SSEC No.86.06.M1.

THE SCHWERDTFEGER LIBRARY
1225 W. Dayton Street
Madison, WI 53706

Final Report to NOAA on

MICROWAVE-INFRARED RAIN ALGORITHMS FOR THE SSM/I

A REPORT

from the space science and engineering center
the university of wisconsin-madison
madison, wisconsin

Final Report to NOAA on

MICROWAVE-INFRARED RAIN ALGORITHMS FOR THE SSM/I

For

Contract No. NA-84-DGC-00240

For the Period

28 September 1984 - 31 March 1986

Submitted by

David W. Martin
Principal Investigator

Barry B. Hinton
Co-Investigator

Space Science and Engineering Center
at the University of Wisconsin-Madison
1225 West Dayton Street
Madison, Wisconsin 53706
(608) 262-0544

June 1986

ABSTRACT

We have studied rain rate estimation using data from the subset of Nimbus-7 SMMR (Scanning Multichannel Microwave Radiometer) channels close in frequency to the channels to be used on the DMSP SSM/I (Special Sensor Microwave Imager) --except for the 85 GHz channel for which there is no SMMR analog. We concluded that SSM/I rain estimates should be within 13% of SMMR estimates in accuracy, and somewhat better if 11 μm IR data were used with it. A potential problem raised in this research is the possibility that the coefficients of the regression equations used to recover rain rates from brightness temperatures are "unstable". A means of dealing with this is suggested.

TABLE OF CONTENTS

MICROWAVE-INFRARED RAIN ALGORITHMS FOR THE SSM/I

1. Introduction

This report summarizes the results of research aimed at further development of methods of inferring rainfall over agricultural lands using remotely sensed data in the microwave portion of the electromagnetic spectrum. Specifically, having noted the success of the methods developed for the five-frequency dual-polarized channels of the SMMR instrument on Nimbus-7, we have assessed the prospects for similar success using the SSM/I instrument to be flown soon on a DMSP spacecraft. Because the SSM/I channels are near the SMMR channels in frequency, we have carried out our work using actual SMMR data. Table 1.1 lists the characteristics of SMMR. However, the SMMR channels are not all represented by analogous SMM/I channels as shown by Table 1.2. For the 85.5 GHz SSM/I channel there is no history of available satellite observations from SMMR, or elsewhere.

Our main objective is to find out whether the performance of SSM/I should be expected to be comparable to SMMR as a rain measuring instrument. The work contained in this report also considers whether 11 μm infrared channel data could be used to enhance the performance of the microwave instrument. This question has received relatively little attention.

In order to improve the flow of the text one of our procedures, variable transformation, has been rationalized in an appendix. Similarly, because our methods are statistical and have generated much tabular material, longer tables not requiring immediate inspection have been placed in an appendix. Most of the tables, however, are at the end of the section in which they are introduced. Figures have been placed similarly.

Table 1.1.
SMMR Characteristics

Freq. GHz.	Polar-ization	Footprint (km by km)		T Acc. Deg. K	T Res. Deg. K
6.63	h and v	121	79	2.0	0.51
10.69	h and v	74	49	2.0	0.72
18	h and v	44	29	2.0	0.89
21	h and v	38	25	2.0	1.01
37	h and v	21	14	2.0	1.23

Table 1.2.
SSM/I Characteristics

Freq. GHz.	Polar-ization	Footprint (Km by km)		T Acc. Deg. K	T Res. Deg. K
19.35	h and v	70	45	1.5	0.80
22.24	v only	60	40	1.5	0.80
37	h and v	38	30	1.5	0.60
85.5	h and v	16	14	1.5	1.10

2. Background

Electromagnetic radiation interacts with precipitation as well as the gaseous components of air --including the moisture required to form precipitation. The interaction processes include emission, absorbtion and scattering. Although the fundamental theoretical principles are well understood, the application to rain clouds is difficult. [See, for example, Ulaby *et al.*, 1981; Ishimaru, 1978; Fang and Chen,1982; Fung, 1982; Ishimaru *et al.*, 1984; and many others] Consequently, it is fair to say, success in theoretical analysis of environmental observations requires artful use of approximations. [See Weinman and Davies, 1978; Savage, 1976, 1978; Wu and Weinman, 1985; Olson, 1985; and others.]

In this situation empirical studies play a dual role. First, They either give a directly useable answer --an observed relation between radiation and rain-- or they may suggest that there is *no simple answer*. Fortunately, the "no-answer" result has not been obtained as Table 2.1 illustrates. Second, the observed relations represent a body of facts which may constrain the theoretical approximations even if the observations are not useful directly.

It is interesting to note that the theoretical results have treated ensembles of precipitaing *model clouds* --systems whose parameters are specified *a priori*. The modelled radiative transfer results are treated somewhat like empirical studies to obtain the desired rain-radiation relationships [*e.g.* Wilheit and Chang, 1980].

Infrared (11 μm window) data has been much studied for rain estimation but without concurrent use of microwave data. IR methods are reviewed by Barrett and Martin [1981]. Microwave rain studies in the SMMR context may be found in Spencer *et al. [1983a,b,c] and Spencer and Santek [1985].*

Table 2.1.
Rain Estimation Skill Exhibited with SMMR Data[1]

Season	No. Records	Variance Explained
Spring	4753	39.1 %
Summer	4663	63.1 %
Fall	2615	56.5 %

[1] Data in the table are from Spencer [1984].

3. General nature of model studies

Because of the interplay of theory and data, it is useful to introduce here a few illustrations of expected relations between microwave temperature and rain rate from one of the newer calculations. This will suggest some approximations and notation to be used below.

Fig 3.1 and Fig 3.2, plotted from the tabulated model results of Wu and Weinman, [1985] are prototypical. The following features are of interest. At low rain rates over a land background the brightness temperature first increases with rain rate to a maximum value which varies with frequency. The maximum brightness temperature is greater at lower frequencies (except near 21 GHz) and the rain rate corresponding to the maximum is higher at low frequencies. Note that these curves assume a constant value of land surface emissivity for all frequencies shown and both v- and h-polarization. Usually there are both frequency and polarization variations in emissivity, so that the low rain rate ends of the curves would not necessarily approach a common origin. These details are not important to the present discussion.

The bottom panel of Fig. 3.1 is an alternate presentation of the same information as given in the top panel. There are several reasons for using the square root of the rain rate as the rain variable. Most rain occurs at low rates. That is, the distribution function is very peaked near zero and has a long tail as shown in Fig 3.3, which also shows how the distribution appears when the square root function is used. Statistical techniques which we shall use below may give undue emphasis to the small number of high rain rate events in the data set if the rain rate itself is used. Also validity tests assume that residuals from fitted curves are normally distributed. This, it turns out, is more nearly the case if a power less than one of the rain rate is used. This will be discus-

sed elsewhere in this report.

At present we are interested in a feature which can be seen clearly in Fig 3.2, that the right hand, or "above maximum temperature" branch of the curve is more nearly linear for the square root function. This affords some simplification of the treatment of the data if this functional variation is a good approximation. On the other hand, analyses which are linear in the rain rate *and* the brightness temperature admit a principle of superposition. Accordingly, an average rain rate could validly be calculated from average brightness temperatures over an area --independently of the size of the area. Superposition eliminates some ambiguity due to the so-called beam filling problem [Spencer, 1986].

In the end, as the reader will discover below, we have elected to mix our analyses --sometimes using the rain rate, sometimes the square root of the rain rate, and in a few instances other the natural logarithm of one plus the rain rate.

Fig 3.4 is an idealization of Fig 3.1 and Fig 3.2 which we shall use below. There is a similar curve for each frequency, f and polarization state, (h or v). The main simplification is that the two branches of the curve are taken to be straight lines in the right panel. C^* is the land brightness temperature, C is the zero-intercept of the high rain rate branch, which has a constant slope, q, in the right panel. T_{MAX} is the maximum brightness temperature which occurs for a low value of the rain rate, ~1 to 3 mm/h for 18 and 37 GHz. The difference between the maximum brightness and the surface brightness is ∂T. The quantities C^*, C, q, and T_{MAX} are all functions of polarization, h or v, and frequency, f. To denote these dependencies we will use a subscript.

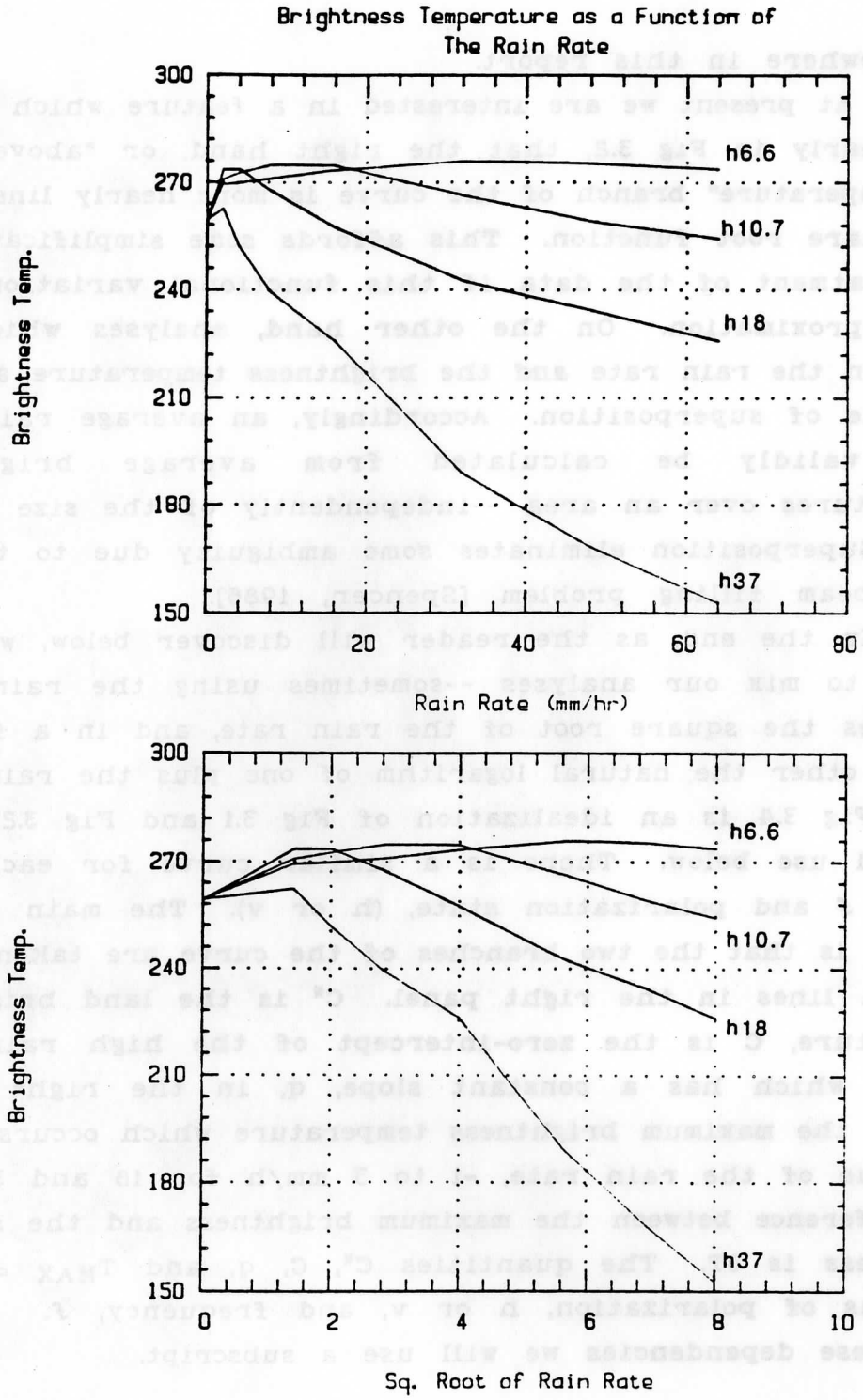


Fig. 3.1 Brightness temperature variations as functions of the rain rate and the square root of the rain rate for several frequencies. All curves shown are for horizontal polarization. Polygonal lines have been drawn between points extracted from tables in Wu and Weinman [1984].

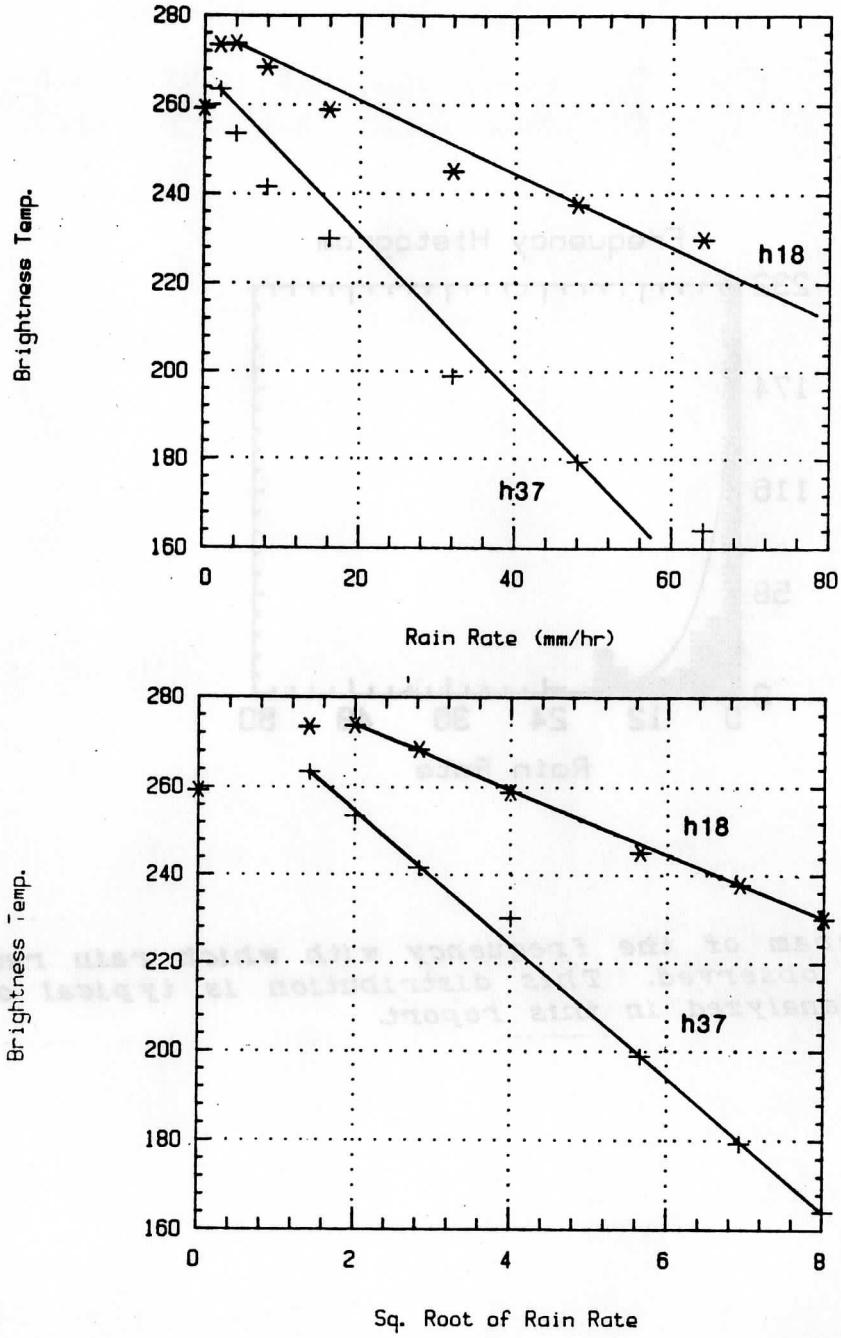


Fig. 3.2 Linear idealizations of brightness temperature functions. The points are from Wu and Weinman [1984]. Stars represent horizontally polarized 18 GHz brightness temperature, crosses 37 GHz horizontally polarized brightness temperature. In the top panel the points are plotted against the rain rate, R. The lower panel shows the same points plotted against the square root of R. Straight lines have been drawn between the maximum brightness temperature points and those corresponding to 48 mm/h to illustrate how well we might expect a linear curve to fit the portions of the curves above the maximum brightness temperature points.

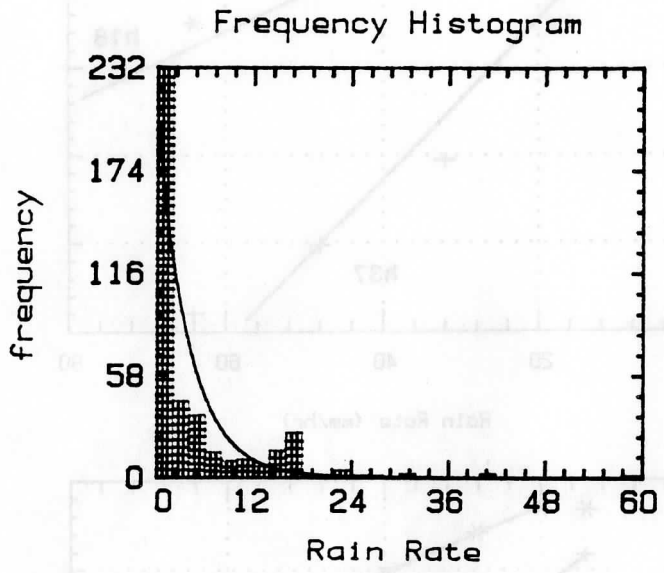


Fig. 3.3 Histogram of the frequency with which rain rate intervals were observed. This distribution is typical of the data files analyzed in this report.

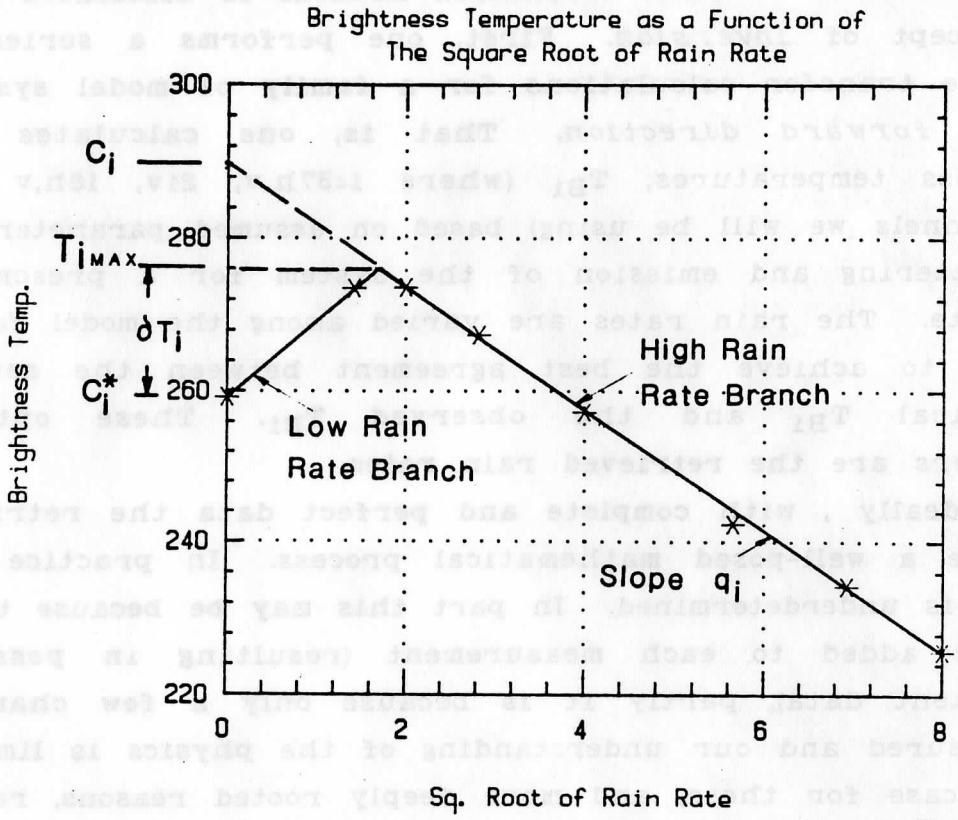


Fig. 3.4 Idealization of brightness temperature variation with the square root of rain rate and parameter defintions. C_i is the intercept at Q=0 of the linear extrapolation of the high rain rate branch of the curve, which has a slope q. C^* is the observed mean brightness temperature for Q=0 and is used to define the intercept of the low rain rate segment which connects it to the maximum brightness temperature, T_{MAX}. The difference between T_{MAX} and C^* is ∂T.

