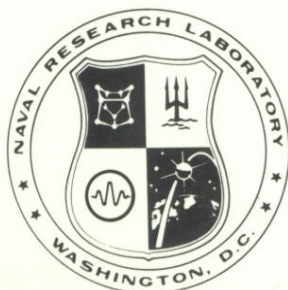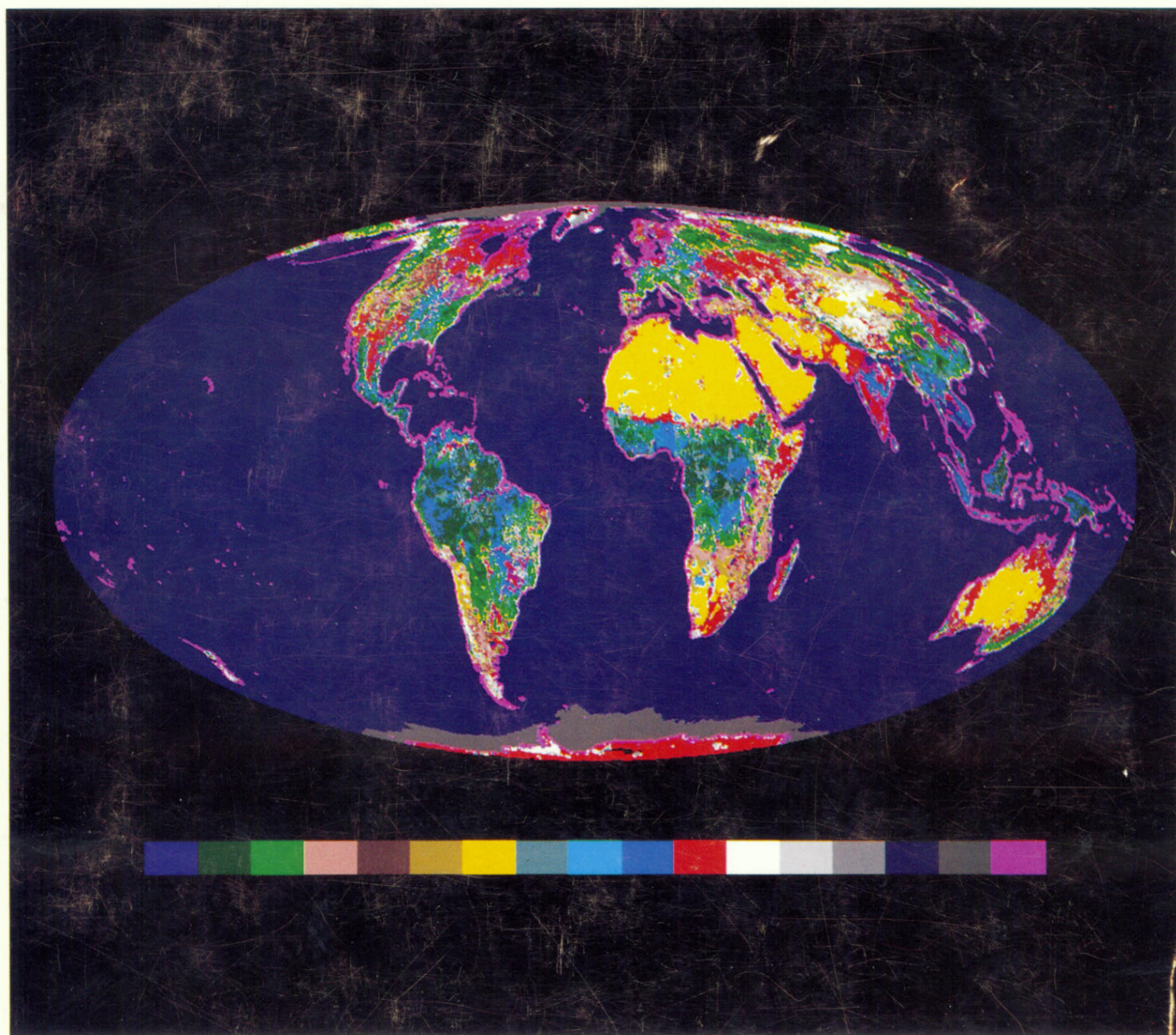# DMSP SPECIAL SENSOR MICROWAVE/IMAGER CALIBRATION/VALIDATION

NAVAL
RESEARCH
LABORATORY
WASHINGTON, D.C.

FINAL REPORT VOLUME II

Cover Picture

The cover picture shows retrieved land surface types using the SSM/I land surface typing algorithm given in Section 9.1. The Mollweide equal-area projection is composed of SSM/I measurements from 64 consecutive revolutions of the DMSP satellite over a four day period starting September 21, 1987. The surface of the earth has been divided into 17 separate surface categories. Each of the categories has been assigned a unique color to produce this image. The land surface types are as follows. Starting at the left end of the color bar, medium blue represents the land surface type of standing water or flooded conditions. Examples of this type can be seen in Bangladesh, Nepal, and Thailand. The dark green represents dense vegetation, as seen in parts of Brazil and central Africa, followed by light green which is dense agricultural/rangeland vegetation and can be seen in Argentina and the east coast of the United States. Dry arable soil is next, shown as beige, followed by moist soil colored brown, semi-arid surfaces tan and desert yellow. Dry arable soil can be seen in Angola, Zambia, and the northern Great Plains of North America. Examples of moist soil can also be seen in the northern Great Plains of North America. Semi-arid conditions exist along the Andes in South America and north of the Kalahari Desert in Africa. Desert surface types are seen in the Sahara, the Arabian Peninsula, Australia and other regions of the world. These are followed by precipitation over vegetation shown as blue gray and seen in small regions in South America and central Africa and precipitation over soil shown as turquoise and seen, for example, in southern Africa. Composite vegetation and water is shown as light blue and is seen in large regions of South America and Africa, composite soil and water/wet soil surface is shown as red and is seen in Canada and other regions. The land surface types finish with three snow types. Dry snow is white and is seen in central Asia, wet snow is light gray which can also be seen in central Asia, and refrozen snow, which is medium gray, can be found in Greenland. The black areas in Greenland and Antarctica are regions where a land surface type could not be identified. The ocean is divided into 2 categories. Dark blue which represents open ocean and dark gray which is sea ice. The final category is violet which designates the coastal regions.

Cover picture by
David J. Spangler

DMSP SPECIAL SENSOR MICROWAVE/IMAGER

CALIBRATION/VALIDATION

DMSP SSM/I Cal/Val Team

Coordinated by

J. P. Hollinger
Radio/IR/Optical Sensors Branch
Naval Research Laboratory
Washington, DC 20375-5000

FINAL REPORT
VOLUME II

20 May 1991

ACKNOWLEDGEMENT

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF TABLES

SECTION 7

WATER VAPOR AND CLOUD WATER VALIDATION

by

John C. Alishouse
National Oceanic and Atmospheric Administration
National Environmental Satellite, Data and Information Service
Washington, D.C.

and

Sheila Snyder[1], Jennifer Vongsathorn[2] and Ralph R. Ferraro
S. M. Systems and Research Corporation
Landover, MD

[1]Present affiliation: Computer Sciences Corp., Calverton, MD
[2]Present affiliation: Vandair Corp., Bethesda, MD

7.0 WATER VAPOR AND CLOUD WATER VALIDATION

7.1 INTRODUCTION

Two of the many useful geophysical parameters that the SSM/I can measure are the amount of water vapor and the amount of cloud liquid water between the ocean's surface and the top of the atmosphere. The water content of the atmosphere is very important for meteorology, climatology, and hydrology. The evaporation of water from the ocean surface and its condensation into clouds and precipitation is an important energy transport mechanism for the dynamics of the atmosphere. The amount of liquid water in clouds affects the incoming and outgoing radiative fluxes. The water that eventually falls as precipitation over land comes from the ocean.

The objectives of this investigation were to validate the initial or Hughes algorithms for total precipitable water and cloud liquid water and, if necessary, derive a new or improved algorithm. The Hughes algorithms (see [1]) are divided into eleven (11) climate codes per hemisphere. Each climate code represents a set of coefficients for a particular latitude zone and season. There are three distinct sets of coefficients for the retrieval of water vapor and nine for the retrieval of cloud liquid water over the ocean. For the retrieval of cloud liquid water over land, there are eleven distinct sets of coefficients.

Validating the Hughes algorithm required the acquisition of surface measurements from a variety of latitude zones and seasons. For the total water vapor validation, radiosonde data from small island stations and the few remaining weather ships was collected. Initially a list of 49 potential stations was compiled with size and latitude being the only considerations. Once the various match-up criteria were invoked, data from only 19 stations were used. The criteria were that the satellite observation and radiosonde must be coincident within 2 hours and 2 degrees of latitude and longitude. These radiosonde observations were collected from National Meteorological Center (NMC) files, integrated to obtain the total precipitable water, and matched with the satellite data. The period of collection of data was from June, 1987 to August, 1988.

The selection of sites for surface measurements is important in that the launching site, be it ship or island, must be sufficiently small so that it does not affect the radiometer measurements.

Initially it was planned that surface observations of cloud liquid water would be obtained from upward looking radiometers and aircraft. Due to delays in the launch of the SSM/I, it was not feasible to acquire aircraft data. Data for the cloud liquid water determinations over the ocean were taken by NOAA/Wave Propagation Laboratory (WPL) personnel from San Nicholas Island as part of Project FIRE and by University of Massachusetts (UMass) personnel from Kwajalein Island. Data over land were taken by NOAA/WPL from the four stations that make up the Colorado remote profiling network.

NESDIS contracted with S. M. Systems and Research Corporation (SMSRC) of Landover, MD, to develop and run the necessary software to do the comparisons and the algorithm refinement.

Since the preparation of material for Volume I of the Cal/Val Team Final Report additional analysis has been done and additional data have been acquired. The continuing process of algorithm development and refinement has led to some minor changes in the coefficients used to derive total precipitable water and completely different algorithms for cloud liquid water. Additionally we are now expressing our results in the more commonly used kg/m^2.

7.2   DATA PROCESSING

The data handling procedures described in this section were devised, encoded, and executed by SMSRC personnel. The details are given in [2].

7.2.1   Data Handling - Total Precipitable Water

1.   Radiosonde Observation (Raob) Collection

Raob reports were collected daily from the NMC ADPUPA files for 00 UTC and 12 UTC for the selected radiosonde stations. The job was submitted operationally beginning June 23, 1987, and ending August 3, 1988. All reports found for the selected stations were picked up without regard to quality of data.

2.   Surface Report Collection

Reports of surface conditions from the radiosonde stations were picked up immediately after the raob collection job finished executing. This job was submitted operationally from July 3, 1987, to August 3, 1988.

3.   Collection of SSM/I Data

Matches between SSM/I data and collected raob reports were predicted using a version of SMIOPS which runs on the NAS 9050. Based on these predictions selected SSM/I revolutions were requested from NRL. When data were received, the tapes are mounted on the system. Those SDR files on tape which contained desired data were read and the data (brightness temperatures) were unpacked and stored in an 'SSM/I temporary file'.

4.   Matching between SSM/I Data and Raob Data

A match program was run which read the SSM/I temporary file and the raob holding file and found the four closest SSM/I footprints to each raob report. All matches had to be within 2 hours and 2 degrees latitude and longitude of the raob station.

5.   Calculation of Total Precipitable Water

A program, which was based on the FNOC code provided to us, read the matches from the match file, and calculates the total precipitable water values from the SSM/I brightness temperatures. Values which were outside specified limits were set to 12.70. Values which could not be calculated due to presence of rain or ice were set to 12.75.

6.   Quality Control

Resulting matches were printed in a summary report. Matches which have SSM/I retrieved or calculated values of total precipitable water which were unreasonable were investigated. If necessary, these matches were removed. Raobs were screened for missing surface pressure values.

7.   Statistics

Means of raob and SSM/I retrieved precipitable water values, the bias, and RMS errors between them were calculated.

8.   Regression

Raob - retrieval matches were used in a regression to obtain total precipitable water as a function of brightness temperatures.

7.2.2 Data Handling - Cloud Liquid Water

1.   Observation Collection

Cloud liquid water observations were received from the various observation sites. The data were then reformatted into the 'SSM/I temporary file' format.

2.   Collection of SSM/I Data

Matches between SSM/I data and cloud liquid water observations were predicted. All SSM/I data within 2 degrees latitude and longitude of the observation sites were collected.

3.   Matching between SSM/I Data and Observations

A match program was run which picked the four SSM/I retrievals which were closest to the observation site. Then the observation which was closest in time to the SSM/I overpass was chosen for the match.

4.   Calculation of Cloud Liquid Water

A program, based on the FNOC code, read the match file and calculated the cloud liquid water value using the SSM/I brightness temperatures.

5.   Statistics

Means and RMS errors between the observed and retrieved cloud liquid water values were calculated.

7.3   TOTAL PRECIPITABLE WATER (WATER VAPOR)

7.3.1 Surface Data Sources

The major source of surface data for the validation of the SSM/I determinations of total precipitable water was the international radiosonde network. The pressure, temperature, and humidity data from the radiosondes were integrated numerically to give a value which could be compared with SSM/I values. It was required that the radiosonde station be a small island or one of the remaining weather ships. Small is defined as less than 18% of the instantaneous field-of-view (IFOV) of the 19 GHz channels. Initially a list of 49 possible stations was prepared. These stations are shown in Table 7.1. Most radiosondes are launched at 0 and 12 UTC, while the SSM/I has an approximately 0600 LST ascending node. Of the 49 possible stations, matches from only 19 were obtained. The stations indicated with an asterisk are the "match-up" stations. To be considered a "match-up" it was required that the radiosonde and satellite measurements be coincident within 2 hours and 2 degrees of latitude and longitude.

The radiosondes measure pressure, temperature, and humidity at various levels in the atmosphere. These measurements are then transmitted worldwide to various meteorological centers, including the U. S. National Meteorological Center. The raobs and selected surface observations were combined with matching SSM/I brightness temperatures to form a data set which could be used to evaluate algorithms for deriving total precipitable water over the ocean.

The total precipitable water was calculated from the equation

$$U = 1/2g \ \Sigma \ (q_i + q_{i+1})(p_i - p_{i+1}) \tag{1}$$

where g = acceleration of gravity, q_i = the mixing ratio of water vapor to dry air at the ith level, and p_i = pressure at the ith level. The units of U are kg/m^2, thus requiring g to be expressed in m/s^2, q_i in kg/kg, and p_i in newtons/m^2.

TABLE 7.1 RADIOSONDE STATIONS						
NAME	NUMBER	LAT	LONG		AREA	IFOV
MACQUARIE IS.*	94998	-54.50	158.95	E	109	4.41
MARION IS	68994	-46.88	37.87	E	388	15.71
GOUGH IS.	68906	-40.35	9.88	W	83	3.36
I.N. AMSTERDAM*	61996	-37.80	77.53	E	62	2.51
KERMADEC IS.	93997	-29.25	177.92	W	34	1.38
NORFOLK IS.	94996	-29.05	167.93	E	34	1.38
AUSTRAL IS.	91958	-27.62	133.33	W	47	1.90
EASTER IS.*	85469	-27.17	109.43	W	117	4.74
TOTEGEGIE	91948	-23.10	134.87	W	31	1.26
COOK ISLES	91843	-21.20	159.82	W	218	8.83
TRINDADE IS.	83650	-20.50	29.32	W	10	.40
TUAMOTU	91944	-18.07	140.95	W	OK	N/A
ST. HELENA	61901	-15.97	5.70	W	122	4.94
PAGO PAGO	91765	-14.33	170.72	W	135	5.47
COCOS IS.*	96996	-12.18	96.83	E	14	.57
ATUONA	91925	-9.82	139.02	W	200	8.10
PENRHYN	91801	-9.02	158.07	W	10	.40
FUNAFUTI	91643	-8.52	179.22	E	2.80	.11
ASCENSION IS.	61902	-7.97	14.40	W	88	3.56
DIEGO GARCIA*	61967	-7.35	72.48	E	152	6.15
MAJURO	91367	7.03	171.38	E	10	.40
KOROR	91408	7.33	134.48	E	8	.32
TRUK	91334	7.47	151.85	E	118	4.78
KWAJALEIN	91366	8.72	167.73	E	16	.65
YAP	91413	9.48	138.08	E	54	2.19
ISLA SAN ANDREAS*	80001	12.58	81.70	W	20.50	.83
TARAWA	91610	13.05	172.92	E	23	.93
BARBADOS*	78954	13.07	59.50	W	431	17.45
JOHNSTON IS.	91275	16.73	169.52	W	1.30	.15
SAN MAARTEN*	78866	18.05	63.12	W	85	3.44
WAKE IS.	91245	19.28	166.65	E	8	.32
ROBERTS FLD.*	78384	19.30	81.37	W	183	7.41
MARCUS IS.	47991	24.30	153.97	E	2.60	.11
ISHIGAKIJIMA*	47918	24.33	124.17	E	215	8.70
MINAMIDAITO JIMA*	47945	25.83	131.23	E	46.6	1.89
CHICHI JIMA	47971	27.08	142.18	E	24.60	1.00
MIDWAY	91066	28.22	177.37	W	15	.61
TANGO	C7T	29	135	E	OK	N/A
KINDLEY FIELD*	78016	32.37	64.68	W	53	2.15
HACHIJA JIMA*	47678	33.12	139.78	E	69.90	2.83
ROMEO	C7R	47	17.00	W	OK	N/A
SHEMYA IS.	70414	52.72	174.10	E	21	.85
COCA*	C7C	52.75	35.50	W	OK	N/A
LIMA*	C7L	57	10.00	W	OK	N/A
ST. PAUL IS.*	70308	57.15	170.22	W	90.60	3.67
MIKE*	C7M	66	2.00	E	OK	N/A
JAN MAYEN	01001	70.93	8.67	W	373	15.10
BJORNOYA*	01028	74.52	19.02	E	179	7.25

In Table 7.1 the entries are the name of the station, its World Meteorological Organization number, latitude with degrees south expressed as minus, longitude, the area in square kilometers, and the percentage of the 19.35 GHz IFOV that the station occupies. We were unable to obtain exact estimates of the areas of some of the islands. We were able to verify that they were small in comparison to the 19 GHz footprints. The weather ships also were assumed to be small.

7.3.2 Comparisons

7.3.2.1 Initial Algorithm

The initial algorithms for retrieving SSM/I geophysical parameters are described in [1] and will be referred to in this section as the Hughes algorithm or more precisely the Hughes algorithms as in reality there are several algorithms. The Hughes algorithm is divided into eleven climate codes for each hemisphere, each of which relates to selected latitude zones and/or seasons. This approach permits "fine tuning" the coefficients for a particular climate, however the boundaries between the latitude zones are "hard" and the climate changes occur instantaneously. This approach can lead to unnaturally large gradients in parameters at these boundaries. In addition to validating the algorithms for each climate code, it was felt that it was necessary to check for the existence of these boundary gradients. Figure 7.1 shows the boundary discontinuities between climate codes. This figure shows data for August 11, 1987 for revs 740, 741, and 742. The land mass in rev 742 is Africa and Saudi Arabia. Deep red denotes flagged areas which are either land or areas of precipitation. Revs 740 and 741 show boundaries at 25 degrees south latitude and the equator. The equatorial boundary shown in yellow to the south and red to the north is particularly noticeable. The boundary at 25 degrees south is noticeable as a line between lighter and darker blue. The sharp boundary near 20 degrees south is a boundary between air masses. Lesser amounts of water vapor are shown in darker blue and increasing amounts are shown in lighter blue, yellow, and light red.

Table 7.2 shows the comparisons for the Hughes algorithms with radiosonde determinations for the latitude zones and also globally. All entries in the table are kg/m^2 or precipitable millimeters. The columns labeled mean show the mean value for all the retrievals and radiosondes for that particular latitude zone. The columns labeled standard deviation (STD DEV) are the natural variance of the sample set. This is the variance exhibited by the total precipitable water in this climate zone. The column labeled rms diff is the rms difference between the SSM/I retrieval and the corresponding radiosonde value. The column labeled bias is the difference between the mean SSM/I retrieval and the mean radiosonde retrieval. A negative bias indicates an underestimate by the SSM/I and a positive bias indicates an overestimate.
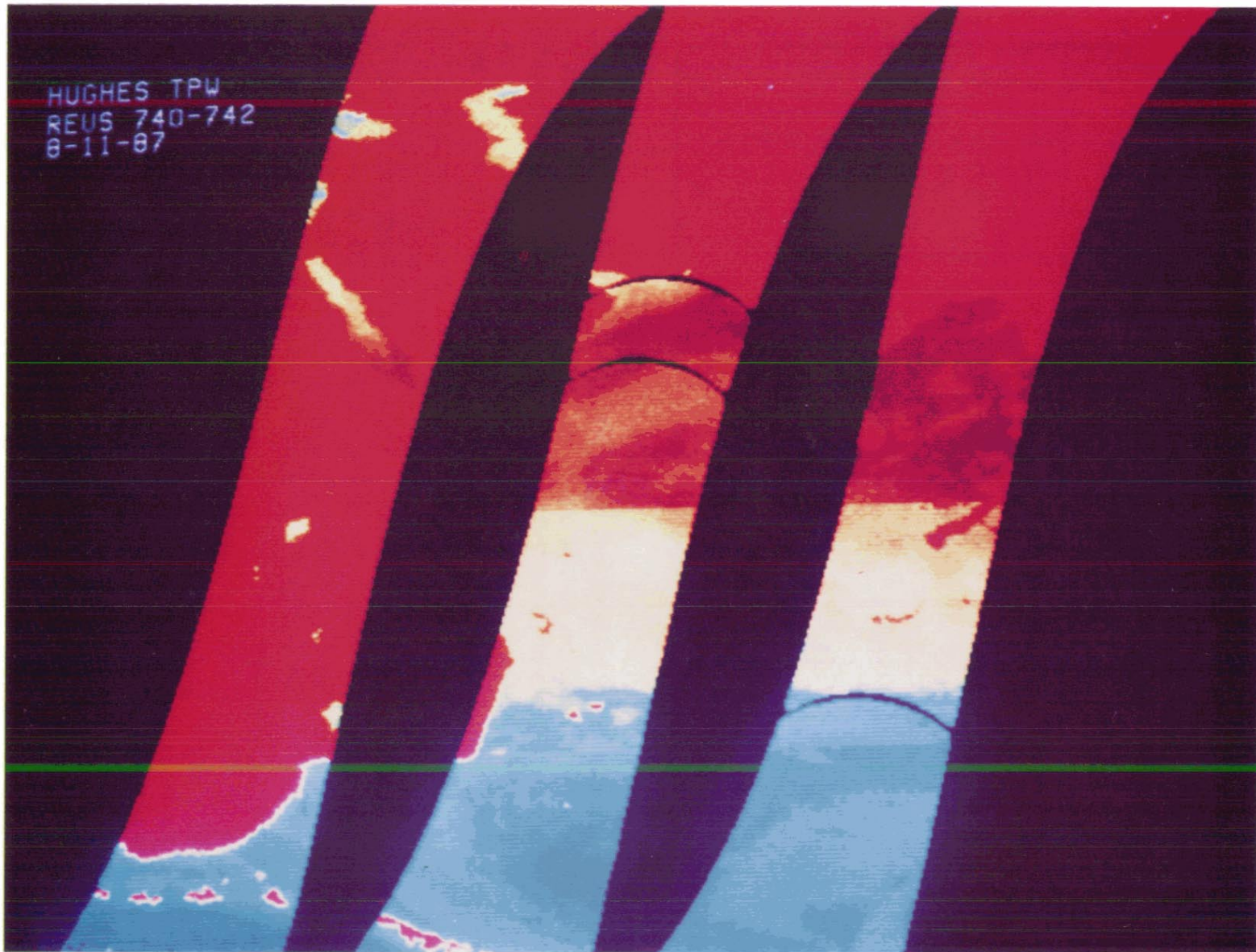
Figure 7.1 – Revs 740-2 showing gradients at the geographic boundaries of the climate codes of the Hughes algorithm. The area shown is the Indian Ocean. Gradients are notable at the equator and at 25 S latitude. Lesser amounts of water vapor are shown in blue and increasing amounts are shown in yellow and red. Solid red areas are flagged values and indicate land or precipitation.

LATITUDE	SAMPLE	MEAN		STD DEV		RMS	BIAS
	SIZE	RET	RAOB	RET	RAOB	DIFF	
ZONE							
60-90 N	209	8.7	11.0	4.3	4.2	3.2	-2.3
55-60 N	37	12.7	13.3	10.6	8.0	3.8	-0.6
25-55 N	59	23.6	27.1	15.0	14.0	5.2	-3.5
20-25 N	35	45.3	38.9	13.8	12.2	7.4	6.4
0-20 N	134	44.2	39.8	11.7	9.0	6.7	4.4
0-20 S	66	53.8	50.4	10.2	9.3	6.3	3.4
20-25 S	0						
25-55 S	47	23.5	22.2	11.1	9.9	3.8	1.3
55-60 S	0						
60-90 S	0						
GLOBAL	587	27.0	26.3	20.1	16.9	5.1	0.7

TABLE 7.2
HUGHES ALGORITHM RETRIEVALS

The algorithm used in the polar regions has a distinct tendency to underestimate the amount of water vapor that is present and the algorithm used in the warm tropics shows a tendency to overestimate the amount of water vapor that is present. These two tendencies effectively cancel each other as the global data set shows a negligible bias. All of the rms differences are larger than the desired $\pm$ 2.0 kg/m^2. Figure 7.2 is a scatter plot of the global data set.

At least two factors are sources of differences between the radiosonde and the SSM/I derived values of total precipitable water. One is errors in the radiosonde determinations of temperature, pressure, and humidity. A coefficient of variance of 0.042 for US radiosondes was obtained in [3]. This translates into an error of 1.1 kg/m^2 for this sample set. The other factor is small scale variability in water vapor. An estimate of this was obtained by comparing the four values derived from the SSM/I with each other. The rms difference between the four samples for each raob match-up is 1.5 kg/m^2. When these two factors are taken into account the rms difference between raobs and the Hughes algorithm becomes 4.7 kg/m^2. The data presented in Table 7.2 and shown in Figure 7.2 are from the trimmed data set. The trimming procedure is discussed in Section 7.3.2.2.
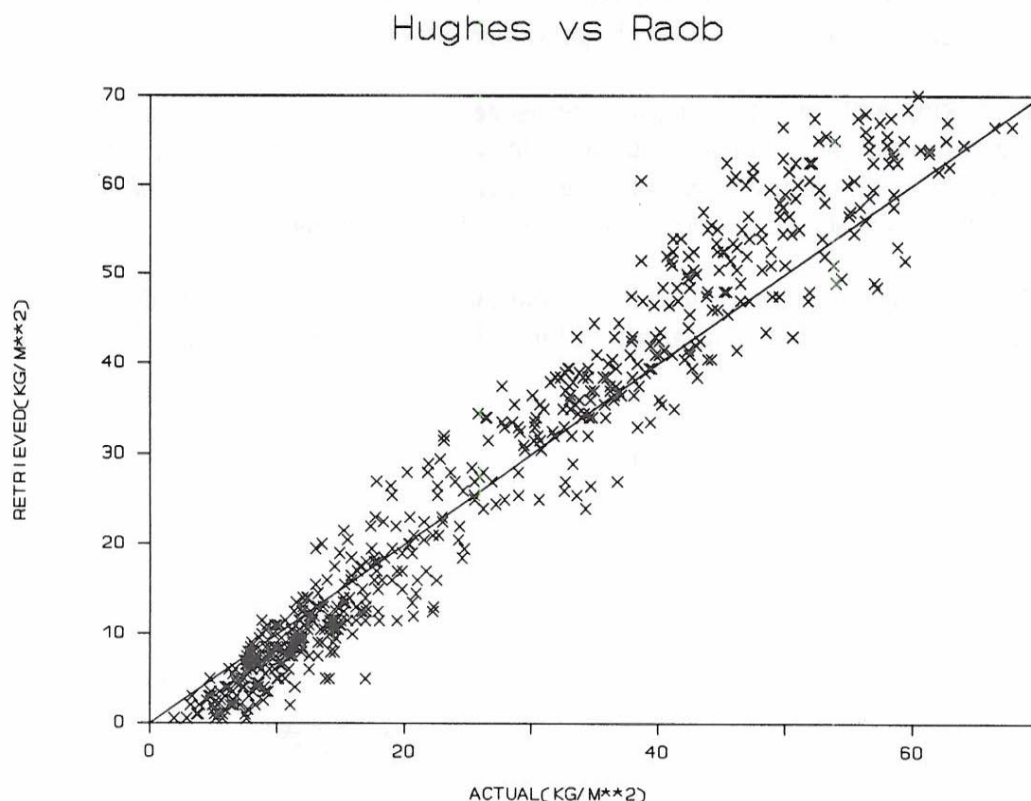
Hughes vs Raob

Figure 7.2 - Retrievals from the Hughes algorithm vs raobs. Units are kg/m^2 or precipitable millimeters. Values are from the trimmed data set and are composited from the climate codes that comprise the Hughes algorithm for total precipitable water.

