
The reason for the variation in the number of days for the different groups is due to the number of days with no rainfall information, ie missing data, which varied from station to station.

Most of the correlations are fairly high with many over 0.7 (66%). Correlations of 0.7 and above are considered reasonable correlations for the prediction of rainfall in the target area (See The Thomson Catchment Cloud Seeding Experiment, Project Design, CSIRO Division of Atmospheric Research, July 1989). A histogram of the various correlations with the Copeton Dam region, from the above table, is given in Figure 6.

A map, Figure 7, was then drawn highlighting areas with correlations greater than 0.8, between 0.7 and 0.8 and those regions with correlations smaller than 0.7. Correlations of 0.8 and above indicate very good predictors of target area rainfall. The aim of the map is to highlight areas of high correlations which could be used to accurately predict the rainfall in the Copeton region. The use of several such "control" areas, which are highly correlated with the Copeton region, increases the power of the test for seeding effect and thus decreases the number of years of seeding needed to obtain the same power as a test without controls. Plotting the correlations also highlighted several stations which appear to be unreliable. For example, station 55082 is surrounded by stations with much higher correlations, casting doubt on the accuracy of its observations. Such bad stations were :

54039, 54113, 54116, 55082, 55050, 55149,
56167, 54077, 54056, 54016, 56029, 54153

Figure 6 : Histogram of Correlations with the Copeton Dam Group



5 SELECTION OF CONTROL AND TARGET GROUP STATIONS

In order to position the control areas correctly the winds at the 3000m MSL on raindays at the Copeton Dam in the period 1964 to 1988 were analyzed using rawinsonde data from Moree. The wind direction at 3000m MSL was considered as this is approximately the level where seeding would occur. It was found (see Figure 8) that approximately 45 percent of the rainfall occurred with winds from the southwest quadrant and 36 percent with winds from the northwest quadrant. Only about 19 percent of the rainfall occurred with winds with an easterly component. This indicates that the control regions for an experiment should be positioned to the west of the Copeton Dam region.

In positioning the control areas it was also necessary to consider how the wind speed varied with wind direction. Figure 9 is a summary of the findings for the Moree region using data from 1964 to 1988. The strongest winds were from the northwest and southwest quadrants and averaged 13 and 12 ms^{-1} respectively. For the eastern quadrants, wind speeds were only about 8 ms^{-1} . Given the average speed for westerly winds, and assuming seeding occurs 0.5-1 hour upwind of the target, seeding should on average occur around 20-50 km upwind of the target area. Seeding will therefore normally not occur upwind of the control area.

Using groups of stations with high correlations within the group five control regions were selected to the west of the target area (the Copeton Dam group given above). The positions of these groups are shown in Figure 10 and the stations within the groups are:

GROUP A : 54012,54027,54036,54048,54049,54057

GROUP B : 53095,54124,54129

GROUP C : 54040,54123

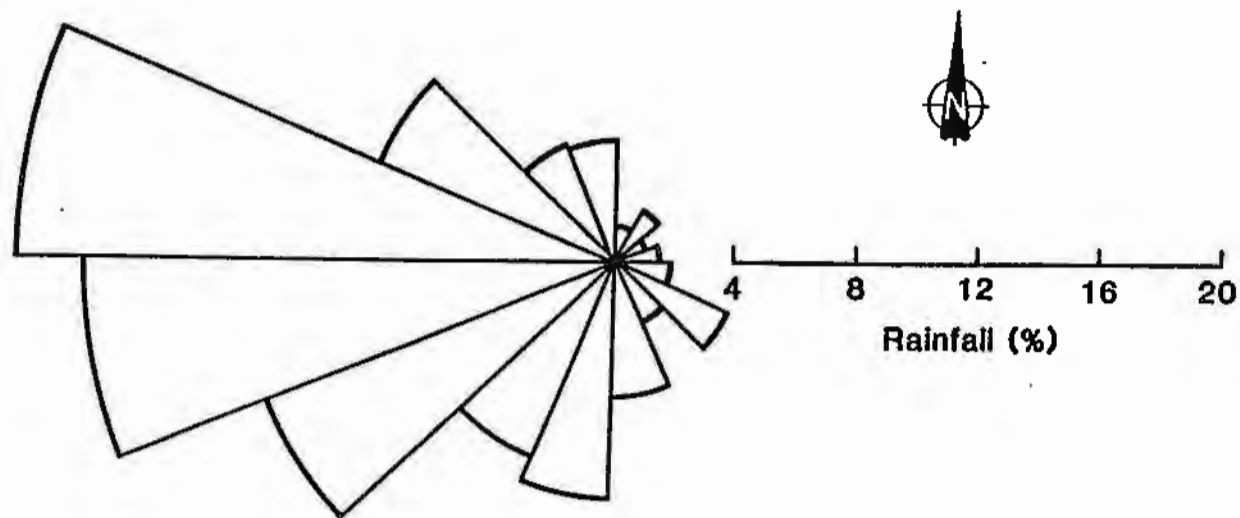
GROUP D : 54021,54119,54125,54138

GROUP E : 54003,54023,54102,54105,54117,55090,55091,55291,55314,56180,56214,56228

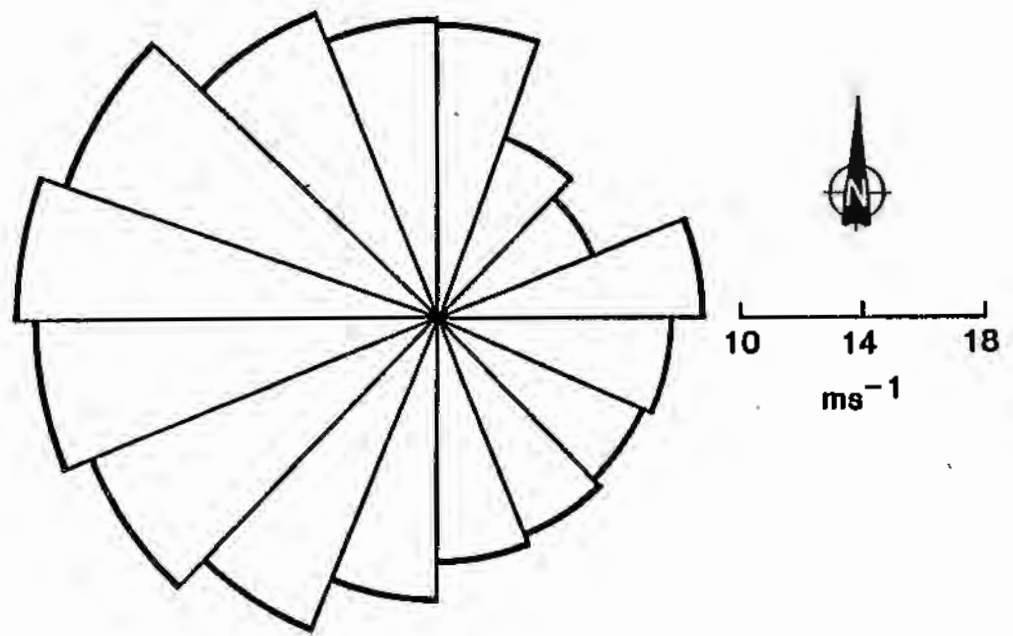
The following table gives the correlations of the control areas with the Copeton Dam (or target) Group.

	Control Area				
	A	B	C	D	E
All days	0.864	0.785	0.836	0.851	0.893
Raindays only	0.803	0.736	0.764	0.770	0.847
Raindays with winds from					
180° - 240°	0.780	0.728	0.762	0.803	0.821
250° - 300°	0.808	0.766	0.808	0.779	0.874
310° - 360°	0.765	0.632	0.595	0.634	0.866

Note the correlations are all fairly high for the case when all days are included and for the case when only raindays are used. The correlations are also fairly high for raindays classified according to wind direction when it is taken into account that only Group E is in the wind range 180° - 240°, Groups D and C are in the range 250° - 300° and Groups A and B are in 310° - 360°. Although the control area gauges do not cover the 250°-300° region very well the high correlations of the control area gauges with the target region indicates that these existing gauges will be adequate for a prospective cloud seeding