4. Retrieval methods

One class of rain estimation methods is associated with the concept of *inversion*. First, one performs a series of radiative transfer calculations for a family of model systems in the *forward direction*. That is, one calculates the brightness temperatures, T_{Bi} (where i=37h,v, 21v, 18h,v for the channels we will be using) based on assumed parameters of the scattering and emission of the system for a prescribed rain rate. The rain rates are varied among the model family members to achieve the best agreement between the set of theoretical T_{Bi} and the observed T_{Bi}. These optimal parameters are the retrieved rain rates.

Ideally , with complete and perfect data the retrieval would be a well-posed mathematical process. In practice the problem is underdetermined. In part this may be because there is noise added to each measurement (resulting in possibly inconsistent data), partly it is because only a few channels are measured and our understanding of the physics is limited. In any case for these, and more deeply rooted reasons, remote sensing solutions are not necessarily unique. This is the reason that constraints must often be imposed --to make sure that the "correct" solution is chosen.

One means of devising these constraints is to obtain data sets covering ranges of the conditions under which one intends to apply the retrievals, and to relate the $\{T_{Bi}\}$ empirically to a set of rainfall measurements. In other words, one obtains $T_{Bi}(R)$, brightness temperature as a function of rain rate, R.

In this report we will discuss empirical brightness temperature vs. rain rate relations from an amalgamation of 25 data sets used previously in studies presented by Spencer [1984]. The over-all set includes three seasons (spring, summer, and fall) taken from portions of 25 SMMR passes over

the United States. It consists of approximately 12 000 data "records" each containing a radar-derived rain rate for "truth" in addition to SMMR brightness temperatures. For further discussion see Spencer's paper cited above.

The radar data was obtained from microfilm records of the displays of operational radars. These images were then digitized on a uniform grid. The data have several non-ideal characteristics which must be noted. Each film record displayed at most seven radar reflectivity levels, each interpreted *nominally* as one of seven rain rates, in mm/h, which are:

0, 4, 17, 42, 85, 147, 190

This ordering of discrete levels results in a standard error of about 60% --even if no other source of error is present. For rates under about 44 mm/h the error is about 2.5 mm/h. Rain rates other than those in the above list enter only through area weighted averaging over the 20 km by 20 km areas of each grid cell. Nevertheless, the "magic numbers" (0, 4, 17, ...) persist in the data. This results in a jagged frequency distribution of rain rates with a notable peak at 17 mm/h. This data set focussed on "rainy" cases. Further, since rain rates above 42 mm/h are rather rare, and there is a very large peak at zero, the shape of the frequency distribution requires attention.

If we smooth out the subsidiary peaks in the distributions we note the features commonly found: the very large peak at zero, mentioned above, and a very long "tail". Finally it is significant that rain rates are non-negative. We show in Appendix A that the "magic number" rain rates are an annoyance, but probably not a serious issue in our statistical interpretation of the data. The peak at zero, the error which grows linearly with rain rate and the non-negativity are of

greater concern, as they could affect the results of the regression procedures we will use as well as significance tests.

The conventional prescription for "pathogenic" data distributions is to transform the original variable, R, to another, Q, for which the distribution has a more desirable shape. (See Appendix A and Johnson and Wichern [1982].) If there is particular interest in regression, we try to find a transformation, such that both Q and the whole set of residuals of Q from the truth data, have normal distributions and further that the variance of the residuals is approximately constant --not a strong function of Q.

For the present data sets two simple transformations were found which meet these objectives somewhat better than the original variable, R (rain rate). These are,

$$Q = \sqrt{R} \quad \text{and} \quad L = \ln(R + 1)$$

Further details may be found in Appendix A. Also, for the additional reasons discussed in Sec. 3, regressions linear in $Q = \sqrt{R}$, rather than R are convenient.

The constants and coefficients obtained by regression for each channel form a "data set" having a length of 25. These may be found in Table B.1, Appendix B. The results are summarized in Table 4.1 and Table 4.2 for the regressions on Q. These 25 determinations can then be analyzed for correlations between the constants C_i and C_i^* and the slopes $q_i = dT_{Bi}/dQ$. The mean values, variations and interrelations among the 125 C_i, C_i^* and q_i in these relations might be taken to circumscribe the allowable radiative transfer model. The relationships of these parameters to others shown in further tables can be seen in Fig. 3.4, which illustrates them schematically. Because of the varying size and validity of the data sets, some of the combined analyses (shown in Table

4.1 and Table 4.2) have been weighted by the standard errors of each parameter.

Even a very quick examination of Tables 4.1 and 4.2 (as well as others to follow) reveals that the parameters introduced in Fig. 3.4 are not "well determined" in the sense that the spreads, as shown by the standard deviations, the interquartile ranges, or the extreme ranges, are large relative to the means. Thus each parameter is characterized by a statistical distribution of values. Note that these distributions are mostly quite skewed or asymmetrical. (Positive skewness denotes a "tail" extending to higher values --to the right in a conventional plot; negative skewness extends to the left --toward lower , or more negative, values. Kurtosis describes how pointed or flat the peak of the distribution is. (Large kurtosis signifies a pointed central maximum. A normal probability curve has a kurtosis of 3.0, to which the "standard values" are referred.)

Even a casual examination of Table B.1, Appendix B, suggests that there are correlations among the q's and the C's. This conjecture is borne out and quantified in Table 4.3, which is a correlation matrix. A simplistic interpretation is that the "brightness temperature of rain" at a particular rain rate has a tendency toward constancy. The land surface brightness temperature, which is partially observed at very light rain rates and at heavier rain rates through incompletely covered pixels, actually varies somewhat independently. However in our curves, which were fit to data subject to the condition that $R \geq 1$ (or $Q \geq 1$), the constant terms in each of the 25 data sets do *not* represent a typical surface emission temperature for that data set as one might expect. Rather, the constants are just the zero-intercept of the *extrapolated* rain rate greater than one mm/h branch of the curve as shown in Sec. 3. In a sense then, because no data were used with Q or $R \leq 1$, the constants reflect the statistical

determination of the quantity

$$C_i = [\ T_{Bi}(Q=1) - q_i(Q=1)\ ]$$

Thus one should *expect* the slope coefficients (q_i) to be correlated with the C's.

The single most interesting feature among Tables 4.1-4.7 relates to $T_{MAX,i}$ (Table 4.5), and is further illustrated in Fig. 4.1 (for i = h37). This variable is more normally distributed than C_i (Table 4.1), C_i^* (Table 4.4), or ∂T_i (Table 4.6) based on the skewness and kurtosis values. In addition, based on the variance values,the relative spread is less.

T_{MAX} has been used in connection with choosing threshhold brightness temperatures for rain/no-rain descions. Thus it is of special interest for this reason, in addition to separating the two branches of the $T_{Bi}(R)$ or $T_{Bi}(Q)$ curves. Consequently, we have briefly investigated whether any natural variations of $T_{MAX,i}$ can be accounted for operationally. For this purpose possible correlations with C_i are of no use, since this quantity is an artifact --not an observable or measurable one. On the other hand, one can quite naturally interpret C_i^* as representing a meaningful geophysical parameter. It is just the *surface brightness temperature.* This can be estimated from surface thermometric temperature and microwave emissivity, ϵ_i. The surface temperature,T_S, may be obtainable from weather reports or from climatology. In some cases T_S might be found from surrounding microwave data from areas known to be non-raining as indicated by the absence of clouds in visible and/or infrared satellite images. Table 4.7 lists the correlations between the T_{MAX}'s and the C_i^*'s which are large enough to warrant further study.

The quantity ∂T_i, displayed in Table 4.6, can be interpreted as an estimate of the characteristic slope of the low

rain rate branch of the $T_{Bi}(R)$, or $T_{Bi}(Q)$ function --on the assumption that these functions can sensibly be approximated as linear and have their maxima near 1mm h^{-1}. This may be a fairly good approximation for 37 Ghz., but is unlikely to be valid at the lower frequencies. The available truth data is inadequate to investigate empirically the detailed shape of this branch of the curve. Radar data contain almost no information on rain rates less than 4 mm/h due to the digital encoding scheme used. Consequently, reliance on theoretical guidance or intuition may be required. Nevertheless ∂T_i is valuable for evaluating one constant associated with the low R branch once the shape of this branch is selected, as it is automatically in the *forward* radiative transfer problem when one selects a specific model.

Table 4.1.
Statistical Properties of the C_i-Values

	h, 37	v, 37	v, 21	h, 18	v, 18
Unweighted statistics					
Variance	63.15	30.99	29.68	93.32	49.34
Std. Deviation	7.95	5.57	5.49	9.66	7.02
Range	32.63	19.70	22.56	39.20	26.23
Skewness	0.41	0.10	0.31	0.34	0.00
Std. Value	0.84	0.20	0.64	0.70	0.00
Kurtosis	2.72	2.09	2.75	2.61	2.26
Std. Value	-0.29	-0.93	-0.26	-0.40	-0.76
Mean	260.00	268.00	262.62	251.15	259.91
Weighted values					
Mean	260.96	268.07	263.66	256.45	262.00

Table 4.2
Statistical Properties of the q_i-Values

	h, 37	v, 37	v, 21	h, 18	v, 18
Unweighted statistics					
Variance	6.55	5.15	2.45	6.10	3.34
Std. Deviation	2.56	2.27	1.57	2.47	1.83
Range	10.54	8.85	6.52	12.12	8.70
Skewness	-0.75	-1.08	-0.53	-1.29	-0.09
Std. Value	-1.53	-2.21	-1.09	-0.53	-1.29
Kurtosis	3.31	3.61	3.00	5.98	4.42
Std. Value	0.31	0.62	2.19	3.04	1.45
Mean	-3.22	-3.23	-0.60	-0.43	-0.05
Weighted values					
Mean	-2.86	-3.08	-0.51	0.07	-0.15

Table 4.3
Correlations Among The C_i and the q_i

		Slopes, q_i					Constants, C_i				
		v37	h37	v21	v18	h18	v37	h37	v21	v18	h18
	v37	1.00	0.97	0.74	0.52	0.10	-.59	-.66	-.48	-.39	-.50
	h37	0.97	1.00	0.77	0.57	0.12	-.61	-.69	-.51	-.44	-.55
q	v21	0.74	0.77	1.00	0.89	0.18	-.51	-.55	-.61	-.52	-.51
	v18	0.52	0.57	0.89	1.00	0.22	-.53	-.52	-.63	-.64	-.57
	h18	0.10	0.12	0.18	0.22	1.00	-.45	-.49	-.53	-.48	-.49
	v37	-.59	-.61	-.51	-.53	-.45	1.00	0.90	0.88	0.92	0.87
	h37	-.66	-.69	-.55	-.52	-.49	0.90	1.00	0.91	0.86	0.96
C	v21	-.48	-.51	-.61	-.63	-.53	0.88	0.91	1.00	0.93	0.94
	v18	-.39	-.44	-.52	-.64	-.48	0.92	0.86	0.93	1.00	0.93
	h18	-.50	-.55	-.51	-.57	-.49	0.87	0.96	0.94	0.93	1.00

Table 4.4
Statistical Properties of the C_i^*-Values

	h, 37	v, 37	v, 21	h, 18	v, 18
			Unweighted values		
Variance	119.66	35.05	43.93	198.39	71.31
Std. Dev.	10.94	5.92	6.63	14.08	8.44
Range	42.31	42.31	26.98	52.99	29.53
Skewness	-0.43	-0.74	-0.42	-0.31	-0.51
Std. Value	-0.87	-1.51	-0.85	-0.62	-1.05
Kurtosis	2.62	2.59	2.75	2.31	2.26
Std. Value	-0.39	-0.42	-0.26	-0.71	-0.75
Mean	249.20	261.11	258.35	242.56	255.04

Table 4.5
Statistical Properties of $T_{MAX,i}=T_{Bi}(R=1)$

	h, 37	v, 37	v, 21	h, 18	v, 18
			Unweighted values		
Variance	41.50	21.29	21.76	72.48	37.00
Std. Dev.	6.44	4.61	4.66	8.51	6.02
Range	28.14	18.28	20.12	35.60	23.60
Skewness	0.61	0.24	0.59	0.44	0.15
STd. value	1.24	0.49	1.21	0.89	0.30
Kurtosis	2.94	2.94	3.10	2.72	2.59
Std. value	-0.06	-0.06	0.10	-0.29	-0.41
Mean	256.78	264.76	262.02	251.57	259.86

Table 4.6
Statistical Properties of $\delta T = T_{Bi}(R=1) - C_i$*

	h, 37	v, 37	v, 21	h, 18	v, 18
			Unweighted values		
Variance	87.90	35.39	33.32	140.15	58.31
Std. Dev.	9.38	5.95	5.77	11.84	7.64
Range	36.38	24.95	22.54	49.61	32.55
Skewness	0.24	0.42	0.48	0.08	-0.10
Std. value	0.49	0.86	0.99	0.16	-0.20
Kurtosis	2.32	2.85	2.72	2.43	2.59
Std. value	-0.69	-0.15	-0.29	-0.58	-0.42
Mean	7.59	3.65	3.67	9.02	4.82

Table 4.7

Correlations of the T_{MAX}'s with the C_i^*'s

		$T_{MAX,\,i}$				
		v, 37	h, 37	v, 21	v, 18	h, 18
C_i^*	v, 37	0.38	0.43	0.47	0.47	0.44
	h, 37	0.25	0.52	0.50	0.39	0.53
	v, 21	0.29	0.50	0.52	0.42	0.51
	v, 18	0.38	0.50	0.52	0.48	0.50
	h, 18	0.27	0.54	0.52	0.40	0.55

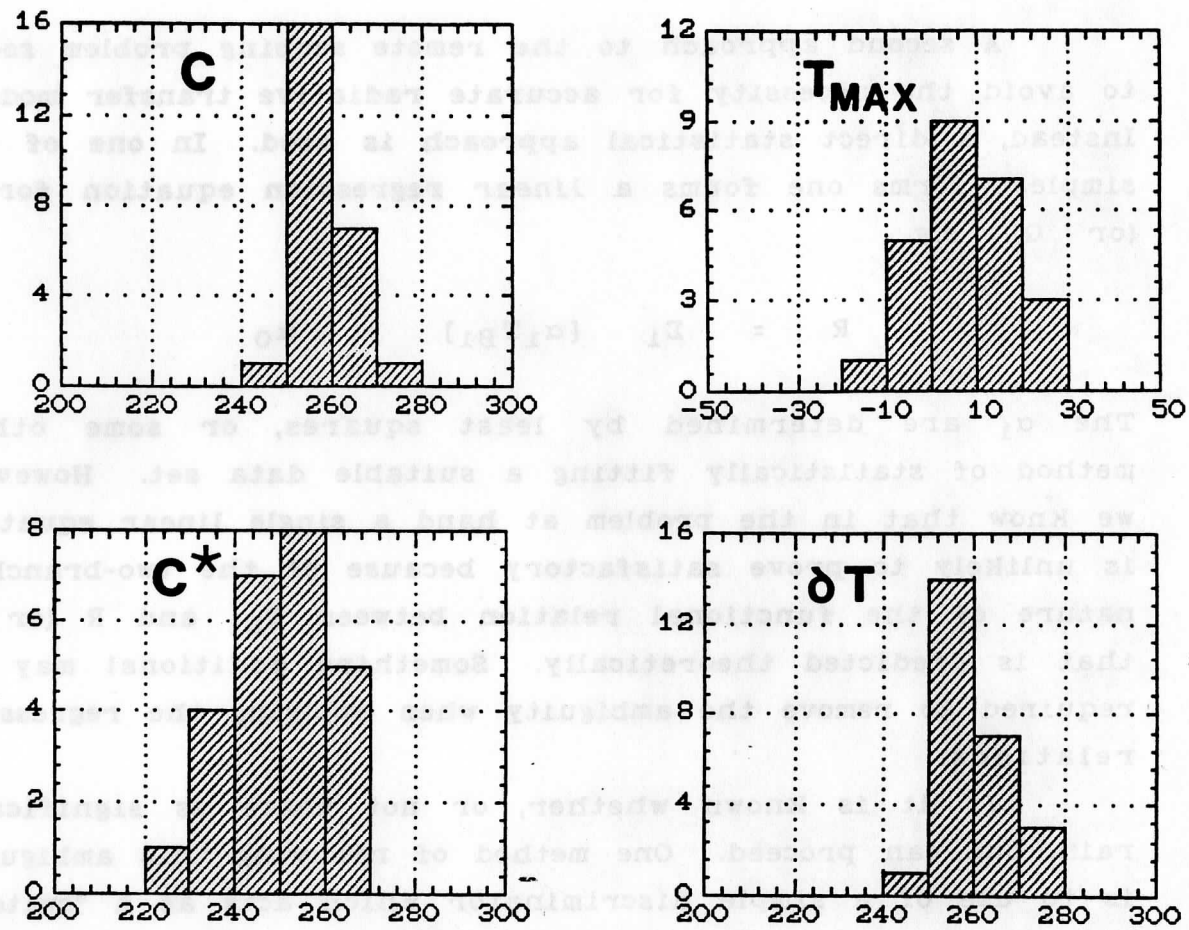


Fig. 4.1 Histograms of the values of the parameters defined in Fig. 3.4 which were recovered by regression on the 25 data sets. Only the results for horizontally polarized 37 GHz are shown. C_i is shown at upper left, C^* at lower left, ∂T at upper right and T_{MAX} at the lower right.

5. Direct or forward regression methods.

A second approach to the remote sensing problem seeks to avoid the necessity for accurate radiative transfer models. Instead, a direct statistical approach is used. In one of its simplest forms one forms a *linear* regression equation for R (or Q). *e.g.*

$$R = \Sigma_i [\alpha_i T_{Bi}] + \alpha_0$$

The α_i are determined by least squares, or some other method of statistically fitting a suitable data set. However, we know that in the problem at hand a single linear equation is unlikely to prove satisfactory because of the two-branched nature of the functional relation between T_{Bi} and R (or Q) that is predicted theoretically. Something additional may be required to remove the ambiguity when applying the regression relations.

If it is known whether, or not there is significant rain, one can proceed. One method of resolving this ambiguity is to use of a simple discriminator which acts as a "switch". Another approach is to apply an additional data set to the problem which does not share the same ambiguity.

For example, it has been found that polarization of the microwave brightness provides useful discrimination [Spencer, 1984]. Using statistical techniques one can construct "classification functions" for use as discriminators as we shall discuss below. To some degree the infrared radiance data available from weather satellites can remove the ambiguity -- at least in one direction. One is quite confident that the absence of cloud indicates that there is no rain. Additionally, cloud height, thickness and appearance are useful for discriminating between convective rain and stratiform rain. The former is more likely to be associated with the higher rain

rate branch, and the latter is often associated with the lower rain rate branch. Because, there is overlap in the rain rates produced by these two types of rain systems, misclassification errors are only modestly reduced, not eliminated.

With the foregoing discussion in mind concerning the limitations of *linear* models (which ignore the double branched nature of the $T_{Bi}(R)$ functions), we shall nevertheless present the results from a sequence of such models. This will be followed by results from models incorporating several forms of discrimination. For this purpose a special data set was constructed. This data set was designed specifically to realistically address the discrimination issue. Because the relative content of difficult cases was intentionally enriched in comparison with other data sets from which results have been reported, the results may appear less impressive.

Sets of registered images were displayed on a video monitor. Each set consisted of an image of T_{Bi} for i={37h, 37v, 21v, 18h, 18v} as well as radar images as digitized by Spencer [1984], and the infrared window channel from GOES. Data were selected over the coverage area on a 2· latitude by 2· longitude grid without regard to rain, cloud or any other conditions. Next, areas were sampled in which the radar indicated significant rain rates. In order to eliminate spurious degrees of freedom, care was taken that these samples were not taken too close together. Presumably cloud physical parameters, such as drop size distribution, thickness, and temperature are correlated over significant distances within a given cloud, so that this extra care may be required to insure independence. After this, data were selected from areas close to and surrounding the significant rain samples, but which had light rain rates. Finally, data at points for which the radar indicated no rain were chosen "intimately" near the rain areas.

Thus, the data set is biased toward the difficult

cases. Consequently, as mentioned above, results from the regression models (*e.g.* standard errors of estimate, correlation coefficients *etc.*) may not appear as impressive as previously. However, because the data set is "tuned" to eliminate the trivial cases the resulting models should perform better in non-trivial cases.

Insight into the general nature of the data is readily obtained from the correlation matrix displayed as Table 5.1. Further insight can be gotten by looking at partial correlations (shown in Table B.2, Appendix B, for several subsets of variables). These measure the relationships between pairs of variables *while controlling for the possible effects of the other variables*. This is useful for uncovering both unsuspected, or "hidden" relationships, and spurious ones. Our purpose is to get insight into the strength of IR data relative to microwave data.