7.3.2.2 Algorithm Improvement

Previous experience with the SMMR instruments on SEASAT and Nimbus 7 shows that it is possible to achieve rms differences between satellite and radiosonde determinations in the range 2.0 to 2.5 kg/m^2 [3] and [4]. In addition the SMMR algorithms are global and do not introduce latitudinal or seasonal discontinuities in the retrieved water vapor maps.

A global linear algorithm was determined to reduce the retrieval errors presented in Table 7.2. A statistical regression between the set of SSM/I brightness temperatures and the corresponding total water vapor as determined from the raobs was used. This preliminary algorithm was presented at the July 1988 Cal/Val Team meeting. This algorithm was based on matches that had been obtained up to that time. The data set was biased in that there was an

over representation of arctic soundings which biased the sample toward small values of water vapor. When additional soundings from the tropics were obtained, it was noted that the preliminary algorithm did not estimate large values of water vapor very well.

In working with such a large and heterogeneous data set, there are many possibilities for errors. To eliminate the erroneous data that had not been eliminated earlier, 2% of the largest positive and 2% of the largest negative differences between raob and retrieval were eliminated (i.e. trimmed) and the statistics were re-computed. This trimming procedure is described in [5].

Using standard regression techniques and the larger, more complete data set, an attempt was made to develop a global, linear algorithm. Figure 7.3 shows the best linear relationship
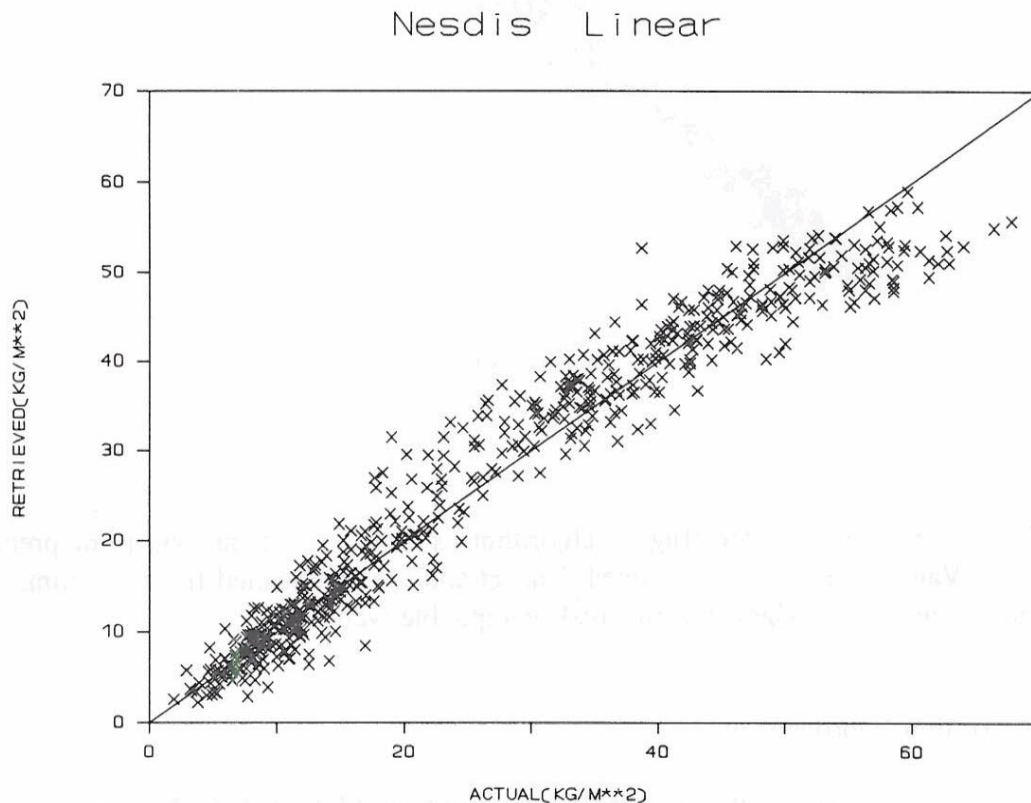
Figure 7.3 - Retrievals from a linear, global algorithm developed at NESDIS vs raobs. Units are kg/m^2 or precipitable millimeters. The linear algorithm has significant non-linearities in the retrievals. It shows a tendency to overestimate at medium values and underestimate at larger values.

for the dependent data set. In Figure 7.3 there appears to be a non-linear relationship between total precipitable water as deduced from radiosondes and that deduced from the SSM/I. The linear algorithm overestimates water vapor in the mid-range and underestimates large values. This observation plus a review of previous work [6-8] led to the consideration of a non-linear algorithm.

The square of the 22 GHz brightness temperature was introduced as a predictor and the regressions were performed as before. Figure 7.4 is a scatter plot of the best four channel non-linear algorithm. This algorithm uses 19V, 22V, 37V, and 22V squared. Thus the equation becomes

$$TPW = b_0 + b_1 T_{B19V} + b_2 T_{B22V} + b_3 (T_{B22V})^2 + b_4 T_{B37V} \tag{2}$$

The coefficients are given in Table 7.11.

Table 7.3 gives the statistics for this algorithm for both the global data set and latitude zones of the Hughes algorithm. When the radiosonde precision and small scale variability of water vapor are taken into account, the rms difference becomes 2.4 kg/m^2.

TABLE 7.3 IMPROVED NON-LINEAR ALGORITHM							
LATITUDE ZONE	SAMPLE SIZE	MEAN		STD DEV		RMS DIFF	BIAS
		RET	RAOB	RET	RAOB		
60-90 N	209	11.0	11.0	3.8	4.2	2.0	0.0
55-60 N	37	14.1	13.3	7.7	8.0	1.9	0.8
25-55 N	59	26.2	27.1	14.4	14.0	3.3	0.9
20-25 N	35	41.3	38.9	11.1	12.2	3.7	2.4
0-20 N	134	40.3	39.8	9.2	9.0	3.5	0.5
0-20 S	66	48.2	50.3	8.2	9.3	4.3	-2.1
20-25 S	0						
25-55 S	47	22.8	22.2	10.2	9.9	2.6	0.6
55-60 S	0						
60-90 S	0						
GLOBAL	587	26.3	26.3	16.7	16.9	3.0	0.0

The non-linear algorithm still shows a tendency to underestimate at the highest water vapor values, but overestimates slightly for the next two largest classes. There should be no
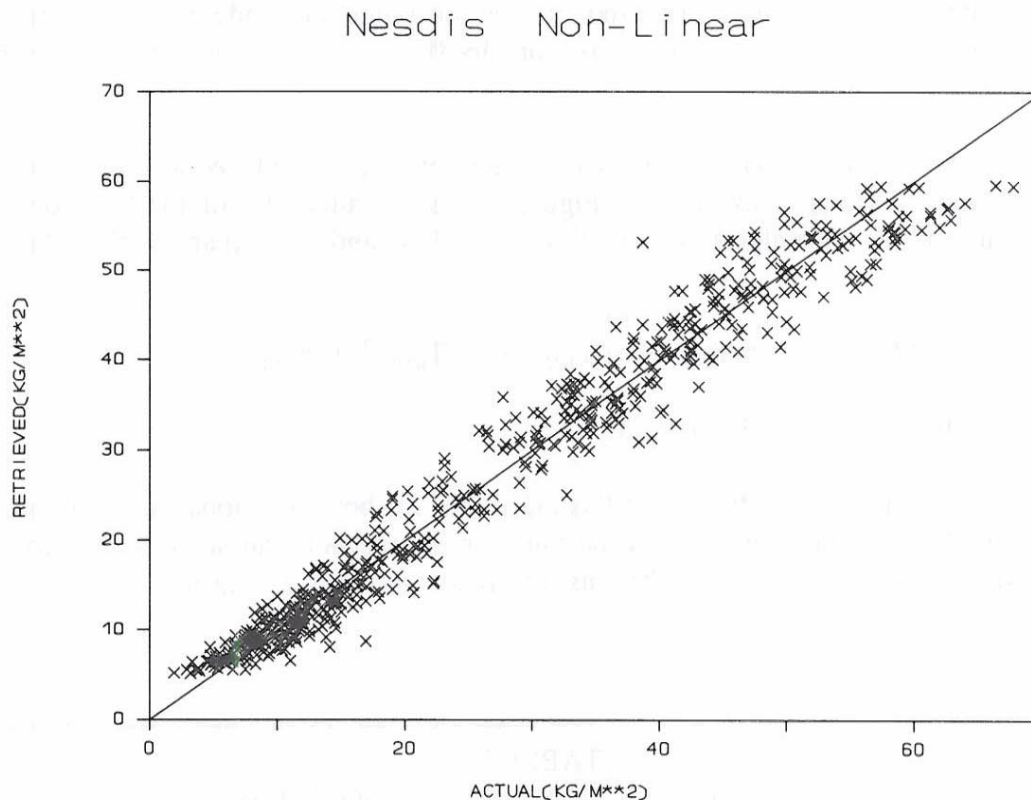
Figure 7.4 - Retrievals from the NESDIS or improved algorithm vs raobs. Units are kg/m² or precipitable millimeters. Values are from the trimmed data set and were derived with a global, non-linear algorithm.

bias in the global data set as it is the dependent data set. Another encouraging fact is that the standard deviations of the raob and predicted (retrieved) sets are about the same globally and in the latitude zones. Figure 7.4 also gives a hint of the tendency to underestimate at large values and also a hint of a tendency to overestimate at the lowest values.

Additionally, we investigated a segmented non-linear algorithm using the square of the 22V brightness temperature with a weighted average in the transition zones. The best of these gives very slightly better results than the global algorithms, but has not been implemented due to coding complexity.

The results presented above are for dependent data that were taken between July, 1987 and April, 1988. We continued to collect raobs until August, 1988. Thus the data taken from

May through July, 1988 constitute an independent data set. Figure 7.5 is a plot of retrieved vs. observed for the independent data set.
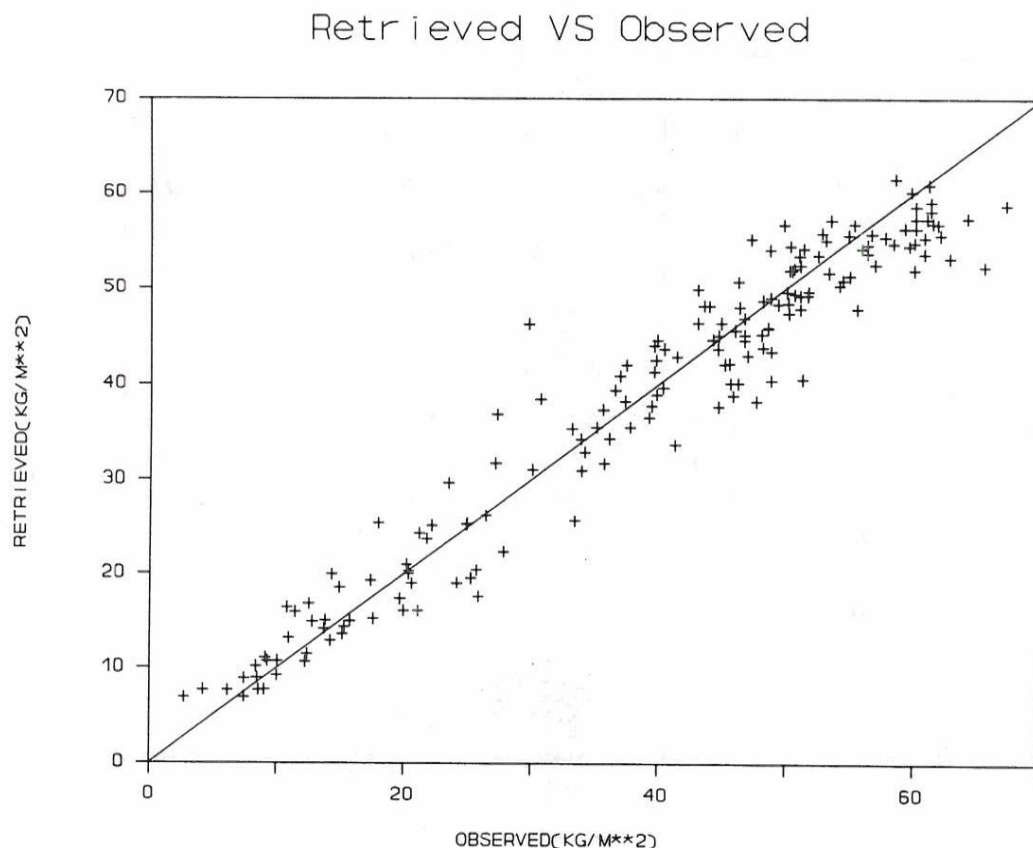
Figure 7.5 - Retrievals from the non-linear algorithm for an independent data set collected May-July, 1988. Statistics are in Table 7.4.

The fractional error {(rms diff/mean) x 100} of the water vapor data shown in Table 7.4 is 10.8% which is very comparable to the dependent data set's fractional error of 11.4%. All entries in Table 7.4 are in kg/m^2.

In addition to the non-linear algorithm used to produce the results thus far presented, two other non-linear algorithms have been suggested [8] and [9]. The non-linear algorithm of [8] uses variables of the form log (T_0 - T_B), where T_B is the brightness temperature and T_0 is a threshold temperature greater than any T_B. Using our dependent data set of 587 observations, we derived coefficients for an algorithm using four logarithmic variables. The rms difference between predicted and observed is 2.97 kg/m^2. The best four channels are 19H, 22V, 37V, and

85H. The 22V is obviously the dominant predictor and other combinations involving 22V do almost as well.

TABLE 7.4 INDEPENDENT DATA SET							
LATITUDE ZONE	SAMPLE SIZE	MEAN		STD DEV		RMS DIFF	BIAS
		RET	RAOB	RET	RAOB		
60-90 N	33	14.6	14.2	5.4	6.3	3.3	0.4
55-60 N	0	0	0	0	0	0	0
25-55 N	19	35.1	36.7	11.3	12.1	5.2	-1.6
20-25 N	16	46.5	45.3	12.4	12.7	4.1	1.2
0-20 N	40	45.6	46.4	7.1	8.8	5.1	-0.8
0-20 S	51	50.3	51.9	7.0	8.3	3.8	-1.6
20-25 S	0	0					
25-55 S	12	21.7	22.0	9.5	11.1	4.0	-0.4
55-60 S	0						
60-90 S	0						
GLOBAL	171	38.2	38.9	16.1	17.1	4.2	-0.7

7.3.3 CONCLUSIONS

Based on comparisons of the SSM/I retrieved total water vapor and total water vapor derived from radiosonde data, the Hughes algorithm does not meet the SSM/I specifications of $\pm$ 2.0 kg/m^2 over the range 0 - 80 kg/m^2. A global rms difference of 4.7 kg/m^2 was observed with zonal rms differences ranging from 2.6 kg/m^2 in the Arctic to 7.2 kg/m^2 in the tropics. In addition due to the intrinsic limitations of the zonal or sequential algorithm at the boundaries, obvious erroneous discontinuities were introduced in the retrieved water vapor maps.

Initial efforts to improve the algorithm centered on deriving a global linear algorithm. When this formulation also proved inadequate to meet specifications, a non-linear algorithm was constructed using a quadratic term for the 22V channel. This algorithm resulted in a large reduction of the rms retrieval errors on a global and zonal basis and removes the discontinuities at the boundaries of the zonal regions. The global rms differences were reduced to 2.4 kg/m^2. It should be noted that due to the flexibility of the SSM/I software, the non-linear algorithm used to generate results in Tables 7.3 and 7.4 may be readily implemented.

A few cautionary remarks are in order. The use of a nonlinear algorithm will undoubtedly increase the sensitivity of water vapor retrievals to cloud water amount, the

presence of precipitation, and the presence of sea ice. It is very important that the precipitation screen given with the algorithm coefficients be used with this algorithm. It is also important that retrievals not be attempted when ice is in the SSM/I field-of-view.

It is recommended that radiosonde data be collected periodically to provide a quality control on the water vapor product. This could be accomplished by the periodic collection of raobs from the stations used in the validation study and comparing retrievals.

As part of our validation effort for total precipitable water, we investigated both linear and non-linear algorithms. Our results indicate that a non-linear algorithm is required for best agreement between observed and derived values.

7.4   CLOUD LIQUID WATER

7.4.1 Surface Data Sources

The sources of data for validation of the cloud liquid water content were upward looking microwave radiometers. Measurements were made by NOAA-WPL personnel and University of Massachusetts (UMass) personnel. The NOAA measurements were made at San Nicolas Island as part of Project FIRE and at the four sites of the Colorado remote profiler network. The four sites are Denver (Stapleton Airport), Fleming, Flagler, and Platteville. The radiometers in the Colorado network are fixed zenith viewing radiometers that operate at 20.6 and 31.65 GHz. The Stapleton airport installation also has four frequencies in the oxygen complex for temperature profile retrievals. All of the stations have Doppler radars for wind speed and direction measurements. These stations are described in [10]. The San Nicolas measurements were made by a portable radiometer that has a steerable beam and is described in [11]. Data from the Colorado network stations which operate in an automated continuous mode were obtained for the periods July 15-October 15, 1987 and January 15-April 15, 1988 to provide for a range of seasons and surface conditions. The San Nicolas Island data were taken between July 2 and July 19, 1987. The accuracy of the NOAA profiler network determinations of cloud liquid water are estimated [12] to be 5.2 E-3 kg/m^2.

The UMass measurements were made using an autocorrelation radiometer operating between 20.5 and 23.5 GHz and an auxiliary radiometer at 37 GHz. These measurements were made at Kwajalein Island. The operation of the autocorrelation radiometer is described in [13]. The Kwajalein data were taken between March 24 and April 7, 1988.

To compensate for the different fields-of-view of the surface based radiometers and the SSM/I, the NOAA data were averaged over a two hour period, one hour on either side of the overpass time. The UMass data were averaged over a one hour time period, one half hour on either side of the overpass time.

Characteristics of the NOAA profiler radiometers are given in Table 7.5. All latitudes are North and all longitudes are West. The effective fields-of-view (EFOV) are the 3 dB beamwidths and the spot sizes are in meters for a distance of 1 km.

TABLE 7.5					
CHARACTERISTICS OF NOAA PROFILING NETWORK					
NOAA PROFILER	NETWORK		EFOV		SAMPLING TIME
SITE	LAT	LONG	DEGREE	SPOT	
DENVER	39.8	105	2.5	44 M	2 MIN
FLEMING	40.6	103	4	70 M	2 MIN
FLAGLER	39.1	103	4	70 M	2 MIN
PLATTEVILLE	40.2	105	4	70 M	2 MIN
SAN NICOLAS	33	119	2.5	44 M	1 MIN

Characteristics of the UMass autocorrelation radiometer (CORRAD) are given in Table 7.6. The precision estimate for the CORRAD is given in [14].

TABLE 7.6	
CHARACTERISTICS OF CORRAD	
RF BANDPASS	20.5-23.5 GHz
TIME DELAYS	-0.2 TO 6.1 NS (0.1 NS STEPS)
FREQUENCY RESOLUTION	160 MHz/100 MHz (3 dB)
RECEIVER NOISE TEMPERATURE	2000 K
NOISE FLOOR (ΔT)	$0.5 \ K/(SEC)^{1/2}$
EFOV	2 DEGREES
FOOTPRINT @ 1 KM	35 M
SAMPLING TIME	10 MIN
KWAJALEIN ISLAND	(8.7 N, 167.7 E)
PRECISION	$8.0E-3 \ kg/m^2$

7.4.2 Comparisons

7.4.2.1 Initial Hughes Algorithm

The initial Hughes algorithm [1] used to retrieve cloud liquid water was a linear, four-channel algorithm that was generated by regression using brightness temperatures calculated from simulated clouds and a radiative transfer model. The algorithm was divided into latitudinal and seasonal segments called climate codes. There were eleven climate codes per hemisphere. Over the ocean there were nine distinct sets of coefficients that used the 19H, 22V, 37V, and 37H channels. Over land there was one set of coefficients per climate code which used the 19V, 19H, 37V, and 85V channels. Another set of coefficients was used to retrieve cloud water over snow. This set of coefficients was used for all climate codes and utilized the 22V, 37H, 85V, and 85H channels.

The latitude zones were the same as those used in the water vapor algorithm. The opposite hemisphere is seasonally adjusted so that seasonal algorithms are used in the appropriate season and latitude zone.

Two special categories of retrievals were created; out-of-limits and indeterminate. All geophysical retrievals were tested to determine whether they were within a physically possible range of values. If they were outside the physically possible range, they were assigned an out-of-limits value, usually 1 less than the maximum number of counts allocated for that parameter. The indeterminate classification implies that certain logical conditions are not being met or that the pixel under consideration may be part ocean and part land (i.e., coastal). The indeterminate category was assigned the maximum count value.

We found that more than 90% of all retrieved values of cloud liquid water were either out-of-limits or indeterminate values. This percentage was found at all test sites and before and after the SSM/I's shutdown during December and January 1987-8. Because of this finding, we decided to improve the algorithm.

7.4.2.2 Improved Algorithm

The approach taken to improve the cloud liquid water algorithm was similar to that employed for the water vapor algorithm development. Surface values and brightness temperatures were matched and standard linear regression techniques were used to find the best set of channels and coefficients. Our retrieval equation is linear in brightness temperature and of the form

$$CLW = a_0 + \Sigma\ a_i * T_{Bi}, \ i = 1,2,...7 \tag{3}$$

where the a_i's are coefficients and the T_{Bi}'s are brightness temperatures. Table 7.7 gives the explicit relationships between channel frequency and polarization and coefficient number.

TABLE 7.7 ALGORITHM CHANNEL AND COEFFICIENT DESIGNATION			
CHANNEL NO	COEFFICIENT	SYMBOL	CHANNEL
	a_0		
1	a_1	T_{B1}	19V
2	a_2	T_{B2}	19H
3	a_3	T_{B3}	22V
4	a_4	T_{B4}	37V
5	a_5	T_{B5}	37H
6	a_6	T_{B6}	85V
7	a_7	T_{B7}	85H

The land and ocean cases were separated and the land cases were further divided into snow and no snow groups. Initial correlations on the entire data set, which consisted of clear and cloudy cases yielded very low correlation coefficients. Next cases, where the CLW content was $< 5.0E-3$ kg/m^2 were excluded from the data set. The channels which gave the best correlation are 19V, 19H, 37V, and 85H. The 85V channel was excluded from the regressions because of its increased noise. The snow data set was analyzed separately.

Using the discriminants 19H - 85H $> 8K$ and maps of weekly snow cover to establish the presence of snow, a set of observations was analyzed for cloud liquid water content. All cases where the CLW content was $< 5.0E-3$ kg/m^2 were excluded. The best results of our attempts to find an improved cloud liquid water algorithm over land are shown in Table 7.8. It is readily apparent from the low values of the correlation coefficients in Table 7.8 that the development of a CLW algorithm for land and snow surfaces will be very difficult at best. It is worth noting the correlation coefficients improve if the presence of clouds can be inferred from other sources.

The cloud liquid water determinations over the ocean yielded better results. Initially it was intended to analyze the San Nicholas Island and Kwajalein Island data separately and then as a combined data set. When the sizes of the two data sets were considered (10 samples per island), the decision was made to analyze them as a combined data set. In addition to the standard linear regression procedures, we performed additional independent statistical analysis as well. In analyzing the data all possible four channel combination were considered as well as a full six channel algorithm. The six channel algorithm gives a slightly higher explained variance or correlation coefficient than any four channel algorithm, however the standard error of the estimate is greater because of the reduced number of degrees of freedom. Table 7.9 gives the correlation coefficient (R) and standard error of the estimate (S.E.E.) for some channel combinations. Figure 7.6 is a plot of satellite versus surface values for the combined oceanic data set using the best four channel algorithm. After completing our analysis of the best four channel algorithm, the 85H channel became quite noisy. The other four channel algorithm shown in Table 7.9 is the best algorithm excluding 85H.

TABLE 7.8

CLOUD LIQUID WATER RESULTS

NO. OBS.	CORR. COEFF.	MEAN(KG/M^2)	RMS DIFF (KG/M^2)
		LAND	
232	0.214	0.005	0.021
		CLW > 0.005	
32	0.445	0.037	0.045
		SNOW	
201	0.185	0.007	0.031
		CLW > 0.005	
41	0.369	0.032	0.063

TABLE 7.9

CLW ALGORITHM RESULTS

CHANNEL	NO. SAMPLES	MEAN	S.E.E.	R
1, 2, 3, 4, 5, 7	20	0.136	0.042	0.892
2, 3, 4, 7	20	0.136	0.039	0.891
2, 3, 4, 5	20	0.136	0.040	0.886
4	20	0.136	0.039	0.871

As part of our statistical analysis, it was noted that the 37V channel alone is a good predictor of cloud liquid water. The statistics for the 37V channel are also included in Table 7.9.