**FIGURE 8: PERCENTAGE OF TOTAL RAINFALL AT
COPETON DAM VS 3000 m MSL WIND DIRECTION
AT MOREE FOR 1964-1988 RAINDAYS**



**FIGURE 9: WIND SPEED VS WIND DIRECTION AT
3000 m MSL AT MOREE FOR 1964-1988 RAINDAYS**

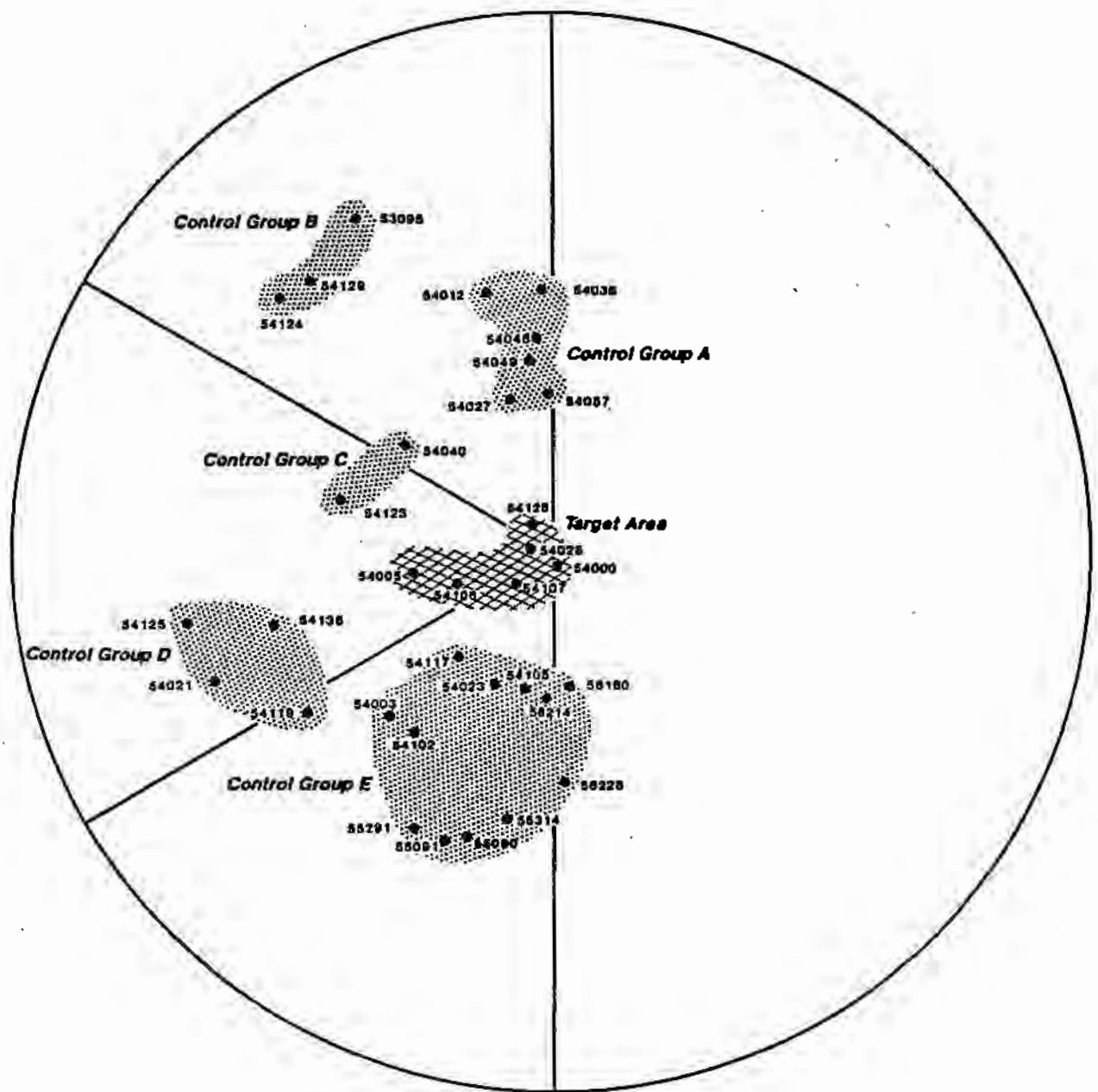


FIGURE 10: LOCATIONS OF CONTROL AND TARGET AREA RAIN GAUGE GROUPS

experiment (This is reinforced in the high chances of detection of a seeding effect found in the rainfall simulation study, see section 6). If the placement of new gauges in this region resulted in extra gauges with high correlations with the target area then the chance of detecting a seeding effect may be even higher.

Now that the target and control area groups have been selected we take a closer look at the adequacy of the existing Bureau of Meteorology gauges which make up these groups.

In the Copeton Dam (target) Group four of the gauges had been operating since 1964 with all gauges operational by 1969. There were very few periods when the gauges did not operate and very few months contained periods when the gauges were not checked daily.

All of the six gauges in Group A were operating from 1964 and few days of no records or days when the gauges were not checked occurred.

All three stations in Group B were operational by 1972 and stations in this group have very good daily records.

Both stations in Group C were operating by 1967, though station 54040 experienced four years of non-operation before 1972. After this period, however, the gauges have good records.

Group D had all four stations operational by 1974. However, station 54125 was fairly unreliable until 1984. Since then it has been extremely good. Otherwise most records are good.

All twelve stations of Group E were operational by 1985. After this time the records are good except for station 55291 which had 5 out of 7 months of data missing in 1990, and station 56180 which had 19 days of accumulated rainfall in the seven month period, April to October, of 1990.

Overall the Bureau of Meteorology gauges seem reliable enough to use in a potential cloud seeding experiment provided that the speed of returns from certain stations can be accelerated. These stations, with data three or months overdue, are :

Target (Copeton) gauges 54106, 54107

Group A 54057

Group C 54040, 54123

Group D 54021

Group E 54117, 55091, 55291, 56214

However the Bureau of Meteorology informs us that none of the gauges is heated, although snow falls are not uncommon in the region. All of the stations are ordinary raingauges except for the following stations which have pluviograph data :

54000 for 2 months in 1971

54036 up until December 1969

54102 from January 1971 to December 1981

54105 from January 1975 to December 1981

54138 from November 1976 to December 1980

6 STATISTICAL SIMULATION OF THE CLOUD SEEDING EXPERIMENT USING HISTORICAL PRECIPITATION DATA

This section describes the numerical simulations of the likely outcomes of a cloud seeding experiment over the Copeton Dam Region. The simulations use historical precipitation data from the Bureau of Meteorology for the period 1964 to 1988.

The data were for the April-October period for all of the stations named in the target and control areas. In addition to the use of the historical precipitation data information was also used on the 3000m MSL wind velocities at Moree.

The rainfall amounts for the target area and for each of the controls were averaged and days for which the target area average rainfall was less than 1mm removed as these were not considered to be raindays.

Not all raindays in the target area will be suitable for cloud seeding. Because of this only a subset of the raindays was used in the simulations. Based on the Tasmanian experiments and those of the Thomson Catchment project 15 suitable days per season were used. This means that a simulated 5-year experiment would consist of 75 raindays, and a 6-year experiment would consist of 90 days. To avoid results reflecting anomalous selections of "suitable" days the results quoted here are the average over three random selections of days to be included in each simulation.

The exact number of suitable raindays per season for the Copeton region can only be determined once synoptic classifications of raindays is made and more information is available of the suitability of the clouds present for seeding.

Three different regression analyses were used for estimating target precipitation from control area precipitation. The significance of a double ratio was then determined by a permutation test for each of the three methods. The outcome of the simulations was the number of experiments out of 100 experiments which showed a significant positive effect of seeding when one was introduced into the data. This shows the power of the experiment, and of the analysis, to detect the introduced effect.

The three regression methods were:

1. Without taking wind direction into account, the target precipitation that would have fallen without seeding is estimated from a combination, $P_{control}$, of the five control areas, P_i ($i=A,B,C,D,E$). That is:

$$P_{control} = a + b_1P_A + b_2P_B + b_3P_C + b_4P_D + b_5P_E$$

where $a, b_1, b_2, b_3, b_4, b_5$ were fitted from all the data selected for the experiment.

2. The target precipitation was estimated from only those control groups which were immediately upwind of the target area on that day. This means that control group E was used when the wind direction was 180-240 degrees, $(C+D)/2$ when the wind direction was 250-300 degrees, and $(A+B)/2$ when the wind direction was 310-360 degrees.

3. The third approach estimated the target precipitation from the five controls, but with a separate regression relationship for each of the three wind directions.

For each of the three regression approaches above, simulations were made of:

5-year and 6-year experiments

1:1 and 2:1 seed:no-seed (unrestricted) randomizations

tests at the 5% and 10% significance levels

multiplicative precipitation increases due to seeding of 10%, 20% and 30%.

The 30 percent precipitation increase on seeded days was taken as an upper limit, as with the Thomson Catchment experiment, as this was the result obtained in the Tasmanian experiments. If the proposed Copeton cloud seeding experiment were to go ahead and the precipitation increase on seeded days was higher than 30 percent then the chances of detection of that effect would be greater than that of the 30 percent increase given in the tables. Tables 1, 2, and 3 give the results obtained in the simulations. Note that each figure in the body of the Tables is the average, over three random selections of rainday subsets to be included in the experiment, of the number of positive statistically significant results obtained in 100 trials of an experiment conducted under the conditions specified.