The simplified exegesis of Table 5.1 and Table B.2 is this: IR is the best single channel, although it is not indespensible.

The linear model results are summarized in Table 5.2. The coefficients and more detailed results are in Table B.3, Appendix B. Table 5.2 strongly suggests that the 18 GHz channels contribute little to the skill of linear models, as we have tested them on this data set oriented toward the discrimination problem. One might surmise that the principal utility of 18 GHz would be in the determination of higher rain rates which are not adequately represented in the data. This could have been anticipated on the basis of the results in Sec. 4 which found an 18 GHz channel sensitivity less than expected on the basis of the Wu and Weinman models.

Table 5.1
Correlation Matrix for Microwave, IR and Rain Data

	R	Q	IR	v37	h37	v21	v18	h18
R	1.00	0.93	-0.42	-0.42	-0.32	-0.28	-0.17	-0.12
Q		1.00	-0.48	-0.43	-0.32	-0.29	-0.18	-0.13
IR			1.00	0.45	0.34	0.39	0.29	0.23
v37				1.00	0.92	0.84	0.71	0.72
h37					1.00	0.81	0.77	0.85
v21						1.00	0.81	0.81
v18							1.00	0.86
h18								1.00

Table 5.2
Performance of Linear Models

Model* Variables	Std. Error of Estimate of Q	R^2 (Adjusted for Degrees of Freedom)
v37, h37, v21, v18, h18	1.350	0.243
v37, h37	1.378	0.211
IR	1.366	0.224
v37, h37, IR	1.306	0.291
v18, h18, IR	1.368	0.223
v37, h37, v21, v18, h18, IR	1.278	0.321

*All models incorporate a constant term.

6. Models with simple discrimination.

Precedent [Spencer *et al.* 1983; Spencer, 1984] as well as preliminary examination of the data suggest that two simple variables should be valid discriminators. The first one is the 37 GHz polarization, P_{37}, defined as (Tv37-Th37). A threshhold function S_{37} is defined as:

$$S_{37}(P_{37}) = \begin{cases} 1 & \text{for } P_{37} \leq P_{thresh} \\ 0 & \text{otherwise} \end{cases}$$

The second discriminator is an infrared threshhold function, $S_{IR}(T\text{-}T_{thresh})$, defined as:

$$S(T_{IR}) = \begin{cases} 1 & \text{for } T_{IR} \leq T_{thresh} \\ 0 & \text{otherwise} \end{cases}$$

The physical concept which underlies S_{37} is that upwelling radiation from land is slightly polarized, from water bodies markedly polarized, while wet land is an intermediate case. Water surfaces--and wet land--tend to look radiatively cold like rain, which depolarizes upwelling radiation, although ice is thought to contribute to polarization by scattering. In the investigations cited above it was concluded that a value of about 15° for P_{thresh} gave satisfactory results for the elimination of "wet" surfaces. In the case of $S(T_{IR})$, a number of investigators have concluded that clouds warmer than 270°K are unlikely to rain and that raining clouds are usually colder than -20°C (253°K). As a specific example consider the rain rate algorithm of Robertson [1985], shown in Fig. 6.1. Roughly speaking the onset of rain is very near 250°K, with only a very slight "tail" extending to about 255°.

This has also been investigated extensively by the present investigators as well as many others [Barrett and Martin 1981]. The effects of S_{37} and $S(T_{IR})$ are illustrated in Fig. 6.2, which should be compared with Fig. 3.3.

We have defined a family of functions resembling the shape of the negative of Robertson's curve. One of these is shown as an example if Fig. 6.3. The equation defining these is:

$$t(T_{IR},T_{thresh}) = S(T_{IR})\cdot(T_{IR}-T_{thresh})$$

Table 6.1 shows how these t-functions correlate with the two precipitation variables Q and R. Surprisingly, for Q the maximum is only approached at 275°K. In reality, the threshhold effect is *nil* in this case--very little data is being excluded. On the other hand, the maximum for R is at about 252°K, very near the expected temperature. Note, however, that even in this case the maximum is not pronounced. Thus, it turns out that threshholding T_{IR} is unlikely to materially change mean or mean square error values of rain rate determinations. It *will* reduce the tendency linear models have for producing negative rain rates. The difference between the R and Q vs. T_{IR} behavior arises because Q gives greater relative weight to low and moderate rain rates. These are more likely to occur at comparatively warm temperatures.

It is also quite interesting to note that the correlation of $S(T_{IR})$ -- which can take on only the values 0 or 1 -- with Q can be as high as 0.41 for T_{IR} values around 250°K. This is shown in Table 6.2. If a sequence of linear models in t-functions is generated for various T_{thresh} values the results of Table 6.3 are obtained. Based on this table we concluded that any T_{thresh} value between 250° and 270° could be adopted with a similar effect. In the examples to be discussed below we have inclined toward the high end of this

range. This is "conservative" in the sense that less data is subject to truncation. In passing we note that the "best" of these equations for R is,

$$R = \begin{cases} 1.1969 - 0.2227 \cdot t(T_{IR},250) & \text{for } T_{IR} \geq 250^{o} \\ 0 & \text{for } T_{IR} \leq 250^{o} \end{cases}$$

Fig. 6.4 shows the distribution of P_{37}-values with rain rate, R, in the data set. Visual inspection of this figure suggests that there is no reason to select a value other than 15°C brightness temperature difference for P_{thresh}.

Now that definite values have been adopted for T_{thresh} and P_{thresh}, we can proceed to investigate a sequence of models with discrimination as mentioned above. The variables will be drawn from the set: {T_{37v}, T_{37h}, T_{21v}, T_{18v}, T_{18h}, T_{IR}, $t(T_{IR},270)$, $S_{37}(P_{37})$}. In this list, and subsequently we omit the subscript "B" from brightness temperature when the meaning is clear. One consistent way to incorporate the effects of $S_{37}(P_{37})$ and $t(T_{IR},270)$ is to multiply each of the other variables (including the constant term) by them, or in effect by their product,

$$S_w = S_{37} \cdot S(T_{IR})$$

The effects of this combined threshholding are illustrated in Fig. 6.5. The regression results are summarized in Table 6.4. Not surprisingly, best results are obtained by using all "channels" in the data set. Further, comparison with Table 5.2 establishes that this type of simple discrimination has *not* (even slightly) improved the results as was anticipated above, in the discussion of Table 6.1. For reasons which will be apparent in Sec. 9, we have not included more detailed results in an appendix as we did for the simple linear regres-

sion models.

Table 6.1
Correlation of The t-Functions
With Rain Variables, Q and R

T_{thresh}	Correlation Coefficent R	Q
230	-0.408	-0.385
235	-0.419	-0.407
240	-0.425	-0.420
245	-0.433	-0.436
250	-0.437	-0.448
255	-0.437	-0.455
260	-0.435	-0.459
265	-0.435	-0.467
270	-0.433	-0.471
275	-0.420	-0.475
280	-0.420	-0.475

Table 6.2
Correlation of $S(T_{thresh})$ with R and Q

T_{thresh}	Correlation Coefficient R	Q
230	0.360	0.375
235	0.353	0.374
240	0.365	0.390
245	0.368	0.408
250	0.354	0.416
255	0.336	0.394
260	0.327	0.405
265	0.324	0.406
270	0.317	0.396
275	0.314	0.392
280	0.294	0.376

Table 6.3
Results from a Sequence of Linear Models for R
Using t-Functions for Various T_{thresh} Values

T_{thresh}	Std. Error of Est.	Corr.
230	6.483	-0.408
240	6.427	-0.425
250	6.386	-0.437
260	6.392	-0.435
270	6.398	-0.433
280	6.444	-0.420

Table 6.4
Performance of Nonlinear Models

Model* Variables	Std. Error of Estimate of Q	R^2 (Adj. for Deg. of Freedom)
v37, h37, v21, v18, h18, S_{37}, $t(T_{ir}, 270)$, $[T_{IR} \cdot S]$	1.303	0.294
V37, H37, V21, V18, H18, s_{37}, $t(T_{IR}, 270)$	1.305	0.292
v37, h37, S_{37}, $t(T_{IR}, 270)$	1.327	0.268
v37, h37, S_{37}	1.370	0.221
v37, h37, v18, h18, S_{37}	1.348	0.244
v37, h37, $t(T_{IR}, 270)$, S_{37}, $[T_{IR} \cdot S]$	1.328	0.267
$t(T_{IR}, 270)$, S_{37}, $[T_{IR} \cdot S]$	1.391	0.196
v18, h18, $t(T_{IR}, 270)$, S_{37}, $[T_{IR} \cdot S]$	1.392	0.195
v18, h18	1.555	0.000
v18, h18, S_{37}	1.509	0.054

*All regression equations include a constant term (in addition to S_{37} when it is shown in the variable list). The microwave channels are all implicitly operated on by S_{37}.

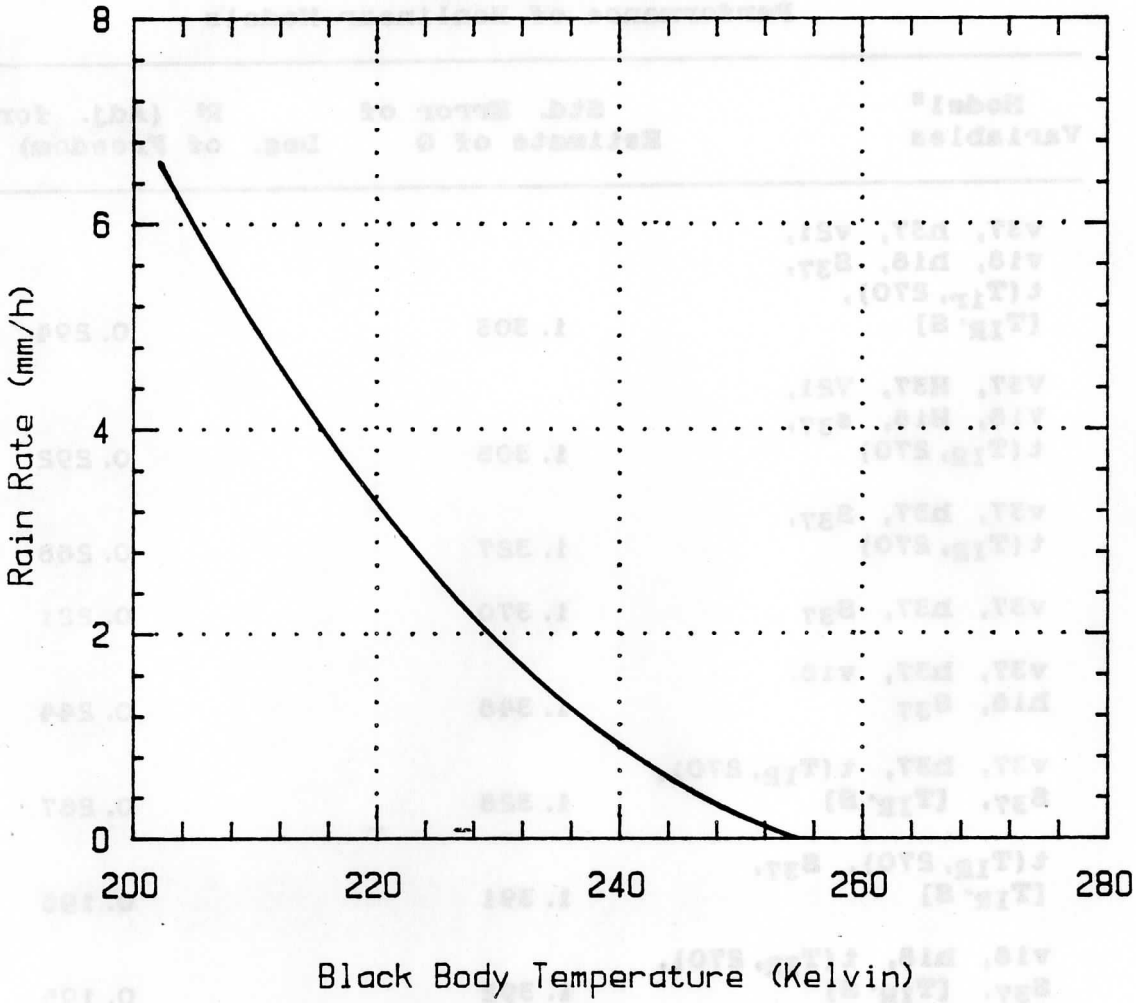


Fig. 6.1 Rain rate as a function of 11 μm IR temperature as given by Robertson [1985].

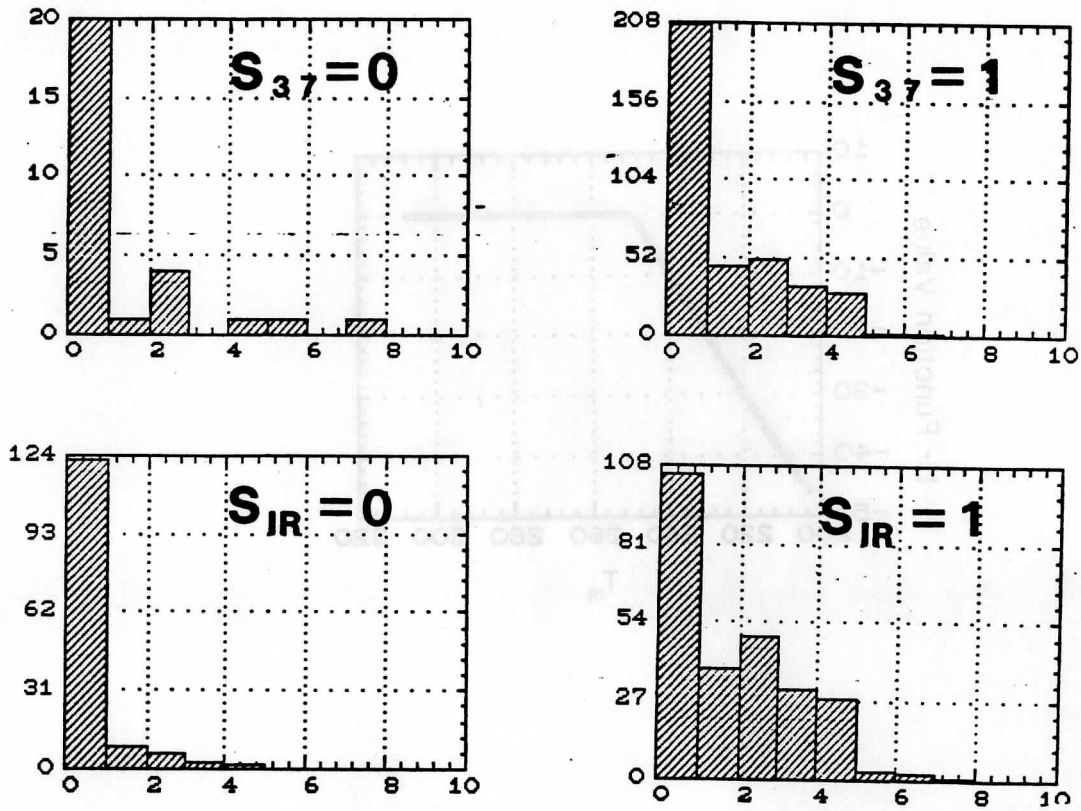


Fig. 6.2 Effects of threshholding. The top left panel illustrates the distribution of values of Q observed in a sample data set when S_{37} = 0, the right panel shows the distribution for S_{37} = 1. (The total data set would be the "sum" of the histograms.) Ideally, fewer Q=0 cases would survive in the right panel and fewer Q>0 in the left. The lower panels, for IR temperature, show more effective removal of the Q=0 cases and less leakage of Q>0 into the Q=0 class (on the left).

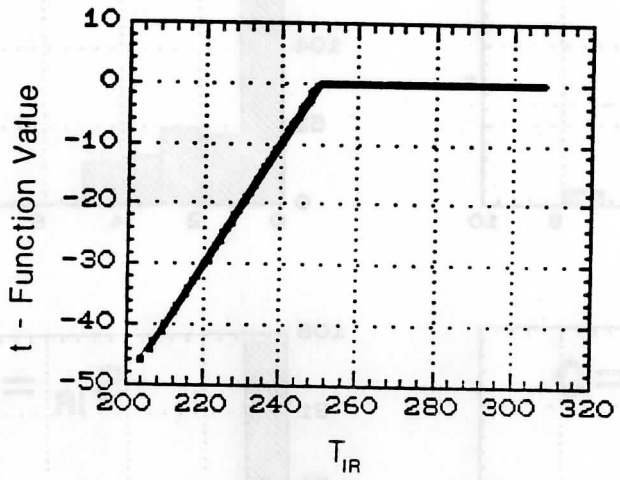


Fig. 6.3 A t-function for a threshhold value of 250 K.

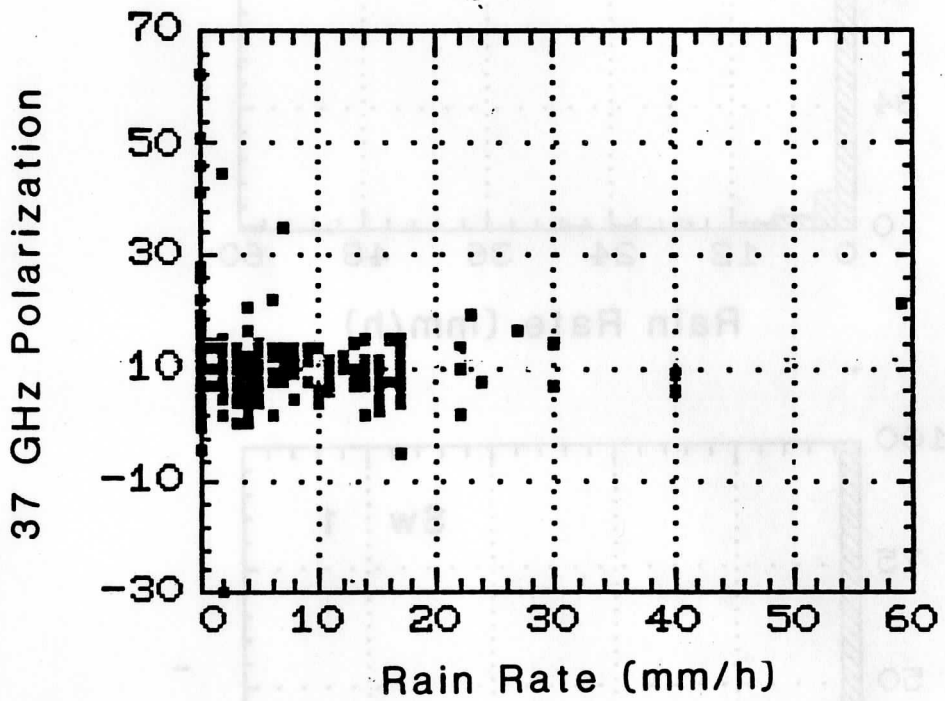


Fig. 6.4 Observed distribution of 37 Ghz polarization as a function of rain rate. This figure in accord with Fig. 6.2 suggests that 37 GHz is not an extremely efficient discriminator.