Because of the limited size of the ocean data set, further statistical analysis was performed. We used the cross-validation and jackknifing techniques to examine our results.
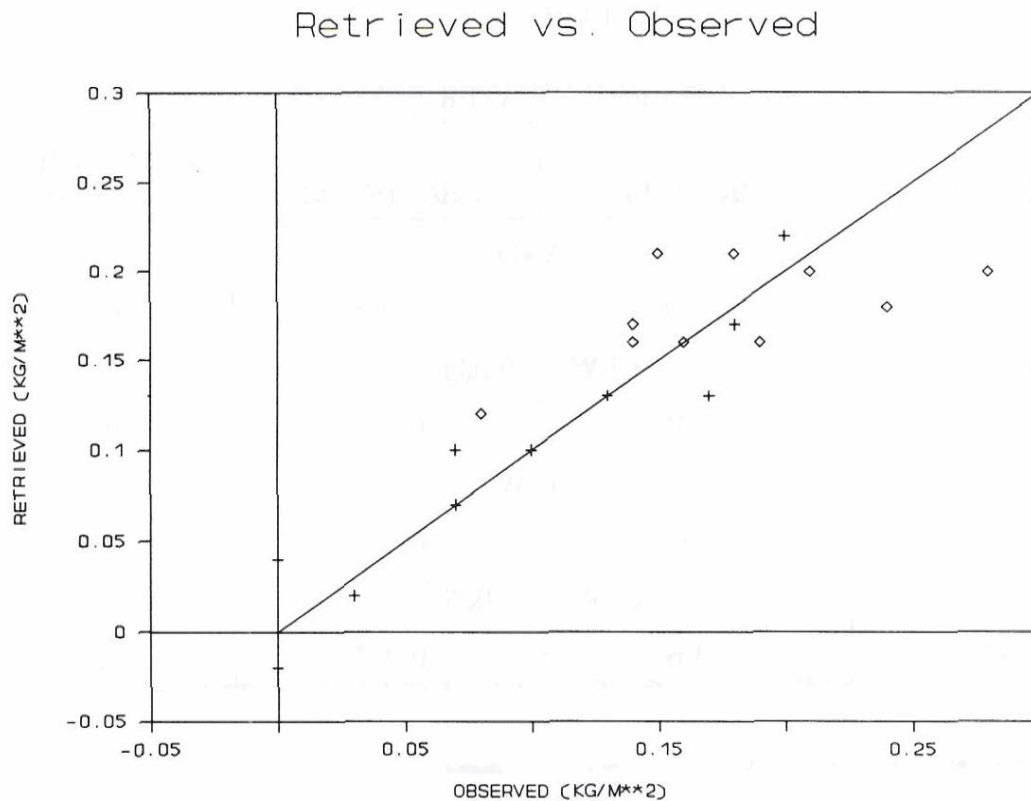
Figure 7.6 - A plot of retrieved vs observed cloud liquid water for the combined San Nicolas and Kwajalein data sets. The units are kg/m². The San Nicolas points are shown as pluses and the Kwajalein points are shown as diamonds. The solid line is the "perfect agreement" line. The retrieved values are from the dependent data set.

These procedures are discussed in [15]. For cross-validation we generated a quasi-independent data set by using 19 of the 20 points as dependent data and predicting the 20th. This was repeated until all 20 points had been predicted independently. We used the same four channels that gave the lowest rms difference for the completely dependent data set.

Using a procedure known as jackknifing [15], we generated another independent estimate of the retrieved mean, standard deviation, and rms difference between the retrievals and the ground based cloud liquid water measurements. The relationship [15] (PARMSTAR) = 20*(-PARMALL) - 19*(PARM)$_i$, where PARMALL is the parameter from the completely dependent data set and PARM is the parameter when it is calculated from a data set of 19 points, was used. The jackknifed values presented in Table 7.10 are averages of 20 such computations of

PARMSTAR for each parameter. Table 7.10 presents a summary of the observed and three retrieved data sets.

	OBS	REGR	CROSS-VAL	JACKKNIFE
TABLE 7.10 STATISTICAL COMPARISONS				
NO OF CASES	20	20	20	20
MINIMUM	0.00	-0.02	-0.04	-
MAXIMUM	0.28	0.22	0.23	-
MEAN	0.136	0.136	0.140	0.136
STD DEV	0.077	0.067	0.068	0.077
RMS DIFF	--	0.035	0.048	0.040

The three retrieved data sets are quite consistent. All have a negligible bias about the mean when compared with the observed data set. The standard deviations of the retrieved values are slightly smaller than the standard deviation of the observed value, a fact not uncommon to regression algorithms. The rms differences are remarkably similar for the three computations. Even though the sample size is small, there are four predictors and 15 degrees of freedom. The improved algorithm seems to be statistically significant and "robust".

7.4.3 Conclusions

In view of the very low correlations between brightness temperatures and cloud liquid water content over land and snow, it is recommended that retrievals of this parameter not be attempted. If an independent way of determining the presence of clouds can be found, it might be possible to devise an algorithm that will give a useful estimate of the cloud liquid water content.

It should be noted that Colorado is not an ideal site for testing cloud liquid water content algorithms because of its altitude and generally dry conditions. The NOAA profiler network is almost the only source of routine measurements of cloud liquid water. It is likely that the SSM/I can detect heavier water clouds over land before the onset of actual precipitation.

It is recommended that a quality control procedure be instituted for the cloud liquid water product as well. Collecting significant amounts of cloud liquid water measurements from surface based systems is a major undertaking. The recommended procedure is to use either OLS data which can be co-located with the SSM/I or GOES visible and infrared data and compare qualitatively where the SSM/I algorithm places clouds and their water content versus the visible and

IR images which should show cloud location over the ocean rather well. It should be obvious if the SSM/I "misses" clouds or places them in clear areas.

The initial recommendation was to use the six channel algorithm because it explained the most variance. Further examination revealed that a four channel algorithm probably gave better results, especially when considering the standard error of the estimate. Our additional analysis also revealed that the 37V channel alone is a very good predictor of cloud liquid water. After most of the analysis for the Cal/Val effort was completed, the noise of the 85H channel increased significantly. As a result, we developed a CLW algorithm that does not use either 85 GHz channel.

Table 7.11 gives the actual coefficients for the channels used in the recommended algorithms. We show the latest total precipitable water algorithm which is non-linear and global. We show algorithms for cloud liquid water. Included are coefficients for a six channel algorithm, coefficients for two four channel algorithms with and without 85H and a single channel algorithm using only 37V. The retrieved parameters will have the units of kg/m^2.

The results presented here have also been presented in Alishouse et al [16] and Alishouse et al [17].

			TABLE 7.11		
		RECOMMENDED ALGORITHM COEFFICIENTS			
CHANNEL	WATER VAPOR	CLW OCEAN 6 CHANNELS	CLW OCEAN W 85H	CLW OCEAN W/O 85H	CLW OCEAN 37V
19V	-0.148596	1.5817E-3	-----	-----	-----
19H	----	5.8075E-3	6.0257E-3	8.4333E-3	-----
22V	-1.829125	-5.6345E-3	-4.8803E-3	-7.5959E-3	-----
22V (SQRD)	6.193E-3	-----	-----	------	-----
37V	-0.36954	2.0097E-2	1.9595E-2	2.0131E-2	1.18122E-2
37H	-----	-7.200 E-4	-----	-5.3066E-3	------
85V	-----	-----	-----	-----	------
85H	-----	-2.7658E-3	-3.0107E-3	-----	------
INTERCEPT	232.89393	-3.31378	-3.14559	-2.838179	-2.45276

Precipitation Screen:

$$\text{If } -11.7939 - 0.02727 \ast T_{B37V} + 0.09920 \ast T_{B37H} \leq 0 \text{ K}$$

then compute water vapor and cloud liquid water over ocean.

7.5 REFERENCES

[1]   J. Hollinger, R. Lo, G. Poe, R. Savage, and J. Peirce, <u>Special Sensor Microwave/Imager Users Guide</u>, Naval Research Laboratory, Washington, DC, 120 pp., 1987.

[2]   R. R. Ferraro and S. A. Snyder, "SSM/I Validation of Total Precipitable and Cloud Liquid Water," Final Report for Contract No. 50-DGNE-6-00083, 105 pp. and appendixes, S. M. Systems and Research Corp., Landover, MD, 1989.

[3]   J. C. Alishouse, "Total precipitable water and rainfall determinations from the SEASAT scanning multichannel microwave radiometer," <u>J. Geophys. Res.</u>, vol. 88, pp. 1929-1935, 1983.

[4]   H. D. Chang, P. H. Hwang, T. T. Wilheit, A. T. C. Chang, D. H. Staelin, and P. W. Rosenkranz, "Monthly distributions of precipitable water from the NIMBUS 7 SMMR data," <u>J. Geophys. Res.</u>, vol. 89, pp. 5328-5334, 1984.

[5]   V. Barnett and T. Lewis, <u>Outliers in Statistical Data</u>, New York: John Wiley & Sons, 365 pp., 1978.

[6]   D. H. Staelin, K. F. Kunzi, R. L. Pettyjohn, R. K. L. Poon, R. W. Wilcox, and J. W. Waters, "Remote sensing of atmospheric water vapor and liquid water with the Nimbus 5 microwave spectrometer," <u>J. Appl. Meteor.</u>, vol. 15, pp. 1204-1214, 1976.

[7]   N. C. Grody, A. Gruber, and W. C. Shen, "Atmospheric water content over the tropical Pacific derived from the Nimbus-6 scanning microwave spectrometer," <u>J. Appl. Meteor.</u>, vol. 19, pp. 986-996, 1980.

[8]   T. T. Wilheit and A. T. C. Chang, "An algorithm for retrieval of ocean surface and atmospheric parameters from the observations of the scanning multichannel microwave radiometer (SMMR)," <u>Radio Sci.</u>, vol. 15, 525-544, 1980.

[9]   F. J. Wentz, "A model function for ocean microwave brightness temperatures," <u>J. Geophys. Res.</u>, vol. 88, 1892-1908, 1983.

[10]   D. C. Hogg, M. T. Decker, F. O. Guiraud, K. B. Earnshaw, D. A. Merritt, K. P. Moran, W. B. Sweezy, R. G. Strauch, E. R. Westwater and C. G. Little, "An automatic profiler of the temperature, wind and humidity in the troposphere," <u>J. Clim. Appl. Meteor.</u>, vol. 22, pp. 807-831, 1983.

[11]   D. C. Hogg, F. O. Guiraud, J. B. Snider, M. T. Decker and E. R. Westwater, "A steerable dual-channel microwave radiometer for measurement of water vapor and liquid in the troposphere," <u>J. Clim. Appl. Meteor.</u>, vol. 22, pp. 789-806, 1983.

[12] P. Ciotti, E. R. Westwater, M. T. Decker, A. J. Bedard, and B. B. Stankow, "Ground based microwave radiometric observations of the temporal variations of atmospheric geopotential height and thickness," IEEE Trans. Geosci. Remote Sensing, vol. GE-25, pp. 600-615, 1987.

[13] C. S. Ruf and C. T. Swift, "Atmospheric profiling of water vapor density with a 20.5-23.5 GHz autocorrelation radiometer," J. Atmos. Ocean. Tech., vol. 5, pp. 539-546, 1988.

[14] C. S. Ruf, unpublished manuscript, 4 pp., Jet Propulsion Laboratory, Pasadena, CA, 1989.

[15] F. Mosteller and J. W. Tukey, Data Analysis and Regression, Reading: Addison-Wesley, 588 pp., 1977.

[16] J. Alishouse, S. Snyder, J. Vongsathorn, and R. Ferraro, "Determinations of Oceanic Total Precipitable Water from the SSM/I," IEEE Trans. Geosci. Remote Sensing, vol. 28, no. 5, pp. 811-816, 1990.

[17] J. Alishouse, J. Snider, E. Westwater, C. Swift, C. Ruf, S. Snyder, J. Vongsathorn, and R. Ferraro, "Determination of Cloud Liquid Water Content Using the SSM/I," IEEE Trans. Geosci. Remote Sensing, vol. 28, no. 5, pp. 817-822, 1990.

SECTION 8

WIND SPEED VALIDATION

Led by

John C. Wilkerson
NOAA/NESDIS
Washington, DC

Supported by

Calvin T. Swift, Mark Goodberlet, Kuo-Huya Hsueh,
John Morris, Christopher Ruf, and David Choi
University of Massachusetts
Amherst, MA

and by

Frank Wentz
Remote Sensing Systems
Santa Rosa, CA

8.0 WIND SPEED VALIDATION

The SSM/I wind speed retrieval algorithm developed by Environmental Research and Technology, Inc. (ERT) for Hughes Aircraft is called the D-matrix algorithm and has the following form [1]:

$$SW = C_{0j} + C_{1j} \cdot T_{B19H} + C_{2j} \cdot T_{B22V} + C_{3j} \cdot T_{B37V} + C_{4j} \cdot T_{B37H}. \tag{8.1}$$

Equation (8.1) is valid only over open ocean where the wind speed, SW, is in m/s and referenced to a height of 19.5 m above the surface. Equation (8.1) also contains the terms T_{Bx} which represent the brightness temperature of frequency/polarization combination "x" and the D-matrix coefficients, C_{ij}, where "j" is the climate code index and varies from 1 to 11. The eleven sets of coefficients (only 9 of which are distinct) used in the original D-matrix algorithm are listed in Table 8.1. Each of the 9 distinct climate codes represents a particular season and latitude band as shown in Table 8.2.

Since microwave radiation at the SSM/I frequencies is heavily attenuated by rain in the earth's atmosphere which masks the wind speed signature generated by waves and foam on the ocean surface, ERT suggested the use of a rain-flag for the purpose of identifying conditions

TABLE 8.1 COEFFICIENTS OF THE ORIGINAL HUGHES D-MATRIX ALGORITHM					
Climate Code	C_0	C_1	C_2	C_3	C_4
1	191.560	.4903	-.4432	-.9199	.3577
2	168.390	.5366	-.4548	-.7656	.2635
3	177.315	.3913	-.2818	-1.0083	.4095
4	147.760	.5077	-.3547	-.7409	.2333
5	127.130	.4788	-.2546	-.7162	.2030
6	163.070	.2923	-.1204	-1.0967	.4612
7	95.994	.6106	-.3034	-.4638	.0192
8	130.420	.3676	-.1508	-.8400	.3056
9	117.590	.4225	-.1899	-.7096	.2081
10	130.420	.3676	-.1580	-.8400	.3056
11	117.590	.4225	-.1899	-.7096	.2081

TABLE 8.2

CLIMATE CODES OF THE HUGHES D-MATRIX ALGORITHM

ZONE	SEASON (NORTHERN HEMISPHERE)	CLIMATE CODE
Tropics (0-20 LAT.)	JUN-NOV	1
	DEC-MAY	2
Low-Lat. Transition (20-25 LAT.)	JUN-NOV	3
	DEC-MAY	4
Mid-Latitude (25-55 LAT.)	SEP-NOV, MAR-MAY	5
	JUN-AUG	6
	DEC-FEB	7
Arctic (55-90 LAT.)	MAY-OCT	8
	NOV-APR	9

under which less accurate wind speed retrievals are produced. The original rain-flag logic is shown below.

IF: $T_{B19H} > 190K$
OR: $[T_{B37V} - T_{B37H}] < 25K$
Then possible rain exists and rain-flag = 1

IF: $[T_{B37V} - T_{B37H}] < 10K$
Then heavy rain exists and rain-flag = 2
Otherwise rain-flag = 0

The accuracy specification for wind speed retrievals under conditions of no rainfall (i.e., rain-flag = 0) was $\pm$ 2 m/s over the range 3 to 25 m/s. Accuracy was not specified for wind

retrievals from cells flagged either 1 or 2. In fact, the original D-matrix algorithm did not attempt to retrieve winds under rain-flag 2 conditions.

8.1 NOAA BUOY SYSTEM AND CRITERIA FOR COMPARISON

Validation of the SSM/I wind speed retrievals was done using the anemometer measured winds of open ocean buoys maintained by the National Oceanic and Atmospheric Administration (NOAA). These buoys record an 8.5 minute average of the wind once every hour with an accuracy of ± 0.5 m/s for winds less than 10 m/s and 5% for winds greater than 10 m/s [2].

In anticipation of SSM/I antenna sidelobes, which could give rise to land contamination of ocean brightness temperatures, only buoys further than 100 km from land were chosen for the validation. The 19 NOAA buoys actually used for the validation are listed in Table 8.3.

The wind speed observations taken by the ocean buoys were at heights of either 5 or 10 meters above the surface. These measurements were converted to equivalent winds at 19.5 meters above the surface [3] so that they could be compared directly to the SSM/I estimates which predict winds at the 19.5 meter level. Converted buoy winds and D-matrix winds were paired only when the SSM/I retrieval was located within 25 km of the buoy position and the SSM/I overpass time was within 30 minutes of the buoy wind speed measurement. Based on the work of Monaldo [4], a spatial difference of 25 km and a temporal difference of 30 minutes between SSM/I and buoy measured wind speeds adds variances of approximately 0.5 m/s and 0.2 m/s, respectively, to the total variance of the comparison. These variances increase the total standard deviation of 2 m/s by less than 10% and therefore contribute only slightly to the overall error. Because a 25 km spatial separation introduces little additional error to the comparison of SSM/I and buoy winds, the SSM/I geolocation problem (see the instrument calibration section of this report) which results in positioning errors of between 5 and 25 kilometers, does not significantly affect the wind speed validation. This comparison criteria also stipulates that only one SSM/I-buoy pair be selected from each SSM/I overpass. Thus the validation data set was composed of independent comparisons.

8.2 REQUIRED NUMBER OF COMPARISONS

The accuracy specification of ± 2 m/s for D-matrix wind speed retrievals can be interpreted in at least two ways. One interpretation is that this is the standard deviation, in an average sense, of the difference between all coincident buoy and SSM/I wind speed measurements. An alternative interpretation is that the standard deviation of such comparisons in any sub-interval of the 3-25 m/s wind speed range must not exceed 2 m/s. The first of these two interpretations can disguise the fact that over certain sub-intervals of the 3-25 m/s wind speed range, the accuracy of the D-matrix prediction may be worse than ± 2 m/s. In fact, a modeled error budget (discussed in section 8.3) predicts that the accuracy is wind speed dependent. It is possible that sub-intervals with accuracies worse than ± 2 m/s could average with sub-intervals having accuracies better than ± 2 m/s to give a resulting overall accuracy of better than ± 2 m/s. This is often true for regression-type algorithms, like the D-matrix, which

		TABLE 8.3		
NOAA BUOYS USED FOR THE SSM/I WIND SPEED VALIDATION				
BUOY I.D.	LATITUDE	LONGITUDE (E)	ZONE	NUMBER OF SSM/I PASSES IN 30 DAYS
51002	17.2	202.2	Tropics	31
51004	17.5	207.4	Tropics	31
51003	19.2	199.2	Tropics	32
51001	23.4	197.7	Low Lat Trans	33
42001	25.9	270.3	Mid Lat	33
42002	26.0	266.5	Mid Lat	33
42003	26.0	274.1	Mid Lat	33
41006	29.3	282.6	Mid Lat	34
41002	32.2	284.7	Mid Lat	35
44004	38.5	289.4	Mid Lat	38
46006	40.8	222.4	Mid Lat	39
44011	41.1	293.4	Mid Lat	40
46002	42.5	229.6	Mid Lat	41
44005	42.7	291.7	Mid Lat	41
46005	46.1	229.0	Mid Lat	43
46004	50.8	224.1	Mid Lat	47
46003	51.9	204.1	Mid Lat	48
46001	56.3	211.7	Arctic	55
46035	57.0	182.3	Arctic	56

tend to make especially good predictions near the overall average wind speed and predictions of less accuracy for wind speeds which are removed from the average wind speed. For this reason, the 3-25 m/s wind speed range of interest was divided into the 6 sub-intervals shown in Table 8.4 and the D-matrix performance was analyzed in each sub-interval. Also shown in Table 8.4 is the number of comparisons out of 1,000 for which the buoy wind speed falls within the particular sub-interval range. These comparison counts are based on the global distribution of winds given by Schroeder [5] which is shown in Figure 8.1.

It is preferable to have a sample size of 30 or more when doing statistical analysis [6] of the data. For wind speed sub-intervals 1, 2, and 3, it appears that this sample size can be obtained by collecting approximately 140 comparisons. Preliminary studies showed that about 15% of the data are rain-flagged and since the comparisons are made only with data which is

GLOBAL WINDS DISTRIBUTION
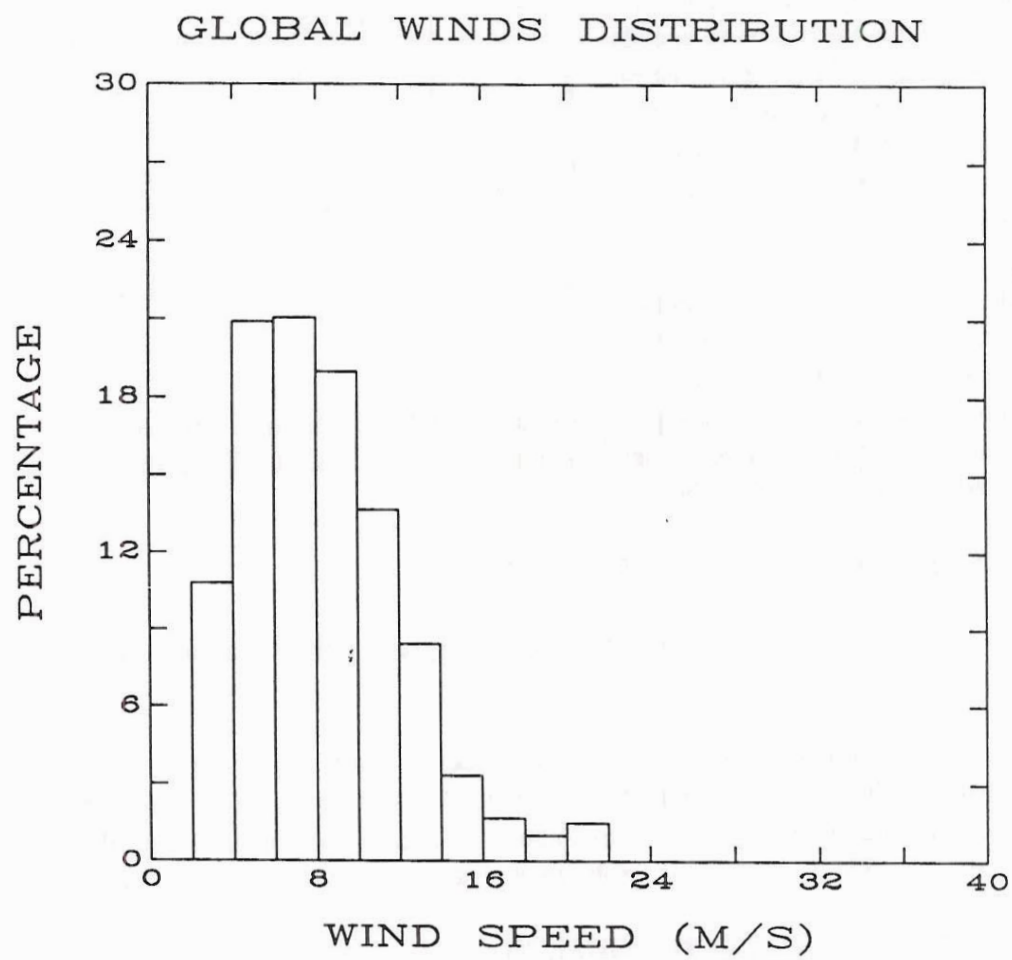
Figure 8.1 - The global distribution of ocean surface winds.

not rain-flagged, the sample size required for each climate code needed to be increased 15% from 140 to 161. Although this relatively small number of comparisons could be obtained in 60 days using three buoys, other factors affect the total required buoy count. These include lost data due to periodic buoy maintenance and the likelihood of encountering wind speeds distributed according to Figure 8.1. These factors were determined from actual climatic summaries [7] prepared by the National Climatic Data Center for the individual data buoys.

		TABLE 8.4	
	D-MATRIX/BUOY WIND SPEED COMPARISONS		
I.D.	Range (m/s)	Comparisons per 1000	
1	3-6	260	
2	6-10	395	
3	10-14	215	
4	14-18	50	
5	18-22	25	
6	22-25	1	

Finally, to complete the validation within one year we must have enough buoys within each of the D-matrix latitude bands and enough SSM/I overpasses per buoy to collect the required 161 comparisons per season. The number of SSM/I overpasses depends on the latitude (LAT) of the buoy and can be approximated by using equation (8.2).

$$\text{SSM/I Overpasses in 30 Days} = 30/\cos(\text{LAT}) \tag{8.2}$$

Equation (8.2) is reasonably accurate up through 60 degrees latitude, above which the error exceeds 15%.

This analysis established that the 19 buoys selected could more than satisfy all but the highest wind speed validation requirements. That is, the low probability of observing winds greater than 15 m/s made it difficult to evaluate the overall performance of the D-matrix algorithm in the range 15-25 m/s. This problem is discussed more fully in section 8.6.

8.3   PRE-LAUNCH VALIDATION MODELING -- ERROR BUDGET

The sources of random errors associated with the comparison of SSM/I wind retrievals and ocean buoy measurements are summarized in the following error budget.

*   Extrapolation noise. (Buoy average at a point differs from the instantaneous spatial average made by the SSM/I).

*   SSM/I instrument noise.

*   Buoy instrument noise.

8-6

* D-matrix algorithm model noise. (Inability of algorithm to model exactly the radiative transfer processes).

* Decorrelation noise. (Spatial and temporal separation of the SSM/I and buoy measurements).

* Translation noise. (Errors in translating the buoy wind measurement to a height of 19.5 m).

* Round-off noise. (Error due to rounding SSM/I winds to the nearest m/s)

The magnitude of these errors (less decorrelation noise and translation noise) is shown in Figure 8.2 over the wind speed range of 3 to 25 m/s for the Climate Code 5 algorithm. Plots for the other 8 versions of the D-matrix algorithm are very similar to the results of Climate Code 5 and are therefore not shown. In generating the extrapolation noise curve of Figure 8.2, the one-dimensional wind field model of Pierson [8] was used as were effective footprint diameters of 55, 49, and 32 km for the 19, 22, and 37 GHz SSM/I channels, respectively. The buoy noise, which was discussed previously, is from Gilhousen [2]. The model noise was specified by Hughes Aircraft in a report by Lo [9].

The instrument noise as specified by Hughes [1] for the 19H, 22V, 37V, and 37H is 0.41, 0.75, 0.38, and 0.39 degrees Kelvin, respectively. The round-off noise is due to the fact that the operational D-matrix algorithm retrievals are rounded off to the nearest whole m/s before being recorded. Although the round-off noise does not contribute significantly to the total error of D-matrix retrievals, subsequent users of the data will introduce an error due to rounding when converting from m/s to either miles/hour or knots (in the case of knots, an average error of 0.7 knots and a maximum error of 1.5 knots will result). The average errors due to spatial and temporal separation of SSM/I and buoy measurements are not included in the plot since they do not contribute significantly to the total. Likewise, errors in converting the buoy wind measurements to a height of 19.5 m are insignificant and are not shown in the plot.