TABLE 1. Percent chance of detecting a prescribed precipitation increase

Regression on all controls (not using wind direction)

	Simulated Precipitation Increase		
	10%	20%	30%
5 years 1:1 randomisation			
5% sig. level	21	49	76
10% sig. level	33	59	88
5 years 2:1 randomisation			
5% sig. level	22	49	69
10% sig. level	34	64	81
6 years 1:1 randomisation			
5% sig. level	26	62	81
10% sig. level	38	73	90
6 years 2:1 randomisation			
5% sig. level	21	56	82
10% sig. level	32	73	88

TABLE 2. Percent chance of detecting a prescribed precipitation increase

Regression using only upwind controls

	Simulated Precipitation Increase		
	10%	20%	30%
5 years 1:1 randomisation			
5% sig. level	4	3	14
10% sig. level	7	14	30
5 years 2:1 randomisation			
5% sig. level	2	10	13
10% sig. level	7	18	27
6 years 1:1 randomisation			
5% sig. level	3	10	18
10% sig. level	10	22	34
6 years 2:1 randomisation			
5% sig. level	2	7	14
10% sig. level	9	17	32

TABLE 3. Percent chance of detecting a prescribed precipitation increase

Regression on all controls (using wind direction)

	Simulated Precipitation Increase		
	10%	20%	30%
5 years 1:1 randomisation			
5% sig. level	35	56	73
10% sig. level	52	72	83
5 years 2:1 randomisation			
5% sig. level	28	48	63
10% sig. level	45	69	81
6 years 1:1 randomisation			
5% sig. level	44	65	80
10% sig. level	62	79	93
6 years 2:1 randomisation			
5% sig. level	37	64	76
10% sig. level	56	80	89

The Tables indicate that:

a) The chances of detecting an effect are higher under regression on all controls than they are using only upwind controls. The results using all controls and wind direction are generally better than using all controls without wind direction (18 of the

24 designs had a greater detection chance with all other designs showing only slightly smaller chances of detection). The results of regression on all controls using winds are particularly good for increase in rainfall of 10-20 percent. The relatively poor results when using only upwind controls is probably due to the fact that most rainfall was associated with winds from 220-290 degrees. This corresponds to a region that has very few control gauges and the gauges which are found in this region generally had lower correlations with the target area.

b) The effects of longer experiment duration, five years vs. six years, and a more balanced ratio of seed to no-seed days, 1:1 vs. 2:1 on improving the chances of success are clearly shown, particularly in the case of regression on all controls using wind direction (Table 3).

c) The most important contributor to improvement of the chances of successful detection of an effect is the magnitude of the effect. Hence, it is necessary to conduct the experiment to maximise the effect of seeding on any occasion when clouds meet the suitability criteria and that the delivery of the seeding material on seeded occasions will achieve accurate targeting of the seeding effect.

It should be kept in mind that the figures in the Tables are estimates, based on limited simulations, of the chances of success. For example, the chance of detecting a 30 percent increase in rainfall, using all controls and wind direction, for a six year experiment with a 2:1 seed:no-seed randomisation, at the 10 percent significance level is 89 percent. However, due to sampling variability there is a 95 percent chance that the true chance of detection lies in the range 87-91 percent.

While remembering that the figures quoted in the Tables are estimates, it appears that the odds are about 4:1 that a 30 percent increase in precipitation will be detected in a 5-6 year experiment and about 2:1 that a 20 percent increase will be detected.

7 CLOUD CONDITIONS

In order to know at what height the clouds should be seeded in a potential cloud seeding project at the Copeton Dam we need to determine the heights of the 0°C, -5°C, -10°C and -15°C levels as well as at what temperature and height the cloud base and the cloud top occur on raindays.

The height of the various temperature levels varied according to month. The lowest heights being the June, July, August period, the colder season, when the heights were fairly constant. At both ends of this period the height of the temperature levels increased as you moved further away from the centre. Figures 11 to 14 show the monthly variation with the use of boxplots. The heights of the various levels are given in dekametres (1 dekametre = 10 metres). Table 4 gives the mean height of the various levels both overall and monthly as well as the variation (standard error) about this number. Again the mean heights are measured in dekametres.