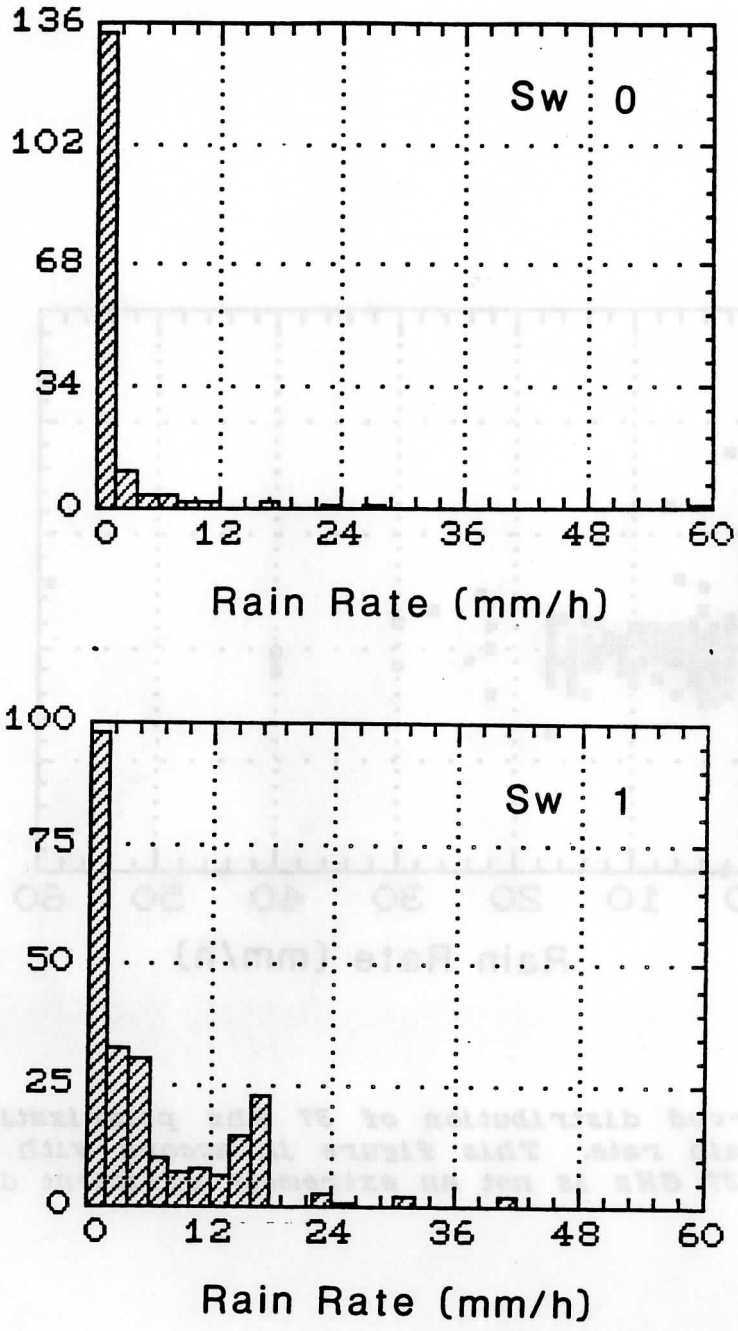


Fig. 6.5 Effects of combining the polarization and IR threshholding. More than half the Q=0 cases would be screened out with relatively little leakage of Q>0 cases into the Q=0 class.

7. Use of statistical classification functions.

In this section we discuss further work on the data set treated above. Each "data vector" ($T_{IR}, T_{37h}, \ldots, T_{18v}$) was assigned to a rain class, or group, according to the scheme,

Group 1: If R is between 0 and 1
Group 2: If R is under 5, & over 1
Group 3: If R exceeds 5

This resulted in the grouped data characteristics shown in Table 7.1. The coefficients of the discriminant functions are shown on the left in Table 7.2. In order to discuss the relative importance of the various variables in the discrimination process, it is simpler to look at the coefficients which would result if each variable were used in its standardized form. That is, if we were to remove the mean and divide by the standard deviation. Thus standardized variables all have the same mean (zero) and standard deviation (one). The corresponding coefficients of the standardized discriminant functions (which appear at the right in Table 7.2) are thus a direct measure of the "power" of an observed variable in influencing the predicted variable.

Because three groups were used there are two discriminant functions. The first one accounts for 97.51% of the variance accounted for by both and has a canonical correlation of 0.52614. Hence, it is moderately efficacious. The second function, on the other hand, is practically useless. Inspecting the standardized form coefficients for the first function, we conclude that three variables are important: T_{IR}, T_{v37}, and T_{h18}. Since there is only one useful function, it must be admitted that it amounts to little more than a linear regression equation with another transformation of the rain variable. Consequently, it does

not appear to be a valuable adjunct for use with a linear regression equation on Q or R as we had hoped. It is redundant. Nevertheless, Fig. 7.1 illustrates the rain classification according to this function.

In working with a number of files of data on discrimination functions *not* incorporating infrared data, we found that the canonical correlations were often somewhat higher than reported here (~0.7 usually). This certainly does not confirm a "negative impact" of infrared data. We suggest four explanations. First, this lesser value of the canonical correlation is a measure of the relative difficulty of our specially prepared data set. Second, although these various data files usually give similar canonical correlations, the discrimination functions themselves presented a surprising variability. This is because the very high correlations among the microwave channels render the determination of the regressions somewhat unstable. Thus, functions having quite dissimilar coefficients could really be substantially similar in terms of a hypothetical "principal component."

A principal component analysis disclosed that the "first" component, Θ_1, accounts for 84.8% of the variance in the standardized 18, 21 and 37 GHz channels. This component is a nearly equally weighted sum of all the channels (when the variables are expressed in standardized form). The second component, Θ_2, accounts for only 7.6% of the variance of the standardized microwave data. But it contains more information on rain (*i.e.* the variable R) than the first principal component. Together these two components account for 75% of the total rain variation explained by all five of these channels. Including a third component raises this to 89%. Thus The five channels are more or less equivalent to two or three independent variables. The first three principal components, in terms of the standardized brightness temperatures, π_i, are:

$$\Theta_1 = .459\pi_{37h} + .442\pi_{37v} + .451\pi_{21v} + .448\pi_{18h} + .437\pi_{18v}$$
$$\Theta_2 = -.333\pi_{37h} - .623\pi_{37v} - .017\pi_{21v} + .414\pi_{18h} + .573\pi_{18v}$$
$$\Theta_3 = .472\pi_{37h} - .075\pi_{37v} - .715\pi_{21v} + .480\pi_{18h} - .175\pi_{18v}$$

A third explanation for variation in the discriminant function coefficients may be the differing sensitivities and calibrations of the radars used in these various subsets of data. A fourth reason may lie in the differing background conditions effected by the underlying surfaces. Although an observed data record for rain almost always shows contrast with a nonraining data record, that contrast (which is what the statistical analysis "sees") depends as much on the nonraining background as on the rain. This deserves further attention.

Table 7.1
Characteristics of Grouped Data

	Group 1	Group 2	Group 3
Group counts			
	227	82	90
Group Means			
T_{IR}	262.03	240.78	230.78
T_{v37}	273.38	268.31	262.30
T_{h37}	263.78	258.48	252.71
T_{v21}	270.62	267.61	264.51
T_{v18}	271.09	269.28	266.01
T_{h18}	261.96	260.23	257.20
Group Standard Deviations			
T_{IR}	29.78	21.02	18.84
T_{v37}	10.30	11.58	11.59
T_{h37}	15.60	14.75	12.68
T_{v21}	9.58	8.90	7.10
T_{v18}	13.56	8.53	7.81
T_{h18}	15.08	13.30	11.12

Table 7.2
Coefficients of Discriminant Functions

Variable	Non-Standardized	Standardized
	First Function	
T_{IR}	0.02580	0.671170
T_{v37}	0.07531	0.818617
T_{h37}	0.00300	0.044420
T_{v21}	-0.00165	-0.014752
T_{v18}	-0.00852	-0.098490
T_{h18}	-0.03267	-0.045480
Constant	-16.31700	0.000000
	Second Function	
T_{IR}	-0.02479	-0.644622
T_{v37}	0.02603	0.282904
T_{h37}	0.08457	1.253340
T_{v21}	-0.05379	-0.480700
T_{v18}	0.06446	0.745065
T_{h18}	-0.07904	-1.100210
Constant	-5.17104	0.000000

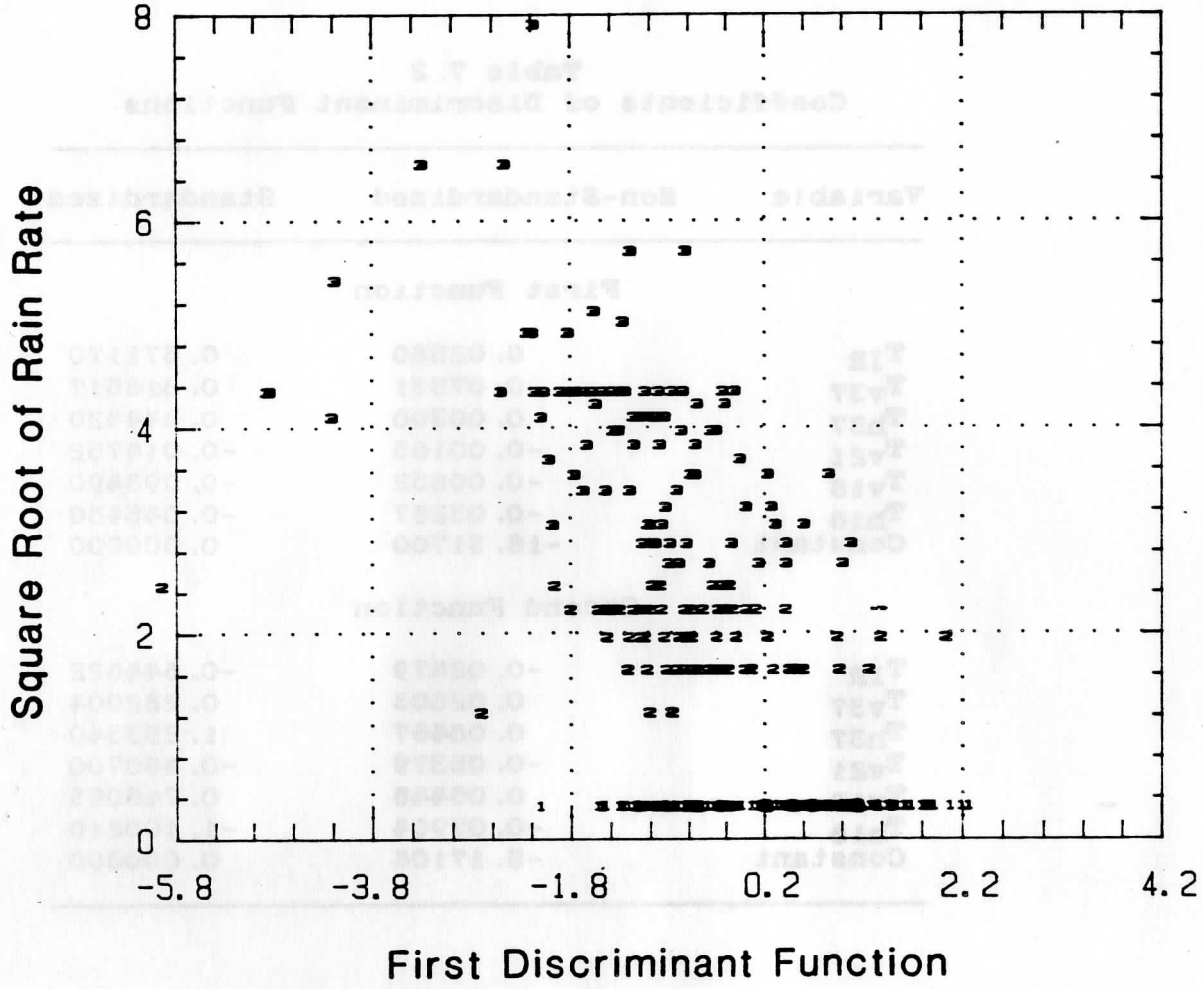


Fig. 7.1 Rain classification by the first discriminant function. Each point is encoded with its rain class. All the Class-1 points lie in the bottom row. The x-axis is the value of the discriminant function. For the rain parameter on the y-axis we have used Q, the square root of the rain rate. Consequently, the rain classes are perfectly dispersed in the y-direction.

8. Comparison of SMMR Performance To the SSM/I Channels

We approach the question of SSM/I's performance in the simplest way. Using the 25 data sets from three seasons (studied previously by Spencer, [1984]), we look at the differences in performances of "best" linear regression models for each data set. *This method was used because it is primarily a test of potential performances of channel differences, rather than a search for the single best algorithm.* We have studied the algorithm question in other sections of this report. In large part the differences in the retrieved regression coefficients for the various files in the total data set are due to the non-uniqueness arising from the inter-channel correlations. Thus, there are usually several algorithms which will work --we choose to compare best ones for individual cases. Similarly, differences in the sets of coefficients from file to file when the same set of channels is used, appear for the reasons stated in Sec. 7 for variability in discrimination functions. Wentz [1985] argues that stable algorithms are best obtained from what he terms "determinstic" systems.

Deterministic systems have *n* relevant physical variables measured by *n* channels which have projections onto the physical variables and a significant degree of channel-to-channel linear independence. On the whole, we believe that 18h, 18v, 21v, 37h, and 37v form such a set for three physical parameters, since we were able to form about three useful principal components in Sec. 7. The two 18GHz channels together provide a measure of the upwelling emission and polarization of the surface, 21v GHz a measure of moisture, while the "sum" of the 37 GHz channels provides the emission of the precipitation sized drops, and their difference provides a measure of the polarization of the overlying ice layer. In this view The SSM/I complement of channels is

adequate--the extra channels on SMMR extraneous. *Even when no extra channels are included, uniformities in one of the physical variables over one of the small data sets (or files), as well as differences in calibrations of the radars used as "truth" may induce variations in algorithm coefficients since the system is then over determined and unstable. Therefore, the pooled data set for determining a universal algorithm should span the entire range of variations in the physical variables.* Even so, as we discuss in Sec. 9, some special treatment may be required. There is a way of removing the overdetermination without throwing out data from any of the channels

The results are summarized in Table 8.1, while the details of the regressions are given in Table B.4, Appendix B. The table shows that the subset of SMMR channels which are analogous to SSM/I channels gives results quite comparable to the whole set of SMMR channels. The bottom line entries of the table suggest differences the order of ten percent in standard error of estimate.

Table 8.2 is based on summer data files, for which GOES 11 μm IR data were included in each record. This table indicates that IR augmented SSM/I results should not be significantly different than SMMR --*even if the new 85.5 GHz channel is not used.*

Recalling the discussion in Appendix A, we note that the variance of the residuals, and therefore the standard error, is not constant. This is especially true for the variable R. Consequently, the standard errors in Tables 8.1 and 8.2 are more representative of the standard error near zero rain rate.

Table 8.1

Summary Comparison of Simulated SSM/I and SMMR Channels

Season	Sq. of Correlation			Std. Error of Est.		
	SSM/I	SMMR	% Diff	SSM/I	SMMR	% Diff
			For The Variable R			
Spring	0.162	0.239	-32	1.524	1.435	6
Summer	0.332	0.349	-5	1.637	1.488	10
Fall	0.445	0.567	-27	1.977	1.626	22
Three Seasons	0.313	0.385	-19	1.713	1.516	13
			For The Variable Q			
Spring	0.166	0.239	-31	0.528	0.492	7
Summer	0.225	0.263	-14	0.533	0.511	4
Fall	0.471	0.491	-4	0.621	0.551	13
Three Seasons	0.287	0.331	-13	0.561	0.518	8

Table 8.2

Summary Comparison of SMMR and Simulated SSM/I Plus IR

	Square of Corr.		Standard Error	
	SMMR	SSM/I+	SMMR	SSM/I+
		For The Variable R		
Weighted Means	0.3955	0.4000	1.538	1.543
Mean Difference		0.00449		0.00473
RMS Difference		0.03506		0.05610
Mean % Difference		1.135		0.308
RMS % Difference		17.7		3.85
		For The Variable Q		
Weighted Means	0.3878	0.42105	0.474	0.458
Mean Difference		-0.03329		0.01581
RMS Difference		0.04623		0.02401
Mean % Difference		-7.904		3.335
RMS % Difference		11.0		5.06

9. Discussion of the stability of the regressions

At several points in the two previous sections we have alluded to the possibility that the mutual correlations of the predictor variables might render the regression coefficients "unstable". We will discuss this problem briefly in this section. The material largely follows Marquardt and Snee [1985]. Their article contains further references.

To counter colinearity of the predictor variables three kinds of procedures have been used:

(a) Stepwise regression including forward selection and backward elimination.

(b) Best subsets regression.

(c) Biased regression, including ridge regression, generalized inverse and latent root regression.

The first two (methods under *(a)* and *(b)*) reduce colinearity simply by eliminating offending variables from the equations. Marquardt and Snee contend that if the variance inflation factors or VIFs (to be defined) are large, use of biased estimators (*i.e. (c)*) is a better alternative than eliminating variables, which (after all) also removes *data.*

In their development, variables are transformed to standardized form by subtracting the mean and dividing by the standard deviation as we discussed in Sec. 7. In the regression equations the coefficients, or rather the *estimates of the coefficients* β_i have an uncertainty, or variance, represented by,

$$\mathrm{Var}(\beta_i) = \sigma^2[R_{ii}]^{-1}/[(n-1)s_i^2],$$

in which σ is the standard deviation of the observation error, R_{ii}^{-1} the i-th diagonal element of the inverse of the correlation matrix R_{ij} of the *predictor variables*, n the sample size and s_i the standard deviation of the i-th (non-standardized) variable. (Double subscripts or square brackets are used to distinguish the correlation matrix from the rain rate variable R.) The $[R]^{-1}$ factors clearly show that the uncertainty in the regression coefficient is an increasing function of the multiple-correlation, or "multicolinearity" of the predictor variables. The factor R_{ii}^{-1} *is* the variance inflation factor. To summarize, the VIF measures the collective impact of the correlations of all j-variables (i not equal to j) on the i-term regression coefficient. Clearly if the variance of β_i is large, the least squares fit is unlikely to perform well on new data--especially data outside the realm of the learning data set. The technique of reducing the ill-conditioning by discarding variables is based on the view that a predictor is either "good" or "no good." Marquardt and Snee suggest that a little bit of all the variables is better than all (of some) or none (of others).

In order to understand their solution, it is necessary to recall that least squares is an *unbiased* estimate. It turns out that one can sometimes obtain a *biased* estimate which has *less variance than least squares.* (See Fig. 9.1.) One method, ridge regression, often achieves a major improvement in over-all variance by allowing a little bias. In effect, one performs a least squares solution for the coefficient vector *not with just the correlation matrix* $[R_{ij}]=[R]$, *but rather with* $[R_{ij}] + \Theta[\delta_{ij}] = [R] + \Theta[I]$ *for* $\Theta \geq 0$.

If $\Theta = 0$, we obtain the usual least squares estimate. The ridge estimate gives *the smallest regression coefficients consistent with a given level of increase in the residual sum of squares.* The variance decreases as Θ increases. The bias

increases as Θ increases. Since the total mean square error is the sum of the variance and the square of the bias, there is some value of Θ for which the sum of squares of the coefficients --the square of the "length of the coefficient vector"-- is a minimum.

Because the predicted variable is given as $[\Sigma_i\ (\beta_i x_i)]$, that is as a linear transform of the β_i, the variance of the predicted variable will be a minimum when the mean square error of the coefficients is a minimum. These remarks follow *only* when standardized variables are used of course.

If it is advantageous to use the ridge regression technique, there is a procedure to aid in the selection of the bias parameter, Θ. This procedure, called the ridge trace, is a plot of the regression coefficients for the standardized variables form of the regression over a range of values of Θ --usually in the interval between 0 and 1. If the predictor variables are highly correlated the VIFs are large and the coefficients will change rapidly with small values of Θ then gradually stabilize and change only slowly with further increase of Θ. The Θ value at which the coefficients first stabilize indicates the desired set of coefficients. Fortunately the exact value is usually not critical as the sums of the squares of the coefficients often have a rather flat minimum as Θ changes.

Now that we have concluded this rather brief discussion, we ask the rhetorical question: Do we need to use ridge regresion? To answer we shall look at R_{ii}^{-1}, the VIFs. In the case of the predicted variable R with no threshholding, they are: 1.33 for T_{IR}, 13.2 for T_{37h}, 10.8 for $T_{37\bar{v}}$, 5.5 for T_{21v}, 7.6 for T_{18h}, and 4.6 for T_{18v}. T_{IR} is very stable, the others are only moderately unstable. The ridge trace in Fig. 9.2 shows fairly rapid changes in four of the coefficients for Θ less than about 0.20 or 0.25. Thus, it may be of some interest to consider the set of regression coefficients for

one of these values (say Θ = 0.25) as a reserve model should the ordinary least squares model appear to fail when presented with "new data". Because the coefficients (Table B.7) relate to standardized variables based on the training data set, this model would be written,

$$R = \Sigma_i [\beta_{i,R} \pi_i']\sigma_R + Ave(R)$$

However the standardized brightness temperatures (π_i) *must* be defined in terms of the mean and standard deviation of the learning data set --the data used to derive the coefficients. Thus we denote them by π_i'. Similarly, Ave(R) is the average value of R in the learning set and σ_R the standard deviation of R on the learning set.