8.4   VALIDATION RESULTS

Performance of the climate code 5 version of the original D-matrix algorithm is shown by the scatter plot in Figure 8.3. The legend shown in the lower right hand corner of the scatter plot is interpreted as follows. The bias and slope data indicates the y-axis intercept and slope of the regression line which has been chosen to minimize the sum of the squares of the horizontal distances from each point to the regression line. The SD is the standard deviation of the quantity, (D-matrix winds minus buoy winds). The line labeled "CORR(R)" is the correlation coefficient [10] between buoy winds and D-matrix winds. Finally, the line labeled "#OBS" gives the number of observations or data points in the scatter plot. Figure 8.3 indicates that the Climate Code 5 D-matrix wind speed retrievals are scaled and biased by 0.85 and 5.7 m/s, respectively. This poor performance of the Climate Code 5 algorithm is typical of the other versions of the original D-matrix algorithm.
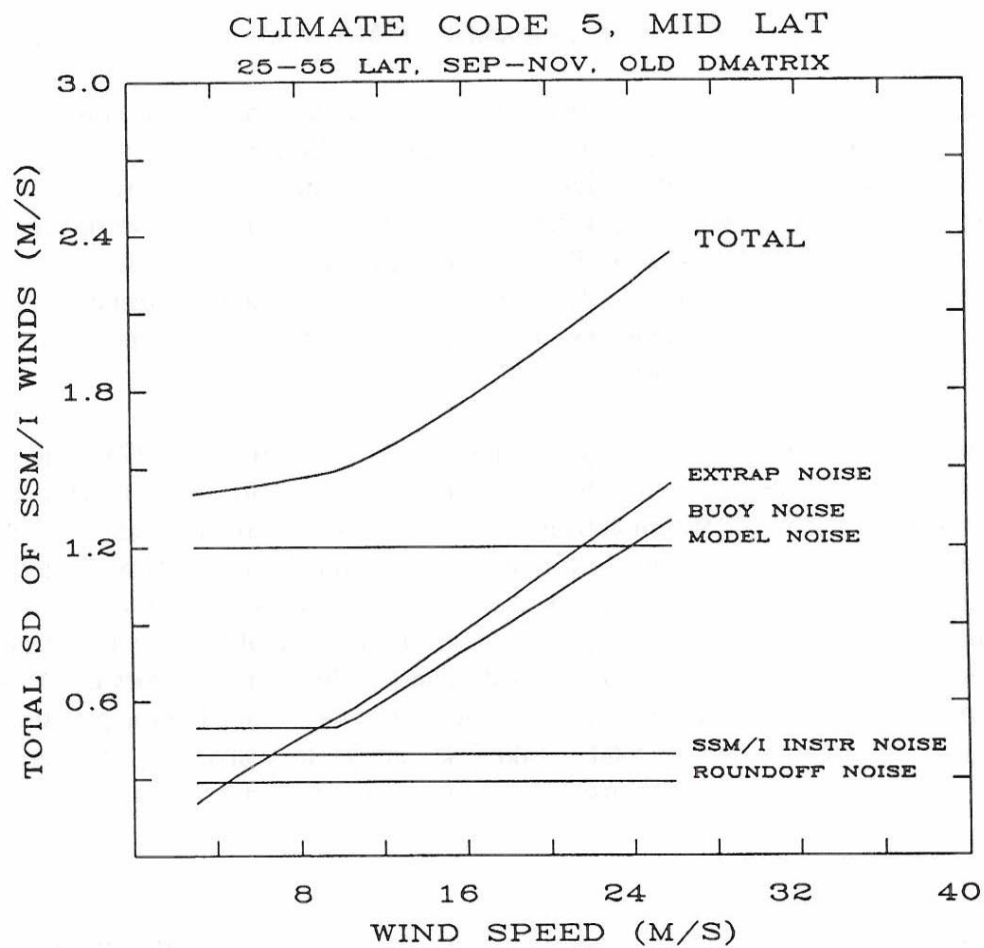
Figure 8.2 - Random errors affecting D-matrix wind speed
retrievals from climate code 5.
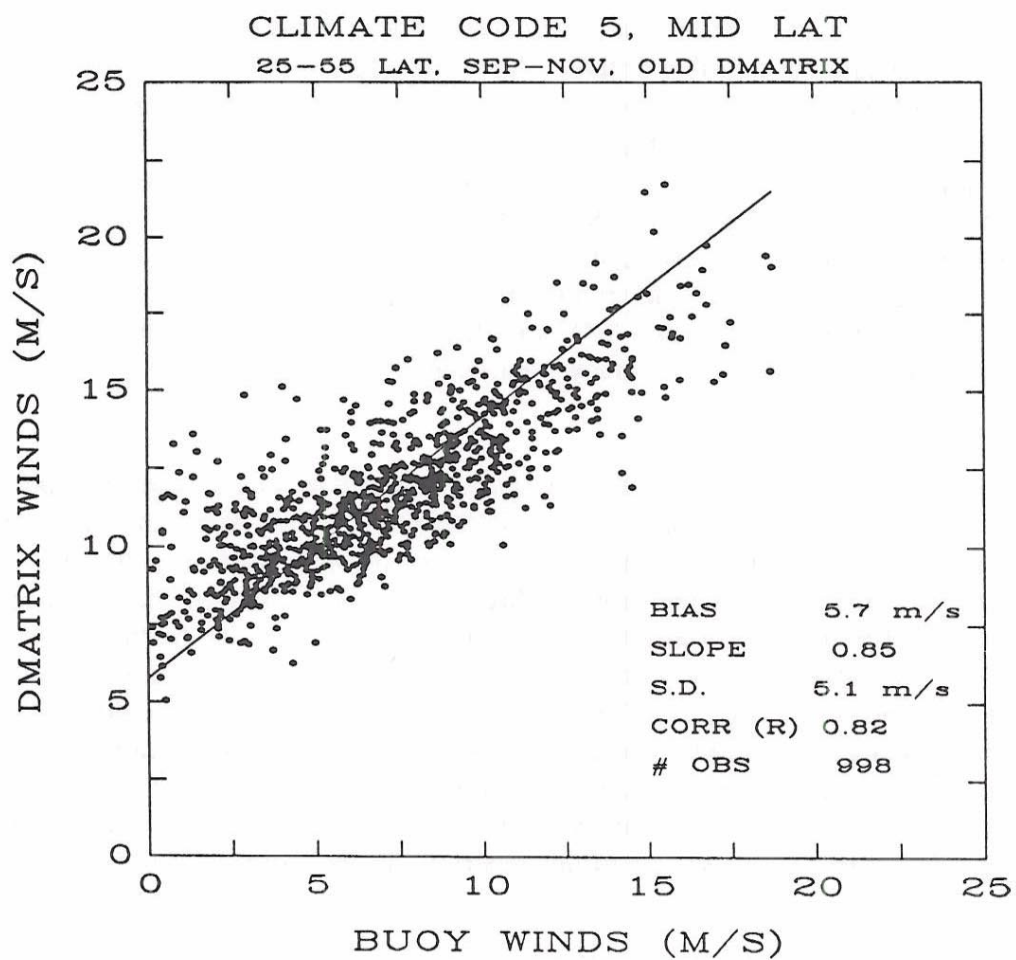
Figure 8.3 - Perfomance of the original D-matrix algorithm for climate code 5.

To correct this problem new coefficients were generated using standard linear regression of buoy wind speed on the coincident SSM/I brightness temperature measurements, T_{B19H}, T_{B22V}, T_{B37V}, and T_{B37H}. Performance of the new Climate Code 5 D-matrix algorithm is shown in Figure 8.4. The regression line associated with this scatter plot now has the desired slope of 1.0 and bias of 0.0 indicating that the scale and bias problems of the original algorithm have been corrected. Despite the apparent good performance of the new algorithms, additional improvements are necessary and will be discussed later in section 8.5.

Before analyzing the retrieval accuracy over various wind speed sub-intervals, it was necessary to re-evaluate the rain-flag criteria. New rain-flag thresholds were determined using residual plots like those shown in Figures 8.5 and 8.6 which indicate the performance of the new D-matrix algorithm as a function of the parameters used to determine rain, which are (T_{B37V} - T_{B37H}) and T_{B19H}. Each of the data sets in the residual plots were then sub-divided into a number of range bins and the standard deviation, SD, and average (also called bias) of the points falling within each bin were calculated. The results of these calculations are shown in Figures 8.7 and 8.8. The rain-flag thresholds were determined from these plots by locating values of the rain-flag parameters for which either the "SD" or "BIAS" curves crossed some predetermined accuracy level. For example, the accuracy requirement for retrievals with rain-flag zero is 2 m/s.

From Figures 8.7 and 8.8, one can see that the algorithm fails to meet this specification when either (T_{B37V}-T_{B37H}) < 50 or T_{B19H} > 150. In this way, entirely new rain-flag criteria were defined. These are summarized in Table 8.5. Note our recommendation to use the four rain-flags 0, 1, 2, and 3, instead of the original three. It is recommended that wind speeds be calculated under all rain-flag conditions and that the associated rain-flag be the user's guide to the accuracy of the retrieval. This practice differs from the operation of the original D-matrix which retrieved winds only under rain-flag 0 and 1 conditions. Finally, it should be pointed out that the term "rain-flag" is somewhat misleading since the rain-flags (except rain-flag 0) indicate any condition (including rain) which leads to reduced retrieval accuracy. The accuracy of the D-matrix retrievals is, in fact, very sensitive to rain since rain rates of less than 1 mm/hr will trip rain-flag 1 [11] (see also the section of this report on the validation of the D-matrix rain-rate algorithm).

Table 8.5 shows the new D-matrix coefficients for all 9 climate codes which were derived using actual SSM/I data from the period 10 July 1987 through 31 March 1988. The measured standard deviation of the difference between buoy winds and D-matrix winds for each of the climate codes under rain-flag 0 conditions is shown in Table 8.6. At least in the average sense, all 9 D-matrix algorithms appear to exceed the accuracy specification of $\pm$ 2 m/s. Also shown in Table 8.6 is the total number of buoy/D-matrix wind comparisons from each climate code and the percentage of these that were tagged with a rain-flag of 1 or higher. Although the results shown in Table 8.6 are quite good, the D-matrix wind speed algorithm has several limitations which are discussed in the following section.
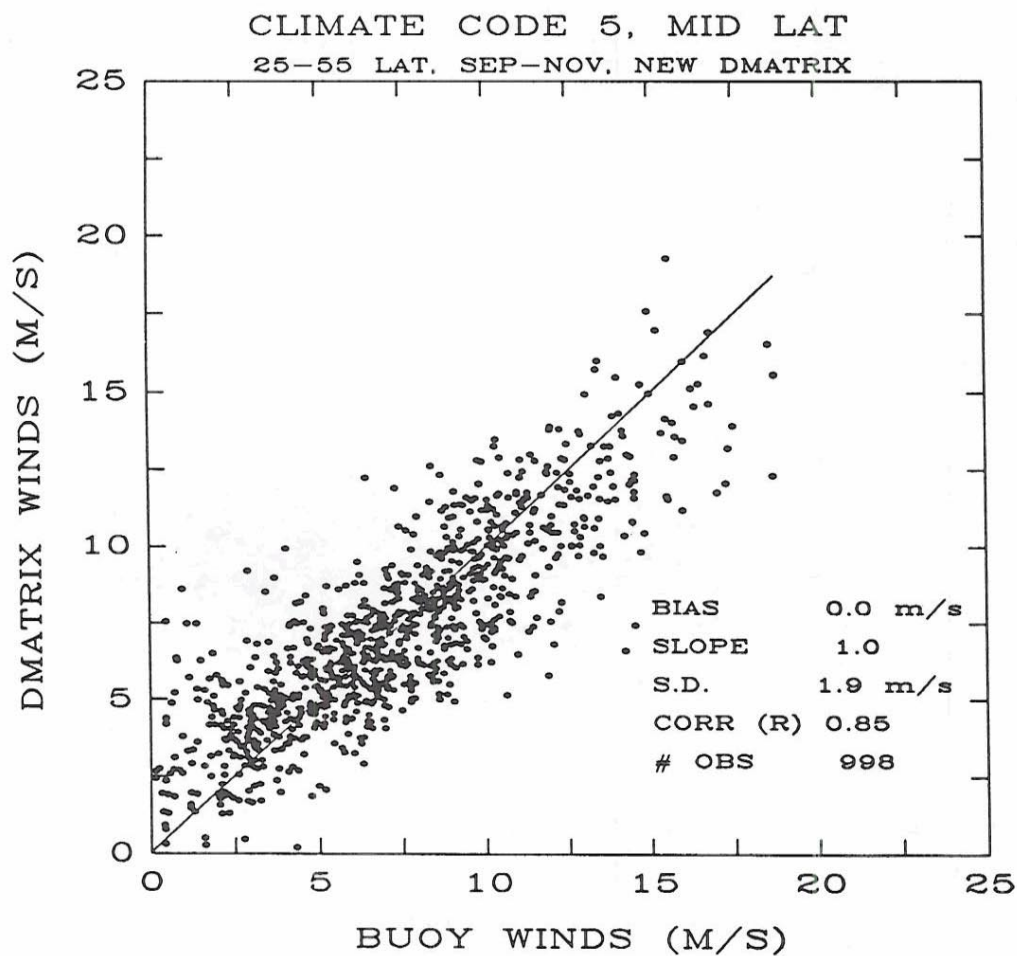
Figure 8.4 - Performance of the revised D-matrix algorithm
for climate code 5.
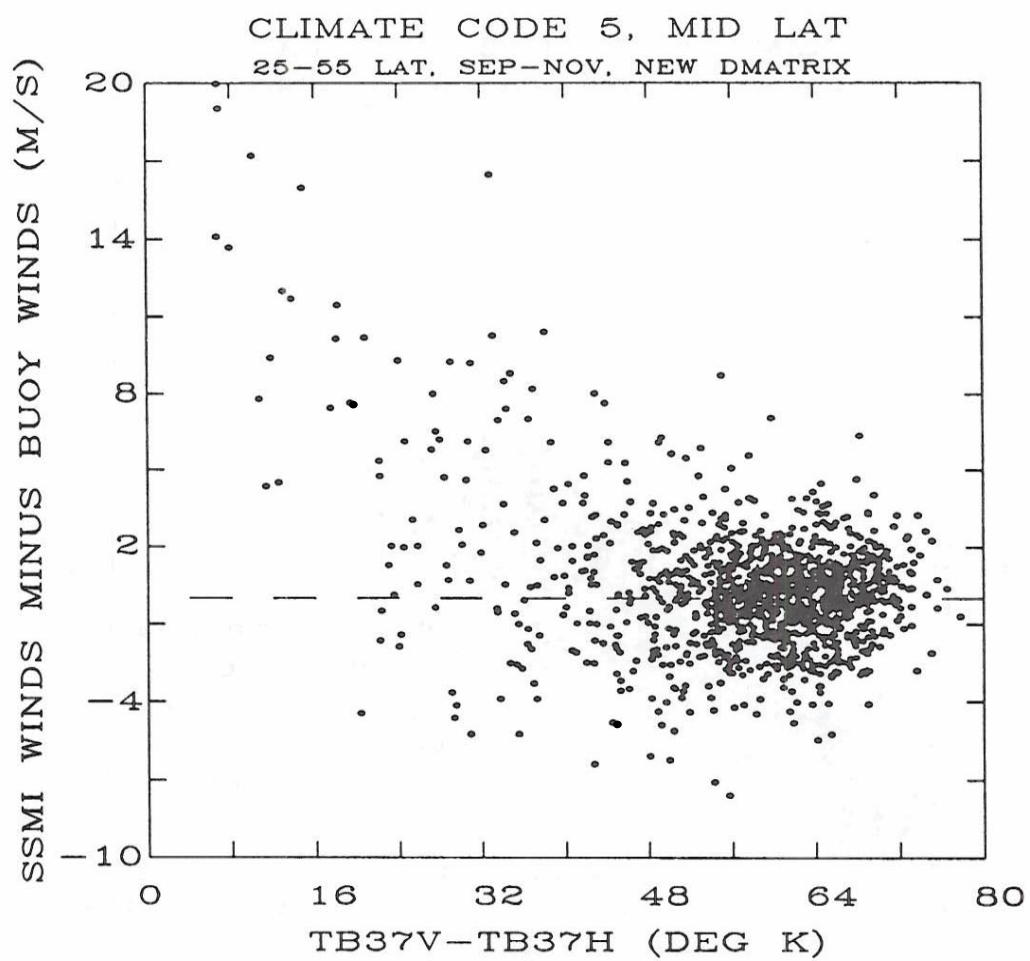
Figure 8.5 - D-matrix residual versus (T_{B37V}-T_{B37H}) for climate code 5.
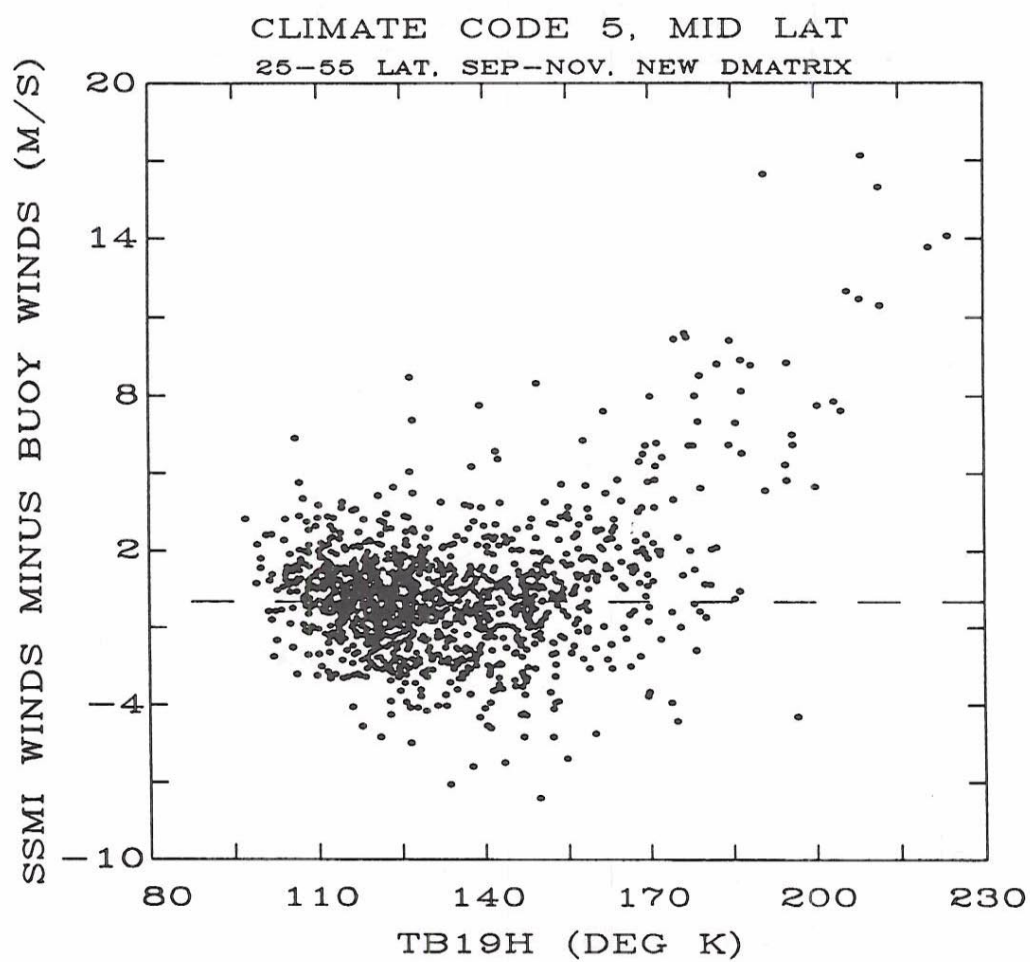
Figure 8.6 - D-matrix residual versus T_{B19H} for climate code 5.

CLIMATE CODE 5, MID LAT
25—55 LAT, SEP—NOV, NEW DMATRIX
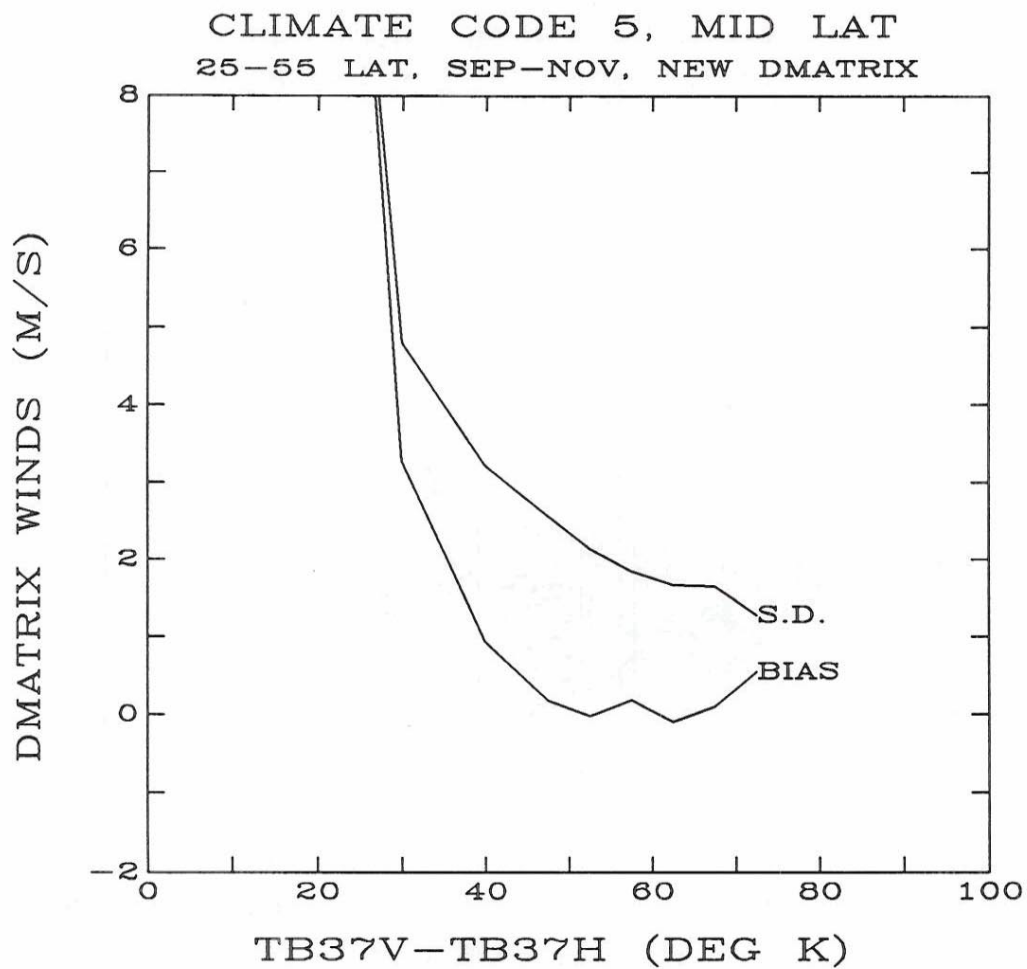
Figure 8.7 - Standard deviation and bias of D-matrix winds as a
function of the rain-flag, $(T_{B37V}-T_{B37H})$, for climate
code 5.
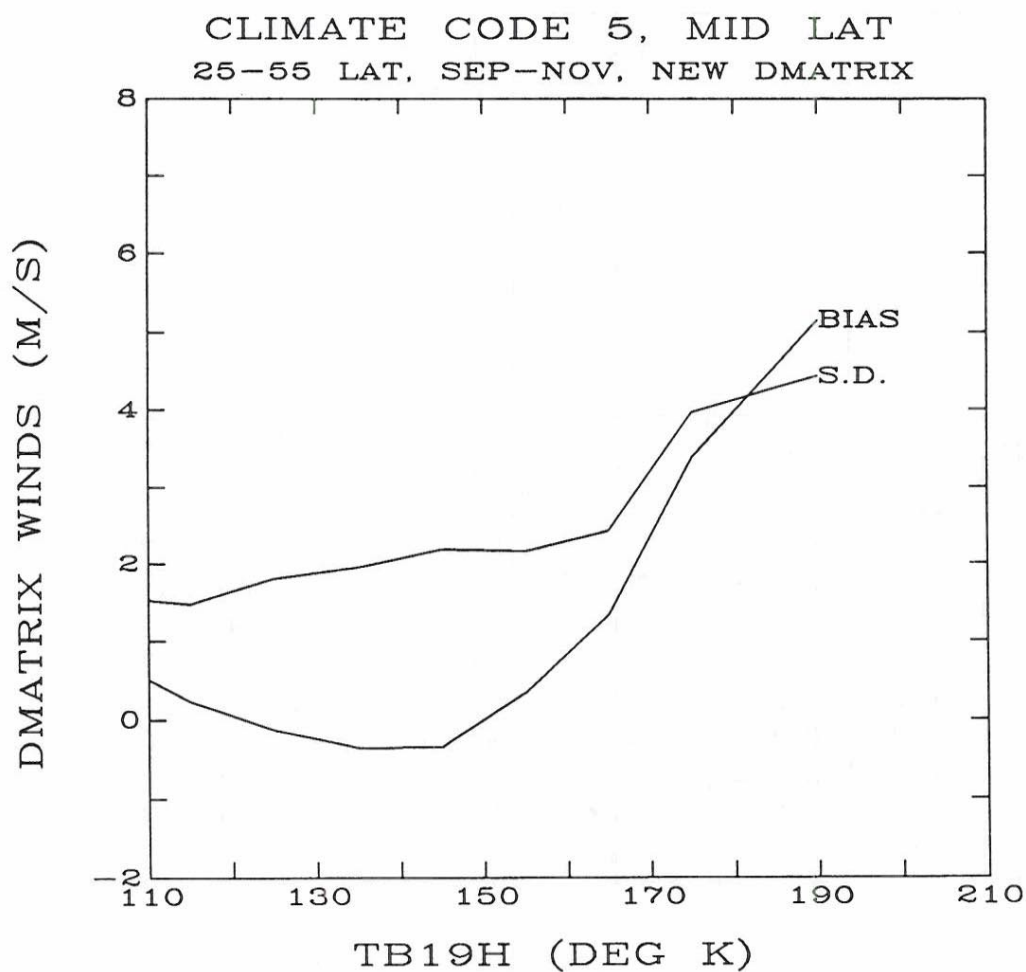
Figure 8.8 - Standard deviation and bias of D-matrix winds as a function of the rain-flag, (T_{B19H}), for climate code 5.