FIGURE 11: Boxplots of Height of 0 degree level overall and monthly

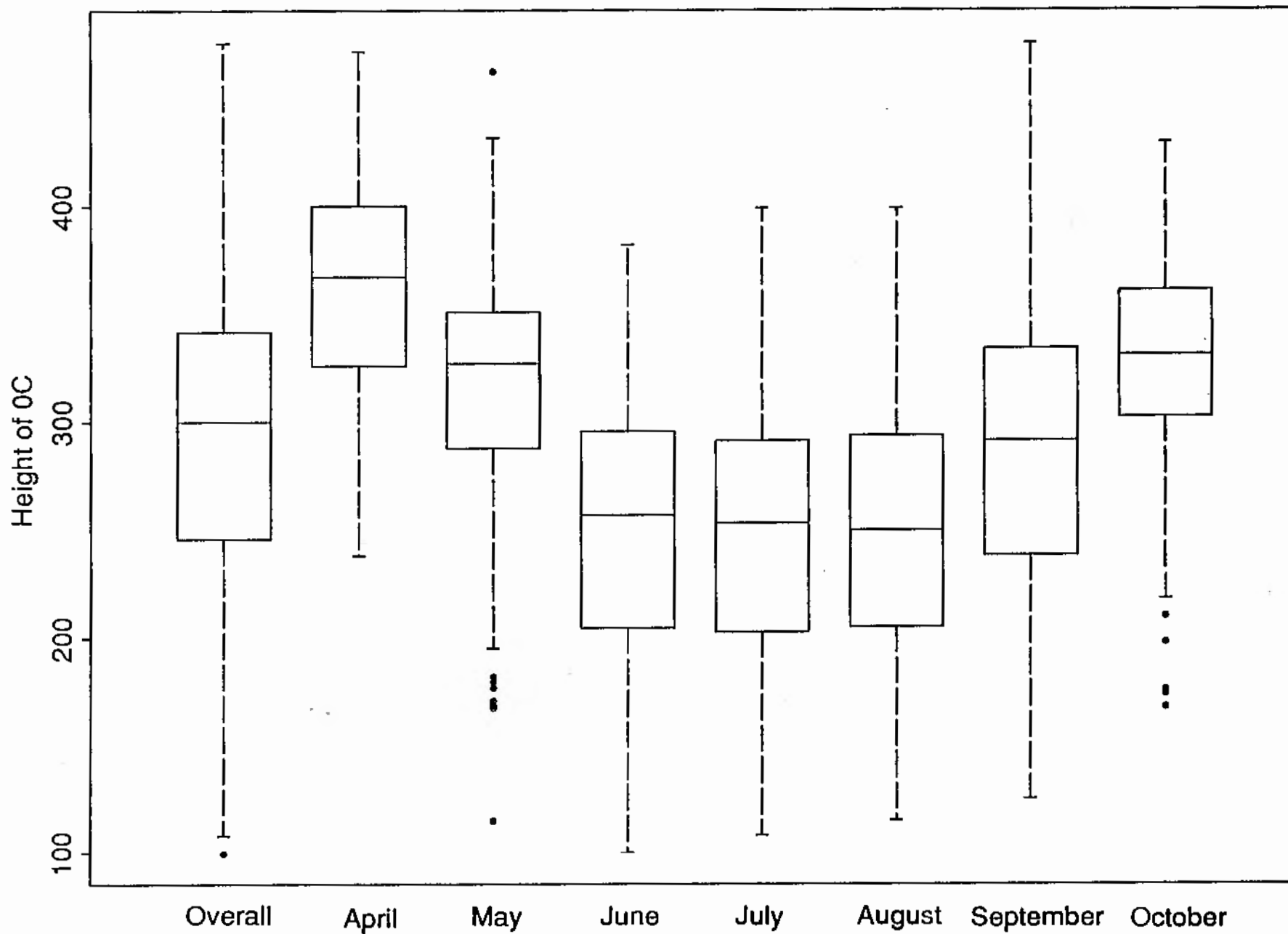


FIGURE 12: Boxplots of Height of -5 degree level overall and monthly

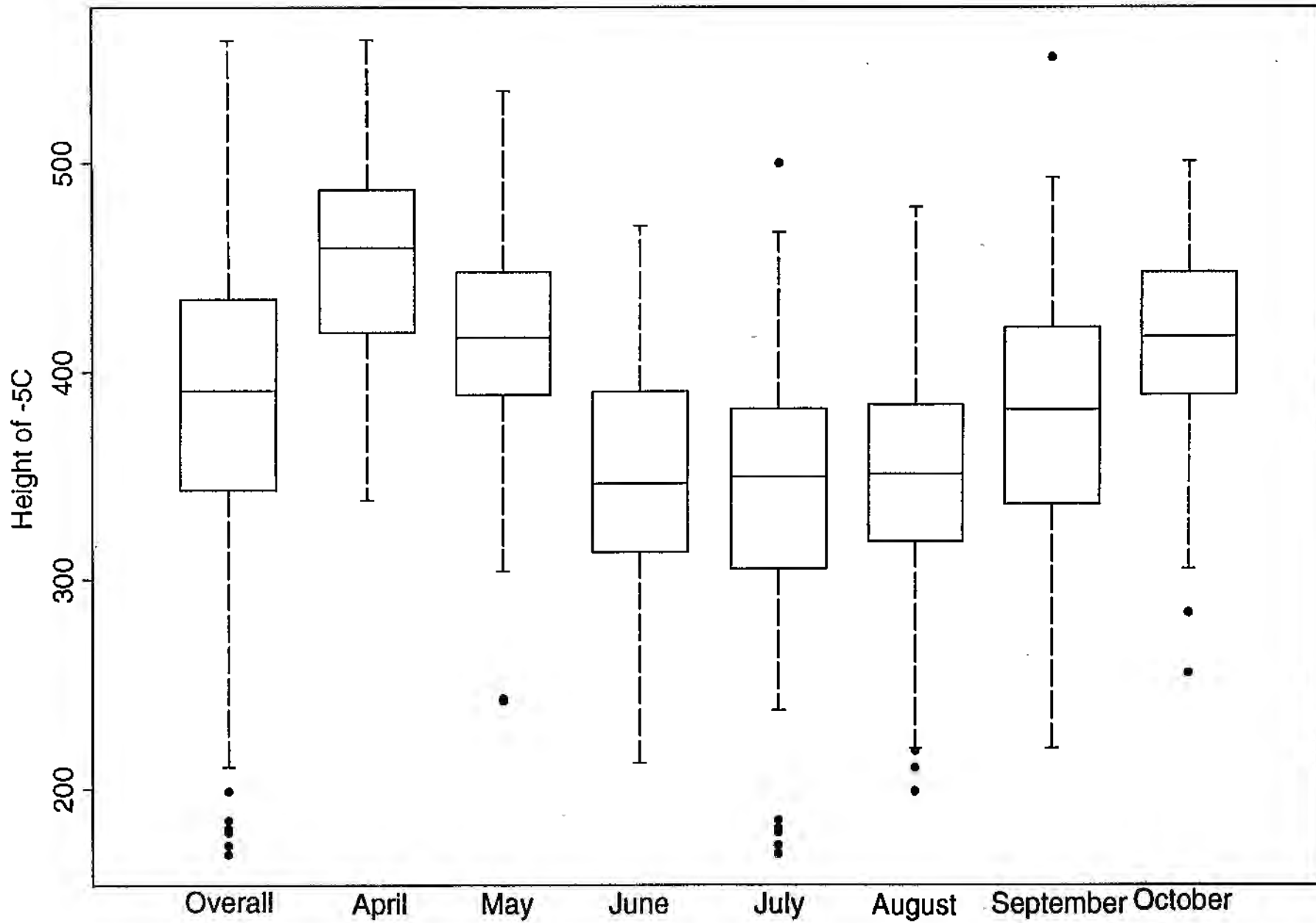


FIGURE 13: Boxplots of Height of -10 degree level overall and monthly

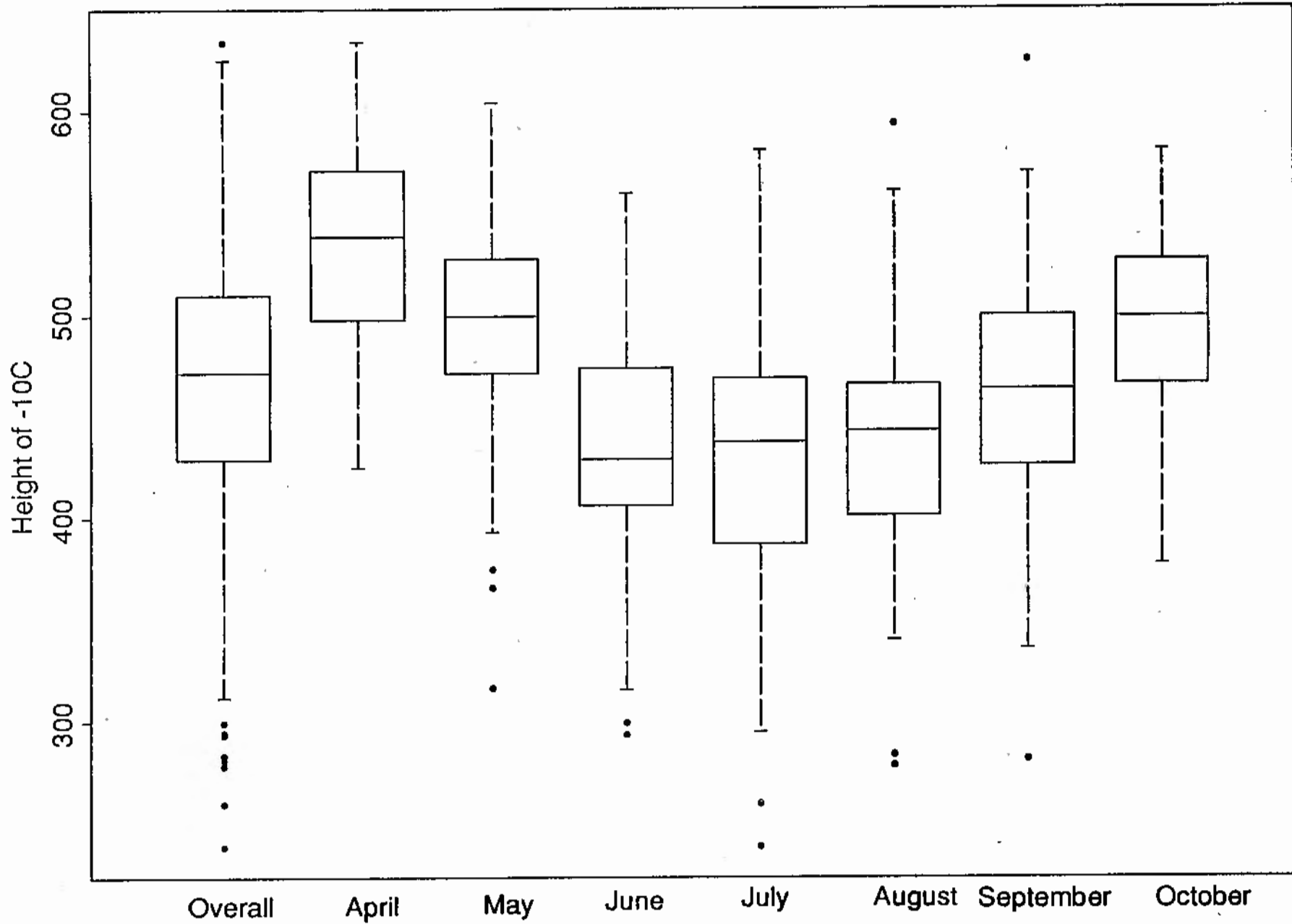


FIGURE 14: Boxplots of Height of -15 degree level overall and monthly

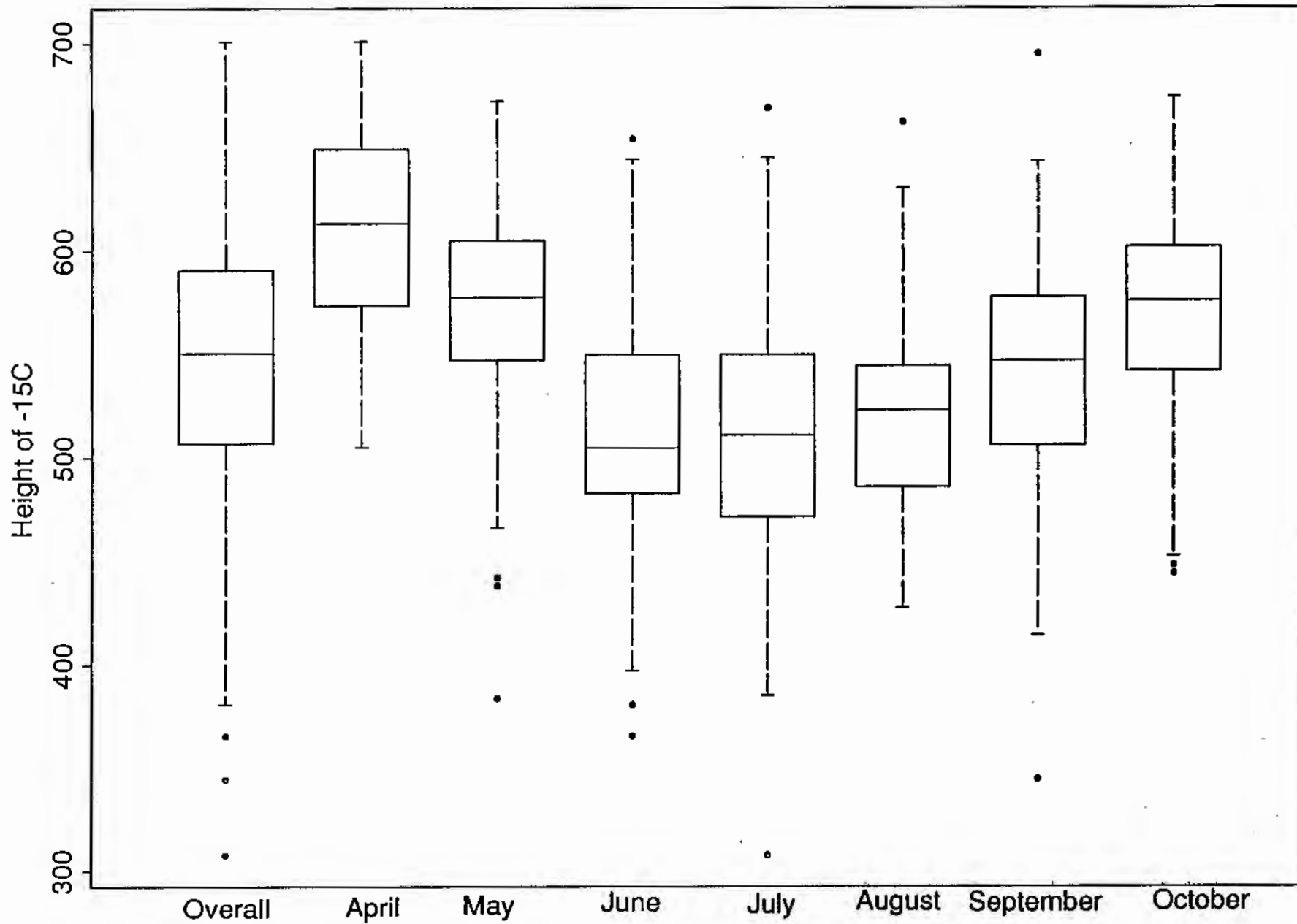


TABLE 4: Heights of the various temperature levels

	Height of 0°C Level	
	MEAN	STANDARD ERROR
Overall	291.76	71.99
April	361.22	49.37
May	318.77	60.85
June	251.37	60.29
July	246.13	66.26
August	249.06	61.55
September	283.98	67.93
October	326.53	50.19

	Height of -5°C Level	
	MEAN	STANDARD ERROR
Overall	386.66	65.47
April	452.20	48.02
May	413.88	50.71
June	348.31	54.42
July	345.01	66.74
August	351.21	53.23
September	377.46	58.64
October	415.97	46.36

	Height of -10°C Level	
	MEAN	STANDARD ERROR
Overall	469.91	61.92
April	532.60	46.95
May	495.71	47.51
June	436.15	54.34
July	432.29	62.92
August	437.02	50.78
September	460.80	55.25
October	494.21	45.07

	Height of -15°C Level	
	MEAN	STANDARD ERROR
Overall	548.86	60.76
April	608.82	45.22
May	572.85	48.51
June	515.31	55.80
July	514.81	63.82
August	520.17	47.53
September	539.74	56.01
October	570.12	45.74

Cloud base and cloud height were calculated using a program SONDTOOL developed by the Bureau of Meteorology. A check was then made of some of the results using F160 aerological diagrams. First two definitions need to be made.

The Lifted Condensation Level (LCL) is the level at which a parcel which rises adiabatically from the ground become saturated. This is the level at which cloud bases would be expected to lie.

The Equilibrium Level is the height at which the temperature of a convective parcel becomes the same as or cooler than its environment. This height should be close to the cloud top height.

There was a total of 485 days with rainfall greater than 1mm in the Copeton Region and which had data on mixing ratios, temperatures, wind directions and speeds for various pressure levels all of which are needed for the SONDTOOL program to calculate cloud base and top height. For 73 percent of the data the convective parcel never became warmer than the environment resulting in no value for the equilibrium level. On these days we would expect a cloud of little depth and hence, when the cloud depths were calculated, these days were given a cloud depth of 0. Table 5 gives a break down according to month of the number of days with complete data as well as the number of days for which equilibrium levels were obtained and the percentage that these days make up.