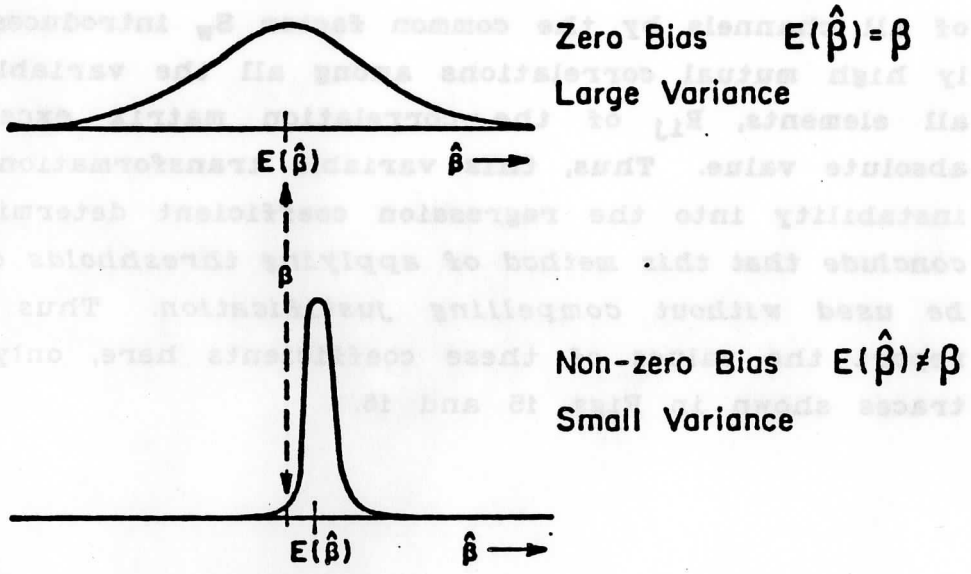
Because the correlation matrix for the predictor variables is the same, the results for Q (also shown in Table B.7.) and R are qualitatively similar. The ridge trace for Q is shown in Fig. 9.3. The Q-model could be expressed as,

$$Q = \Sigma_i [\beta_{i,Q} \pi_i']\sigma_Q + Ave(Q)$$

When the ridge trace procedure was tried on the variables multiplied by the threshhold functions as discussed in Sec. 6, it was found that the coefficients underwent very large and rapid changes for extremely small values of Θ. Initial values of the coefficients were well outside the range -3 to +3. When standardized variables are being used both the predictors and the predicted variable have typical magnitudes of 1. Coefficients greater than +3 or less than -3 suggest that the predictions are based on a relatively small differences among relatively large numbers. Terms in the regression equation are nearly canceling each other. This itself

indicates instability for small perturbations of the coefficients.

Investigation revealed that $[R_{ii}]^{-1}$ attained very large values (~2000) for several channels. That is, the matrix was ill conditioned. The reason for this is that multiplication of all channels by the common factor S_w introduces artificially high mutual correlations among all the variables, so that all elements, R_{ij} of the correlation matrix exceed 0.99 in absolute value. Thus, this variable transformation introduces instability into the regression coefficient determination. *We conclude that this method of applying threshholds ought not to be used without compelling justification.* Thus we do not report the values of these coefficients here, only the ridge traces shown in Figs. 15 and 16.



MSE = Mean Square Error = Variance + $(\text{Bias})^2$

Fig. 9.1 Schematic illustrating variance and bias. [After Marquardt and Snee, 1985].

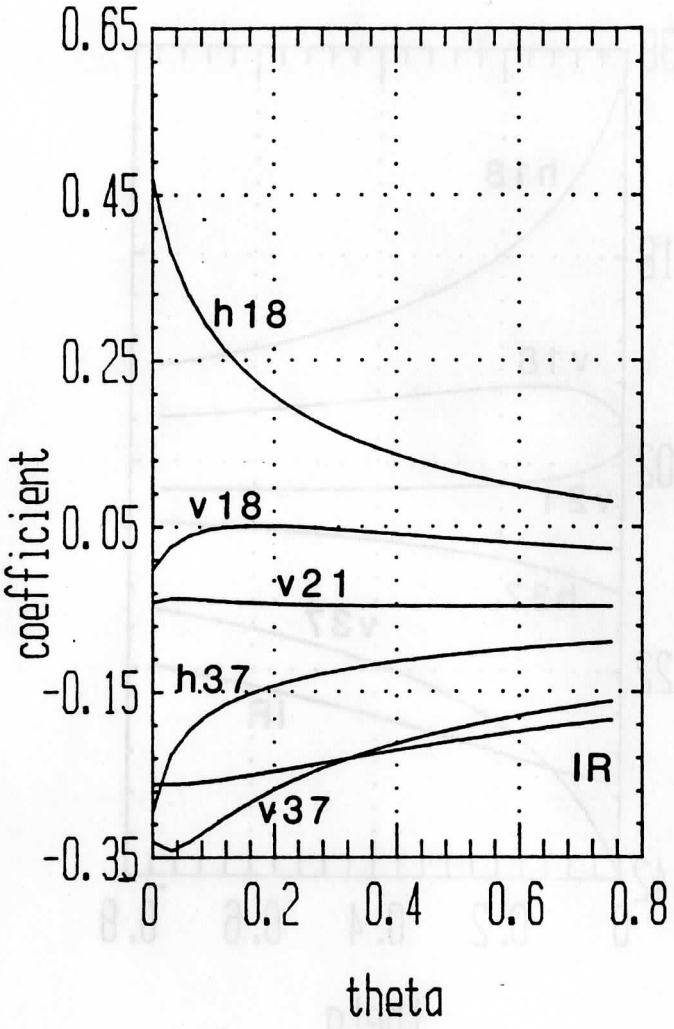


Fig. 9.2 Ridge trace for R. The values of regression coefficients for standardized variables are plotted against the ridge parameter θ. These are also tabulated in Table B.7.

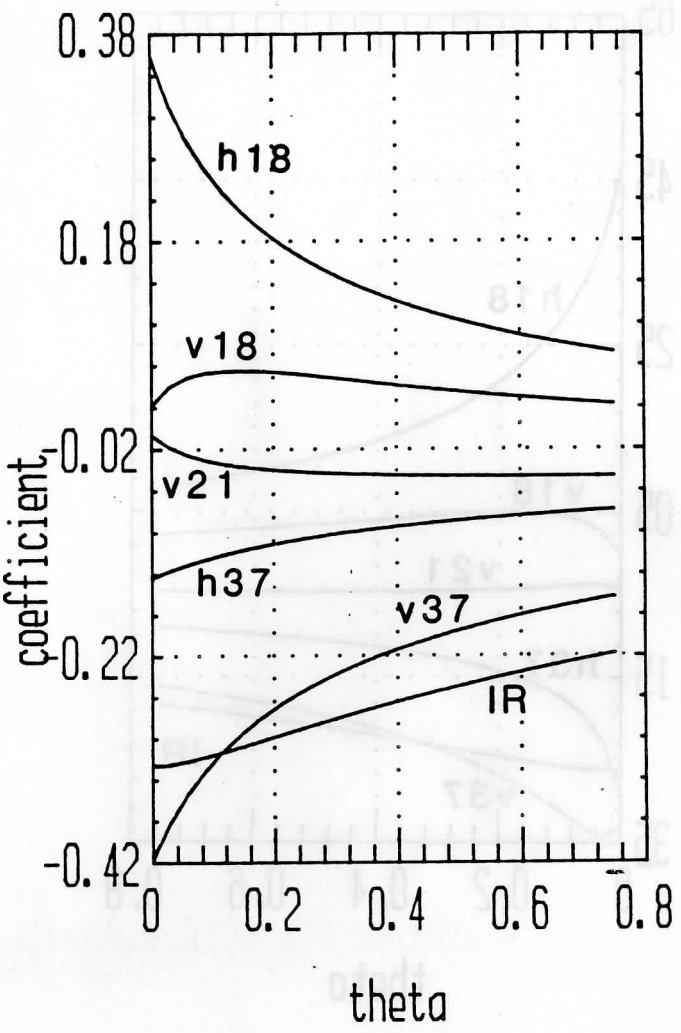


Fig. 9.3 Ridge trace for Q. Note the similarity in the pattern of behavior to that in Fig. 9.2. Values are listed in Table B.7.

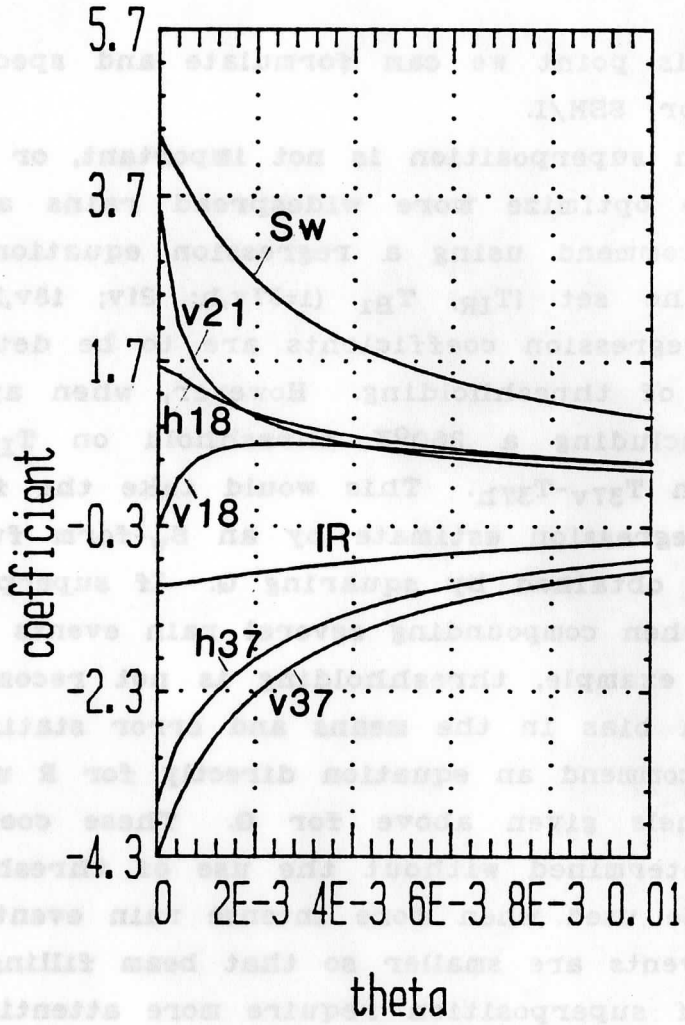


Fig. 9.4 Ridge trace for R when threshholding is incorporated into the variables before regression. Note the magnitudes of the coefficients exceeding 3 and the marked changes in the coefficients for very small θ values.

10. Recommendations: An algorithm for SSM/I

At this point we can formulate and specify a *set* of algorithms for SSM/I.

When superposition is not important, or when the objective is to optimize more widespread rains at lower rain rates, we recommend using a regression equation for Q with analogs of the set $\{T_{IR}, T_{Bi}$ (i=37v,h; 21v; 18v,h)$\}$ as variables. The regression coefficients are to be determined without the use of threshholding. However, when appropriate we recommend including a 260°K threshhold on T_{IR} and a 15° threshhold on $T_{37v}-T_{37h}$. This would take the form of multiplying the regression estimate by an S_w-form function. Rain rate is then obtained by squaring Q. If superposition is important, as when compounding several rain events for a seasonal mean for example, threshholding is not recommended as it will induce a bias in the means and error statistics.

We recommend an equation directly for R using the same set of channels given above for Q. These coefficients also are to be determined without the use of threshholding. This form is to be used when more intense rain events are encountered, the events are smaller so that beam filling and the requirements of superposition require more attention. As for Q, threshholding may be used to suppress negative rain rates for a single event, but must not be used when averaging data. When used, threshholding is accomplished by multiplication by S_w posteriorly as for Q. This recommendation is made with some reluctance due to non-normality of the R residuals and other matters discussed in Appendix A.

We have two real alternatives with respect to choosing the coefficient sets: least squares regression or ridge regression. We display below the least squares coefficients as "first choice" --but not without hesitation. For one thing, we know that the 18 GHz coefficients are not well determined

and it is demonstrable that these variables are not contributing much to predictive skill (Table 5.1). In addition since these channels are correlated with the others, they contribute to instability. Ridge regression could preserve the small contribution made by the 18 GHz channels and maintain stability. So, we also display coefficients for ridge regression as a "back-up" set should least squares clearly fail.

The two threshhold functions referred to above, are combined into a single S_w defined as,

$$S_w = S_{37}(P_{37}) \cdot S(t_{IR})$$

The least squares expression for Q is,

$$Q = [10.216 - 0.0067T_{IR} - 0.0477T_{37h} - 0.0274T_{37v} + 0.0192T_{21v} + 0.0260T_{18h} - 0.0032T_{18v}]S_w$$

The least squares expression for R is,

$$R = [48.644 - 0.0523T_{IR} - 0.1841T_{37h} - 0.1420T_{37v} + 0.1375T_{21v} + 0.0807T_{18h} - 0.0126T_{18v}]S_w$$

The ridge regression expressions for Q and R are given above. The values of the coefficients for $\Theta = 0.25$ are:

$\beta_{h37,Q} = -.1059$	$\beta_{h37,R} = -.1331$
$\beta_{v37,Q} = -.2533$	$\beta_{v37,R} = -.2512$
$\beta_{v21,Q} = -.0431$	$\beta_{v21,R} = -.0461$
$\beta_{h18,Q} = +.1621$	$\beta_{h18,R} = +.1825$
$\beta_{v18,Q} = +.0503$	$\beta_{v18,R} = +.0481$

$\beta_{IR,Q}$	= -.2898	$\beta_{IR,R}$	= -.2400
Ave(Q)	= +1.203	Ave(R)	= +3.968
σ_Q	= +1.589	σ_R	= +7.090

The expressions for the standardized temperature variables are:

$$\pi_{h37}' = (T_{h37}-260.398)/15.446$$
$$\pi_{v37}' = (T_{v37}-269.837)/11.749$$
$$\pi_{v21}' = (T_{v21}-268.622)/9.2610$$
$$\pi_{h18}' = (T_{h18}-260.531)/14.017$$
$$\pi_{v18}' = (T_{v18}-269.571)/11.709$$
$$\pi_{IR}' = (T_{IR}-250.614)/29.262$$

It is hardly necessary to point out that due to the SMMR-SSM/I frequency differences, differences in antenna beamwidth, polarization isolation and sensitivity the given sets of coefficients are departure points. One should expect modification as new data sets become available.

11. Concluding Discussion.

A great deal of statistical research material was generated in this program which has not been explicitly discussed. Such a discussion would be tedious, at best. Therefore, we simply note a few highlights from the material presented together with subjective opinions based on all the work.

The background of this work was the discovery that SMMR data from Nimbus-7 "worked". Simple forward linear regressions produced results well beyond initial expectations over land for convective rain. The opportunity of extending the several year time series of Nimbus-7 data with the SSM/I on a DMSP spacecraft suggested that capabilities of the corresponding microwave channels should be studied. We found, using the subset of SMMR channels simulating these --except for the new 85 GHz capability--that one should expect SSM/I to achieve a standard error of estimate within 8% of SMMR for Q and 13% for R. If IR is used as well, SSM/I's performance should equal that of SMMR.

We found that simple "switch-type" screening methods should be introduced posteriorly, and do not effect significantly improved accuracy relative to straight linear regression methods if much data are to be averaged. They *do* alleviate the embarrassing problem of "negative rain rates" produced by these methods and our recommended algorithms for Q and R use them.

Finally, there is some doubt about the stability of the least squares coefficients and this is at least partly related to the retention of rather poorly defined coefficients for the 18 GHz channels. Those testing algorithms should be acutely aware of this question. It might be well to test ridge regression versions of algorithms simultaneously and adopt this method should the evaluation warrant it.

List of Figures

Fig. A.3 Normal probability plot of the function L, the natural logarithm of (1+R).

Fig. A.4 Normal probability plots of residuals from regression equations using R, Q and L as dependent variables.

Fig. A.5 Plot of residuals of R, Q and L against predicted values.

Fig. A.6 Distribution of residuals of R and Q for predicted values after adding a randon number to each residual.

Fig. A.7 Rain rate against rain rate displaying error bars.

List of Tables

References

Barrett, E.C. and D.W. Martin, 1981: *The Use of Satellite Data in Rainfall Monitoring.* Academic Press, London, 352 pp.

Fang, D.J. and C.H. Chen, 1982: Propagation of centimeter/millimeter waves along a slant path through precipitation. *Radio Sci. 17,* 989-1005.

Fung, A.K., 1982: A review of volume scatter theories for modelling applications. *Radio Sci. 17,* 1007-10017.

Ishimaru, A., 1978: *Wave Propagation and Scattering in Random Media. Volume 1: Single Scattering and Volume Transport Theory.* Academic Press, New York, N.Y.

Ishimaru, A., D. Lesselier and C. Yeh, 1984: Multiple scattering calculations for nonspherical particles based on the vector radiative transfer theory. *Radio Sci. 19,* 1356-1366.

Johnson, R.A. and D.W. Wichern, 1982: *Applied Multivariate Statistical Analysis.* Prentice-Hall, Engelwood Cliffs, N.J.

Marquardt, D.W. and R.D. Snee, 1985: Developing empirical models with multiple regression: Biased estimation techniques, in *Probability, Statistics and Descision Making in the Atmospheric Scienses.* A.H. Murphy and R.W. Katz, eds., Westview Press, Boulder CO.

Olson, W.S. 1985: Estimation of Rainfall Rates in Tropical Cyclones by Passive Microwave Radiometry. PhD. Dissert., University of Wisconsin, Madison, WI. 200 pp.

Robertson, F.R., 1985: IR precipitation in the South Pacific durning FGGE SOP-1. *Preprint Volume, 26th Conference on Hurricanes and Tropical Meteorology* (14-17th of May, 1985, Houston), American Meteorological Society.

Savage, R. 1976: The Transfer of Thermal Microwaves through Hydrometeors. PhD. Dissert, University of Wisconsin, Madison WI. 147 pp.

Spencer, R.W., W.S. Olson, W. Rongzhang, D.W. Martin, J.A. Weinman and D.A. Santek, 1983a: Heavy thunderstorm observed over land by the Nimbus-7 Scanning Multichannel Microwave Radiometer. *J. Climate Appl. Meteor., 22,* 1041-1046.

Spencer, R.W., D.W. Martin, B.B. Hinton, and J.A. Weinman, 1983b: Satellite microwave radiances correlated with radar rainrates over land. *Nature, 304,* 141-143.

Spencer, R.W., 1984: Satellite passive microwave rain rate measurement over croplands during spring, summer, and fall. *J. Climate Appl. Meteor.,23,* 1553-1562.

Spencer, R.W. and D.A. Santek, 1985: Measuring the global distribution of intense convection over land with passive microwave radiometry. *J. of Climate Appl. Meteor., 24.* 860-864.

Spencer,R.W. 1986: A satellite passive 37 GHz scattering based method for measuring oceanic rain rates. Submitted to *J. Climate and Appl. Meteorol.*

Tsang, L., 1984: Scattering of electromagnetic waves from a half space of nonspherical particles. Radio Sci., 19 966-974.

Ulaby, F.T., R.K. Moore and A.K. Fung, 1981: *Microwave Remote Sensing: Active and Passive. Volumel: Microwave Remote Sensing Fundamentals and Radiometry.* Addison-Wesley Advanced Book Program, Reading Mass. 456 pp.

Weinman, J.A. and R. Davies, 1978: Thermal microwave radiance from horizontally finite clouds of hydrometeors. *J. Geophys. Res. 83,* 3099-3107.

Wilheit, T.T. and A.T.C. Chang, 1980: An algorithm for retreival of ocean surface and atmospheric parameters from observations of the scanning multichannel microwave radiometer. *Radio Sci. 15,* 525-544.

Wu, R., and J.A. Weinman, 1984: Microwave radiances from precipitating clouds containing aspherical ice, combined phase, and liquid hydrometeors. *J. Geophys. Res., 89* 7170-7178.

Appendix A
Normality of Data and Residual Distributions

Many statistical techniques, including significance tests assume that each observation comes from a normal distribution. If the number of samples in the data set is larger than 25 or so, histograms are helpful in assessing the nature of the distribution as we did in Sec. 4 (Fig. 4.1). For larger sample sizes one can construct special plots to judge the assumption of normality, called quantile-quantile plots [Johnson and Wichern, 1982] or normal probability plots. These are plots of a sample quantile against the quantile one would expect *if* the data were normally distributed. When the points lie along a straight line, or nearly so, normality is a reasonable assumption, otherwise it is suspect.

If normality is doubtful, or actually ruled out, there is some peril in proceeding as if it held. An alternative is to transform the data to achieve a description of the problem in terms of variables more nearly conforming to the normality assumption. Johnson and Wichern [1982] suggest power transformations. These are generated by replacing x, the observed value of the variable X, with x^{α}. There is a special definition for $\alpha=0$: x is replaced by ln(x). For values of α less than 1, large values of x are pulled in and small values stretched out. On the other hand for values of α greater than 1, large x values are extended, small values compressed. The choice of α may be dictated by simplicity, or ease in interpretation in part. Thus one might favor $\alpha=0.5$ --even if $\alpha=0.509$ were slightly better.