<table>
<tr><td colspan="4" style="text-align:center">TABLE 8.5</td></tr>
</table>

TABLE 8.5

NEW RAIN-FLAG CRITERIA AND COEFFICIENTS

Rain Flag	Criteria	Accuracy
0	$(T_{B37V} - T_{B37H}) > T_0$ and $T_{B19H} < T_1$	< 2 m/s
1	$(T_{B37V} - T_{B37H}) < T_0$ or $T_{B19H} > T_1$	2 - 5 m/s
2	$(T_{B37V} - T_{B37H}) < T_2$	5 - 10 m/s
3	$(T_{B37V} - T_{B37H}) < T_3$	> 10 m/s

Climate Code	C_0	C_1	C_2	C_3	C_4	T_0	T_1	T_2	T_3
1	211.22	0.5090	-0.3703	-1.1944	0.4458	50	175	25	20
2	202.87	0.1316	-0.2455	-1.3138	0.8080	50	175	25	20
3	195.18	0.2996	-0.2363	-1.2266	0.5776	50	175	25	20
4	172.72	0.3908	-0.3130	-1.0396	0.4926	50	175	25	20
5	158.63	0.4224	-0.2439	-0.9839	0.3725	50	165	30	25
6	161.45	0.2964	-0.1613	-1.0632	0.4524	50	165	30	25
7	151.04	0.5994	-0.3274	-0.9137	0.2977	50	165	30	25
8	137.72	0.7330	-0.4208	-1.7533	0.1804	50	130	35	30
9	109.93	0.8695	-0.4710	-0.6008	0.1158	50	130	35	30

8.5   D-MATRIX LIMITATIONS

Wind speed residual plots were again used to study limitations of the D-matrix algorithm. Plotting the residual as a function of buoy measured wind speed demonstrates the D-matrix performance over sub-intervals of the 3-25 m/s range. Figure 8.9 shows the plot for Climate Code 5 which is typical of all 9 climate code versions of the D-matrix. Dividing the region of Figure 8.9 into a number of range bins and calculating the SD and bias (i.e., average) of the points falling within each bin results in the "interpreted" residual plot shown in Figure 8.10. This figure shows that the accuracy of the D-matrix retrievals is best near the global average
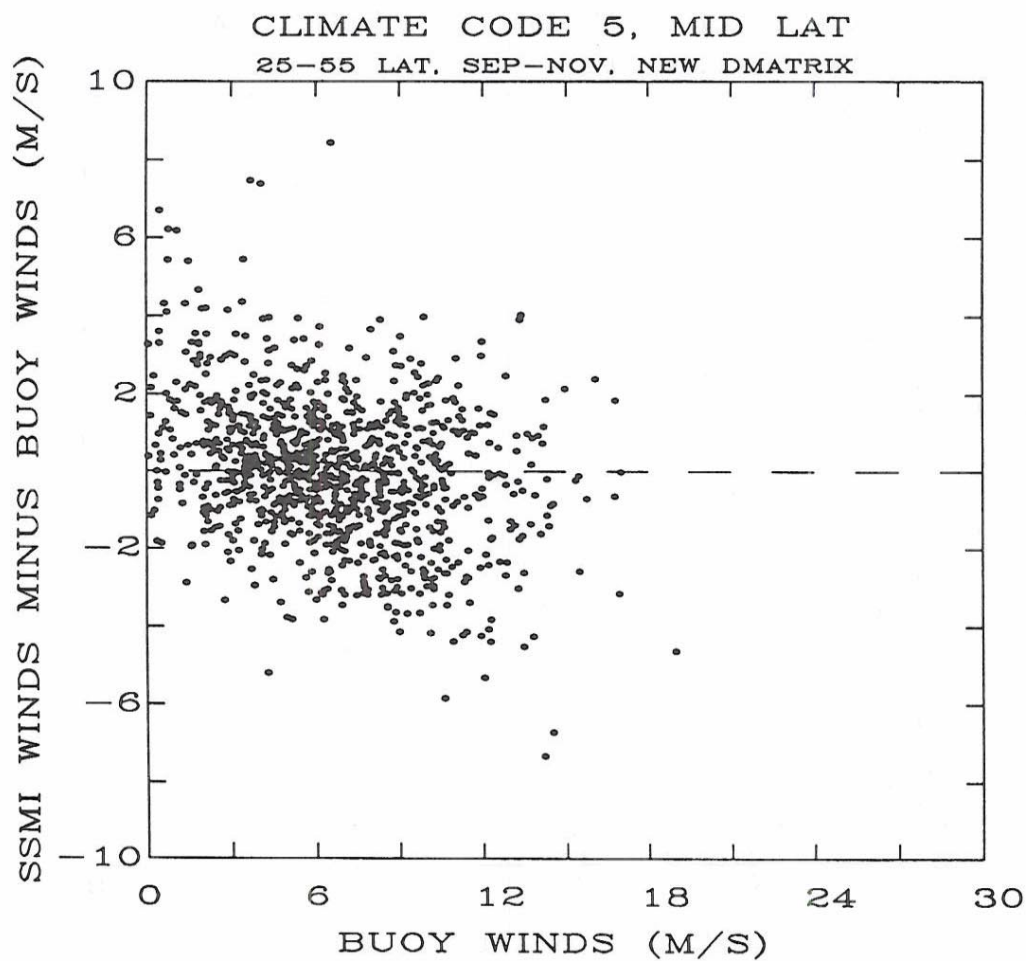
Figure 8.9 - D-matrix residual versus buoy winds for climate code 5.
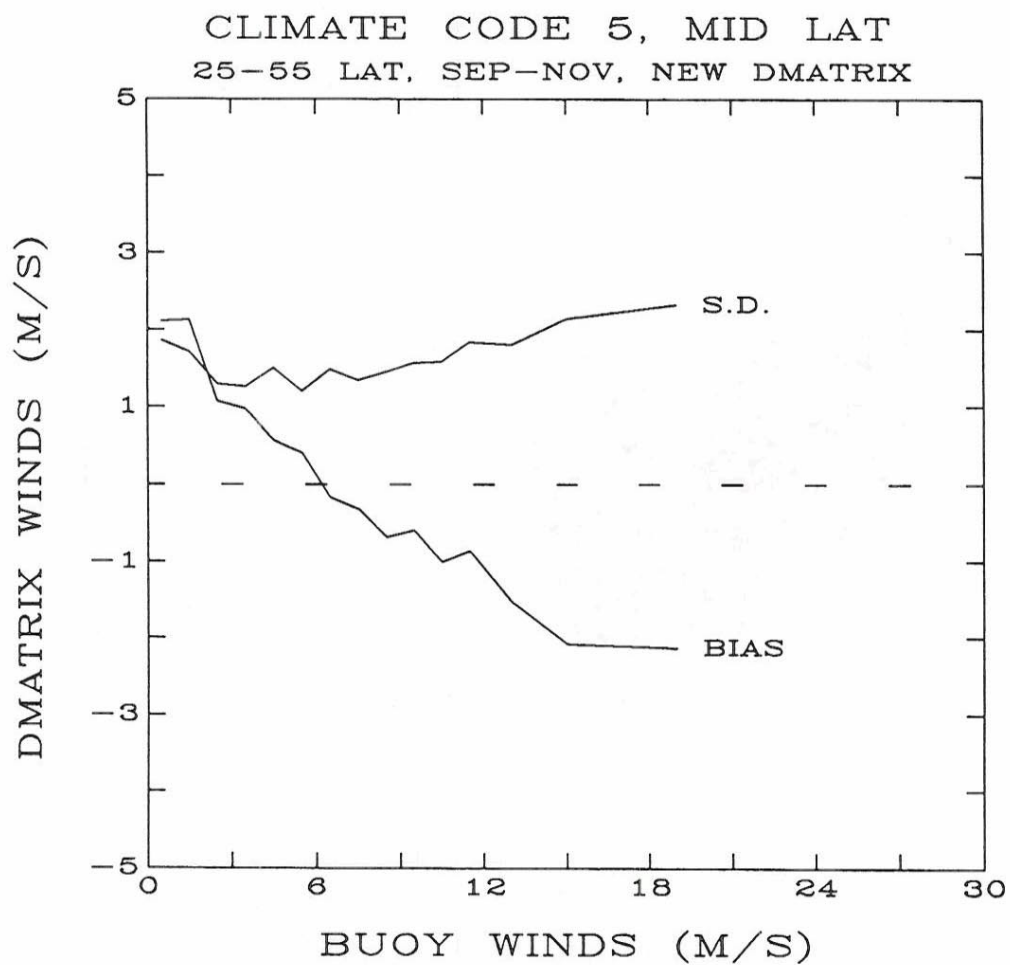
Figure 8.10 - Standard deviation and bias of D-matrix winds as a function of buoy winds for climate code 5.

TABLE 8.6

PERFORMANCE OF REVISED D-MATRIX ALGORITHM

Climate Code	S.D. m/s	Percentage Rain Flagged	Number of Comparisons
1	1.5	17	376
2	1.4	10	63
3	1.5	8	109
4	1.5	9	43
5	1.8	13	1296
6	1.5	12	643
7	1.9	18	516
8	1.8	19	279
9	1.6	9	277

wind speed of 7 m/s and becomes worse for predictions away from 7 m/s. Note that the trend of the SD curve agrees quite well with the pre-launch error budget model described in figure 8.2. Also note from the bias curve of Figure 8.10 that the high wind speed (>15 m/s) retrievals are biased low by more than 2 m/s.

Although the retrieval accuracy is met across the climate code boundaries, the discontinuity of the retrieved winds across these boundaries is disturbing. This is illustrated in the global chart (see Figure 8.11) of SSM/I wind speeds for the period January - February 1988. The average discontinuity across each latitude band boundary was also calculated using actual SSM/I data. The results are summarized in Table 8.7.

The accuracy of the wind speed retrievals deteriorates rapidly in rain as was indicated by Figure 8.7. This is not so much a problem with the algorithm as it is a problem with the frequencies used by the SSM/I. Microwave radiation at 19, 22 and 37 GHz is heavily attenuated by water vapor and rain in the earth's atmosphere, effectively masking the wind speed signature generated by ocean surface foam and waves. This attenuation significantly affects the ability of the SSM/I to retrieve accurate winds in and around typhoons and hurricanes where rain and heavy clouds are prevalent. Figure 8.12 shows the rain-flagged areas of typhoon Wynne as it appeared on July 25, 1987 at approximately 2040Z. According to aircraft reconnaissance data collected by the Air Force/Navy Joint Typhoon Warning Center, the boundary enclosing the rain-flag 3 area corresponds roughly to the 25 m/s wind speed radius of this storm. Visually observed winds from the aircraft near the storm center were reported to be as high as 60 m/s.
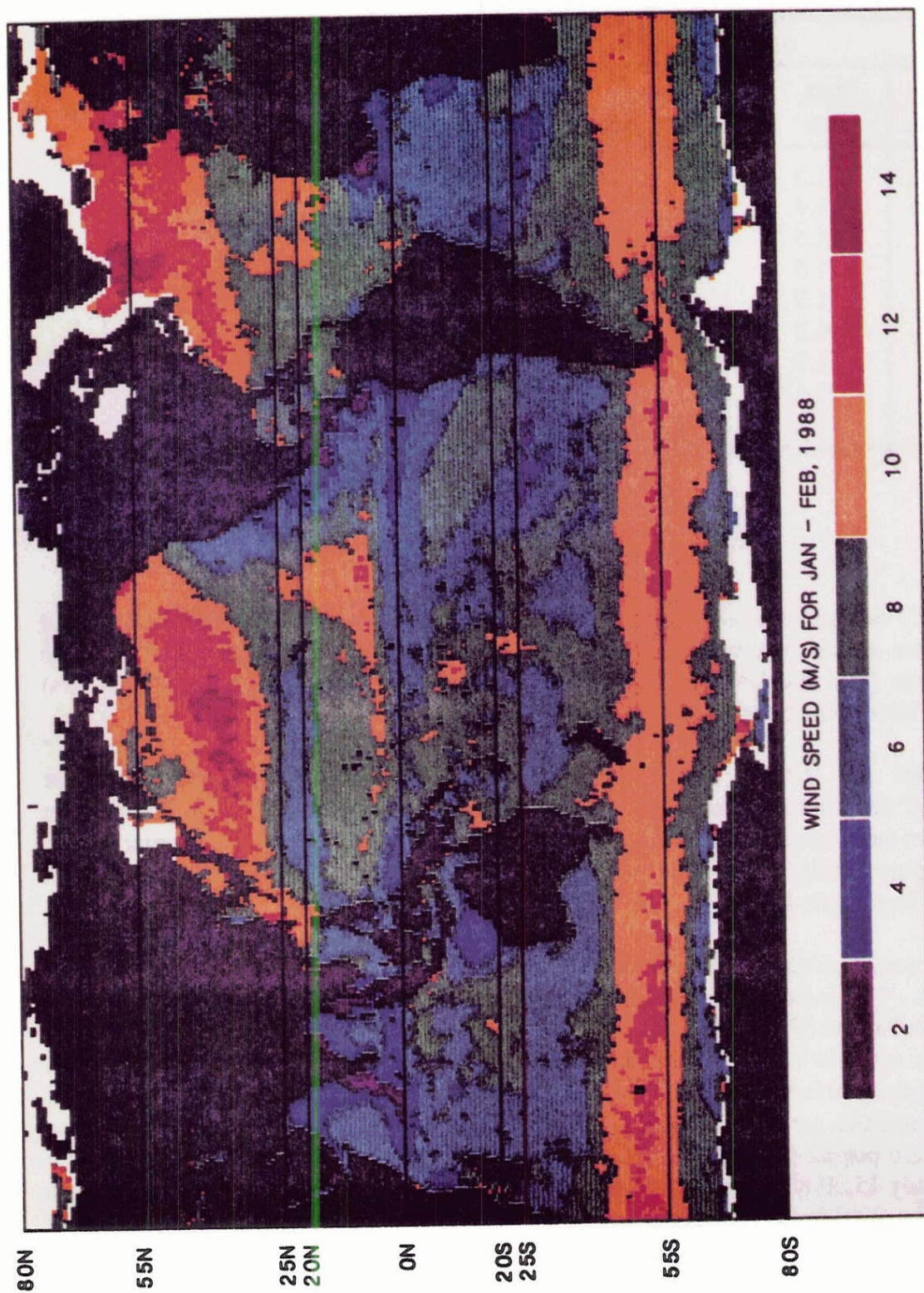
Figure 8.11 - Global color-coded chart of winds produced with the operational D-matrix algorithm for the period January-February 1988. Land is shown in black and sea ice is shown in white. Seven colors are used to display the wind speed in 2 m/s intervals, going from 2 to 14 m/s. Discontinuities are visible at the transition lines, especially at 20N, 25N, 20S and 25S.

WIND SPEED (M/S) FOR JAN – FEB, 1988

2   4   6   8   10   12   14

80N   55N   25N ON   20S   55S   80S
           20N   25S
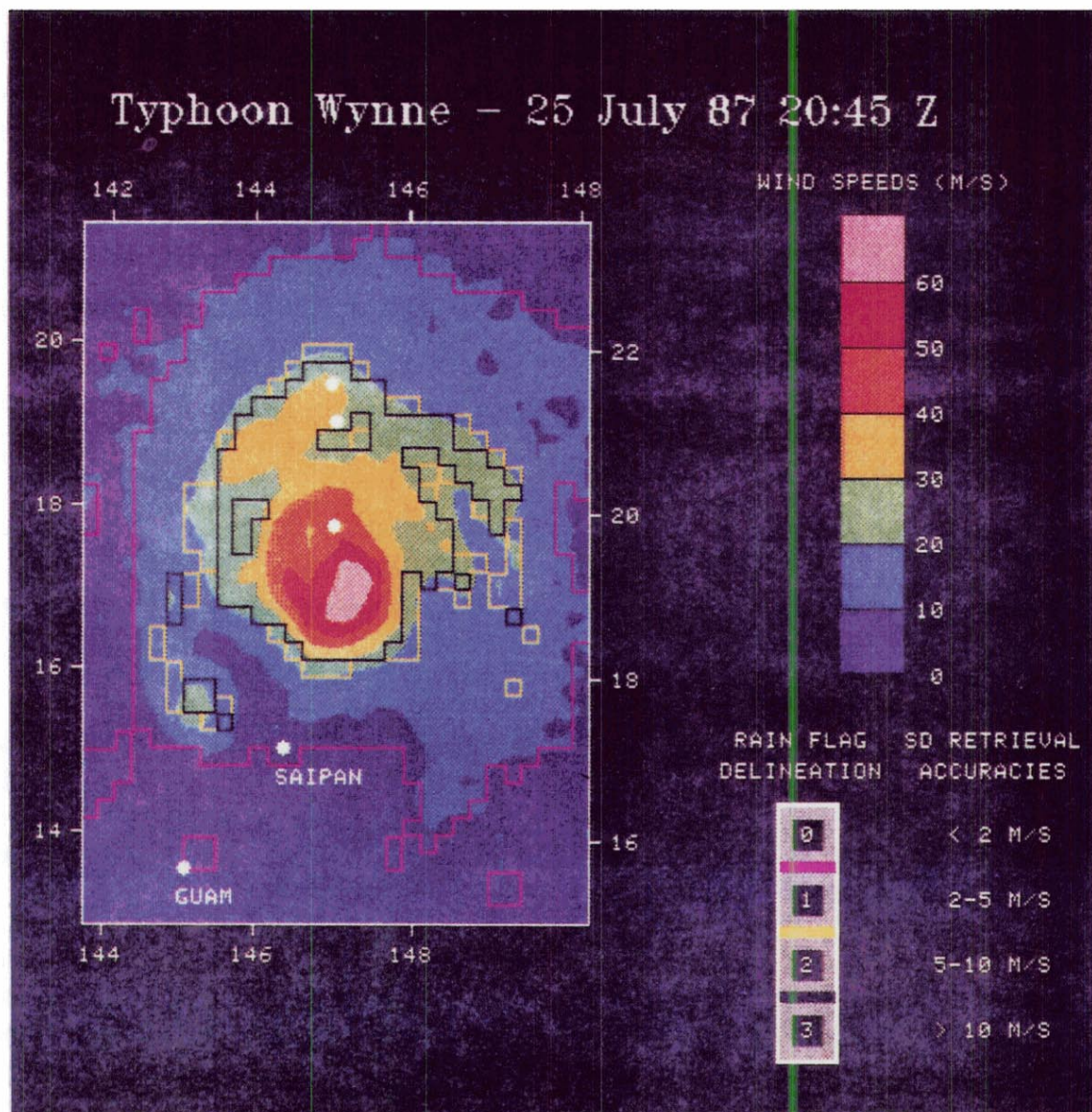
Figure 8.12 - Rain-flagged areas of typhoon Wynne as it passed over the Mariana islands on June 25, 1987 at 2045Z. Lines delineating rain-flagged areas within the storm are shown with their standard deviation retrieval accuracies.

Climate Codes	1	2	3	4	5	6	7	8	9
TABLE 8.7									
WIND SPEED DISCONTINUITY ACROSS THE D-MATRIX ZONAL BOUNDARIES									
1	0	0.1/1.4	0.5/0.6	---	---	---	---	---	---
2		0	---	1.2/1.8	---	---	---	---	---
3			0	---	1.9/1.6	2.0/1.5	---	---	---
4				0	1.4/1.2	---	0.8/2.1	---	---
5					0	---	---	0.4/1.3	1.9/2.9
6						0	---	0.5/1.5	---
7							0	---	1.3/1.9
8								0	---
9									0
Average (m/s)/Standard Deviation (m/s)									

An apparent SSM/I scan position bias in the D-matrix winds has been observed using the residual plot shown in Figure 8.13. A pitch, yaw and roll error of the SSM/I is believed to be partly responsible for this phenomena. This question is discussed further in another section of this report which addresses the geolocation problem. When the geolocation problem is solved, a slight adjustment of the D-matrix coefficients may be necessary.

In concluding this section, it should be noted that two and possibly three serious limitations of the 9-version original D-matrix algorithm warrant use of an alternate algorithm. As will be shown in the next section, both the high wind bias and zonal discontinuity problems can be partially solved using an alternate D-matrix type algorithm which utilizes a single set of coefficients, instead of nine, without a loss in the specified ± 2 m/s accuracy.

8.6   IMPROVED ALGORITHM

A single D-matrix algorithm, valid at all latitudes and during all seasons was developed and found to meet the ± 2 m/s accuracy specification under rain-flag 0 conditions. This global wind speed algorithm was developed using 900 randomly selected SSM/I buoy pairs (100 from each of the 9 climate codes). Out of this total, only 708 matched pairs (rain-flagged either 0 or 1) were retained to develop the new algorithm. In this way, the coefficients for the algorithm
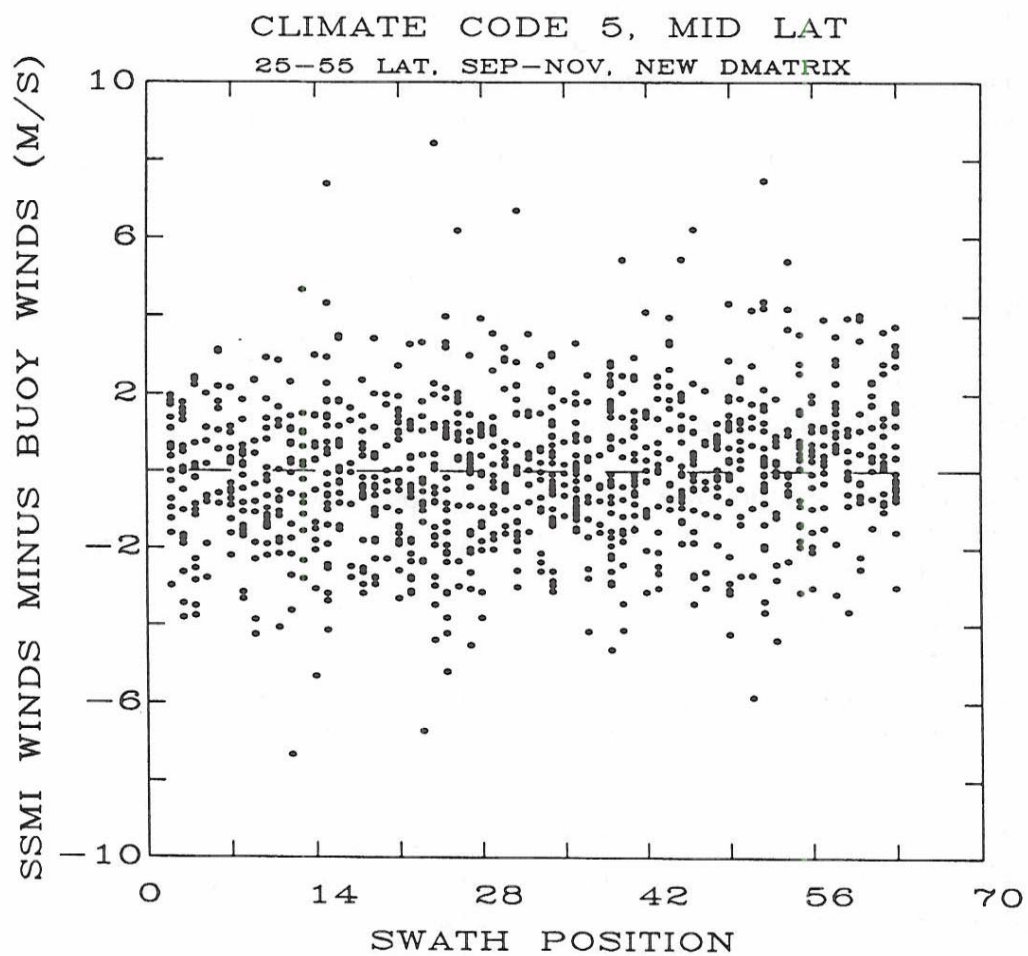
Figure 8.13 - D-matrix residual versus SSM/I swath position for climate code 5.

were generated using some data affected by rain (rain-flag 1 data), making the global algorithm somewhat tolerant of rain.

A weighted linear regression [10] of the buoy wind speeds on the coincident SSM/I brightness temperatures of T_{B19V}, T_{B22V}, T_{B37V}, and T_{B37H}, was done using the data set described above. The reason for using T_{B19V} instead of T_{B19H} will be discussed later in this section. The weights used in the regression were set equal to one over the square root of the wind speed density function (see Figure 8.1), evaluated at the particular buoy wind speed. This type of weighting has the effect of making all wind speed ranges equally important in the creation of the new algorithm. In contrast, the unweighted regression used previously tends to emphasize those wind speed ranges with the greatest amount of data and de-emphasize the ranges where little data was collected. This is precisely why the original D-matrix performed well near the global average wind speed of 7 m/s and performed poorly (both in terms of SD and bias) in the high (> 15 m/s) range.

Performance of the alternate global D-matrix algorithm, under rain-free conditions, is shown in Figure 8.14. The data used in this figure is comprised of withheld data taken from all 9 of the original D-matrix climate codes. In other words, the global wind speed algorithm was generated using one set of data and tested on another independent set. From Figure 8.14, the retrieval SD is found to be 2.0 m/s which meets the $\pm$ 2 m/s accuracy specification.

Although the regression line in Figure 8.14 shows slight errors in bias and slope, true performance of the alternate global wind speed algorithm is best illustrated by the interpreted residual plot shown in Figure 8.15. These results show that much of the high wind speed bias associated with the original D-matrix retrievals has been removed by the weighted regression technique. The sensitivity of the global wind speed algorithm to rain has not improved significantly as revealed by Figures 8.16 and 8.17. The feasibility of special D-matrix algorithms designed for use under rainy conditions will be addressed later in this section.

It is useful to know what SSM/I channels are most important in the retrieval of wind speeds. This aides in the construction of new algorithms and indicates what retrieval accuracies are possible should an SSM/I channel become inoperative. To this end, the 708 matched pairs of data previously described were again used to create the best global multichannel regression algorithms where the number of channels varied from 1 to 5. The results are summarized in Table 8.8 where the SD shown indicates the relative retrieval accuracy.