TABLE 5 : Number of Days With Complete Data Sets

	Total Number of Days	With Equilibrium Level Values	Percentage
Overall	485	131	27
April	70	18	26
May	60	12	20
June	69	10	14
July	49	3	6
August	59	13	22
September	66	22	33
October	112	53	47

The depth of clouds in the Copeton region is lowest over the cooler months of June, July and August with greater depths being reached in April and October. Figure 15 illustrates this using hoxplots. Note that the hox for July does not really represent the true cloud depths expected for this month as there were only three observations available. Table 6 gives the mean and the standard errors of the cloud depths both overall and monthly for all of the data including the occasions when no equilibrium value was obtained. Table 7 gives the same information excluding the days of zero cloud depth. The large discrepancy in the values is due to the large number of days for which no equilibrium value was available resulting in many days with zero cloud depth. This lowers the mean, in this case, by a large amount.

FIGURE 15: Boxplots of Cloud Depth (m) Overall and Monthly

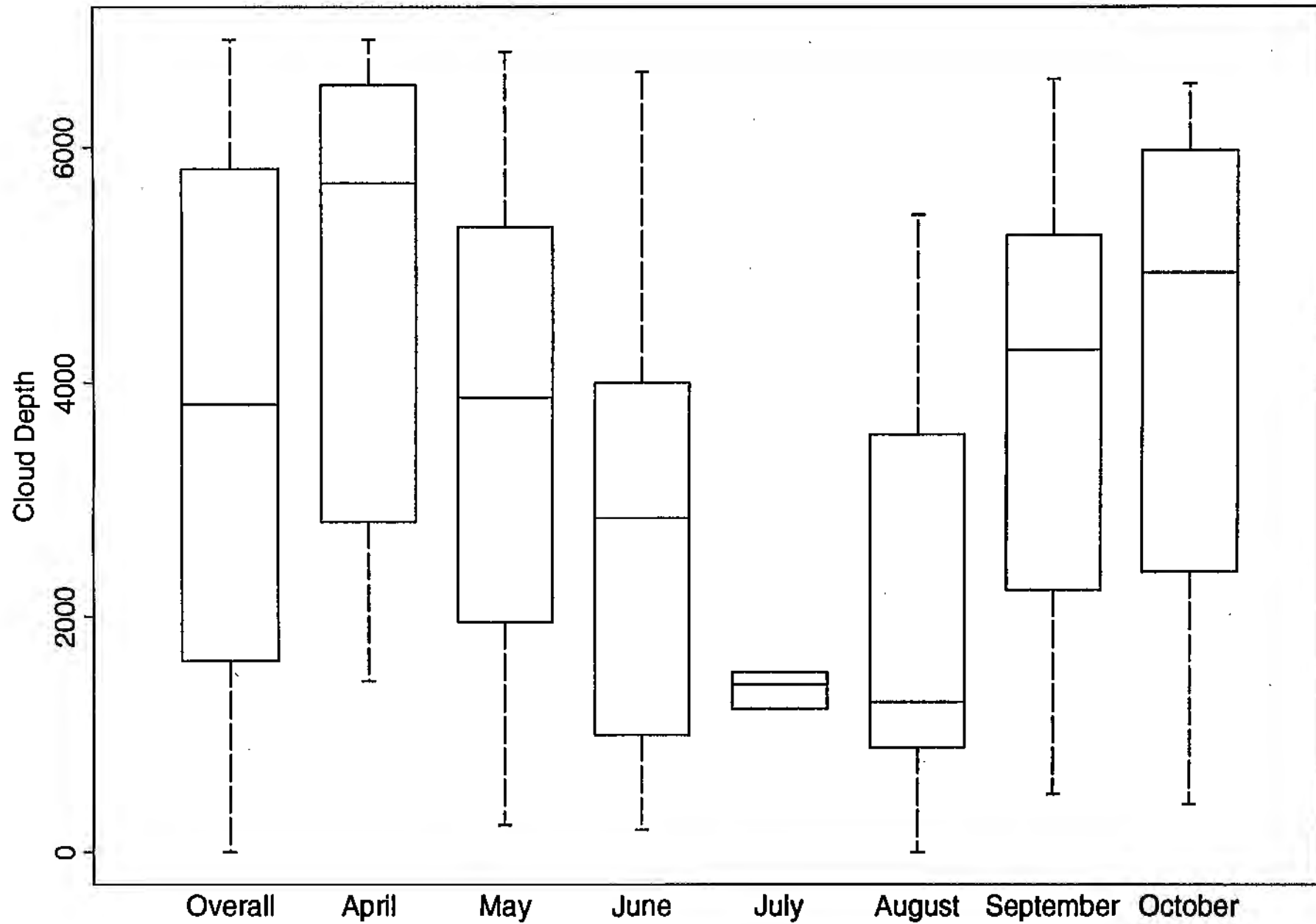


TABLE 6 : Mean and Standard Error of Cloud Depths Using All Days

	Mean	Standard Error
Overall	1008.46	1991.03
April	1198.96	2289.41
May	744.98	1760.74
June	388.72	1185.62
July	84.96	337.63
August	437.54	1158.47
September	1291.36	2151.14
October	1950.42	2503.21

TABLE 7 : Mean and Standard Error of Cloud Depths Excluding Days of Zero Depth

	Mean	Standard Error
Overall	3733.61	2122.74
April	4662.61	2043.10
May	3724.92	2125.85
June	2682.20	1945.70
July	1387.67	156.03
August	1985.77	1776.51
September	3874.10	1959.79
October	4121.64	2062.64

Of the total of 485 days with data available only 11 days had the LCL at or below 0°C. This implies that around 2 percent of the 485 raindays had conditions conducive to snow falling. Hence, it is unlikely that the precipitation would be snow, particularly around the Moree region. This agrees with Figure 3. The height of the cloud base varied according to month with the colder months having lower cloud bases. Tables 8 and 9 give the mean and standard errors of the cloud base with and without the days of zero cloud height.