It follows that examinations of the distributions of our rain rate variable, R, and its transforms are desirable. In Fig. A.1 we have normal probability plots of R. R^2, $\sqrt{R}$

and $\sqrt[4]{R}$. All these show two effects clearly. First, there is a "step" or discontinuity just above zero rain rate. In the data set from which these samples are drawn about half the cases are for R=0. This is typical in our data base. Second, there is a small step near 17 mm/h. If we draw a reference line from the 50 cumulative percent point to the 95 or 99 cumulative-percent points, we can see that the R greater than zero parts of the data are more normally distributed for $\sqrt{R}$ than for the others.

Johnson and Wichern [1982] suggest a systematic means of choosing a power transformation using the Box-Cox family of transforms, defined as

$$x^{(\alpha)} = \begin{cases} (x^{\alpha} - 1)/\alpha & \text{for } \alpha \text{ not equal } 0 \\ \ln(x) & \text{for } \alpha \text{ equal to } 0 \end{cases}$$

Fig. A.2 shows normal probability plots of Box-Cox transformed rain rates for $\alpha = \{0, 0.2, 0.4, 0.6, 0.8, 1.0\}$. These plots show that the best exponent is somewhere in the interval 0.4 to 0.8, probably close to 0.6. This suggests replacing R by the variable $Q = \sqrt{R}$. Although not in this transform family, we found that similar results were obtained for $L = \ln(R + 1)$ as shown in Fig. A.3.

The L-transformation is best understood as the transform that renders the radar digitization error approximately constant, since ∂R_{meas} is $(0.6\ R_{meas})$ except near R=0.

After we formulate a regression model and estimate it's coefficients we should consider the validity and adequacy of the model. Each residual is an estimate of the error of the model. Usually it is *assumed* that the errors are drawn from a normal population with a zero mean and constant variance. Residuals, on the other hand, can actually have unequal variances and non-zero correlations even though

their mean is zero. However, *if* the correlations are small and the variations of the variance not too large, we can still treat them as samples drawn from an approximately normal distribution. One means of detecting significant departures from this approximation is to plot the residuals against the predicted values. Another is a normal probability plot of the residuals [Johnson and Wichern, 1982].

Fig. A.4 illustrates the character of the residuals (see Sec. 7) when the three variables $\sqrt{R}$, R and ln(R + 1) are fit with regression equations. It is obvious that the residuals of both $\sqrt{R}$ and ln(R + 1) are more normally distributed than the residuals of R.

Now consider the structure of the residuals when plotted against the predicted values as suggested above. These are shown in Fig. A.5. We note that the lower left corner of each distribution is "missing". This is caused by the constraint, felt at low predicted rain rates, that only rain rates greater than zero are allowed. Thus at zero rain rate only positive residuals are possible. The banding in the figures is caused by the interaction of two factors: (1) The "peakiness" of the frequency distribution of various rain rates and (2) the assumption that the radar data are "true". A significant portion of the variance should be attributed to the "truth data" which is known to have a variance from the desired parameter that is the same order of magnitude as the regression equation.

If we destroy the "peakiness" and the constraint that $R \geq 0$ by adding a normally distributed random number, r, with a standard deviation less than or equal to the radar measurement error, we should get about the same regression equation, but the residuals will be "unstacked" so we can visually estimate their average distribution. Fig. A.6 shows before-after comparisons for R and $\sqrt{R}$. The randon num-

ber, r, has a standard deviation of 2.5 for R and 1.5 for $\sqrt{R}$.

In both cases the new regression coefficients were covered by the old coefficients plus or minus the uncertainty in the coefficients. Naturally, as suggested above, the added noise reduced the correlation coefficients and increased the standard errors. However the point is that the new residuals are satisfactorily distributed. We conclude, therefore, that the "peakiness" and constraint that R is greater than zero were not doing violence to conclusions drawn from the original regressions.

Consider Fig. A.7. The upper left panel shows radar rain rate with "error bars" as a function of rain rate. Actually these are just ±60% of the rate itself (see Sec. 4). The lower right panel shows error bars for $Q = \sqrt{R}$ from an actual regression *--on the assumption that the variance is independent of Q.* If this were the case rain rates above a few mm/h from the regression would be more accurate than the radar data! This absurdity is illustrated at the lower rightwhere the attributed error bars for $\sqrt{R}$ are shown as vertical lines with those for radar shown as horizontal lines.

This clearly suggests a need for better truth, for careful interpretation of goodness of fit parameters, or if radar data must be used, *for a transformation which renders the variance of the (so-called) truth more nearly constant.*

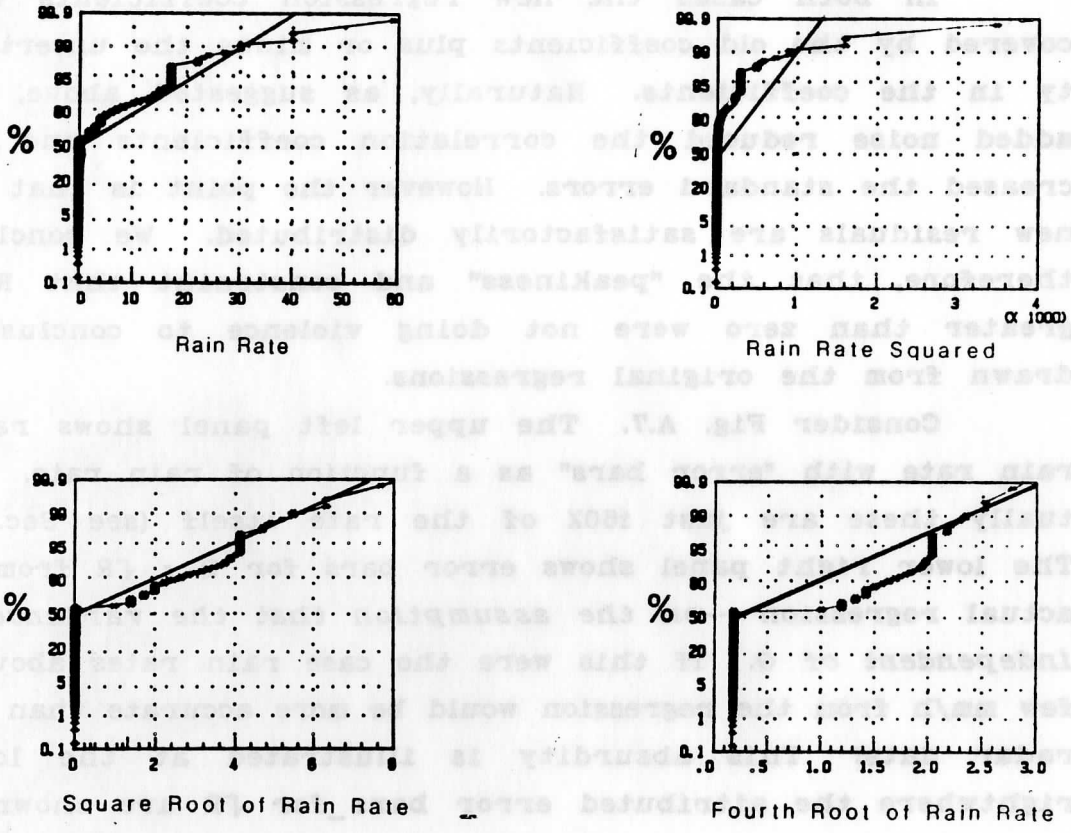


Fig. A.1 Normal probability plots of R, R^2, $\sqrt{R}$, and $\sqrt[4]{R}$. The vertical axis is the cumulative fraction of the distribution expressed in percent.

BOX-COX TRANSFORMED VARIABLES, $X^{(\alpha)}$ FOR X = R

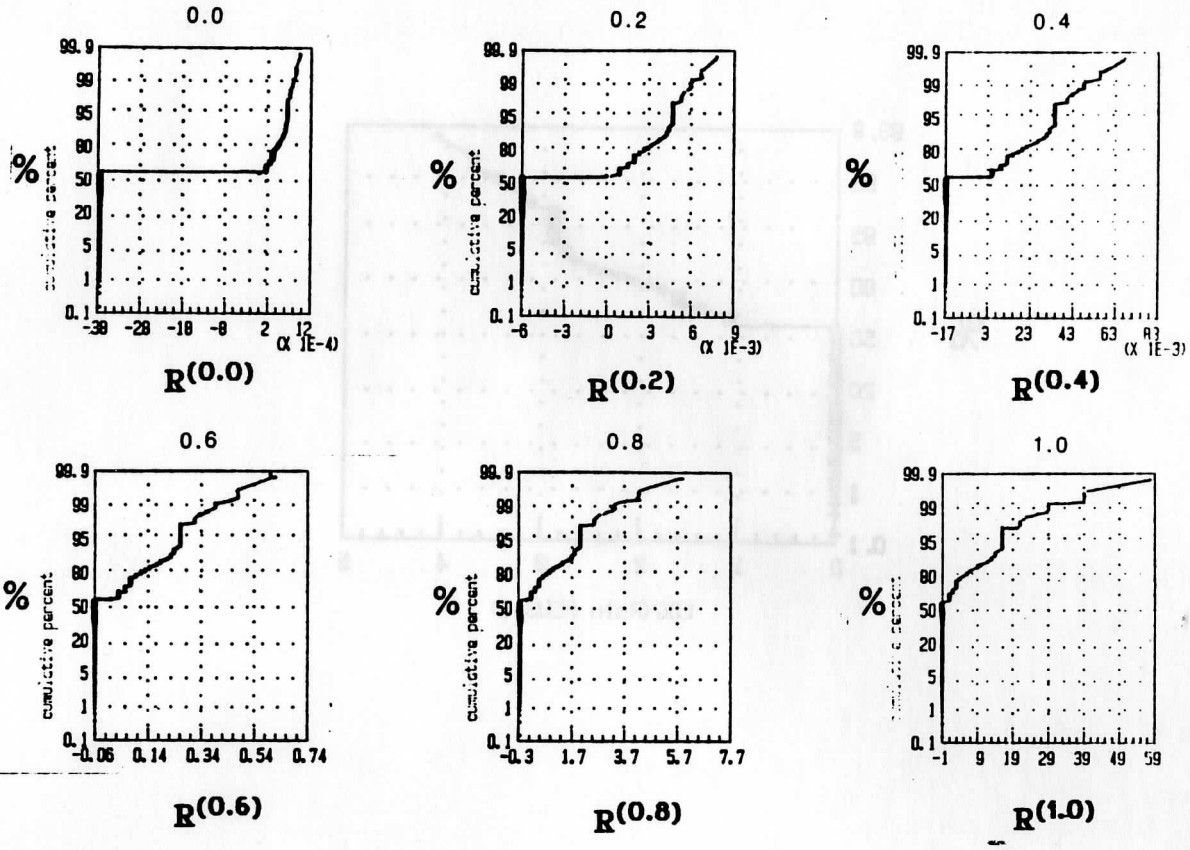


Fig. A.2 Normal probability plots of Box-Cox transformed rain rates. The vertical axis is the cumulative fraction of the distribution expressed in percent.

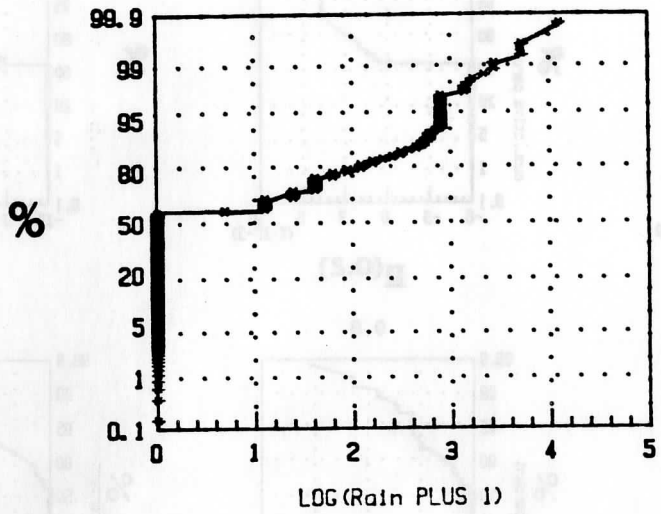


Fig. A.3 Normal probability plot of the function L, the natural logarithm of (1+R). The vertical axis is the cumulative fraction of the distribution expressed in percent.

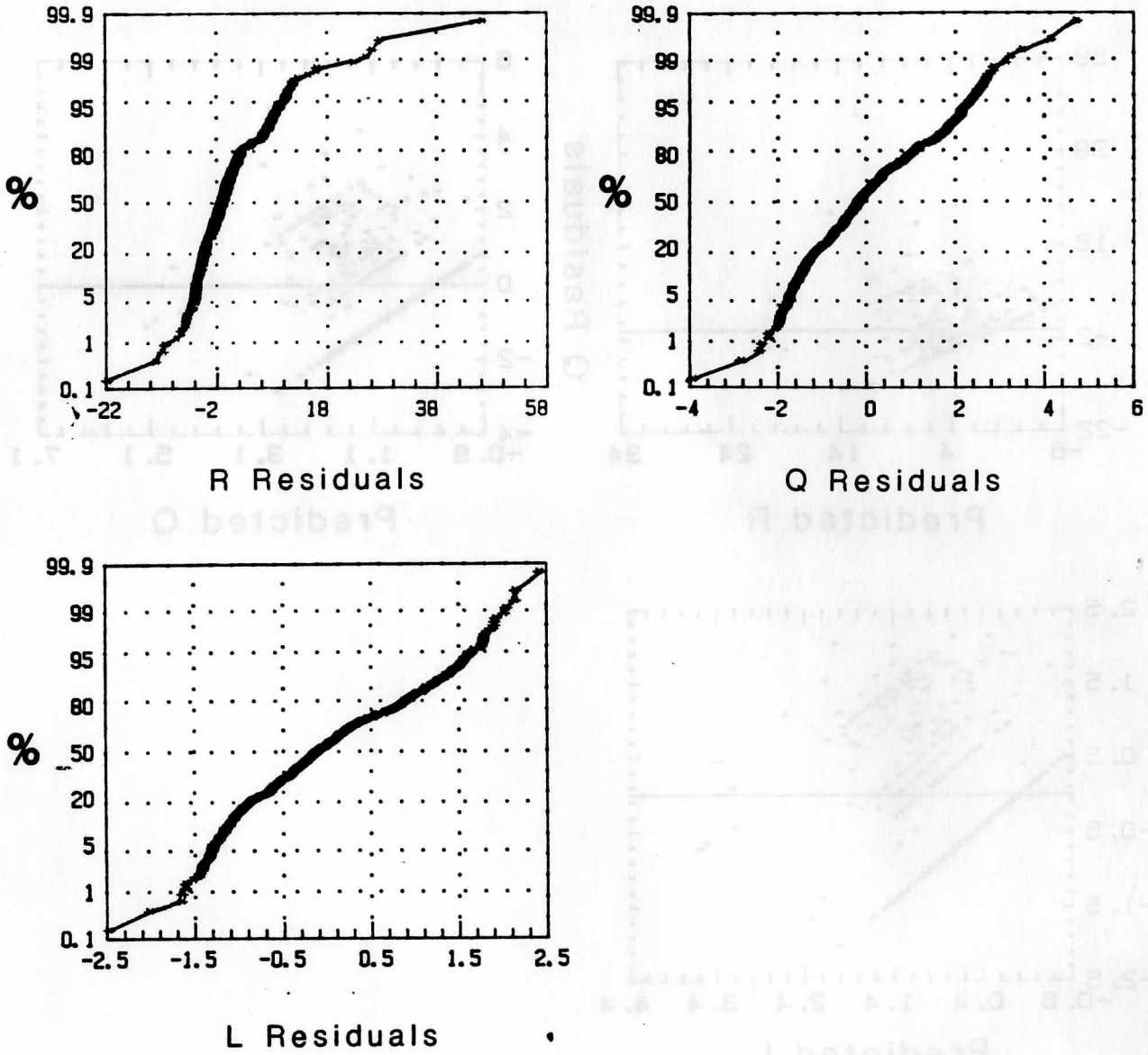


Fig. A.4 Normal probability plots of residuals from regression equations using R, Q and L as dependent variables. The vertical axis is the cumulative fraction of the distribution expressed in percent.

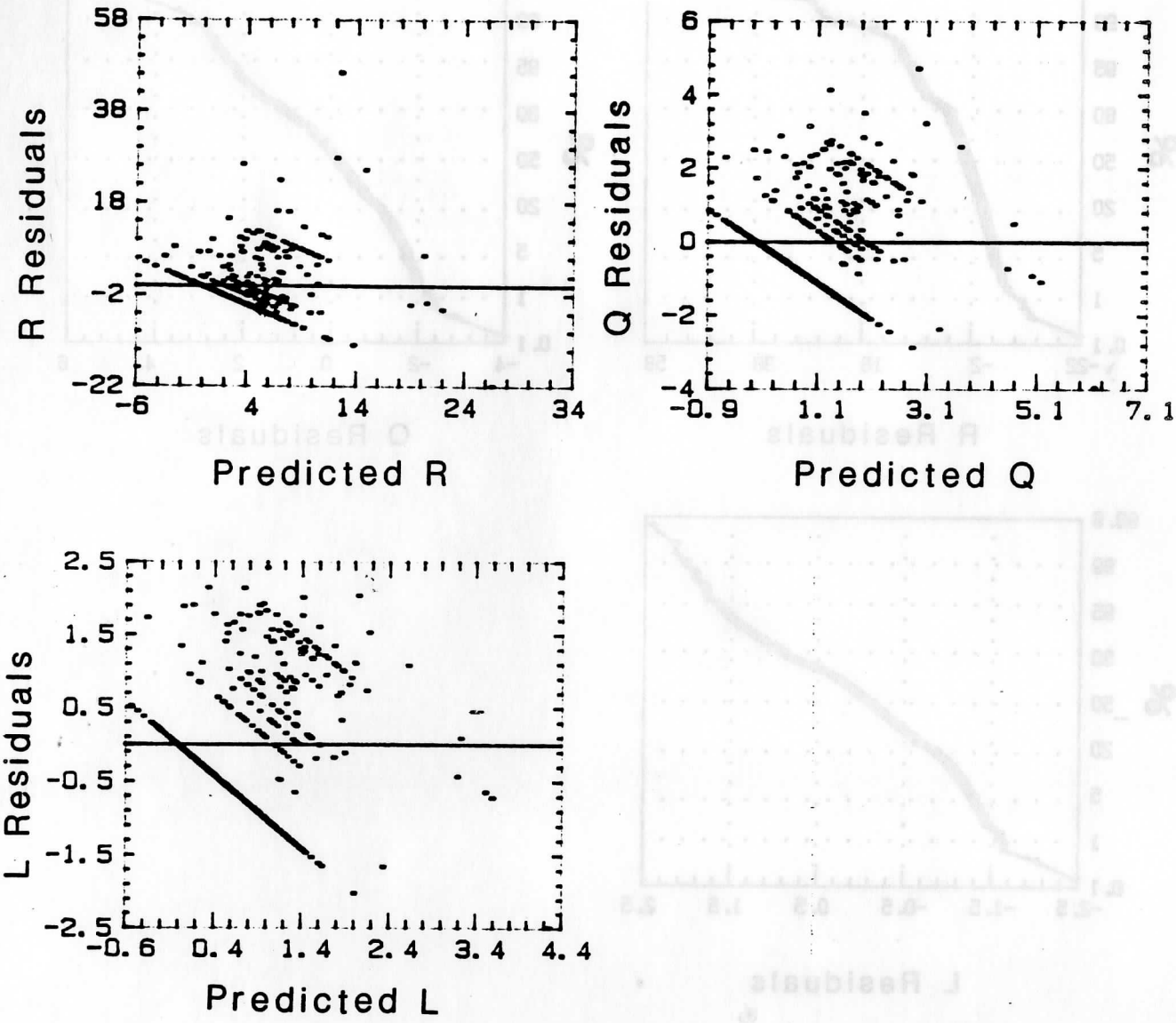


Fig. A.5 Plot of residuals of R, Q and L against predicted values.