It is interesting to note that the best 4-channel algorithm (the proposed alternate global algorithm) does not use the same four channels as the original D-matrix algorithm. The proposed global algorithm uses T_{B19V} instead of the T_{B19H} channel employed by the original D-matrix algorithm. If T_{B19H} had been chosen instead of T_{B19V}, the performance would have been slightly worse with an SD of 2.1 m/s under rain-flag zero conditions. As in the original algorithm, the alternate global algorithm also uses the 4-channel D-matrix since it represents a good compromise between calculation efficiency and retrieval accuracy.
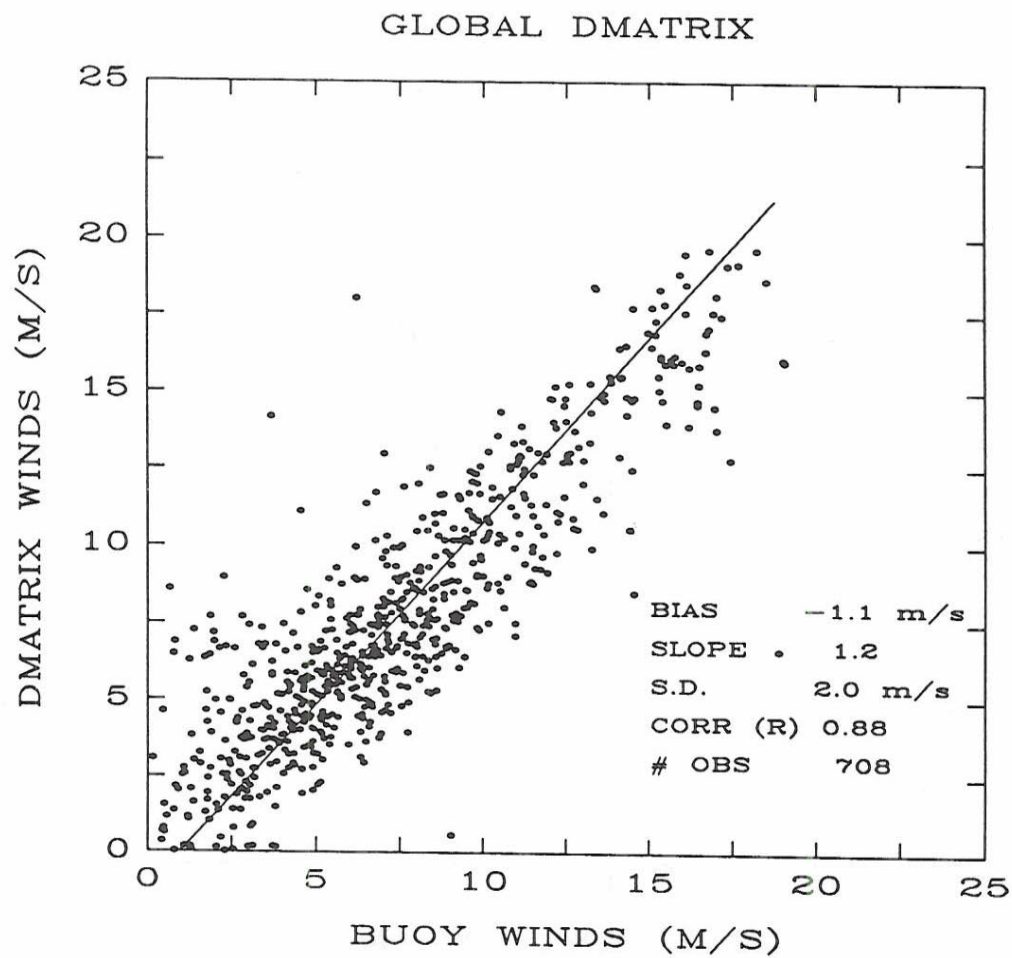
GLOBAL DMATRIX

Figure 8.14 - Performance of the global D-matrix algorithm.
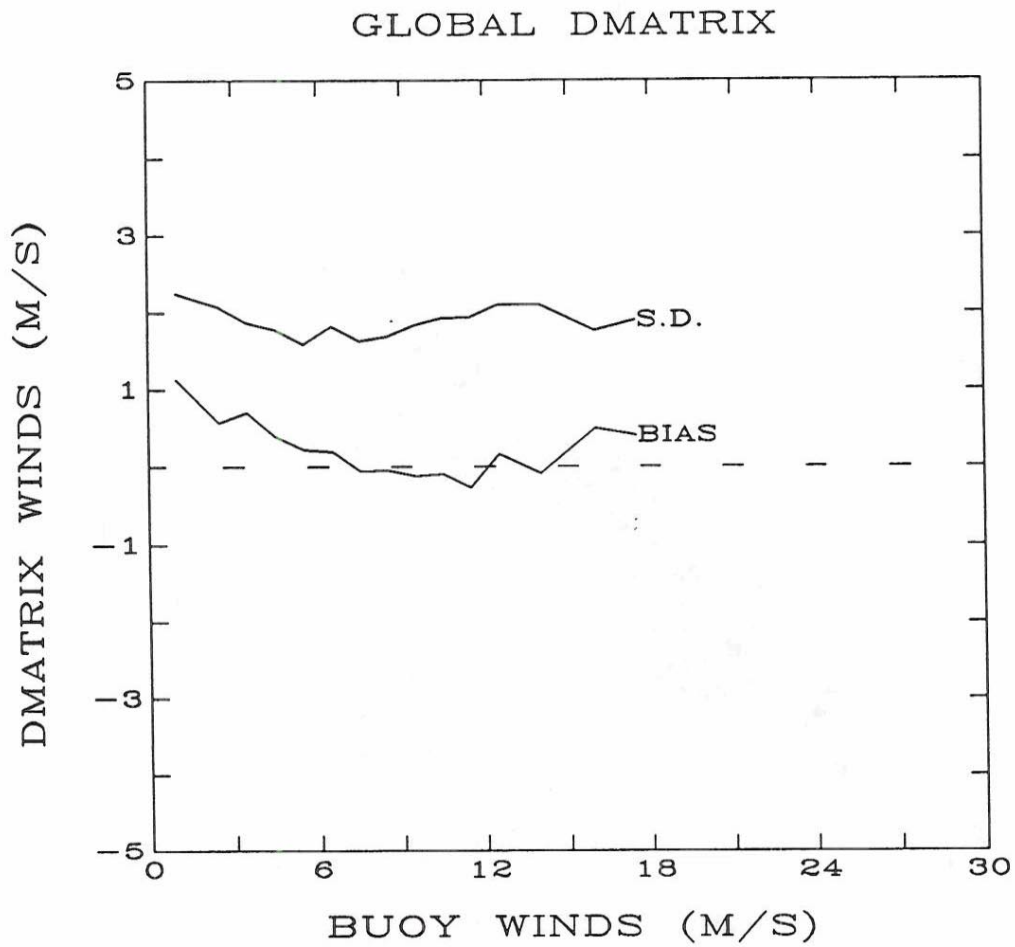
GLOBAL DMATRIX

Figure 8.15 - Standard deviation and bias of the global D-matrix
winds as a function of buoy winds.
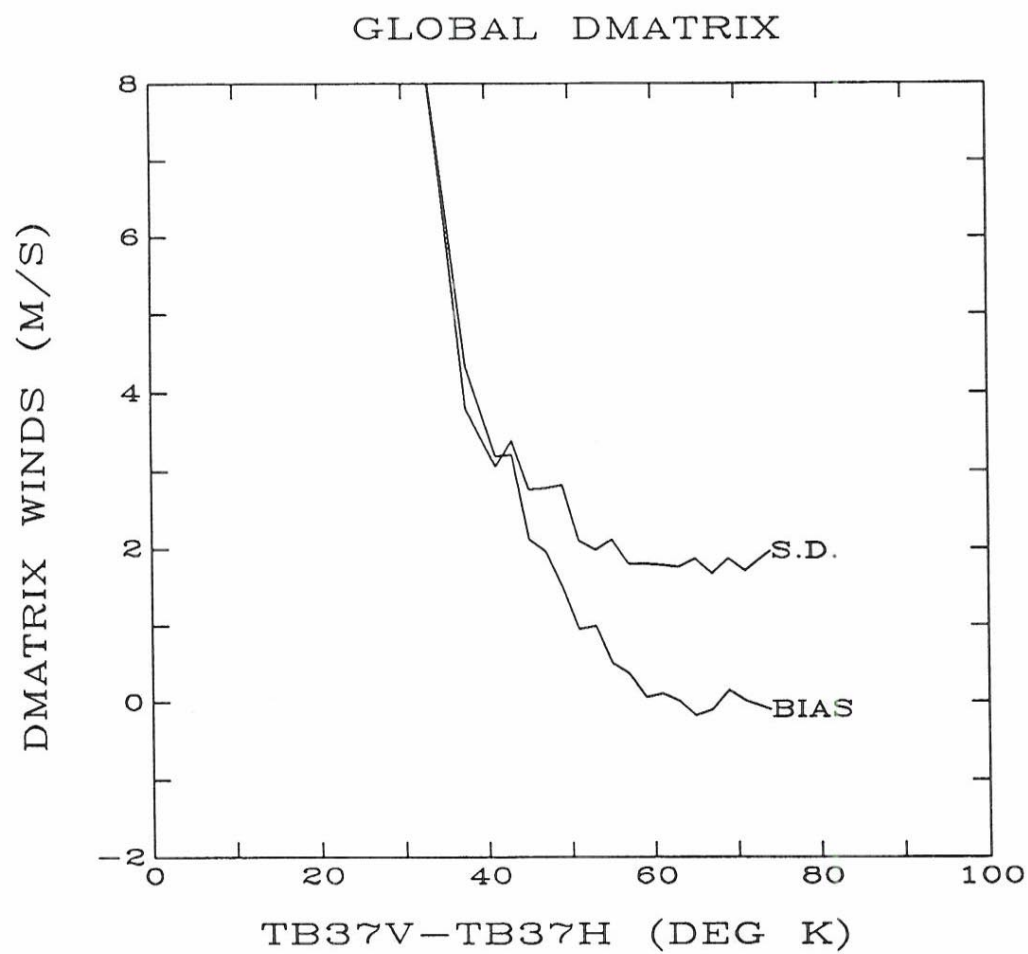
GLOBAL DMATRIX

Figure 8.16 - Standard deviation and bias of the global D-matrix
winds as a function of the rain-flag, $(T_{B37V}-T_{B37H})$.
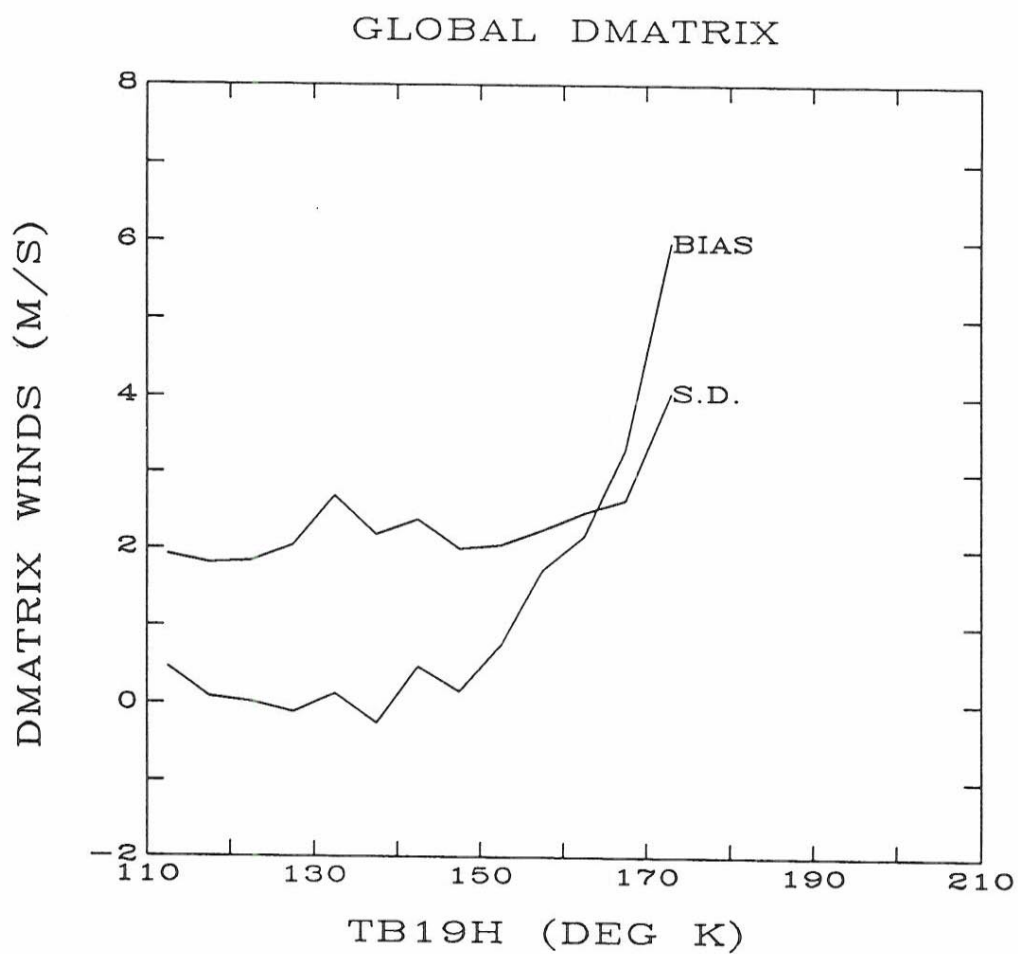
GLOBAL DMATRIX

Figure 8.17 - Standard deviation and bias of the global D-matrix
winds as a function of the rain-flag, (T_{B19H}).

TABLE 8.8

COEFFICIENTS AND RELATIVE PERFORMANCE
OF THE BEST MULTICHANNEL D-MATRIX ALGORITHM

NO. OF CHAN.	COEFFICIENTS						S.D
	CON	19V	19H	22V	37V	37H	
1	44.38	...	...	-0.1495	...	...	5.0
2	195.07	...	...	...	-1.5341	-0.9144	2.5
3	237.57	0.2613	...	...	-2.0413	1.0092	2.3
4	147.90	1.0969	...	-0.4555	-1.7600	0.7860	2.0
5	148.25	1.0233	0.0678	-0.4692	-1.6859	0.7371	2.0

Should one of the four selected channels become inoperative, a 3-channel or 4-channel algorithm can be constructed which would perform as indicated in Table 8.9. All algorithms coefficients in this table were generated from the same data set used to make the global algorithm.

In an attempt to get more accurate retrievals under rain-flagged conditions, special rain D-matrix algorithms were created and tested. These algorithms were constructed using a data set containing SSM/I-buoy pairs that were rain-flagged either 1, 2 or 3. The results are shown in Table 8.10. Note that the low-frequency channels (19 and 22 GHz) were identified as being "best" for the 1 and 2-channel algorithms indicating that they are less attenuated by the rain than are the high-frequency channels. The SD of the rain D-matrix retrievals under rain-flag 1, 2 and 3 conditions appear quite good. However, the results are misleading as indicated by Figure 8.18. This figure shows that the best rain D-matrix algorithm is simply predicting a near constant wind speed of approximately 10 m/s. The correlation coefficient associated with Figure 8.18 is 0.53, indicating that the algorithm can account for only about 25% of the variance in buoy wind speeds. The global D-matrix performance on the same data set is shown in Figure 8.19. A fair number of the global D-matrix retrievals in rain are quite good. This is expected since the algorithm was constructed using data that was rain-flagged either 0 or 1. Figure 8.19 also shows that the rain-flagged retrievals are typically biased high and the correlation coefficient of 0.27 indicates that the global wind speed algorithm performs poorly in rain as did the special rain D-matrix algorithms. Based on this analysis, it can be concluded that a special rain D-matrix
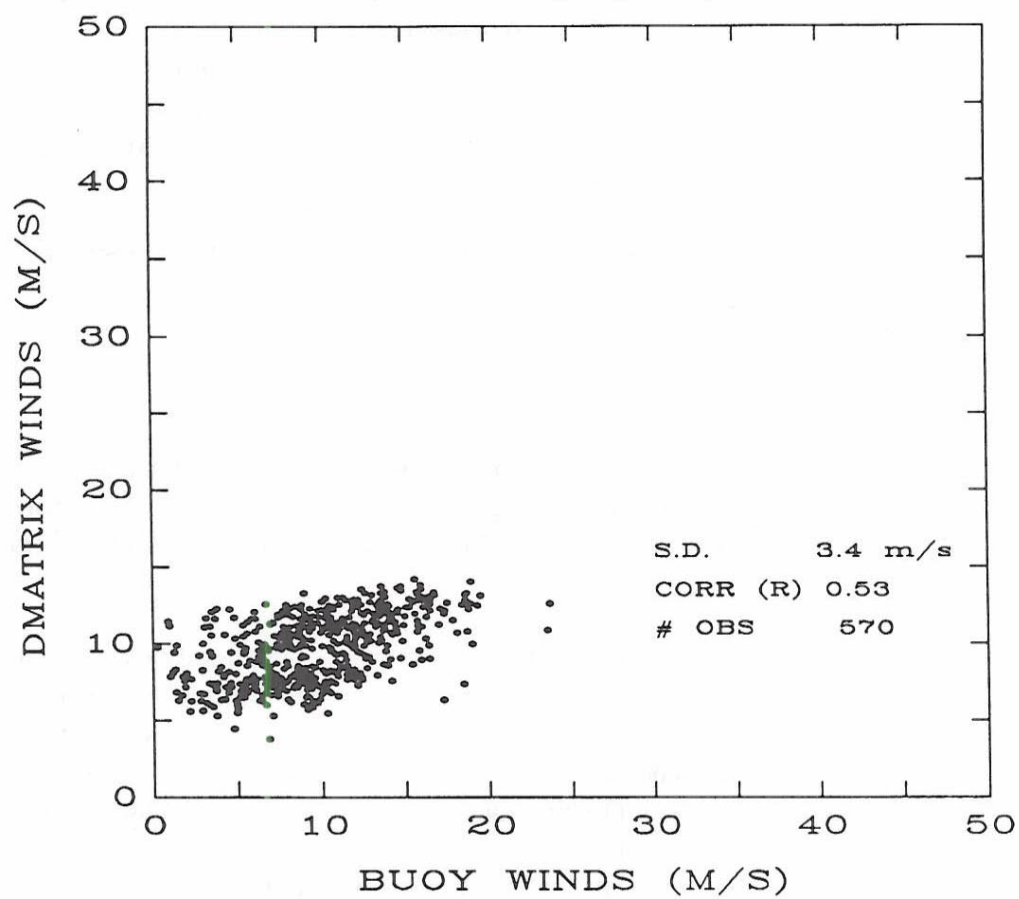
Figure 8.18 - Performance ot the special 5-channel D-matrix
algorithm designed for retrievals under rain-flag
1, 2 and 3 conditions.
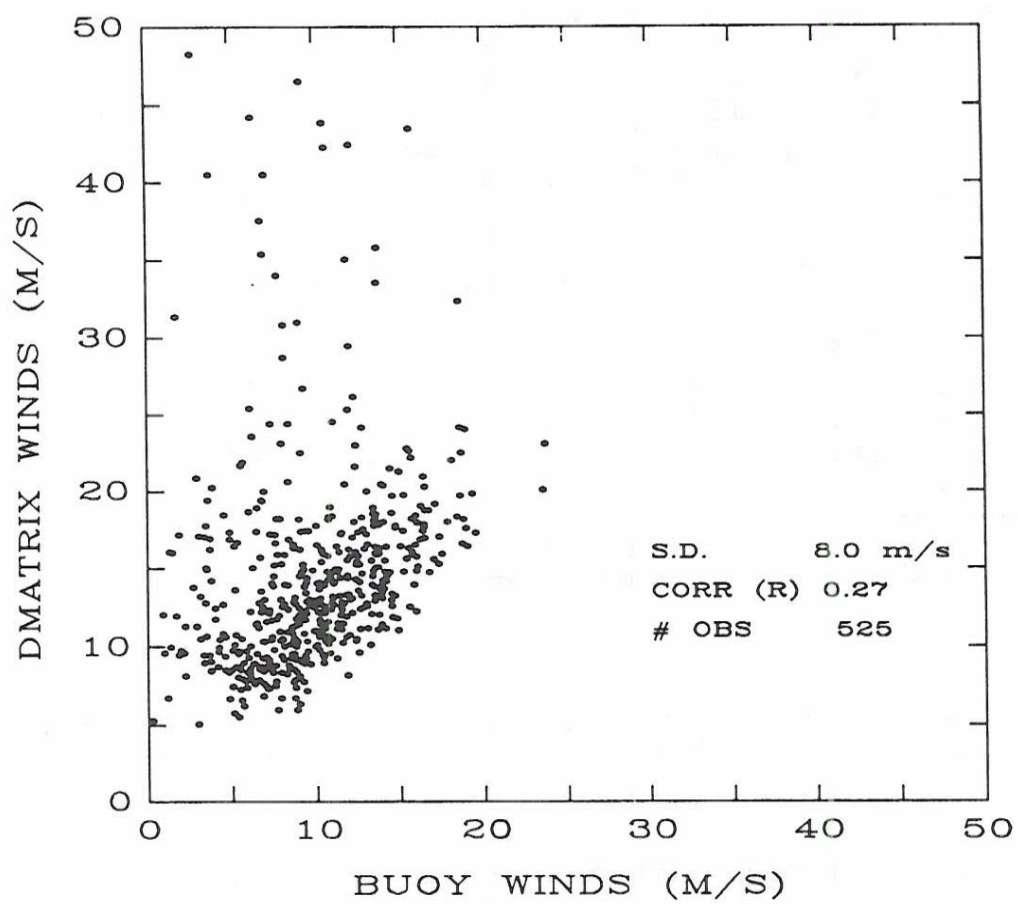
Figure 8.19 - Performance ot the global D-matrix algorithm
under rain-flag 1, 2 and 3 conditions.

| ALGOR. | COEFFICIENTS | | | | | | S.D. |
ID	CON	19V	19H	22V	37V	37H	
(3-Channel Algorithms)							
1	198.66	X	...	0.0072	-1.5642	0.9227	2.4
2	237.58	0.2613	...	X	-2.0413	1.0092	2.3
3	-47.46	0.8133	...	-0.6816	X	0.2988	3.3
4	-113.06	1.8278	...	-1.1173	0.0419	X	3.7
(Revised 4-Channel Algorithms)							
5	165.86	X	0.7208	-0.4729	-0.9091	0.2983	2.1
6	213.29	1.0437	-0.5325	X	-2.5612	1.3443	2.3
7	93.68	-0.1989	1.1056	-0.7511	X	-0.1703	2.5
8	124.65	0.1256	0.9607	-0.7236	-0.5050	X	2.3

TABLE 8.9

GLOBAL WIND SPEED ALGORITHMS WHICH CAN BE USED IF THE SSM/I LOSES A CHANNEL

"X" = Lost Channel

algorithm is not required and that the global D-matrix algorithm should be used to calculate winds under all conditions. It should be pointed out that 45 of the data points in Figure 8.18 do not appear in Figure 8.19 because the D-matrix values were above 50 m/s.

Although the D-matrix wind speed retrievals meet specifications under rain-free conditions, it has been suggested that an iterative-type algorithm might improve retrieval accuracy. Unlike the D-matrix algorithm, the iterative algorithms are based on a physical model which accurately predicts the effect that both wind speed and rain have on the measured brightness temperature. Since the rain-dependence is known, its contribution to the total brightness temperature can be effectively subtracted out making a more accurate wind speed retrieval possible under rain-free and light rain conditions. However, a fundamental limit on the retrieval accuracy of any wind speed algorithm is determined by the fact that microwave radiation at the selected SSM/I frequencies is heavily attenuated by rain. More specifically, microwave radiation emitted from the ocean surface, which contains information from which wind speed is inferred, must pass through the water laden atmosphere before being measured

TABLE 8.10

MULTICHANNEL D-MATRIX WIND SPEED ALGORITHM
FOR RAIN-FLAG 1, 2, AND 3 CONDITIONS

NO. OF CHAN.	CON	19V	19H	22V	37V	37H	S.D.
1	32.98	...	...	-0.0979	...	...	3.6
2	81.87	-0.5612	0.2895	...	...	...	3.5
3	78.02	-0.6193	0.3045	...	0.0597	...	3.5
4	86.37	-0.8860	0.4861	...	0.2471	-0.1270	3.4
5	79.72	-0.8291	0.4689	-0.0324	0.2862	-0.1477	3.4

by the SSM/I. If this important signal is attenuated to a level below the SSM/I instrument noise then accurate wind speed retrievals are no longer possible. The rain rate at which accurate SSM/I wind speed retrievals begin to degrade, regardless of the algorithm, seems to be about 2 mm/hr.

SSM/I wind speed retrieval accuracy in tropical storms, typhoons and hurricanes is limited not only by the rain associated with these storms but by the spatial resolutions of the 19, 22 and 37 GHz channels (55, 49, and 32 km). Wind speed gradients in the core regions of a storm are typically on the order of 2 m/s per kilometer and can persist over a distance of 25 km or more. Any SSM/I wind speed retrieval under these conditions would be a gross underestimation of the highest winds present in the resolution cell.

In an attempt to gather additional high wind speed data for the validation, D-matrix retrievals were compared with aircraft reconnaissance observed wind speeds in the 15-25 m/s range near typhoons Betty, Cary, Thelma, Vernon and Wynne. The reconnaissance flights were made during the typhoon season of 1987 by aircraft from the Air Force/Navy Joint Typhoon Warning Center. From this large set of data, less than 15 SSM/I-aircraft data comparisons met the criteria of being within 25 km and within 30 minutes of one another. Since only a few of the 15 match-ups were for winds exceeding 20 m/s, the results are considered statistically insignificant and are not shown. However, further analysis of this nature is needed to validate the high-wind (>15 m/s) performance of the D-matrix algorithm.

8.7   CONCLUSIONS

Although wind speed retrievals from the original versions of the D-matrix algorithm did not meet the accuracy specification of $\pm$ 2 m/s, regeneration of the D-matrix coefficients using standard linear regression resulted in an algorithm whose retrievals did meet specifications.

TABLE 8.11
THE RECOMMENDED GLOBAL D-MATRIX ALGORITHM

$$SW = 147.90 + 1.0969 \cdot T_{B19V} - 0.4555 \cdot T_{B22V} - 1.7600 \cdot T_{B37V} + 0.7860 \cdot T_{B37H}$$

RAIN FLAG	CRITERIA	ACCURACY
0	$T_{B37V} - T_{B37H} > 50$ AND $T_{B19H} < 165$	< 2 m/s
1	$T_{B37V} - T_{B37H} < 50$ OR $T_{B19H} > 165$	2 - 5 m/s
2	$T_{B37V} - T_{37H} < 37$	5 -10 m/s
3	$T_{B37V} - T_{37H} < 30$	> 10 m/s

An improved global D-matrix algorithm with a single set of coefficients has been developed which meets retrieval accuracy specifications but does not have the zonal discontinuity and high wind speed bias limitations found in the original 9-version D-matrix algorithm. Coefficients and rain-flag criteria for the global algorithm are given in Table 8.11.

The rain-flag criteria was revised to be more restrictive and the global algorithm now uses four rain-flags, 0 thru 3, which indicate retrieval accuracy SD's of <2 m/s, 2-5 m/s, 5-10 m/s and >10 m/s, respectively. In light of this redefinition it is perhaps more appropriate to use the term accuracy flag instead of rain-flag. Approximately 85% of the time, all forms of the D-matrix algorithm can be expected to retrieve ocean surface winds with an accuracy of $\pm$ 2 m/s. The remaining 15% of the time, the scene will be rain-flagged and retrieval accuracies will be worse than $\pm$ 2 m/s.

8.8 REFERENCES

[1]   J. Hollinger, R. Lo, G. Poe, R. Savage, and J. Peirce, Special Sensor Microwave/Imager User's Guide, Naval Research Laboratory, Washington, DC, 120 pp., 1987.

[2]   D. B. Gilhousen, "An Accuracy Statement for Meteorological Measurements Obtained from NDBC Moored Buoys," in <u>Proc. MDS '86 Marine Data Syst. Int. Symp.</u>, Marine Tech. Soc., New Orleans, LA, April 30 - May 2, 1986.

[3]   W. T. Liu and T. V. Blanc, "The Liu, Katsaros, and Businger (1979) Bulk Atmospheric Flux Computational Iteration Program in FORTRAN and BASIC", Naval Research Laboratory, Memo. Rep. 5291, Washington, DC, 1984.

[4]   F. M. Monaldo, "Expected Differences between Buoy and Radar Altimeter Estimates of Wind Speed and Significant Wave Height and their Implications on Buoy-Altimeter Comparisons," <u>J. Geophys. Res.</u>, vol. 93, no. C3, pp. 2285-2302, 1988.

[5]   L. C. Schroeder and J. L. Sweet, "Merge and Archival of SEASAT-A Satellite Data and In situ at Selected Illuminated Sites Over the Ocean," NASA Langley Tech. Memo., Hampton, VA, 1986.

[6]   S. A. Book, <u>STATISTICS</u>, New York: McGraw-Hill, 1977.

[7]   D. B. Gilhousen, M. J. Changery, R. G. Baldwin, T. R. Karl, and M. G. Burgin, "Climatic Summaries for NDBC Data Buoys," National Data Buoy Center Pub., NSTL, MS, 1986.

[8]   W. J. Pierson, "The Measurement of the Synoptic Scale Wind Over the Ocean," <u>J. Geophys. Res.</u>, vol. 88, 1983.

[9]   Hughes Aircraft Company, "Special Sensor Microwave/Imager SSM/I Critical Design Review," vol. 2, Ground Segment Report, pp. 117, Los Angeles, CA, 1980.

[10]   N. R. Draper and H. Smith, <u>Applied Regression Analysis</u>, 2nd ed. New York: Wiley, 1981.

[11]   W. S. Olson, private communication, University of Wisconsin, Madison, WI, 1988.

[12]   W. L. Jones, et al, "Airborne Microwave Remote-Sensing Measurements of Hurricane Allen," <u>SCIENCE</u>, vol. 214, 1981.

SECTION 9

LAND PARAMETER ALGORITHM VALIDATION AND CALIBRATION

by
Marshall J. McFarland
Professor
Department of Agricultural Engineering
Texas A&M University

and
Christopher M. U. Neale
Assistant Professor
Department of Agricultural and Irrigation Engineering
Utah State University

contributors:
Matthew Batchelor, Dennis Hill, Robert Miller
Jeffrey Miller, Richard Miller, Sanjay Salem, Susan Steinberg
Departments of Agricultural Engineering, Meteorology,
and Electrical Engineering
Texas A&M University

and
Bruno G. Gerard
Department of Agricultural and Irrigation Engineering
Utah State University

9.0   LAND PARAMETER ALGORITHM VALIDATION AND CALIBRATION

9.0.1   Data Storage and Software Development

The calibration and validation of algorithms to retrieve land parameters from the SSM/I passive microwave data required the development of databases for several land areas of the world. For example, test areas over tropical jungles, deserts, and agricultural areas were established for the development of land surface type classification rules. Additional test areas were developed in the United States where ground truth was available from surface observations.