TABLE 8 : Mean and Standard Error of Cloud Base for All Days

	Mean	Standard Error
Overall	945.11	412.30
April	1013.10	422.86
May	763.90	295.38
June	690.33	272.17
July	754.71	270.54
August	799.46	327.49
September	1049.26	432.20
October	1255.29	385.25

TABLE 9 : Mean and Standard Error of Cloud Base Excluding Days of Zero Depth

	Mean	Standard Error
Overall	826.55	371.31
April	678.94	340.21
May	538.58	223.03
June	464.00	173.54
July	575.00	247.44
August	690.23	217.58
September	795.86	267.79
October	1070.70	357.00

The clouds of the colder months (June, July and August) are generally lower and have less depth. It should also be noted that the temperatures of set heights are likely to be slightly lower in the Copeton Dam region than at Moree due to differences in terrain. Figure 3 illustrates this as it is more likely that snow will fall in the Copeton region than at Moree.

Comparing the values found in Table 9 and those for the -10°C level in Table 4 (remembering that the values in Table 4 are in dekametres and those in Table 9 are in metres) we find that the -10°C level is well above the cloud base. It is, on average, 3873m above the cloud base. Looking then at Figure 15 we see that the -10°C level occurs near the top of most of the rain clouds over Moree. This is important in determining flight paths in potential cloud seeding experiments in the Copeton area.

8 CONCLUSION

The very high correlations of the Copeton Dam region with surrounding gauges stations indicates a good potential for cloud seeding. Should cloud seeding go ahead there are a number of individual gauges and groups of gauges, 27 individual gauges (in the 5 control groups), which could act as very good predictors of the rainfall in the Copeton Dam region.

Most of the days resulting in rainfall in the Copeton Dam region had winds from the northwest and southwest (81 percent of all raindays) with speeds of around $12-13 \text{ m s}^{-1}$. All of the control area gauges selected are to the northwest and southwest of the target area resulting in good predictors of the amount of rainfall in the target area on non-seeded occasions.

The Bureau of Meteorology gauges seem to be reliable enough for use in a potential cloud seeding experiment provided that the returns of rainfall data from the farmers operating the gauges can be accelerated.

Snow falls to the west of the Copeton Dam region are very unlikely in the course of a prospective cloud seeding experiment in these regions. The cloud base height, on most occasions, was around 826m and the temperature at this height was generally above 0°C . The -10° level mostly occurred near the top of the clouds at around 4699m.

Statistical simulations of a potential cloud seeding experiment, using historical data, suggest that the best method of estimating the target area rainfall which would have occurred had seeding not been carried out is regression on all controls using wind direction. Under this model the odds are about 4:1 that a 30% increase in precipitation would be detected in a 5-6 year experiment and about 2:1 that a 20% increase would be detected. The chance of detecting a prescribed precipitation increase is greater with increasing length of the experiment and the larger the effect. It is also higher for a 1:1 seed:no-seed experiment than for a 2:1 seed:no-seed experiment.

APPENDIX 1 - GAUGE LOCATIONS

GAUGE NUMBER	LATITUDE (deg min S)	LONGITUDE (deg min E)	NAME
53048	2928	14951	Moree
53095	2902	15027	North Star (Wolonga)
53108	2948	15009	Terry Hie Hie Post Office
54000	3002	15057	Bundarra (Mundawah)*
54003	3023	15037	Barraba Post Office
54005	3004	15039	Upper Bingara*
54012	2915	15045	Coolatai (Orana)
54014	2955	15022	Bingara (Derra Derra)
54016	2927	15047	Delungra (Craigmore)
54021	3018	15016	Barraba (Mount Lindsay)
54023	3018	15049	Barraba (Neranghi)
54027	2933	15053	Delungra (Pearsby Hall)
54028	2958	15055	Copeton (Carinya)*
54032	2912	15036	Coolatai (Willunga)
54036	2915	15054	Wallangra Station
54039	3000	15047	Bingara (Keera)
54040	2941	15037	Koloona (Alkira)
54043	2921	15109	Ashford (Llanarth)
54044	2919	15109	Ashford (Tarrangower)
54045	2921	15117	Ashford (Springvale)
54046	2924	15125	Ashford (Burrabogie)
54048	2922	15055	Graman (Maneroo)
54049	2925	15054	Graman (Ulupna)
54056	2928	15056	Graman Post Office
54057	2931	15058	Cherry Tree Hill (Kulki)
54063	2930	15119	Nullamanna (Severn Vale)
54065	2935	15115	Nullamanna (Lone Pine)
54069	2936	15109	Dinton Vale Composite
54073	2939	15114	Nullamanna (Belmore)
54074	2940	15101	Oakwood (Lochiel)
54077	2941	15104	Oakwood (Glendowne)
54078	2940	15109	Inverell (Wandera)
54082	2938	15124	Kings Plains (Woodlands)
54102	3023	15040	Barraba (Rosevale)
54104	2924	15114	Pindari Dam
54105	3020	15056	Bundarra (Granite Hghts)

GAUGE NUMBER	LATITUDE (deg min S)	LONGITUDE (deg min E)	NAME
54106	3004	15046	Bingarra (Cromarty)*
54107	3006	15054	Bundarra (Kurrajong)*
54113	3009	15058	Bundarra (Ross)
54114	3006	15101	Bundarra (Beverley)
54116	3018	15042	Barraba (Treborane)
54117	3014	15043	Cobbadab (Rocky Glen)
54119	3022	15028	Barraba (Camposanto)
54122	2921	15035	Warialda (Croppa)
54123	2949	15023	Bingara (Derra Lea)
54124	2917	15016	Crooble (Crooble Station)
54125	3008	15009	Caroda (Rosebery Park)
54127	3013	14959	Narrabri (Woodlands)
54128	2953	15054	Copeton Dam*
54129	2913	15022	Croppa Creek (Rawdon)
54130	2908	15018	Croppa Creek Post Office
54138	3009	15023	Upper Horton (Doorah)
54153	3008	15008	Caroda (Paleroo)
54159	2930	15108	Bukkulla
54160	2909	15118	Bonshaw (Applewood)
55000	3055	15052	Attunga (Garthowen)
55004	3053	15109	Bendemeer (Charles st)
55050	3056	15038	Somerton Post Office
55054	3105	15051	Tamworth Airport
55055	3058	15028	Carroll (The Ranch)
55058	3026	14955	Turrawan (Wallah)
55076	3031	15014	Boggabri (Gumera)
55082	3028	15057	Orabah (Warrabah)
55090	3038	15044	Manilla Upper (Carmarthen)
55091	3039	15045	Manilla (Strathfield)
55102	3044	15104	Watsons Creek (Fairlands)
55103	3042	15107	Watsons Creek (Tilmunda)
55105	3048	15052	Attunga (Brooklyn)
55109	3048	15107	Bendemeer (Glenclair)
55118	3055	15042	Somerton (Clermont Park)
55119	3053	15040	Somerton (Lonou)
55120	3056	15047	Attunga (The Pines)

GAUGE NUMBER	LATITUDE (deg min S)	LONGITUDE (deg min E)	NAME
55122	3051	15055	Attunga (Minderoo)
55136	3058	15121	Woolbrook (Danglemah Road)
55138	3101	15039	Somerton (Kallaroo)
55140	3101	15043	Somerton (Glen Burn)
55143	3102	15104	Moonbi (Bellevue)
55146	3102	15120	Woolbrook (Back Creek Station)
55149	3105	15041	Winton (Nioka)
55157	3105	15039	Winton (Dalblair)
55201	3050	15021	Kelvin (Newery Park)
55265	3032	15001	Boggabri (Merriendi)
55272	3051	15015	Gunnedah (Colstoun South)
55274	3047	15026	Kelvin (Carellan)
55276	3053	15029	Keepit Dam WRC
55279	3105	15046	Tamworth (Inverness)
55291	3038	15040	Upper Manilla Post Office
55296	3046	15115	Bende Meer (Babalasta)
55314	3034	15054	Manilla (Tarwonga)
55316	3059	15119	Woolbrook (Glenview)
55319	3037	15035	Upper Manilla (Borah Downs)
55320	3103	15111	Limbri (Jo-car)
55324	3105	15116	Curlewis (Wongalee)
56002	3031	15140	Armidale Radio Station 2AD
56007	3000	15118	Tingha (Crystal Hill)
56011	2944	15144	Glen Innes Post Office
56013	2942	15142	Glen Innes Agric. Research Station
56017	2947	15107	Inverell Post Office
56018	2947	15105	Inverell Soil Cons Services
56026	2946	15154	Red Range (Victoria Street)
56029	2927	15129	Emmaville (Strathbogie)
56036	3004	15131	Wandsworth (Strabanne)
56042	3039	15118	Uralla (Lana)
56057	3005	15123	Tenterden (Balook)
56065	3043	15140	Uralla (Mihi)
56068	3023	15145	Armidale (Belraven)
56069	3039	15145	Armidale (Tiverton)
56075	3102	15126	Walcha Road (Boxley)