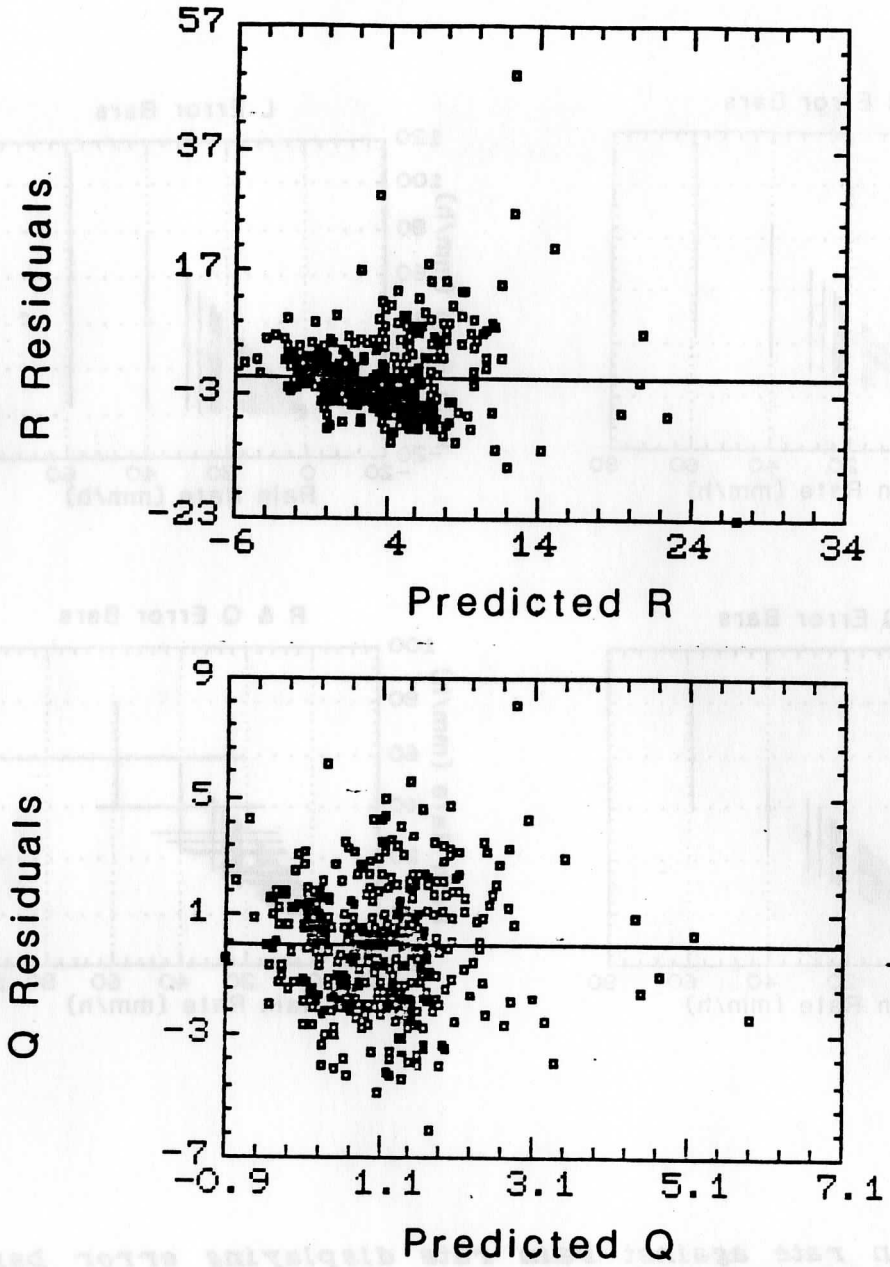


Fig. A.6 Distribution of residuals of R and Q for predicted values after adding a randon number to each residual. The random numbers were grawn from gaussian distributions with zero means and standard deviations of 2.5 and 1.5 for R and Q respectively.

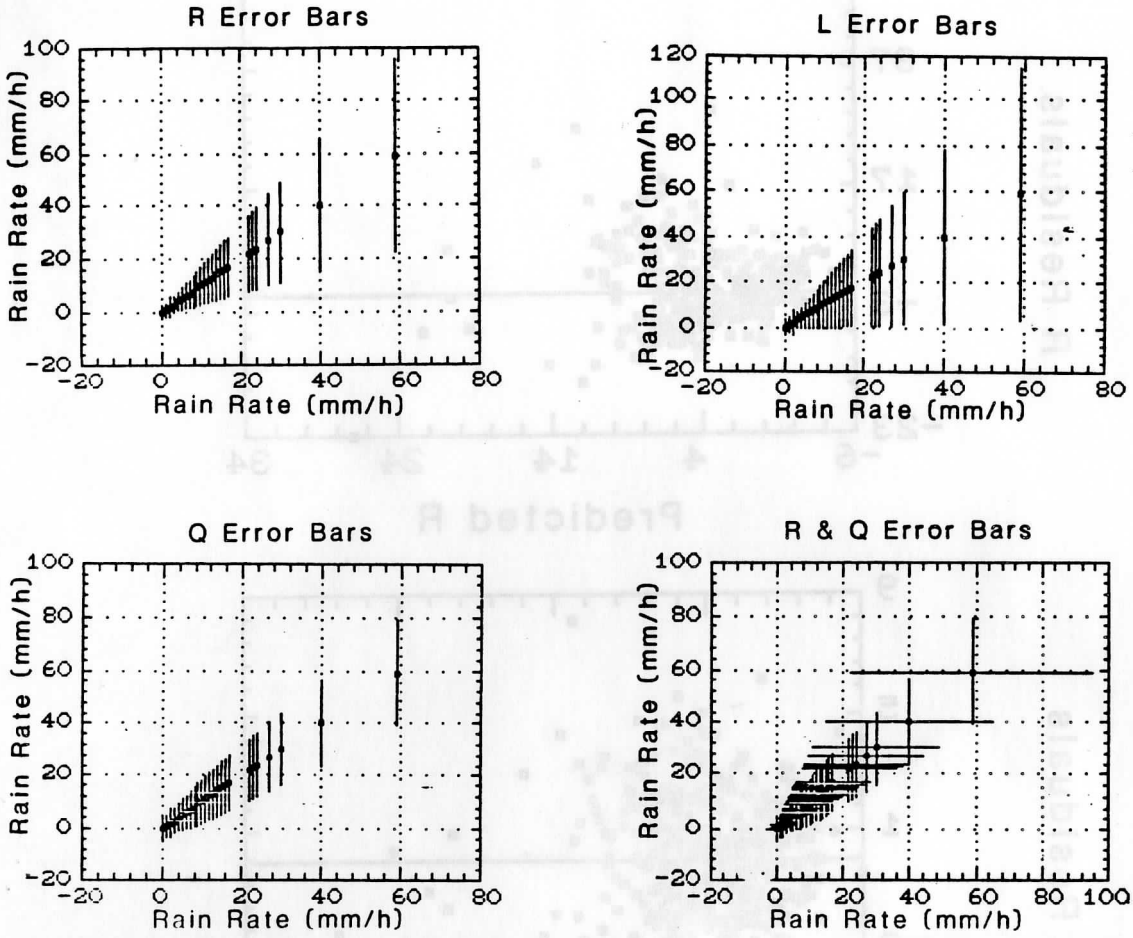


Fig. A.7 Rain rate against rain rate displaying error bars. The error bars are attributed to radar (upper left panel), calculation of R from Q regression assuming a constant variance for Q (lower left), and calculation of R from L, assuming a constant variance for L (upper right). The lower right panel displays both the R (horizontal) and Q (vertical) error bars together.

Appendix B

Extended Results of Regressions and Partial Correlations

Table B.1

Regression Results for Brightness Temperature As a Function of Rain Rate

C_{h37}	$\sigma_{c,h37}$	q_{h37}	$\sigma_{q,h37}$	C_{v37}	$\sigma_{c,v37}$	q_{v37}	$\sigma_{q,v37}$
ch37	stch37	qh37	qsh37	cv37	stcv37	qv37	qsv37
262.039	3.07916	-6.02441	1.54418	269.537	1.8864	-4.75994	0.94601
255.073	3.48825	-2.60512	1.89479	266.55	2.36307	-2.67848	1.2836
266.45	2.42157	-5.6389	1.12395	275.343	2.13695	-5.64922	0.99183
264.639	3.04878	-1.40308	2.07443	275.519	2.36445	-1.73533	1.60881
243.819	2.19646	0.79400	1.17482	259.716	1.20062	-0.71195	0.64216
259.679	1.90476	-0.61564	1.13152	270.855	1.34055	-1.93953	0.79635
256.515	2.44774	-3.57894	1.03358	268.312	1.51713	-3.34446	0.64062
257.567	9.94842	-2.48226	6.98909	266.336	4.97856	-1.35294	3.4976
269.633	3.11652	-8.32454	1.53763	272.4	2.89831	-8.23871	1.42996
253.071	5.56703	-2.59343	2.0847	259.641	4.36627	-2.96533	1.63505
272.731	1.31494	-5.32916	0.59317	275.307	1.15904	-4.99333	0.52285
265.983	1.80228	-3.43091	0.80739	269.811	1.56558	-3.32621	0.70135
276.449	2.63479	-3.69774	1.36665	278.393	2.35653	-3.53354	1.22231
250.76	8.7125	0.55185	4.82002	258.695	6.69635	-0.34245	3.70463
275.067	2.95804	-9.75508	1.14897	276.174	2.81827	-9.19176	1.09468
255.596	3.38133	-4.93107	0.91276	261.705	2.76201	-5.12191	0.74558
265.412	0.76901	-2.06214	0.32337	268.058	0.69061	-1.89192	0.2904
257.7	1.81925	-3.39417	0.88955	267.605	1.16279	-3.1619	0.56856
252.675	1.86192	-0.66408	0.69454	263.293	1.26904	-1.18893	0.47338
252.801	1.16742	-0.79100	0.48732	262.362	0.93788	-0.75528	0.39151
261.338	1.20772	-2.84678	0.54633	269.935	0.84239	-2.52545	0.38107
258.163	1.49697	-5.76108	0.66007	270.78	0.62594	-5.86238	0.276
255.597	2.42794	-2.15948	1.21206	264.371	1.85717	-2.09373	0.92712
254.792	1.60491	-1.53103	0.81344	264.158	1.16398	-1.54304	0.58996
256.639	1.34811	-2.31099	0.61664	265.039	1.07953	-1.88432	0.49379

Table B.1 is continued on the following page.

Table B.1 Continued.

C_{h18}	$\sigma_{c,h18}$	q_{h18}	$\sigma_{q,h18}$	C_{v18}	$\sigma_{c,v18}$	q_{v18}	$\sigma_{q,v18}$
ch18	stch18	qh18	qsh18	cv18	stcv18	qv18	qsv18
256.387	4.64975	-7.75462	2.33182	265.644	2.81744	-5.53315	1.41292
248.013	4.95268	-0.39619	2.69026	260.688	2.93075	-1.10212	1.59196
253.383	1.7116	0.70572	0.79441	263.002	1.09212	0.22525	0.50689
260.392	3.08166	1.05347	2.0968	267.618	2.03235	0.10584	1.38284
231.792	3.73198	2.69903	1.99608	249.216	2.16653	1.46279	1.15879
253.783	2.02284	1.0476	1.20167	267.181	1.463	-0.55870	0.86903
246.626	3.90264	-1.82774	1.64793	260.56	2.53963	-2.09736	1.07238
247.92	11.6915	0.92577	8.21367	259.077	7.38296	1.23886	5.18677
257.646	1.58202	-0.80122	0.78053	261.038	1.35194	-1.04506	0.66702
240.04	6.8148	2.53635	2.55196	247.47	5.53693	1.81256	2.07343
268.476	1.34425	-1.36917	0.6064	271.143	1.00364	-1.40576	0.45275
259.059	1.88416	-0.82122	0.84407	264.197	1.60749	-1.15914	0.72012
270.992	2.5119	-0.90500	1.3029	273.705	2.30477	-0.83136	1.19547
241.491	11.881	2.42814	6.57295	250.36	9.5523	1.17242	5.28462
267.37	1.52898	-2.74025	0.59388	268.631	1.19016	-2.64181	0.46228
243.407	3.47981	1.7351	0.93934	249.774	2.87733	1.06151	0.77677
259.935	0.63858	-0.20704	0.26852	263.191	0.46721	-0.32407	0.19646
251.006	2.56356	-1.01841	1.2535	261.828	1.7717	-1.19285	0.86630
240.554	2.54697	3.9124	0.95009	253.592	1.58909	1.98849	0.59277
242.946	1.85539	1.35409	0.77451	252.976	1.15498	1.41858	0.48213
251.939	1.78077	2.58648	0.80556	260.721	1.20238	1.78395	0.54391
243.572	1.97603	0.98764	0.8713	256.957	1.25486	0.09009	0.55331
246.813	2.00556	4.36368	1.0012	255.242	1.4649	3.16334	0.78121
246.992	2.86262	2.21232	1.45092	256.005	1.53505	1.21313	0.77803
247.966	1.93393	0.16606	0.88461	258.025	1.4823	-0.19155	0.67803

Table B.1 is continued on the following page.

Table B.1 continued.

C_{v21}	$\sigma_{c,v21}$	q_{v21}	$\sigma_{q,v21}$
cv21	stcv21	qv21	qsv21
266.343	2.30702	-4.47542	1.15695
262.644	2.70209	-2.03477	1.46776
266.341	0.98493	-0.84518	0.45714
266.123	2.15405	1.03422	1.46564
251.994	1.63021	1.31	0.87195
266.512	1.3241	-0.43208	0.78658
262.316	1.83693	-1.57075	0.77566
261.459	5.0827	0.3903	3.57077
262.896	1.53774	-2.04593	0.75868
256.021	3.58899	0.21122	1.34398
272.731	0.84332	-1.78215	0.38042
266.607	1.2688	-1.49062	0.5684
274.553	1.86495	-1.12452	0.96733
260.958	5.29483	0.43518	2.92926
270.655	1.26861	-3.51818	0.49275
258.015	1.73667	-0.678	0.4688
266.072	0.41461	-0.67155	0.17434
260.94	1.20603	-0.31518	0.58971
256.665	1.25231	1.10884	0.46714
255.878	0.8113	1.4201	0.33866
262.059	0.76567	0.64051	0.34636
262.127	1.259	-2.21177	0.55514
256.889	1.48223	2.046	0.73994
259.162	0.82192	0.04384	0.41659
259.606	1.39908	-0.53352	0.63996

Table B.1 is concluded on the following page.

Table B.1 concluded.

T_{MAX} Values

h37	v37	v21	h18	v18
cranh37	cranv37	cranv21	cranh18	cranv18
241.825	258.856	255.549	232.119	251.449
240.363	259.835	255.52	231.342	251.916
236.903	255.417	251.427	228.505	247.888
237.274	256.891	250.871	227.299	247.354
222.026	247.824	241.847	209.714	236.472
254.086	265.916	263.242	249.881	262.203
249.837	265.234	260.819	241.216	258.49
252.208	264.405	260.73	246.361	259.395
264.021	267.221	265.879	262.434	266
242.799	251.461	252.337	234.846	243.413
262.784	268.079	268.829	260.126	265.29
264.317	268.044	266.49	260.545	264.904
255.291	261.442	260.04	252.477	258.38
261.734	266.412	267.646	259.385	263.708
264.34	266.896	266.545	262.7	265.046
240.383	249.193	249.853	230.855	239.833
255.326	260.458	261.227	251.725	256.791
256.705	267.186	263.041	252.426	263.417
236.979	255.928	251.633	225.107	246.987
251.797	263.048	258.193	245.754	257.426
249.298	262.252	257.655	240.522	254.549
255.395	264.861	260.282	250.361	259.593
238.632	256.834	253.046	230.411	248.075
254.72	265.403	261.315	248.974	259.194
240.901	258.641	254.81	228.89	248.282

Table B.2
Partial Correlations Relating Q, R and L to T_{IR_i} and T_{B_i}

	Q	IR	v37	h37
Q	- 1.00	-0.32	-0.23	-0.13
IR		- 1.00	0.27	-0.16
v37			- 1.00	0.91
h37				- 1.00

	Q	IR	v18	h18
Q	- 1.00	-0.45	-0.05	0.03
IR		- 1.00	0.15	-0.03
v18			- 1.00	0.85
h18				- 1.00

	Q	IR	v21
Q	- 1.00	-0.41	-0.13
IR		- 1.00	0.29
v21			- 1.00

	Q	IR	v37	h37	v21
Q	- 1.00	-0.33	-0.26	0.11	0.12
IR		- 1.00	0.21	-0.17	0.11
v37			- 1.00	0.76	0.36
h37				- 1.00	0.20
v21					- 1.00

	R	IR	v37	h37
R	- 1.00	-0.26	-0.22	-0.11
IR		- 1.00	0.30	-0.18
v37			- 1.00	0.91
h37				- 1.00

	R	IR	v18	h18
R	- 1.00	-0.39	-0.06	0.04
IR		- 1.00	0.15	-0.03
v18			- 1.00	0.85
h18				- 1.00

Table B.2, Partial correlations, continued

	R	IR	v21
R	- 1.00	-0.35	-0.14
IR		- 1.00	0.31
v21			- 1.00

	R	IR	v37	h37	v21
R	- 1.00	-0.27	-0.25	0.08	0.12
IR		- 1.00	0.24	-0.19	0.10
v37			- 1.00	0.75	0.36
h37				- 1.00	0.21
v21					- 1.00

	L	IR	v37	h37
L	- 1.00	-0.33	-0.23	-0.14
IR		- 1.00	0.27	-0.15
v37			- 1.00	0.91
h37				- 1.00

	L	IR	v18	h18
L	- 1.00	-0.45	-0.05	0.03
IR		- 1.00	0.15	-0.03
v18			- 1.00	0.85
h18				- 1.00

	L	IR	v21
L	- 1.00	-0.42	-0.13
IR		- 1.00	0.31
v21			- 1.00

	L	IR	v37	h37	v21
L	- 1.00	-0.34	-0.26	0.11	0.12
IR		- 1.00	0.21	-0.17	0.10
v37			- 1.00	0.76	0.36
h37				- 1.00	0.20
v21					- 1.00

Table B.3

Linear Regression on Rain Rate and Square Root of Rain Rate

The tables on the following pages are reproductions of the output of the regression computing program. The notation varies slightly from that used in the balance of the report. However, the notation is reasonably obvious. For example, either th37 or Th37 might represent T_{h37}. The dependent variables "Rain" and "qrain" represent rain rate (R) and square root of the rain rate (Q) respectively.