The Hughes Early Orbital Display System (HEODS) software was used for the selection of satellite overpasses for areas and dates of interest. The SSM/I data of interest were furnished on 9 track, 6250 bpi magnetic tape from the Naval Research Laboratory. Data from the tapes were downloaded to files on a clustered computer system consisting of a VAX-8300, a VAX-8650, and a VAX-8800. The Sensor Data Records (SDRs) were pre-processed to strip them of inter-record blank spaces, header records, and other non-relevant information. These files, saved in the conical scan format for the test areas, were denoted as SCAN files. The SCAN files for the test areas were backed up on magnetic tape. Environmental Data Records (EDRs) were processed in a similar manner. Images were created from SCAN files for visual screening on an International Imaging System operated as a peripheral on a VAX-750.

The SDR SCAN files, which consisted of the seven channels of microwave brightness temperatures and the latitude-longitude tags for each pixel, were then loaded into a relational database, RDB. Latitude and longitude, the "relation" of the RDB, was used to facilitate the development of the ground truth data base coincident with the SSM/I database. With the specification of latitude and longitude coordinates, all SSM/I and ground truth data could be assembled as one file.

Data sets of coincident SDRs and ground truth were extracted from the RDB in the form of one-half degree latitude and longitude cells. These CELL files contained the average of all data with a latitude-longitude location in the cell. Some CELL files were also created for one-quarter degree latitude-longitude boundaries. SPOT files were created by matching the closest SDR file of seven channels of brightness temperatures with a specified latitude and longitude.

The primary source of ground truth consisted of climatological data from the National Oceanic and Atmospheric Administration (NOAA) cooperative observer network. The Summaries of the Day Elements (TD3200) from the reporting stations were provided on 9 track, 6250 bpi magnetic tape from NOAA/NESDIS, Asheville, NC. These daily elements included maximum and minimum air temperatures, rainfall or water equivalent of snow, daily snow depth, and total snow accumulation. GOES satellite imagery from the Department of Meteorology at Texas A&M University was used to visually screen the data for cloud and synoptic weather conditions.

9.1 LAND SURFACE TYPE CLASSIFICATION SCHEME

9.1.1 Rationale for New Classification Rules

The EXTLND module, described in the SSM/I User's Guide [1], is a subset module of the entire environmental parameter extraction software. Surface types over land are classified within EXTLND using SDR brightness temperatures so that appropriate parameter extraction algorithms are used. Initial analysis of EDR's resulting from the original EXTLND algorithms indicated numerous misclassifications with respect to land surface types. One of the most common misclassifications was the indication of rain when no rain or clouds were present in the scene. This was due to a flag within the original EXTLND logic which compared the brightness temperatures in the 37 GHz and 19 GHz channels. If the 37V brightness temperatures were less than 19V brightness temperatures, a heavy rain event was classified. However, over naturally occurring surfaces such as vegetation, bare soil, and deserts, the brightness temperatures at 37V GHz were frequently found to be less than those at 19V GHz.

In addition to misclassification within the EXTLND logic, it was imperative that surface types be differentiated prior to the creation of calibration/validation databases. The reasons were:

1. The calibration/validation project required parameter extraction algorithms over different surface types. Some extraction algorithms are mutually exclusive, such as surface moisture and snow parameters, but require the proper identification of those conditions. Other surface types such as standing water, do not require the extraction of surface parameters. In addition, during the course of an annual cycle for an agricultural region, such as winter wheat production areas, a natural change in the land surface type occurs. Surface conditions would begin with dry snow in the middle of the winter. The snow would undergo morphological changes and additional accumulation in the snow accumulation phase. With the onset of warmer weather, the snow would enter the ripening phase, again with pronounced responses in the microwave frequencies and polarizations. With complete snow melt, a flooded or wet soil surface may occur. Spring tillage or greenup of winter grains would be associated with arable land, with varying degrees of soil moisture. Increases in vegetation canopy density would decrease the response to soil moisture, but theoretically should increase the accuracy of the land surface temperature retrieval. From harvest to snow accumulation, the cycle continues with bare soil, developing canopy of the winter wheat, frozen and unfrozen soils, and snow accumulation. Rains and varying atmospheric water vapor and liquid water contents occur throughout the entire year.

2. Over land, there may be a large variability of natural surface types within an SSM/I footprint. These include different degrees of vegetation cover, topographic characteristics, and the presence of water bodies such as lakes and reservoirs. Water bodies can increase the noise in parameter extraction regression data sets for surface moisture and land surface temperature if included. As they have a distinct detectable signature in the 85.5 GHz channels, their

9-2

classification and removal from the data sets would ultimately increase the confidence in parameter retrievals.

3. The surface moisture retrieval algorithm was based on an apparent emissivity (19H/37V). The degree of vegetation cover within a footprint affects the sensitivity of this variable with respect to moisture at the soil surface, thus requiring further categorization.

9.1.2 Land Surface Type Classification Methodology

9.1.2.1 Observations

The approach used in the development of the classification rules can be considered a combined physical/statistical method. Channel brightness temperature and polarization differences along with statistically determined threshold values were used to form the rules. For a particular surface type, the channel combination or polarization difference selected had a microwave physical basis. The basic land surface types developed were selected to function with the land surface parameter extraction algorithms being validated in a parallel effort and presented in Sections 9.2, 9.3 and 9.4 of this report. Additional land surface classes are possible, but would probably be subsets of the major classes presented herein or anomalous cases.

9.1.2.2 Methodology

The CLIPS expert system environment, created by NASA, was used to develop the land surface type classification algorithms. An expert system environment was selected for this purpose because it facilitated the addition, removal, or modification of rules as well as brightness temperature and polarization difference thresholds without the necessity of recompiling the software code. The CLIPS shell and the rules for classification were embedded within a main program module written in the C programming language.

The initial set of classification algorithms incorporated the logic and thresholds of the original EXTLND module, described in the SSM/I Users Guide [1]. These classification algorithms, and their subsequent modifications, were used to classify various land surface types. Images of the classifications were used in conjunction with geographical and natural resource maps to determine the accuracy of the classification scheme.

Training areas were selected for the various surface types in different regions of the world and the United States. For example, control areas in the Amazon and Congo jungles were used to identify the characteristic microwave signature of dense vegetation in the SSM/I channels. Control areas in the Sahara and Sonoran deserts were used to identify the desert signatures. A summary of the main training (control) areas is shown in Table 9.1.

SDR data from several orbits over these training areas were grouped according to overpass time, cloud condition, and season. The SDR data used consisted of the seven brightness temperatures of the A scan concentric footprints. The value for the 85.5 GHz

channels assigned to the concentric footprint consisted of an average of the surrounding eight 85.5 GHz footprints from A and B scans. Several combinations of SDR brightness temperatures with respect to frequency and polarization differences were calculated. These combinations are shown in Table 9.2.

Basic statistics of brightness temperatures and polarization differences were obtained for each surface type. These basic statistics included mean, standard deviation, mode, skewness, distribution type etc. A set of new rules, identified through the statistical analysis, were developed based on brightness temperatures, brightness temperature combinations, and polarization differences. New rules were added to the expert system module and tested against independent data sets. These data sets were either for different geographical areas with similar characteristics or for different seasons.

Another source of ground truth information for the validation of classification rules was the major land resource area (MLRA) classifications of the Soil Conservation Service [2]. These classifications grouped areas with similar characteristics with respect to topography, natural vegetation, land use, climate, soils and water resources.

TABLE 9.1   A SUMMARY OF SOME CONTROL AREAS USED IN SURFACE TYPE IDENTIFICATION

Control Area	Surface Type	Location	Boundaries
A	Dense Vegetation	Amazon Jungle	NW Corner: 2°S 54°W SE Corner: 4°S 52°W
B	Dense Vegetation	Congo Jungle	NW Corner: 1°S 20°E SE Corner: 3°S 23°E
C	Dense Vegetation	Amazon Jungle	NW Corner: 5°S 69°W SE Corner: 8°S 66°W
Amazon Basin	Dense Vegetation	Amazon	NW Corner: 0°S 64°W SE Corner: 10°S 50°W
MLRA #130	Dense Vegetation	Appalachian	NW Corner: 36.3°N 83°W SE corner: 35.3°N 82°W
Appalachian Forest	Dense Vegetation	Appalachian	NW corner: 40°N 87°W SE corner: 33°N 80°W
Central Plains	Mixed Vegetation & Soils	United States	NW Corner: 50°N 105°W SE Corner: 32°N 95° W
MLRA #30	Semi-Arid Veg. & Soils	Mojave Desert California	NW corner: 35.5°N 118°W SE corner: 34.5°N 116°W
Sahara	Desert	Sahara, Libya	NW corner: 16°N 18°E SE corner: 14°N 21°E

TABLE 9.2   COMBINATIONS OF MICROWAVE BRIGHTNESS TEMPERATURES USED
FOR THE CHARACTERIZATION OF LAND SURFACE TYPES

(T22V - T19V)	[a]	(T19V + T37V)/2 - (T19H + T37H)/2	[b]
(T19V + T37V)/2		(T37V - T19V)	[c]
(T85V - T37V)	[d]	(T85H - T37H)	[e]
(T37V - T37H)	[f]	(T37H - T19H)	[j]
(T19V - T19H)	[i]	(T85V - T85H)	[k]

Letters in brackets [ ] indicate how the combination is referred to
throughout the text.

9.1.3 Development of Classification Rules

9.1.3.1 Dense Vegetation

Emission by a vegetation canopy consists of contributions from the vegetation layer as well as from the underlying soil surface [3]. At the SSM/I channel frequencies (19.35 GHz and greater), vegetation canopies can be treated as semi-infinite mediums with respect to emission properties. According to Ulaby et al. [3], the brightness temperature of a weakly scattering media above a semi-infinite medium can be simplified to:

$$T_{b,can}(\theta,p) = (1+\Gamma_s(\theta,p)/L(\theta))(1-1/L(\theta))(1-a)T_v + (1-\Gamma_s(\theta,p)T_s/L(\theta)) \qquad (1)$$

where:

Γ_s = the air-soil reflectivity
$L(\theta)$ = the loss factor of the vegetation canopy
a = single-scattering albedo of vegetation
T_v = physical temperature of the vegetation layer
T_s = physical temperature of the soil surface
θ = incidence angle
p = polarization index equal to v or h

The loss factor $L(\theta)$ depends on the height of the vegetation layer, the incidence angle and the microwave frequency. For frequencies above 10 GHz, the optical thickness is large and $L(\theta) >> 1$. Equation (1) can then be approximated by:

9-5

$$T_{b,can}(\theta) = (1-a)T_v \qquad\qquad (2)$$

This implies that the canopy brightness temperature is independent of the incidence angle θ, and of antenna polarization if a is isotropic. This has certainly been true for the SSM/I frequencies in the case of very lossy canopies such as dense jungle. Brightness temperature polarization differences at all frequencies have been very small for pixels over the Amazon and Congo jungles. However a dependence of brightness temperature with frequency has been observed which implies that equation (2) is only a first-order approximation of the emission from vegetation for the SSM/I frequencies.

Table 9.3 summarizes the main statistics for selected channel combinations over dense vegetation control areas. These locations, selected from natural resource maps, avoided large rivers and lakes. Figure 9.1 shows a histogram of average brightness temperature polarization differences in the 19.35 and 37.0 GHz frequencies for combined ascending and descending overpass data over control areas A and B in the Amazon region. The distribution was close to a normal distribution with a mean polarization difference of 0.67 K (combination [b] in Table 9.3). Brightness temperatures in the 37.0 V GHz channel were on the order of 4 K lower than in the 19.35 V GHz channel while the 85.5 V GHz brightness temperatures (T_B's) were around 2 K higher than in the 37.0 V GHz channel. No physical explanation was found for this "dip" in the 37 GHz brightness temperatures over dense vegetation.

By selecting twice the standard deviation as the upper and lower limits for the normal distribution of brightness temperature combinations shown in Table 9.3, 96% of all occurrences will fall between those limits. Based on these results for the three control areas over dense vegetation, the upper limit of average brightness temperature polarization difference in the 19.35 GHz and 37.0 GHz (combination [b]) was set at 1.9 K. The lower limit was around -0.4 K using the same rationale. Although true negative polarization differences are physically impossible from horizontal surfaces a small amount of such cases were observed in the SSM/I data over dense vegetation. This could be due to random noise within the individual channels. As the energy being emitted from dense vegetation is essentially depolarized, it is possible that the brightness temperatures in the horizontal channels can become greater than in the vertical channels on some occasions, but still be within the acceptable variability of the instrument. A second possible explanation involves the structure of jungle vegetation. Microwave energy emitted from dense vegetation will be isotropic. If any predominant orientation is present in the vegetation, the emitted energy will have polarization differences. For a tropical rain forest, the tall, vertical tree trunks could provide the predominant orientation. If this were the case, the frame of reference for the horizontal and vertical polarizations would reverse. The largest component of the emitted radiation would be in a plane perpendicular to the vertical tree trunks. For the frame of reference of the vertically and horizontally polarized brightness temperatures of the SSM/I, the horizontally polarized brightness temperature could exceed the vertically polarized brightness temperature. Based on the combinations shown in Table 9.3 as well as statistics for single channels, the rule to classify dense vegetation becomes (all thresholds and temperatures in Kelvin):

TABLE 9.3 STATISTICAL ANALYSIS RESULTS FOR SELECTED BRIGHTNESS TEMPERATURE COMBINATIONS OVER DENSE VEGETATION CONTROL AREAS

CONTROL AREAS A & B

Combination		Mean (K)	SD (K)	Lower Limit (K)	Upper Limit (K)	Overpass Calendar Date
22V - 19V	[a]	-1.74	1.13	-4.00	0.52	(1988) 56D, 57D
$\frac{(19V+37V)}{2} - \frac{(19H+37H)}{2}$	[b]	0.67	0.53	-0.39	1.73	58D, 59A
37V - 19V	[c]	-4.28	0.70	-5.68	-2.88	(1987) 176D, 178D
85V - 37V	[d]	2.46	1.23	0.00	4.92	222A, 242A 243A, 250A
85H - 37H	[e]	2.58	1.03	0.52	4.64	

CONTROL AREA C

Combination		Mean (K)	SD (K)	Lower Limit (K)	Upper Limit (K)	Overpass Calendar Date
$\frac{(19V+37V)}{2} - \frac{(19H+37H)}{2}$	[b]	0.54	0.49	-0.44	1.52	58D
85V - 37V	[d]	3.26	1.28	0.70	5.82	180A, 222A 231A
85H - 37H	[e]	3.04	0.87	1.30	4.78	231A

AMAZON BASIN

Combination		Mean (K)	SD (K)	Lower Limit (K)	Upper Limit (K)	Overpass Calendar Date
$\frac{(19V+37V)}{2} - \frac{(19H+37H)}{2}$	[b]	0.67	0.54	-0.41	1.75	252D
85V - 37V	[d]	3.04	0.90	1.24	4.84	
85H - 37H	[e]	2.98	0.77	1.44	4.52	

Combination		MLRA #130 Mean	SD	Appalachian Forests Mean	SD	
$\frac{(19V+37V)}{2} - \frac{(19H+37H)}{2}$	[b]	1.57	0.29	1.49	0.40	246D
						247D 225A
37V - 19V	[c]	-4.22	0.77	-3.92	0.78	227A 244A
85V - 37V	[d]	1.89	1.01	1.89	1.03	245A 248A
85H - 37H	[e]	2.19	0.80	1.96	1.10	250A

Dense Vegetation
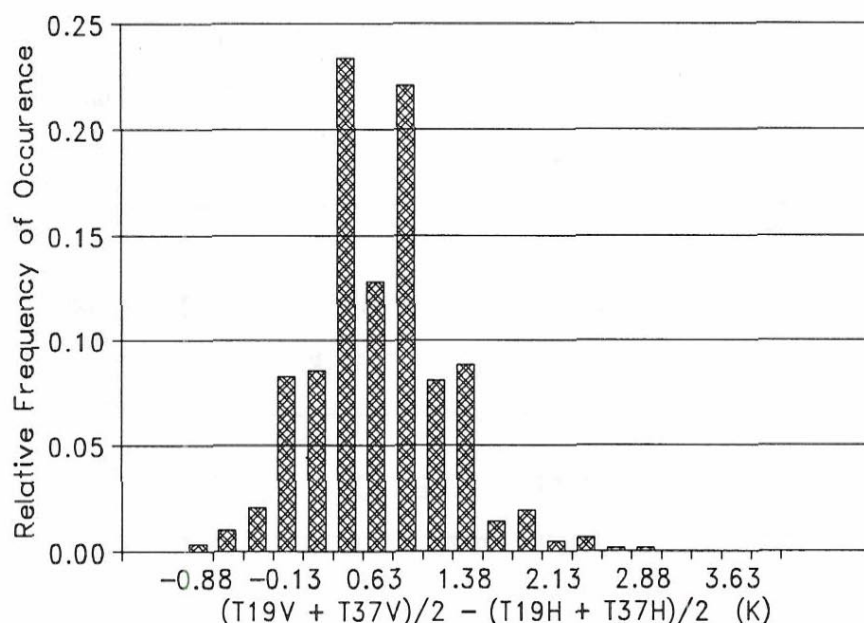
Figure 9.1 Distribution histogram of average polarization in the 19.35 GHz and 37.0 GHz channels over dense vegetation.

$$22V - 19V <= 4.0 \qquad \text{[a]}$$
$$((19V + 37V)/2.0) - ((19H + 37H)/2.0) <= 1.9 \qquad \text{[b]}$$
$$85V - 37V >= -1.0 \qquad \text{[d]}$$
$$85H - 37H < 4.5 \qquad \text{[e]}$$
$$19V > 262 \qquad \text{[g]}$$

Conditions [a] and [e] check for the presence of large quantities of water on the surface within the footprint and will be discussed later. Condition [b] is the check for low polarization differences, the characteristic microwave signature of dense jungle vegetation. Condition [d] is a precipitation flag and is based on the lower limit for this channel combination shown in Table 9.3. The adjusted threshold of -1 K is suggested instead of 0 K (lower limit for the distribution shown in Table 9.3) to ensure that only precipitating clouds are classified for the precipitation over vegetation rule, also discussed later. Condition [g] is a check for above freezing temperatures in the vegetation canopy (for a single-scattering albedo of 0.04).

Table 9.3 also presents brightness temperature combination statistics for additional (and independent) orbits over control area C of the Congo Jungle and a scene encompassing a large portion of the Amazon basin. For the latter scene, data corresponding to footprints sensing rivers and other non-vegetation classifications were removed from the data set. These results are not significantly different from those of control area A and B. Results from the independent data confirm that this rule properly classifies dense vegetation situations.

In the United States, the closest vegetative covers to dense tropical jungles are found in the hardwood forests of the Appalachian mountains. SSM/I data obtained over a small area of the Appalachian mountains as well as the Major Land Resource Area (MLRA) #130 are also shown in Table 9.3. The MLRA 130 resource region consists of dense forests of different oak varieties, white pine, hemlock, red spruce, balsam fir and several species of understory vegetation [2]. The polarization differences were about 1 K greater than for the dense jungle, and resulted from a lower density canopy.

9.1.3.2   Dense Agricultural Crops and Rangeland Vegetation

This rule applies in situations where soil is totally or partially covered by vegetation within an SSM/I footprint. Such occurrences are common in agricultural regions with crops at different stages of growth or canopy cover; on rangeland with grasses and shrub type vegetation at peak growth, or on combinations of these. This category of vegetation is still considered dense with respect to surface soil moisture retrievals. As discussed in section 9.3 of this report, the sensitivity to surface moisture is very small for average polarizations in the 19.35 GHz and 37.0 GHz of less than 4 K, rendering retrievals physically impossible. Examples of such regions are:

1.   Agricultural areas and grasslands of the Central Plains of the U.S. and some rangeland of the western U.S. at peak vegetation cover.

2.   The "cerrado" vegetation region of central Brazil. These are savanna type areas with extensive grasslands mixed with small trees and shrubs.

3.   The Savanna regions of Africa at peak vegetation cover.

The green vegetation density, which can be quantified by the Leaf Area Index (LAI), will vary considerably throughout the year in these regions, according to season. The peak LAI for an agricultural region in the Central Plains can occur during the months of May through August, depending on the latitude and type of vegetation, and if the vegetation is growing under natural precipitation or irrigation (crops). Vegetation densities in grasslands and savannas will also vary according to the precipitation amount and distribution throughout the year.

Table 9.4 shows the mean and standard deviations for some of the main SSM/I channel combinations required for characterizing dense agricultural and rangeland vegetation.

TABLE 9.4   MEAN AND STANDARD DEVIATIONS FOR BRIGHTNESS TEMPERATURE
COMBINATIONS OVER DENSE CROPLAND/RANGELAND VEGETATION
COVER IN DIFFERENT AREAS OF THE WORLD

Combination		Central Plains		Cerrado Region		African Savannas	
		Mean (K)	SD (K)	Mean (K)	SD (K)	Mean (K)	SD (K)
22V - 19V	[a]	-1.12	1.24	-0.12	1.03	-2.49	0.98
$\frac{(19V+37V)}{2} - \frac{(19H+37H)}{2}$	[b]	3.20	0.63	3.57	0.37	2.84	0.7
37V - 19V	[c]	-3.54	0.83	-4.22	0.69	-4.42	0.85
85V - 37V	[d]	1.52	0.96	2.32	0.75	2.08	0.71
85H - 37H	[e]	2.32	1.03	3.20	0.52	2.93	0.77

The dense agricultural and rangeland vegetation can be classified using the following rule:

$$22V - 19V <= 4.0 \qquad [a]$$
$$1.9 < ((19V + 37V))/2.0 - ((19H + 37H)/2.0) <= 4.0 \qquad [b]$$
$$85V - 37V >= -1.0 \qquad [d]$$
$$85H - 37H < 4.5 \qquad [e]$$
$$19V > 262.0 \qquad [g]$$

The 4 K upper threshold for the average polarization in the 19.35 GHz and 37.0 GHz channels, though slightly lower than the mean plus twice the standard deviation limit for that distribution (approximately 4.2 K), is the polarization above which sensitivity to surface moisture begins to occur (see Section 9.3).

9.1.3.3   Soil Rules

Passive microwave emission from a water surface is highly polarized, with an emissivity of about 0.4 for the 19.35 GHz horizontally polarized channel. The emission from bare soil is also polarized, but to a lesser extent, with higher emissivities (typically 0.9 and above in the horizontal channels for a dry surface). The influence of water in an essentially bare soil is to depress the brightness temperatures and to increase the polarization difference. If vegetation is present, the vegetative scattering decreases the polarization difference. Therefore, the soil rules were developed to classify a dynamic combination of bare soil, water in or on the soil

surface, and different degrees of vegetation cover. The natural vegetation cover varies as a function of season and the soil water content. The water present on the soil surface varies as a function of the rainfall and the hydrologic response. The net effect is a broad range of brightness temperatures and polarization differences within this class. These ranges are a function also of frequency due to the variation of the real component of the dielectric constant with frequency. The dielectric constant of water is higher at longer wavelengths (lower frequencies). The depth of the emitting layer is also greater at the longer wavelengths.

An SSM/I footprint with decreasing vegetation cover is characterized by average polarization differences at 19.35 GHz and 37 GHz ranging from 4 K to 19 K, with polarization increasing as more soil is "radiometrically" visible. This range was divided into two large sub-groups. The arable soil classification with average polarization differences from 4 K to 9.8 K (using 19.35 and 37.0 GHz channels) and the semi-arid classification with average polarization differences ranging from 9.8 to 19.0 K. This was done because most SSM/I footprints in the latter group were identified from MLRA regions found in the western United States (Arizona, Nevada, Utah, and California) in which a semi-arid, desert climate is predominant, and vegetation is sparse.

During the development of the surface moisture retrieval algorithms it was determined that the two large polarization sub-groups mentioned above needed to be further broken down according to vegetation density. This was due to the effect of vegetation density on the sensitivity to surface soil moisture and the need for different retrieval equations according to this sensitivity. This will be further explained below under the moist soil rule.

9.1.3.3.1   Dry Arable Soil

The average 19.35 and 37.0 GHz brightness temperature polarization differences are shown in Figure 9.2 for the Central Plains of the United States under the arable soil heading. Statistics are shown in Table 9.5 for the summer and winter season separately. Footprints influenced by rain, snow, water or dense vegetation cover were removed. The larger influence of vegetation during the summer season due to natural vegetation and crop cover is evident by the lower mean of combination [b]. During the winter, the soil is mostly bare which results in significantly larger polarization differences in the 19 GHz and 37 GHz channels. The upper and lower limits contained in Table 9.5 for both seasons were used to define the range of polarization for arable soil. Thus, the classification rule for dry arable soil is:

$$22V - 19V <= 4.0 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad [a]$$
$$4.0 < ((19V + 37V)/2.0) - ((19H + 37H)/2.0) <= 9.8 \qquad\quad [b]$$
$$37V - 19V >= -6.5 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad [c]$$
$$-5.0 <= (85V - 37V) < 0.5 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad [d]$$
$$85H - 37H < 4.2 \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad [e]$$

Conditions [a] and [e] check for flooded conditions or large water bodies and will be discussed later. Condition [b] classifies the area in terms of brightness temperature polarization

differences, condition [c] is a snow flag which will be discussed later, and condition [d] is a rain and surface moisture flag which also will be discussed later.

TABLE 9.5 BRIGHTNESS TEMPERATURE COMBINATION VALUES FOR THE CENTRAL PLAINS STATES OF THE U.S. CORRESPONDING TO THE ARABLE SOIL CLASSIFICATION

Box size: NW corner: 50°N 105°W   SE corner: 32°N 95°W

SUMMER SEASON

Combination		Mean (K)	SD (K)	Lower Limit (K)	Upper Limit (K)	Calendar Date and Node
19V - 22V	[a]	-1.90	1.31	-4.52	0.72	
$\dfrac{(19V+37V)}{2}$ - $\dfrac{(19H+37H)}{2}$	[b]	5.72	1.25	3.22	8.22	
37V - 19V	[c]	-3.59	0.94	-5.47	-1.71	(1987) 211A 222A
85V - 37V	[d]	-1.17	1.15	-3.47	1.13	235A 244A
85H - 37H	[e]	1.30	1.61	-1.92	4.52	253A

WINTER SEASON

Combination		Mean (K)	SD (K)	Lower Limit (K)	Upper Limit (K)	Calendar Date and Node
22V - 19V	[a]	-1.47	0.94	-3.35	0.41	
$\dfrac{(19V+37V)}{2}$ - $\dfrac{(19H+37H)}{2}$	[b]	7.34	1.27	4.80	9.88	
37V - 19V	[c]	-4.06	0.99	-6.04	-2.08	(1988)
85V - 37V	[d]	-2.17	1.25	-4.67	0.33	55A 59A
85H - 37H	[e]	-0.87	1.79	-4.45	2.71	

9.1.3.3.2   Semi-Arid Rule

The semi-arid classification corresponds to areas where natural vegetation is sparse and of a desertic type. A typical example of this type of environment is MLRA region #30 [2], the Sonoran Basin and Range. Most of this area is government owned and consists of thin stands of desert vegetation, mostly Bursage, Joshua tree, juniper, yucca, and cactus. Grasses grow only in years with favorable moisture conditions. The histogram of the brightness temperature polarization difference distribution is shown in Figure 9.2. Table 9.6 contains the statistics for the main channel combinations.