GAUGE NUMBER	LATITUDE (deg min S)	LONGITUDE (deg min E)	NAME
56088	2930	15141	Emmaville (Severn Valley)
56094	2933	15200	Dundee (Wattle Dale)
56096	2936	15138	Reddestone (Redbank)
56098	2935	15157	Dundee (Karinga)
56101	2939	15137	Wellingrove (Glenbrook)
56102	2938	15141	Reddestone (Lonewood)
56111	2947	15121	Elsmore (Danthonia)
56115	2942	15150	Glen Innes (Lyndale)
56121	2953	15125	Elsmore (Paradise Creek)
56123	2953	15128	Elsmore (Paradise Station)
56128	2950	15131	Swan Vale (Numeralla)
56134	2955	15119	Tingha (Bonnie View)
56139	2957	15136	Ben Lomond (Koala)
56161	3010	15153	Guyra (Gowen Brae)
56167	3000	15105	Stanborough (Rockfield)
56170	3005	15112	Tingha (Mooki)
56176	3008	15115	Bundarra (Bimbadeen)
56180	3018	15105	Bundarra (South Winscombe)
56198	2957	15112	Tingha (Darbys Road)
56205	2954	15158	Pinkett (Benbookra)
56214	3019	15100	Bundarra (Willowbank)
56228	3030	15107	Kingston Post Office
56229	3013	15141	Guyra Hospital
56233	2947	15129	Swan Vale (Pieta)

* denotes stations within the Copeton Dam Group.

APPENDIX 2 - EXPLANATION OF TERMS AND DEFINITIONS

To illustrate the following terms the data set given below is used.

1 3 4 6 10

Mean : The mean is the average of a set of numbers. For the above example the mean is

$$(1+3+4+6+10)/5 = 4.8$$

Median : The median is the "middle" number of a set of ORDERED numbers. In the case of an even number of points the average of the two middle numbers is used. In general, if we have n observations, then the median is the $(n+1)/2$ th number. For the above example the median is 4.

Quantile : The quantiles are commonly made up of the minimum value, the maximum value, the median, the point to which 25% of the data is smaller than the number and 75% is greater than it, and the point at which 75% is smaller than the number and 25% is greater than the number.

For the above example the quantiles are

1.0 2.5 4.0 7.0 10.0

ie the minimum number is 1.0

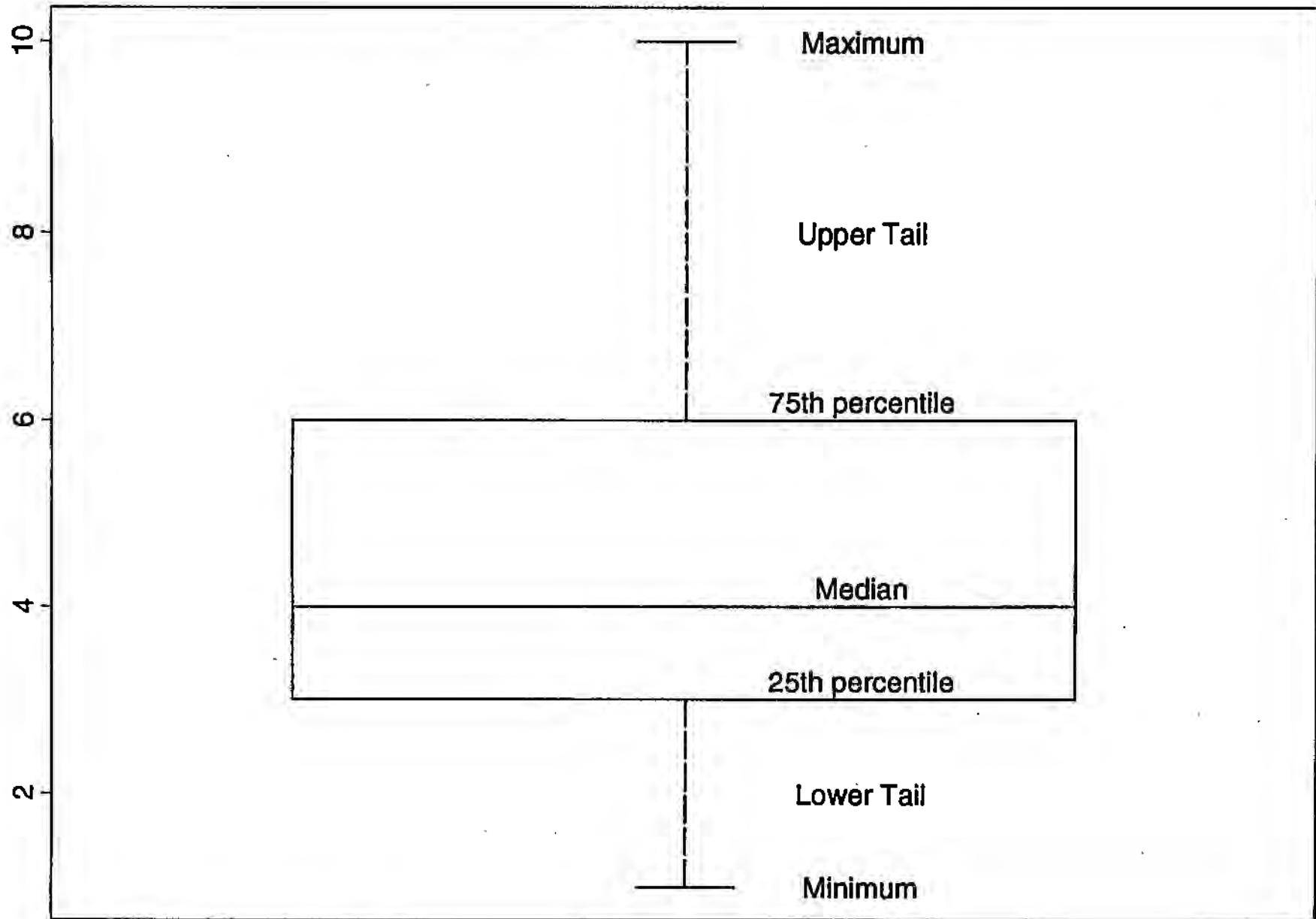
the maximum number is 10.0

the median is 4.0

25% of the data is smaller than 2.5 (75% greater than) (Also known as the 25 percentile)

75% of the data is smaller than 7.0 (25% greater than) (Also known as the 75 percentile)

FIGURE 16: Boxplot of Example Data Set



APPENDIX 3 - REGRESSION ESTIMATES

For the regression of all controls not using wind directions we obtained the following estimates:

Parameter	Estimate	Standard Error
<i>a</i>	8.476	1.954
<i>b</i> ₁	0.216	0.032
<i>b</i> ₂	0.079	0.029
<i>b</i> ₃	0.185	0.024
<i>b</i> ₄	0.026	0.024
<i>b</i> ₅	0.591	0.035

That is, the regression equation for all controls, not using wind direction, is given by

$$P_{control} = 8.476 + 0.216P_A + 0.079P_B + 0.185P_C + 0.026P_D + 0.591P_E$$

The value obtained for $P_{control}$ is used as an estimate for the amount of rainfall that would have fallen in the target area had seeding not taken place.

For the regression of all controls using wind directions the estimates are:

Winds from 180°-240°

Parameter	Estimate	Standard Error
<i>a</i>	11.203	2.968
<i>b</i> ₁	0.267	0.058
<i>b</i> ₂	0.003	0.060
<i>b</i> ₃	0.191	0.046
<i>b</i> ₄	0.211	0.043
<i>b</i> ₅	0.314	0.056

Winds from 250°-300°

Parameter	Estimate	Standard Error
<i>a</i>	8.487	2.784
<i>b</i> ₁	0.165	0.043
<i>b</i> ₂	0.072	0.039
<i>b</i> ₃	0.295	0.035
<i>b</i> ₄	-0.044	0.035
<i>b</i> ₅	0.673	0.049

Winds from 310°-360°

Parameter	Estimate	Standard Error
<i>a</i>	5.752	5.200
<i>b</i> ₁	0.259	0.073
<i>b</i> ₂	0.072	0.062
<i>b</i> ₃	0.015	0.045
<i>b</i> ₄	-0.015	0.057
<i>b</i> ₅	0.757	0.119

APPENDIX 4 - THE DOUBLE RATIO

One way to investigate whether cloud seeding has any effect on rainfall in the target region is to use a double ratio analysis with least-squares regression. This approach has been used before in the Tasmania II and Thomson Catchment experiments.

The double ratio, d , is calculated as

$$d = \frac{\sum_S T_i / \sum_S C_i}{\sum_U T_i / \sum_U C_i}$$

where:

T_i is the target rainfall for day i

C_i is the control rainfall for day i (an estimate of the target area rainfall had seeding not taken place)

\sum_S is the sum over all seeded days

\sum_U is the sum over all nonseeded days

The value of C_i represents the expected value of the target rainfall had seeding not occurred. Hence, to calculate the double ratio the value of C_i needs to be computed first using one of the three regression methods mentioned in Section 6 and Appendix 3 (for more details on how this was actually performed see Appendix 5).