DEP. VARIABLE: qrain

IND. VARIABLES: tv37
th37
tv21
tv18
th18

CONSTANT: YES
WEIGHTS:
FLAGS: NO

MODEL FITTING RESULTS

VARIABLE	COEFFICIENT	STND. ERROR	T-VALUE	PROB(>|T|)
CONSTANT	16.006617	2.233513	7.1666	.0000
tv37	-0.086306	0.018304	-4.7151	.0000
th37	-0.002869	0.015788	-.1817	.8559
tv21	-0.009031	0.017159	-.5263	.5989
tv18	-0.002018	0.012321	-.1638	.8700
th18	0.047037	0.0133	3.5366	.0005

0 CASES WITH MISSING VALUES WERE EXCLUDED.

ANALYSIS OF VARIANCE FOR THE FULL REGRESSION

SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F-RATIO	PROB(>F)
MODEL	241.68620	5	48.33724	26.51913	.00000
ERROR	716.33324	393	1.82273		
TOTAL (CORR.)	958.01944	398			

R-SQUARED = 0.252277
R-SQUARED (ADJ. FOR D.F.) = 0.242764
STND. ERROR OF EST. = 1.35009

NUMBER OF RESIDUALS = 399
SAMPLE AVERAGE = -8.28695E-15
SAMPLE VARIANCE = 1.79983
SAMPLE STANDARD DEVIATION = 1.34158
COEFF. OF SKEWNESS = 0.677913 STANDARDIZED VALUE = 5.52822
COEFF. OF KURTOSIS = 3.52388 STANDARDIZED VALUE = 2.13607
DURBIN-WATSON STATISTIC = 0.928339

DEP. VARIABLE: qrain

IND. VARIABLES: tv37
th37
tv21
tv18
th18
tir

CONSTANT: YES
WEIGHTS:
FLAGS: NO

MODEL FITTING RESULTS

VARIABLE	COEFFICIENT	STND. ERROR	T-VALUE	PROB(>|T|)
CONSTANT	13.445897	2.147448	6.2613	.0000
tv37	-0.054878	0.01793	-3.0607	.0024
th37	-0.014862	0.015049	-.9876	.3240
tv21	-0.001237	0.016284	-.0760	.9395
tv18	0.002695	0.011685	.2307	.8177
th18	0.039955	0.012633	3.1626	.0017
tir	-0.017221	0.002524	-6.8215	.0000

0 CASES WITH MISSING VALUES WERE EXCLUDED.

ANALYSIS OF VARIANCE FOR THE FULL REGRESSION

SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F-RATIO	PROB(>F)
MODEL	317.69661	6	52.94943	32.41518	.00000
ERROR	640.32283	392	1.63348		
TOTAL (CORR.)	958.01944	398			

R-SQUARED = 0.331619
R-SQUARED (ADJ. FOR D.F.) = 0.321388
STND. ERROR OF EST. = 1.27806

NUMBER OF RESIDUALS = 399
SAMPLE AVERAGE = -5.88623E-15
SAMPLE VARIANCE = 1.60805
SAMPLE STANDARD DEVIATION = 1.26841
COEFF. OF SKEWNESS = 0.547609 STANDARDIZED VALUE = 4.46382
COEFF. OF KURTOSIS = 3.18782 STANDARDIZED VALUE = 0.785795
DURBIN-WATSON STATISTIC = 0.966068

DEP. VARIABLE: qrain

IND. VARIABLES: tv18
th18
tir

CONSTANT: YES
WEIGHTS:
FLAGS: NO

MODEL FITTING RESULTS

VARIABLE	COEFFICIENT	STND. ERROR	T-VALUE	PROB(>|T|)
CONSTANT	9.037154	1.592643	5.6743	.0000
tv18	-0.011888	0.011634	-1.0192	.3087
th18	0.005965	0.009598	.6234	.5333
tir	-0.024464	0.002455	-9.9653	.0000

0 CASES WITH MISSING VALUES WERE EXCLUDED.

ANALYSIS OF VARIANCE FOR THE FULL REGRESSION

SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F-RATIO	PROB(>F)
MODEL	218.90001	3	72.96667	38.99483	.00000
ERROR	739.11943	395	1.87119		
TOTAL (CORR.)	958.01944	398			

R-SQUARED = 0.228492
R-SQUARED (ADJ. FOR D.F.) = 0.222633
STND. ERROR OF EST. = 1.36791

NUMBER OF RESIDUALS = 399
SAMPLE AVERAGE = -5.15909E-15
SAMPLE VARIANCE = 1.85708
SAMPLE STANDARD DEVIATION = 1.36275
COEFF. OF SKEWNESS = 0.613867 STANDARDIZED VALUE = 5.00594
COEFF. OF KURTOSIS = 3.14643 STANDARDIZED VALUE = 0.597053
DURBIN-WATSON STATISTIC = 0.888711

DEP. VARIABLE: qrain

IND. VARIABLES: tv37
th37
tir

CONSTANT: YES
WEIGHTS:
FLAGS: NO

MODEL FITTING RESULTS

VARIABLE	COEFFICIENT	STND. ERROR	T-VALUE	PROB(>|T|)
CONSTANT	17.537506	1.683567	10.4169	.0000
tv37	-0.072821	0.015167	-4.8012	.0000
th37	0.029609	0.010945	2.7052	.0071
tir	-0.017327	0.002565	-6.7558	.0000

0 CASES WITH MISSING VALUES WERE EXCLUDED.

ANALYSIS OF VARIANCE FOR THE FULL REGRESSION

SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F-RATIO	PROB(>F)
MODEL	284.02424	3	94.67475	55.48485	.00000
ERROR	673.99520	395	1.70632		
TOTAL (CORR.)	958.01944	398			

R-SQUARED = 0.29647
R-SQUARED (ADJ. FOR D.F.) = 0.291127
STND. ERROR OF EST. = 1.30626

NUMBER OF RESIDUALS = 399
SAMPLE AVERAGE = 1.04709E-14
SAMPLE VARIANCE = 1.69346
SAMPLE STANDARD DEVIATION = 1.30133
COEFF. OF SKEWNESS = 0.577462 STANDARDIZED VALUE = 4.70906
COEFF. OF KURTOSIS = 3.08958 STANDARDIZED VALUE = 0.36527
DURBIN-WATSON STATISTIC = 0.913247

DEP. VARIABLE: qrain

IND. VARIABLES:

tir

CONSTANT: YES
WEIGHTS:
FLAGS: NO

ENTER DESIRED FACTORS, THEN
PRESS ENTER TO FIT OR F KEY.
*1HELP 2ANOVA 3PLTPRD 4COND 5RESIDS 6CMPEFF 7 8 9 10QUIT
PRINT MON JAN 27 1986 08:34:00 AM VERSION 1.1 REC:OFF

MODEL FITTING RESULTS

VARIABLE	COEFFICIENT	STND. ERROR	T-VALUE	PROB(>|T|)
CONSTANT	7.57298	0.590667	12.8211	.0000
tir	-0.025208	0.002341	-10.7679	.0000

0 CASES WITH MISSING VALUES WERE EXCLUDED.

ANALYSIS OF VARIANCE FOR THE FULL REGRESSION

SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F-RATIO	PROB(>F)
MODEL	216.55336	1	216.55336	115.94823	.00000
ERROR	741.46608	397	1.86767		
TOTAL (CORR.)	958.01944	398			

R-SQUARED = 0.226043
R-SQUARED (ADJ. FOR D.F.) = 0.224093
STND. ERROR OF EST. = 1.36663

NUMBER OF RESIDUALS = 399
SAMPLE AVERAGE = -1.18184E-15
SAMPLE VARIANCE = 1.86298
SAMPLE STANDARD DEVIATION = 1.36491
COEFF. OF SKEWNESS = 0.631307 STANDARDIZED VALUE = 5.14815
COEFF. OF KURTOSIS = 3.18392 STANDARDIZED VALUE = 0.749928
DURBIN-WATSON STATISTIC = 0.870912

DEP. VARIABLE: qrain

IND. VARIABLES: tv37
th37

CONSTANT: YES
WEIGHTS:
FLAGS: NO

MODEL FITTING RESULTS

VARIABLE	COEFFICIENT	STND. ERROR	T-VALUE	PROB(>|T|)
CONSTANT	19.465317	1.750232	11.1216	.0000
tv37	-0.111286	0.014829	-7.5046	.0000
th37	0.04539	0.01128	4.0240	.0001

0 CASES WITH MISSING VALUES WERE EXCLUDED.

ANALYSIS OF VARIANCE FOR THE FULL REGRESSION

SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F-RATIO	PROB(>F)
MODEL	206.14566	2	103.07283	54.28683	.00000
ERROR	751.87378	396	1.89867		
TOTAL (CORR.)	958.01944	398			

R-SQUARED = 0.215179
R-SQUARED (ADJ. FOR D.F.) = 0.211215
STND. ERROR OF EST. = 1.37792

NUMBER OF RESIDUALS = 399
SAMPLE AVERAGE = 1.46177E-14
SAMPLE VARIANCE = 1.88913
SAMPLE STANDARD DEVIATION = 1.37446
COEFF. OF SKEWNESS = 0.779154 STANDARDIZED VALUE = 6.35381
COEFF. OF KURTOSIS = 3.4147 STANDARDIZED VALUE = 1.6909
DURBIN-WATSON STATISTIC = 0.891631

DEP. VARIABLE: Rain

IND. VARIABLES: Tir
Tv37
Th37
Tv21
Th18
Tv18

CONSTANT: YES
WEIGHTS:
FLAGS: NO

MODEL FITTING RESULTS

VARIABLE	COEFFICIENT	STND. ERROR	T-VALUE	PROB(>|T|)
CONSTANT	56.22227	10.046432	5.5961	.0000
Tir	-0.063602	0.01181	-5.3853	.0000
Tv37	-0.20009	0.083881	-2.3854	.0175
Th37	-0.134466	0.070404	-1.9099	.0569
Tv21	-0.031804	0.076181	-.4175	.6766
Th18	0.23627	0.059103	3.9976	.0001
Tv18	-0.001192	0.054664	-.0218	.9826

0 CASES WITH MISSING VALUES WERE EXCLUDED.

ANALYSIS OF VARIANCE FOR THE FULL REGRESSION

SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F-RATIO	PROB(>F)
MODEL	5992.0876	6	998.6813	27.9342	.0000
ERROR	14014.489	392	35.751		
TOTAL (CORR.)	20006.576	398			

R-SQUARED = 0.299506
R-SQUARED (ADJ. FOR D.F.) = 0.288784
STND. ERROR OF EST. = 5.97923

DEP. VARIABLE: Rain

IND. VARIABLES: Tir
Tv18
Th18

CONSTANT: YES
WEIGHTS:
FLAGS: NO

MODEL FITTING RESULTS

VARIABLE	COEFFICIENT	STND. ERROR	T-VALUE	PROB(>|T|)
CONSTANT	37.676827	7.503113	5.0215	.0000
Tir	-0.097468	0.011565	-8.4276	.0000
Tv18	-0.071035	0.05495	-1.2927	.1969
Th18	0.037871	0.045075	.8402	.4013

0 CASES WITH MISSING VALUES WERE EXCLUDED.

ANALYSIS OF VARIANCE FOR THE FULL REGRESSION

SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F-RATIO	PROB(>F)
MODEL	3602.1587	3	1200.7196	28.9120	.0000
ERROR	16404.418	395	41.530		
TOTAL (CORR.)	20006.576	398			

R-SQUARED = 0.180049
R-SQUARED (ADJ. FOR D.F.) = 0.173821
STND. ERROR OF EST. = 6.44439

DEP. VARIABLE: Rain

IND. VARIABLES: Tir
Tv37
Th37

CONSTANT: YES
WEIGHTS:
FLAGS: NO

MODEL FITTING RESULTS

VARIABLE	COEFFICIENT	STND. ERROR	T-VALUE	PROB(>|T|)
CONSTANT	75.874235	7.941904	9.5537	.0000
Tir	-0.064839	0.012099	-5.3592	.0000
Tv37	-0.316427	0.071548	-4.4226	.0000
Th37	0.114158	0.051632	2.2110	.0276

0 CASES WITH MISSING VALUES WERE EXCLUDED.

ANALYSIS OF VARIANCE FOR THE FULL REGRESSION

SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F-RATIO	PROB(>F)
MODEL	5008.1615	3	1669.3872	[illegible]	.0000
ERROR	14998.415	395	37.971		
TOTAL (CORR.)	20006.576	398			

R-SQUARED = 0.250326
R-SQUARED (ADJ. FOR D.F.) = 0.244632
STND. ERROR OF EST. = 6.16203

DEP. VARIABLE: Rain

IND. VARIABLES: Tir

CONSTANT: YES
WEIGHTS:
FLAGS: NO

MODEL FITTING RESULTS

VARIABLE	COEFFICIENT	STND. ERROR	T-VALUE	PROB(>|T|)
CONSTANT	29.446932	2.785017	10.5733	.0000
Tir	-0.101668	0.011038	-9.2108	.0000

0 CASES WITH MISSING VALUES WERE EXCLUDED.

ANALYSIS OF VARIANCE FOR THE FULL REGRESSION

SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F-RATIO	PROB(>F)
MODEL	3522.6064	1	3522.6064	84.8385	.0000
ERROR	16483.970	397	41.521		
TOTAL (CORR.)	20006.576	398			

R-SQUARED = 0.176072
R-SQUARED (ADJ. FOR D.F.) = 0.173997
STND. ERROR OF EST. = 6.44371

DEP. VARIABLE: Rain

IND. VARIABLES: Tv37
Th37

CONSTANT: YES
WEIGHTS:
FLAGS: NO

MODEL FITTING RESULTS

VARIABLE	COEFFICIENT	STND. ERROR	T-VALUE	PROB(>|T|)
CONSTANT	83.088318	8.096317	10.2625	.0000
Tv37	-0.460369	0.068597	-6.7112	.0000
Th37	0.173211	0.052178	3.3196	.0010

0 CASES WITH MISSING VALUES WERE EXCLUDED.

ANALYSIS OF VARIANCE FOR THE FULL REGRESSION

SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F-RATIO	PROB(>F)
MODEL	3917.5972	2	1958.7986	48.2121	.0000
ERROR	16088.979	396	40.629		
TOTAL (CORR.)	20006.576	398			

R-SQUARED = 0.195815
R-SQUARED (ADJ. FOR D.F.) = 0.191754
STND. ERROR OF EST. = 6.37407

DEP. VARIABLE: Rain

IND. VARIABLES: Tv37
Th37
Tv21
Th18
Tv18

CONSTANT: YES
WEIGHTS:
FLAGS: NO

MODEL FITTING RESULTS

VARIABLE	COEFFICIENT	STND. ERROR	T-VALUE	PROB(>|T|)
CONSTANT	65.679897	10.238079	6.4153	.0000
Tv37	-0.316165	0.083903	-3.7682	.0002
Th37	-0.090172	0.07237	-1.2460	.2135
Tv21	-0.060591	0.078654	-.7704	.4416
Th18	0.262429	0.060965	4.3046	.0000
Tv18	-0.018599	0.056479	-.3293	.7421

0 CASES WITH MISSING VALUES WERE EXCLUDED.

ANALYSIS OF VARIANCE FOR THE FULL REGRESSION

SOURCE	SUM OF SQUARES	DF	MEAN SQUARE	F-RATIO	PROB(>F)
MODEL	4955.2441	5	991.0488	25.8769	.0000
ERROR	15051.332	393	38.299		
TOTAL (CORR.)	20006.576	398			

R-SQUARED = 0.247681
R-SQUARED (ADJ. FOR D.F.) = 0.238109
STND. ERROR OF EST. = 6.18858

Table B.4
Comparison of Simulated SSM/I- vs. SMMR-Channels

File No.	Sample Size	Sq. of Correlation SSM/I	SMMR	% Diff	Std. Error of Est. SSM/I	SMMR	% Diff
			For the Variable R				
1.1	666	0.231	0.354	-35	1.437	1.317	9
1.2	584	0.152	0.182	-16	1.084	1.065	2
1.3	380	0.470	0.559	-16	1.193	1.088	10
1.4	605	0.088	0.130	-32	0.498	0.469	6
1.5	606	0.146	0.237	-41	1.998	1.889	6
1.6	389	0.150	0.245	-39	0.896	0.844	6
1.7	695	0.373	0.431	-13	2.255	2.149	5
1.8	439	0.039	0.100	-61	0.576	0.557	3
2.1	313	0.635	0.632	0	1.231	1.175	0
2.2	472	0.429	0.466	-8	2.285	2.210	3
2.3	384	0.453	0.539	-16	1.844	1.691	9
"	384	0.134	0.139	-4	0.730	0.727	0
2.4	549	0.027	0.033	-18	0.122	0.121	1
2.5	353	0.416	0.463	-10	2.182	2.094	4
2.6	496	0.082	0.105	-22	1.082	1.068	1
2.7	439	0.287	0.375	-23	1.441	1.35	7
2.8	589	0.693	0.736	-6	1.250	1.160	8
2.9	235	0.762	0.762	0	5.034	5.041	0
2.10	269	0.296	0.314	-6	2.794	2.758	1
3.1	432	0.553	0.750	-26	1.903	1.423	34
3.2	465	0.468	0.523	-11	2.245	2.126	1
3.3	356	0.481	0.629	-24	2.624	2.219	18
3.4	361	0.432	0.548	-21	2.175	1.940	11
3.5	40	0.881	0.914	-4	0.797	0.678	18
3.6	316	0.535	0.594	-10	1.228	1.148	7
3.7	354	0.263	0.354	-26	1.061	0.993	7
3.8	283	0.395	0.484	-18	2.450	2.264	8

Weighted Means

Season	Sq. of Corr.			Std. Error		
Spring	0.162	0.239	-32	1.524	1.435	6
Summer	0.332	0.349	-5	1.637	1.488	10
Fall	0.445	0.567	-27	1.977	1.626	22
Three Seasons	0.313	0.385	-19	1.713	1.516	13

Table B.4 continued

File No.	Sample Size	Sq. of Correlation SSM/I	SMMR	% Diff	Std. Error of Est. SSM/I	SMMR	% Diff
For the Variable Q							
1.1	666	0.292	0.391	-25	0.512	0.474	30
1.2	584	0.181	0.207	-13	0.446	0.439	2
1.3	380	0.432	0.522	-25	0.392	0.360	1
1.4	605	0.116	0.154	-25	0.287	0.281	2
1.5	606	0.186	0.302	-38	0.772	0.715	8
1.6	389	0.164	0.252	-35	0.438	0.415	6
1.7	695	0.394	0.451	-13	0.690	0.656	5
1.8	439	0.058	0.148	-61	0.362	0.344	5
2.1	313	0.519	0.518	0	0.506	0.506	0
2.2	472	0.420	0.454	-7	0.623	0.604	3
2.3	384	0.479	0.546	-12	0.581	0.542	7
"	384	0.134	0.139	-4	0.730	0.727	0
2.4	549	0.062	0.076	-18	0.111	0.110	1
2.5	353	0.374	0.424	-10	0.718	0.688	4
2.6	496	0.072	0.103	-30	0.427	0.420	2
2.7	439	0.300	0.372	19	0.585	0.554	6
2.8	589	0.669	0.687	-3	0.334	0.324	3
2.9	235	0.808	0.809	0	0.815	0.818	0
2.10	269	0.333	0.355	-6	0.805	0.792	2
3.1	432	0.562	0.709	-21	0.608	0.496	23
3.2	465	0.479	0.540	-11	0.664	0.624	6
3.3	356	0.546	0.699	-22	0.762	0.620	23
3.4	361	0.415	0.559	-26	0.758	0.659	11
3.5	40	0.789	0.824	-4	0.361	0.329	10
3.6	316	0.556	0.616	-10	0.422	0.392	8
3.7	354	0.305	0.402	-24	0.401	0.372	8
3.8	283	0.505	0.612	-17	0.704	0.622	13

Weighted Means

Season	Sq. of Corr.			Std. Error		
Spring	0.166	0.239	-31	0.528	0.492	7
Summer	0.225	0.263	-14	0.533	0.511	4
Fall	0.471	0.491	-4	0.621	0.551	13
Three Seasons	0.287	0.331	-13	0.561	0.518	8

Table B. 5
Comparison of SMMR and Simulated SSM/I Plus IR

File No.	No. Cases	Square of Corr. SMMR	Square of Corr. SSM/I+	Standard Error SMMR	Standard Error SSM/I+
		For The Variable R			
2. 1	313	0. 632	0. 667	1. 235	1. 175
2. 2	472	0. 466	0. 460	2. 210	2. 223
2. 3	384	0. 539	0. 478	1. 691	1. 800
"	384	0. 139	0. 150	0. 727	0. 723
2. 4	549	0. 033	0. 047	0. 121	0. 121
2. 5	353	0. 463	0. 447	2. 094	2. 125
2. 6	496	0. 105	0. 121	1. 068	1. 059
2. 7	439	0. 375	0. 444	1. 349	1. 273
2. 8	589	0. 736	0. 701	1. 159	1. 234
2. 9	235	0. 762	0. 761	5. 039	5. 041
2. 10	269	0. 314	0. 357	2. 758	2. 670
Weighted Means		0. 3955	0. 4000	1. 538	1. 543
Mean Difference			0. 00449		0. 00473
RMS Difference			0. 03506		0. 05610
Mean % Difference			1. 135		0. 308
RMS % Difference			17. 7		3. 85
		For The Variable Q			
2. 1	313	0. 518	0. 618	0. 506	0. 450
2. 2	472	0. 454	0. 475	0. 604	0. 593
2. 3	384	0. 546	0. 539	0. 542	0. 547
2. 35	384	0. 147	0. 165	0. 320	0. 316
2. 4	549	0. 076	0. 094	0. 110	0. 109
2. 5	353	0. 424	0. 413	0. 688	0. 695
2. 6	496	0. 103	0. 136	0. 420	0. 412
2. 7	438	0. 372	0. 460	0. 554	0. 514
2. 8	589	0. 687	0. 739	0. 324	0. 296
2. 9	235	0. 808	0. 821	0. 818	0. 789
2. 10	269	0. 355	0. 389	0. 792	0. 771
Weighted Means		0. 3878	0. 4211	0. 474	0. 458
Mean Difference			-0. 03328		0. 01581
RMS Difference			0. 04623		0. 02401
Mean % Difference			-7. 904		3. 335
RMS % Difference			11. 0		5. 06

8909180876b
b89091808766a

Table B. 6

Percent of days with significant regression coefficients

Const.	h37	v37	v21	h18	v18	Ind. Var. / Case
93	81	100	74	11	37	R IR not used
91	82	100	45	9	36	R IR used
93	85	96	74	19	26	Q IR not used
91	82	91	55	18	18	Q IR used