If seeding has no effect then we would expect d to have a value near 1.0 with departures from 1.0 indicating the effect of seeding. The statistical significance of departures from 1.0 is assessed by a permutation test using 1000 re-randomizations.

APPENDIX 5 - PROGRAM USED IN SIMULATIONS

The bulk of the simulations was performed using a computer language called S-PLUS. The code is given below as well as comments as to what is occurring at each step. This code does not pretend to be the most efficient method of calculating the double ratios.

```
# The first step is to read in the file, rain.dat, containing all the
# rainfall averages for the target area and the control areas. It also
# contains wind direction, wind speed and two indicators that, together,
# act as an indicator of which sector is upwind of the target region.

data <- scan("rain.dat",list(aa1=0,aa2=0,aa3=0,aa4=0,aa5=0,aa6=0,w1=0,w2=0,
                             pp1=0,pp2=0))

# Here the "=0" indicates the variables to be read in are numeric.

a1<-data$aa1 # This is all the target area rainfall averages
a2<-data$aa2 # These are the averages for control group A
a3<-data$aa3 # These are the averages for control group B
a4<-data$aa4 # These are the averages for control group C
a5<-data$aa5 # These are the averages for control group D
a6<-data$aa6 # These are the averages for control group E
wd<-data$w1  # These are the wind directions
ws<-data$w2  # These are the wind speeds
p1<-data$pp1 # An indicator variable. 1 if winds from 250-360 degrees,
             # 0 otherwise.
p2<-data$pp2 # An indicator variable. 1 if winds from 180-240 or 310-360
             # degrees, 0 otherwise.

# Next we generate 75 (ie 5-years worth) variables between 1 and 712.
# Once rounded, this will give us a list of 75 days from the total of
# 712 possible days to use in the current cloud seeding simulation.
# If we wished to conduct a simulation for a 6-year experiment we
# simply replace the 75 by 90.

r1<-runif(75,min=0.5,max=712.5)
r1<-round(r1)

# We then collect all the data that we need for the 75 days chosen
# from the data input, rain.dat.

t<-a1[r1]    # These are the target area rainfall averages for the 75 days
gA<-a2[r1]  # These are the group A rainfall averages for the 75 days
gB<-a3[r1]  # These are the group B rainfall averages for the 75 days
gC<-a4[r1]  # These are the group C rainfall averages for the 75 days
gD<-a5[r1]  # These are the group D rainfall averages for the 75 days
gE<-a6[r1]  # These are the group E rainfall averages for the 75 days
```

```

wd1<-wd[r1] # The wind directions for the 75 days
p1<-p1[r1] # Indicator variables as above.
p2<-p2[r1] # Indicator variables as above.

# The next step is to randomly generate a sequence of 0's and 1's
# representing the seed:no seed decision. 1=seed, 0=no seed. Here
# we have the case of a 1:1 seed:no-seed randomization, with no
# restrictions. This line must be modified for the case of a 2:1
# seed:no-seed randomization.

s1<-runif(75,-0.5,1.5)
s1<-round(s1)

# We now increase the target area rainfall average by a prescribed
# amount to simulate cloud seeding. Here the increase is 10 percent.

tar<-s1*t*0.1 + t

# and we estimate the expected target area rainfall had seeding not
# occurred using regression method 1.

et1<-8.476 + 0.216*gA + 0.079*gB + 0.185*gC + 0.026*gD + 0.591*gE

# The double ratio for this regression method is then calculated

d<-(sum(s1*tar)/sum(s1*et1))/(sum((1-s1)*tar)/sum((1-s1)*et1))

# Using regression method 2 we need to create vectors representing
# the expected target area rainfall had seeding not occurred taking
# wind direction into account.

et2<-rep(0,75) # set et2 equal to 75 lots of 0
et3<-rep(0,75)

et2a<-gE # expected target rainfall for 180-240 degree winds
et2b<-(gD+gC)/2 # expected target rainfall for 250-300 degree winds
et2c<-(gA+gB)/2 # expected target rainfall for 310-360 degree winds

x<-1:75 # make a vector consisting of the numbers 1 to 75

z1<-(1-p1)*p2 # equals 1 if winds from 180-240 degrees, 0 otherwise
z2<-p1*(1-p2) # equals 1 if winds from 250-300 degrees, 0 otherwise
z3<-p1*p2 # equals 1 if winds from 310-360 degrees, 0 otherwise

xa<-x[z1==1] # record the index number of days with winds from 180-240 degrees
xb<-x[z2==1] # record the index number of days with winds from 250-300 degrees

```

```

xc<-x[z3==1] # record the index number of days with winds from 310-360 degrees

# create one vector of expected target rainfalls using the appropriate wind
# direction values

et2[xa]<-et2a[xa]
et2[xb]<-et2b[xb]
et2[xc]<-et2c[xc]

# calculate the corresponding double ratio

d2<-(sum(s1*tar)/sum(s1*et2))/(sum((1-s1)*tar)/sum((1-s1)*et2))

# Now we do the corresponding calculations for regression method 3, ie.
# for regression on all controls using wind direction.

et3a<-11.203 + 0.267*gA + 0.003*gB + 0.191*gC + 0.211*gD + 0.314*gE
et3b<- 8.487 + 0.165*gA + 0.072*gB + 0.295*gC - 0.044*gD + 0.673*gE
et3c<- 5.752 + 0.259*gA + 0.072*gB + 0.015*gC - 0.015*gD + 0.757*gE

et3[xa]<-et3a[xa]
et3[xb]<-et3b[xb]
et3[xc]<-et3c[xc]

d3<-(sum(s1*tar)/sum(s1*et3))/(sum((1-s1)*tar)/sum((1-s1)*et3))

# To test the significance of the double ratios calculated above we
# need to randomly permute the seed:no-seed vector, s1, and recalculate
# the double ratios. This needs to be done 1000 times and is
# achieved using the subroutine perm.s.

source("perm.s") # similar to complying perm.s

test <- perm(s1,tar,et1,et2,et3,p1,p2) # the result of the call to perm.s
# is placed in test

dr <-test[,1] # The permutation results for regression method 1.
dr2 <-test[,2] # The permutation results for regression method 2.
dr3 <-test[,3] # The permutation results for regression method 3.

dr<-sort(dr) # values in dr are sorted numerically
dr2<-sort(dr2)
dr3<-sort(dr3)

sink("sim.out") # open an output file sim.out

```

```

d # write the results for regression method 1
dr

d2 # write the results for regression method 2
dr2

d3 # write the results for regression method 3
dr3

sink() # close the file

```

The following is the subroutine, or S-PLUS function, that permutes the vector of seed:no-seed decisions. Firstly temp is assigned the vector of zeros. A for loop then permutes s1 and puts the values in s2. The double ratios are calculated for each of the three regression methods and the results added row wise to a matrix temp. The matrix, minus the initial row of zeros, is then output as the result of the function call.

```

perm<-function(s1,tar,et1,et2,et3,p1,p2) {
  temp<-c(0,0,0,0)
  for(i in 1:1000) {
    s2<-sample(s1)
    dr<-(sum(s2*tar)/sum(s2*et1))/(sum((1-s2)*tar)/sum((1-s2)*et1))
    dr2<-(sum(s2*(1-p1)*p2*tar)/sum(s2*(1-p1)*p2*et2))/(sum((1-s2)*
      (1-p1)*p2*tar)/sum((1-s2)*(1-p1)*p2*et2))
    dr3<-(sum(s2*(1-p1)*p2*tar)/sum(s2*(1-p1)*p2*et3))/(sum((1-s2)*
      (1-p1)*p2*tar)/sum((1-s2)*(1-p1)*p2*et3))
    temp1<-c(dr,dr2,dr3)
    temp<-rbind(temp,temp1)
  }
  temp<-temp[2:1001,..]
  temp
}

```

APPENDIX 6 - SUMMARY OF STEPS TAKEN IN SIMULATIONS

Below is a summary of the steps taken in the simulation analysis of historical precipitation data for the Copeton Dam region.

1. All raindays from the period 1964 to 1988, that also had data on the wind direction, were uniquely numbered starting with the number 1.
2. These numbers were then used to randomly generate a subset of 15 days per year for the required number of years of the experiment (5 or 6 years).
3. Using these numbers the dates and hence the rainfall and wind directions were known for each of the raindays generated as the subset.
4. The random selection of whether a day was to be seeded or not seed was then carried out according to a 1:1 or 2:1 randomisation.
5. The target rainfalls were then increased by the desired amount (10, 20 or 30 percent). An estimate of the amount of rainfall which would have fallen if cloud seeding had not occurred was calculated using the regression methods.
6. Using the double ratio test it was calculated whether the seeding effect was significant at the desired level.
7. The above steps were repeated 100 times to give a percentage chance of detection of the seeding effect.
8. All the steps were then repeated three times to give an average chance of detection of a seeding effect and to reduce the effects of variability.
9. The number of years of the experiment, the seeding ratio or the percentage increase due to seeding were then changed and steps 1 to 9 repeated until all combinations were exhausted.