Polarization Dependence of Different
Surface Types
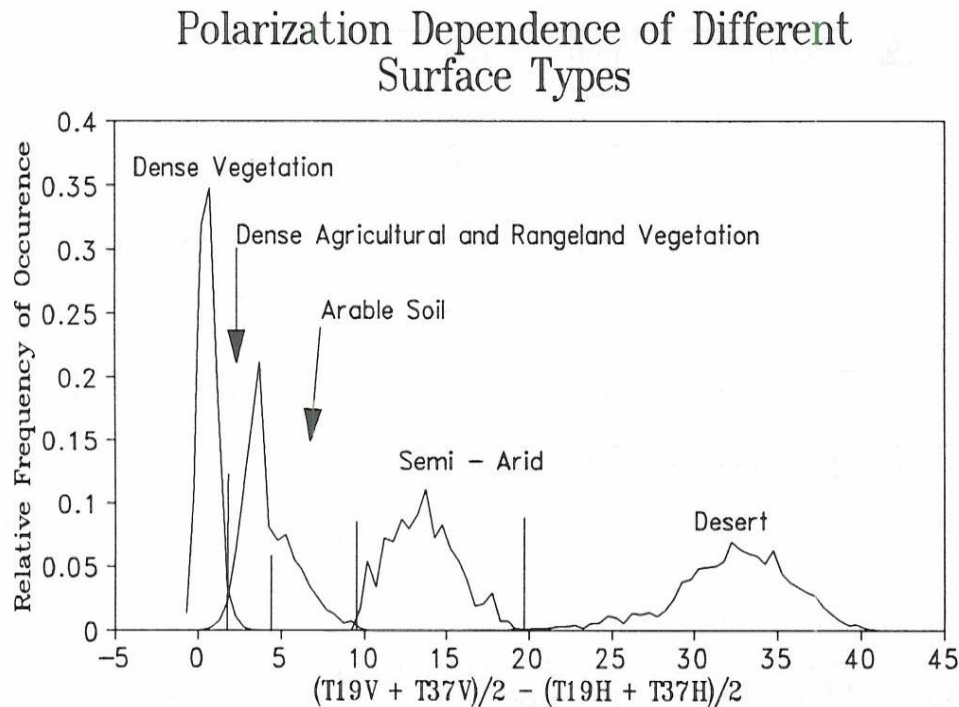
Figure 9.2 Polarization dependence in the 19.35 GHz and 37.0 GHz channels to different surface types.

Based on the estimated upper and lower limits, the threshold values for the channel combinations which best classify semi-arid conditions are:

$$22V - 19V <= 4 \qquad [a]$$
$$9.8 < ((19V + 37V)/2.0) - ((19H + 37H)/2.0) < 19.7 \qquad [b]$$
$$85V - 37V < 0.5 \qquad [d]$$
$$85H - 37H < 6.0 \qquad [e]$$
$$37H - 19H < -1.8 \qquad [j]$$

9-13

TABLE 9.6   STATISTICAL ANALYSIS RESULTS FOR SELECTED BRIGHTNESS
TEMPERATURE COMBINATIONS OVER SEMI-ARID AREAS, MLRA
REGION #30, BASED ON DATA FROM 19 ORBITS

Combination		Mean (K)	SD (K)	Lower Limit[1] (K)	Upper Limit[1] (K)
22V - 19V	[a]	-2.08	3.60	-9.28	5.12
$\frac{(19V+37V)}{2}$ - $\frac{(19H+37H)}{2}$	[b]	13.61	2.02	9.57	17.65
37V - 19V	[c]	-8.00	1.30	-10.60	-5.40
85V - 37V	[d]	-3.17	2.18	-7.53	1.19
85H - 37H	[e]	2.03	2.35	-2.67	6.73
37H - 19H	[j]	-4.04	1.15	-6.34	-1.74

[1]Limits are 2 standard deviations from the mean.

Condition [a] is the check for large water bodies and flooded conditions. The lower limit
for the polarization difference (combination [b]) was 9.57 K while the upper limit for arable soil
(Table 9.5) was 9.88 K during the winter season. An intermediate value of 9.8 K was used as
the dividing threshold between the dry arable soil and semi-arid classes. Large negative values
can sometimes be observed in the vertical polarization channel differences of combinations [c]
and [d]. This is the result of a surface scattering phenomena caused by smooth bare soil which
could be confused with atmospheric scattering or scattering due to snow cover. The upper limit
of combinations [d] and [e] are the thresholds between dry and moist soil and will be discussed
later. Condition [j] is also a moisture flag which differentiates dry soil surfaces from wet snow
surfaces.

9.1.3.3.3   Desert Rule

Deserts are characterized by very large brightness temperature polarization differences
in all channels. The distribution histogram of average polarization differences in the 19.35 GHz
and 37 GHz channels is shown in Figure 9.2. The statistics for several orbits over the control
area are shown in Table 9.7. Polarization differences in the 19.35 GHz channel were, in some

cases, above 40 K, and in the upper 30s (K) for the 37.0 GHz channel. These extreme polarization differences are caused by very smooth, sandy surfaces in the Sahara desert and the total absence of vegetation cover. The relatively high dielectric constant of quartz, the dominant component of desert sand, undoubtedly contributes.

TABLE 9.7   STATISTICS FOR SELECTED BRIGHTNESS TEMPERATURE
COMBINATIONS OVER THE SAHARA DESERT CONTROL AREA

Combination		Mean (K)	SD (K)	Lower Limit[1] (K)	Upper Limit[1] (K)	Calendar Date of Overpass
22V - 19V	[a]	-3.15	1.14	-5.43	-0.87	
$\frac{(19V+37V)}{2} - \frac{(19H+37H)}{2}$	[b]	32.24	3.46	25.48	39.32	233A
37V - 19V	[c]	-8.49	1.22	-10.93	-6.05	328A
85V - 37V	[d]	-9.26	2.24	-13.74	-4.78	176D/178D 224D/232D
85H - 37H	[e]	5.77	2.33	1.11	10.43	233D

[1]Limits are 2 standard deviations from the mean.

The classification rule is:

$$22V - 19V <= 2.0 \qquad \text{[a]}$$
$$((19V + 37V)/2.0) - ((19H + 37H)/2.0) >= 19.7 \qquad \text{[b]}$$
$$85H - 37H > -1.0 \qquad \text{[e]}$$
$$19V > 268 \qquad \text{[g]}$$

Condition [b] is the primary discriminator for deserts with 19.7 K being the upper limit for the semi-desertic regions considering three standard deviations from the mean (Table 9.6). Brightness temperatures in the vertical polarization channels decreased with increasing frequency, with large negative values occurring for combinations [c] and [d] in Table 9.7. These large negative values could be confused with scattering due to heavy rain or snow cover. For this reason, combination [e] is used as an additional check. If 85H - 37H is greater than -1 K, the decrease in brightness temperature in the vertical polarization channels is due to surface phenomena and not atmospheric scattering. Combination [g] also ensures a snow free surface.

9.1.3.4   Classification of Surface Water and Soil Moisture

The short wavelengths of the SSM/I sensor are not suited for soil moisture retrievals due to their small penetration depth in soils and consequently small moisture sensing depth. In addition, there is a considerable loss of sensitivity to surface moisture due to vegetation cover. However, under sparse or incomplete vegetation cover, an assessment can be made of the quantity of water retained on the surface after a heavy rainfall event as well as moisture in the immediate soil surface layer down to a few millimeters. In Section 9.3, a specific quantification of this surface moisture is conducted using an Antecedent Precipitation Index (API) as a surrogate variable.

The main SSM/I channels used for surface moisture retrievals are the 19.35 H GHz and the 37.0 V GHz channels in the form of a normalized brightness temperature T19H/T37V. However, the 85.5 GHz channels have turned out to be excellent for identifying the presence of water bodies within the SSM/I footprints. As the proportion of moist soil and surface water within an SSM/I 3 db footprint increases, the emissivity of the surface layer decreases resulting in lower brightness temperatures. Relative changes are first observed between the 85.5 GHz and the 37.0 GHz channels in both polarizations: the T_b's decrease in both channels but to a greater extent at 37.0 GHz due to the fact that both the permittivity and the dielectric loss factor of water are smaller at 85.5 GHz than at 37.0 GHz [4]. It is important to note that these relative changes in T_b's between the two channels are occurring because the resolution of the 85.5 GHz channels (approximately 14 km) was decreased to that of the 37.0 GHz channels (approximately 33 km) as a result of the averaging scheme. In this way, both channels were sensing approximately the same proportions of water, soil and vegetation in the concentric footprint scenes.

Moist soil surfaces and footprints containing larger water bodies therefore are differentiated from dry surfaces with the 85.5 V - 37.0 V and 85.5 H - 37.0 H channel combinations. Several classification rules were developed to identify surface moisture and surface water bodies (flooded soil, moist soil surface, composite water and soil or wet soil surface, composite water and vegetation). The classification of footprints containing water bodies such as reservoirs, lakes etc. and their removal from the parameter retrieval algorithm regression data sets, decreases the introduced noise and increases the retrieval accuracy of parameters such as land surface temperature over soil and vegetation, assuming the same classification scheme is used operationally. This is because the brightness temperature of a footprint containing a water body would not be lower due to the lower physical temperature of the soil or vegetation but as a result of the contamination by a surface with completely different microwave emission properties.

9.1.3.4.1   Moist Soil Surface

Moist soil surfaces are differentiated from dry arable soils using a threshold value of 0.5 K for combination [d]. This value was approximately the upper limit for this combination under dry arable soil conditions (Table 9.5). In order to differentiate moist soil surfaces from very

wet soil surfaces or footprints containing larger water bodies, an upper threshold value of 4.0 K is used for combination [d] and 4.2 K for combination [e], 85H - 37H. The rule is:

$$22V - 19V <= 4.0 \qquad\qquad [a]$$
$$4.0 < ((19V + 37V)/2.0) - ((19H + 37H)/2.0) < 19.7 \qquad [b]$$
$$37V - 19V >= -6.5 \qquad\qquad [c]$$
$$0.5 <= (85V - 37V) < 4.0 \qquad\qquad [d]$$
$$85H - 37H < 4.2 \qquad\qquad [e]$$

where combination [c] is a snow identifier.

The moist soil surface rule was tested along with other moisture sensing rules by stratifying 0.25 degrees latitude/longitude grid cells according to API_{15} (based on available water for evaporation of 15 mm) values between 0 and 10 mm and greater than 10 mm as well as the number of days since the last precipitation event in each of those classes. These variables are defined and explained in the methodology of Section 9.3 of this report. Table 9.8 summarizes the results.

As expected, 90.5 % of the dry soil classifications (DS) had API_{15} values of less than 10 mm in those grid cells with an average of 2.7 mm. On the other hand 9.5 % of the dry soil classifications had API_{15} values greater than 10 mm with an average value of 16.8 mm. Most of these cases were probably due to localized precipitation events that did not entirely wet the soil surface of the large SSM/I footprint.

Moist soil surface classifications (MS) occurred for API_{15} values less than 10 mm, 83 % of the time with an average value of 3.2 mm. The remaining 17 % of the values above 10 mm had an average API_{15} value of 16.4 mm.

These results are similar to the dry soil classification rule, indicating that spatial distribution of moisture is obviously a factor. In addition, this rule will also sometimes classify footprints which contain small bodies of water such as small lakes and reservoirs as moist soil.

9.1.3.4.2   Composite Soil and Water or Wet Soil

The development of this rule became necessary to identify footprints with locally flooded soil, lakes, large rivers, and other surface waters. These land footprints with water as a component of the land surface would have a passive microwave signature that is a combination of land and water. Because water has a much lower emissivity and a much higher polarization difference than other land surfaces, the resulting brightness temperatures would be very difficult to interpret in terms of physical surface temperature. This rule was developed to exclude footprints with a water component in the signature from the calibration/validation regression database, thus increasing the accuracy of the algorithms.

Several dozen cases were examined over the course of our calibration/validation effort. Comparisons were conducted between brightness temperatures from footprints on the border of large lakes or which contained small water bodies, with non-contaminated surrounding footprints. Surface physical temperatures were compared as well when available. Possible SSM/I footprint geolocation errors were also taken into consideration in this analysis. Both 85.5 GHz channels were sensitive to the presence of water in the footprint, especially the 85.5 H channel. The 85.5 H - 37.0 H T_B difference is a small positive or negative number for a mixed soil and vegetation scene (Tables 9.3, 9.4 and 9.5). With a certain proportion of water, the emissivity is lowered in both channels, but to a greater extent at 37.0 H GHz due to the higher dielectric constant of water at 37.0 GHz. A threshold value of 4.2 K was determined for this channel combination. Over deserts, values greater than 4.2 K have been commonly observed (Table 9.7). Therefore the 85.5 V - 37.0 V combination is checked as well to prevent misclassifications.

$$22V - 19V <= 4.0 \qquad\qquad [a]$$
$$6.4 <= (19V + 37V)/2.0 - (19H + 37H)/2.0 \qquad\qquad [b]$$
$$37V - 19V >= -6.5 \qquad\qquad [c]$$
$$85V - 37V >= 0.5 \qquad\qquad [d]$$
$$85H - 37H >= 4.2 \qquad\qquad [e]$$

Combination [a] is the check for flooded surfaces. Condition [b] identifies large polarization differences due to water in the 3 dB footprint. To differentiate between large polarization differences due to water and those associated with barren deserts, condition [d] is applied.

Footprints classified by this rule, were tested against ground truth API_{15} values gridded at 0.25 degree latitude/longitude cells for the central plains area of the U.S. during 1987. Fifty orbits were included in the analysis. The results are also shown in Table 9.8. and indicate that footprints with high API_{15} values are also classified by this rule. The results were more evenly distributed with 58 % of the grid cells having API_{15} values less than 10 mm (average of 4.0 mm) and 42 % having values greater than 10 mm (average of 21.3 mm). It is probable that most of the grid cells with moisture values less than 10 mm were a result of contamination by water bodies while for API_{15} values greater than 10 mm, most of the classifications resulted from a wet soil surface.

Footprints with a wet soil surface have a similar microwave signature to dry soil footprints contaminated by large water bodies. This presents a problem for the use of the surface moisture retrieval algorithms which should be applied to retrieve moisture when the API_{15} is high but not to the latter case. The solution is to maintain previous surface classifications over a geographic location in a dynamic database and use additional logic to differentiate between these cases. In Section 9.3 of this report, a dynamic database scheme is proposed to work in conjunction with the surface moisture classification rules above.

TABLE 9.8   STRATIFICATION OF CLASSIFIED 0.25 DEGREE GRID CELLS
ACCORDING TO SURFACE MOISTURE VALUES (API$_{15}$) AND NUMBER
OF DAYS SINCE LAST PRECIPITATION EVENT

Classification Rule	API$_{15}$ < 10 mm Number of days since last precipitation event						API$_{15}$ > 10 mm					
	1	2	3	4	5	>5	1	2	3	4	5	>5
DS NO:	251	248	313	392	297	868	68	82	55	32	11	2
API$_{15}$:	4.1	4.2	4.0	3.3	2.7	1.1	18.1	18.8	15.0	13.6	12.4	11
MS NO:	287	313	367	373	302	997	138	134	115	77	45	20
API$_{15}$:	4.7	4.6	4.5	4.4	3.5	1.2	18.6	16.1	16.1	14.8	13.4	17.6
WS NO:	462	342	317	286	180	625	707	429	266	130	38	14
API$_{15}$:	5.4	5.3	5.1	4.4	4.1	1.4	24	20.6	18	17.2	17.3	16.6
WV NO:	549	430	403	299	271	1328	703	376	236	128	36	18
API$_{15}$:	5.3	5.3	4.6	4.1	3.7	1.1	25.1	22.3	19.5	19.6	19.3	22.2

DS = Dry Arable Soil
MS = Moist Soil Surface
WS = Wet Soil Surface or composite soil and water
WV = Vegetation with wet soil background or composite vegetation and water
NO: Number of occurences

9.1.3.4.3   Composite Dense Vegetation and Water

This rule classifies footprints with mixed dense vegetation and water. It is similar to the composite soil and water rule, but with different threshold values. Dense vegetation has a strong unpolarized signature with usually warm brightness temperatures. On the other hand, water has a low emissivity, thus colder brightness temperatures, and a highly polarized signature. Depending on the proportions of water and vegetation as well as the density of the vegetation, the average polarization in the 19.35 GHz and 37.0 GHz channels (combination [b]) will vary. By observing numerous cases the upper threshold value of 6.4 K was determined, allowing for greater polarizations induced by water in the footprint scenes. The rule is:

$$22V - 19V <= 4.0 \qquad [a]$$
$$((19V + 37V)/2.0) - ((19H + 37H)/2.0) < 6.4 \qquad [b]$$
$$85V - 37V >= -1.0 \qquad [d]$$
$$85H - 37H >= 4.5 \qquad [e]$$
$$37V > 257 \qquad [h]$$

The threshold value for combination [e] was based on observations of vegetation/river footprints in the Amazon jungle and is approximately the upper limit obtained for this combination over the dense vegetation control areas (Table 9.3). Combination [d] is the precipitation flag and condition [h] is a snow flag.

This rule was also tested along with the other moisture sensing rules in Table 9.8. Results indicate that 69 % of the grid cells were classified as such, having an API_{15} value of less than 10 mm (average 3.3 mm). The remaining 31 % of the grid cells resulted in an average API_{15} value of 22.9 mm, indicating that the rule is also sensing vegetation with a wet soil background.

9.1.3.4.4   Flooded Soil

Large amounts of water on the soil surface due to a heavy precipitation event, flooding due to heavy rain or melting snow or the presence of large natural lakes and reservoirs, will lower the brightness temperatures at all frequencies due to the high permittivity of water. Brightness temperatures at 22.235 V GHz will be greater than at 19.35 V GHz because the microwave emissivity of water increases with frequency and both channels have approximately the same 3 db footprint size. In addition, the 22.235 GHz channel is sensitive to water vapor. A threshold value of 4 K was determined for the difference between the 22.235 V GHz and the 19.35 V GHz brightness temperatures based on observations of large lakes and reservoirs and areas flooded by large precipitation events. This condition [a] is checked within all classification rules. If the surface is classified as flooded, no parameter retrieval algorithms are applied.

9.1.3.5   Classification of Precipitation Events

9.1.3.5.1   Rain Over Vegetation Rule

Precipitating or convective type clouds within an SSM/I footprint over vegetation will have a drastic effect on the 85.5 GHz brightness temperatures. Clouds containing large water droplets and/or ice will scatter radiation at smaller wavelengths resulting in lower brightness temperatures at 85.5 GHz than at the smaller frequency (longer wavelength) channels. This is especially true over warm, dense tropical vegetation.

Numerous storms were identified through SSM/I data and confirmed by visual analysis of GOES imagery and/or by checking NOAA precipitation charts over the United States. The microwave signature of a large thunderstorm over the Amazon jungle is shown in Table 9.9. The brightness temperature combination data for clear conditions on calendar days 180 and 231 were similar to the expected microwave signature over dense vegetation shown in Table 9.3. The polarization difference (combination [b]) was higher because footprints which had other classifications, i.e footprints containing surface water, but which fell within the selected area, were included.

TABLE 9.9 EFFECT OF LARGE PRECIPITATING STORM CLOUDS OVER DENSE
VEGETATION ON SELECTED BRIGHTNESS TEMPERATURE
COMBINATIONS

Amazon Jungle, South America. Box boundaries:
NW corner: 6.5°S 59°W   SW corner: 8°S 57°W

Combination		CD180 Asc. Mean (K)	SD (K)	CD222 Asc. Mean (K)	SD (K)	CD231 Asc. Mean (K)	SD (K)
$\frac{(19V+37V)}{2} - \frac{(19H+37H)}{2}$	[b]	1.17	1.02	1.26	0.87	1.46	0.96
37V - 19V	[c]	-3.71	0.69	-8.70	5.81	-3.93	0.76
85V - 37V	[d]	5.82	1.17	-15.39	18.40	3.87	1.02
85H - 37H	[e]	6.02	1.68	-17.33	20.15	4.69	1.46

With the presence of storm clouds on day 222, the temperatures in the 85.5 GHz
channels were depressed below the brightness temperatures in the 37.0 GHz channel. The
hydrometeors were also sufficient in size and quantity to scatter microwave radiation at the
longer wavelengths of the 37.0 GHz channels, as indicated by the decrease in the mean value
of the 37.0 V GHz - 19.35 V GHz brightness temperature difference (combination [c]). The
non-uniform nature of the precipitation within the selected area can be seen by the very large
increase in the standard deviation for combinations [c], [d] and [e]. This can be visualized in
Figure 9.3 where the 85.5 V GHz - 37.0 V GHz distribution histogram is plotted for the selected
area, for the overpasses with and without precipitation. For vegetated surfaces, a threshold
value of -1 K was determined for combination [d] as a precipitation flag. This value is
approximately 2.5 standard deviations from the mean value over dense vegetation areas with
no precipitation shown in Tables 9.3 and 9.4. The rule is:

22V - 19V < = 4.0   [a]
((19V + 37V)/2.0) - ((19H + 37H)/2.0) < = 4.0   [b]
85V - 37V < -1.0   [d]
19V > 268.0   [g]

Dense Vegetation — Amazon Jungle
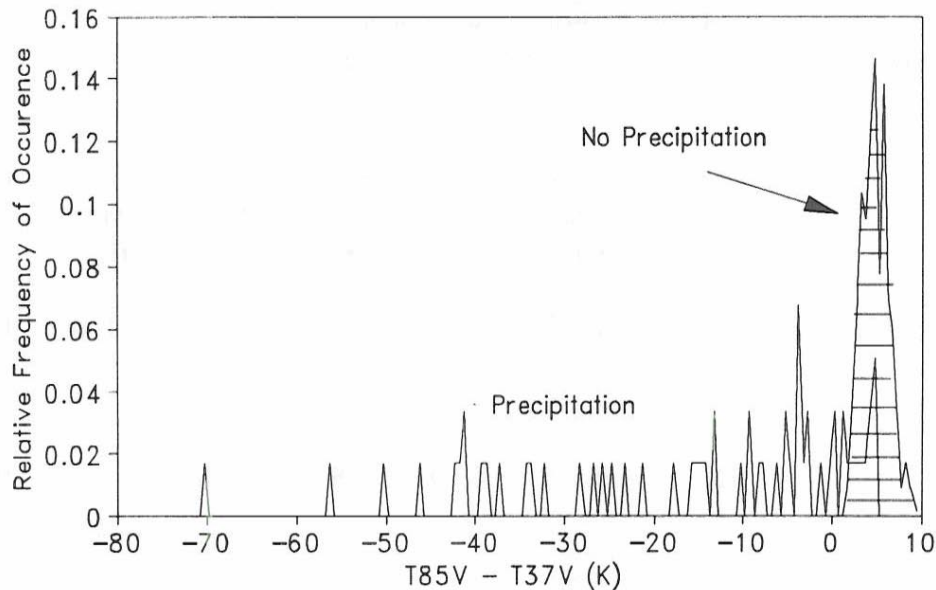
Figure 9.3   Effect of heavy precipitation on the 85V - 37V brightness temperature difference over dense vegetation.

9.1.3.5.2   Rain Over Soil Rule

The detection of precipitation over soil is similar to the detection of precipitation over vegetation. Scattering by hydrometeors in the atmosphere decreases the brightness temperatures in the 85.5 GHz channels more than in the other SSM/I frequencies. However, the background microwave emission by soil is polarized and the relative brightness temperature differences among frequencies are different. Several storms were identified using SSM/I data over the United States test regions. The storms were checked against National Weather Service radar charts, when available, to confirm the locations of thunderstorm cells and occurrence of precipitation. An example is shown in Table 9.10 for a squall line occurring over Oklahoma and Texas on Day 228, 1987. Combination values for Day 227 are typical average signatures for pooled footprints containing mostly dense rangeland and agricultural vegetation and dry arable soil classifications (Tables 9.4 and 9.5). The thunderstorm activity on Day 228 resulted

in considerable scattering in the 85.5 GHz channels (combinations [d] and [e]) and some scattering in the 37.0 GHz channel (combination [c]).

Threshold values for flagging precipitation with combinations [c], [d] and [e] were determined based on the lower limits for these combinations after pooling the data shown in Table 9.5 (dry arable soil) and the study of numerous confirmed precipitation cells over the central plains of the U.S.. The rule is:

$$22V - 19V <= 4.0 \qquad \qquad \text{[a]}$$
$$((19V + 37V)/2.0) - ((19H + 37H)/2.0) > 4.0 \qquad \text{[b]}$$
$$37V - 19V < -3.0 \qquad \qquad \text{[c]}$$
$$85V - 37V < -5.0 \qquad \qquad \text{[d]}$$
$$85H - 37H < -4.1 \qquad \qquad \text{[e]}$$
$$19V > 268.0 \qquad \qquad \text{[g]}$$

Combination [g] is a snow flag, to differentiate the signature caused by hydrometeors, from scattering caused by snow which also depresses the brightness temperatures of the shorter wavelength (higher frequency) SSM/I channels.

TABLE 9.10 EFFECT OF A SQUALL LINE ON BRIGHTNESS TEMPERATURE COMBINATION VALUES OVER MOSTLY ARABLE SOIL FOOTPRINTS IN THE CENTRAL PLAINS OF THE UNITED STATES

Central plains of United States. Approximate box size:
NW corner: 37°N 102°W SE corner: 32°N 95°W

Combination		CD228 Desc. With Precipitation Mean (K)	SD (K)	CD227 Desc. No Precipitation Mean (K)	SD (K)
$\dfrac{(19V+37V)}{2} - \dfrac{(19H+37H)}{2}$	[b]	6.28	2.05	5.45	1.43
37V - 19V	[c]	-9.39	4.74	-3.92	0.99
85V - 37V	[d]	-21.71	13.69	0.70	0.99
85H - 37H	[e]	-21.06	15.03	2.49	0.86

9.1.3.6   Classification of Snow Covered Surfaces

Microwave emissions from snow covered surfaces depend on several factors. These include: (1) the underlying surface type, (2) the moisture content of the underlying soil and if the water is frozen or in liquid form, (3) the depth of the snowpack, (4) the density of the snowpack, (5) the shape and size of the snow crystals and, (6) the liquid water content of the snowpack. Thus, the classification of snow is complicated as the microwave signature from a snowpack with constant depth can vary with snow morphology, snow ripeness and cycles of melting and re-freezing under spring weather conditions. Therefore, the interpretation of microwave signatures from a snow covered surface at any point in time would benefit from the history of previous weather and snowpack conditions.

The characterization of snow signatures and their relationships with parameters such as snow wetness, snow depth and water equivalent has been studied by many authors such as [5], [6], and [7]. Other research concerning snow microwave properties has also been described by [4].

Specific research on snow classification has been done by Kunzi et al. [8] in the development of snow extent, snow depth and water equivalent algorithms for SSMR. Schanda et al. [9] proposed a snow classification scheme based on several years of observations which included classes such as winter snow, wet spring snow and dry, refrozen spring snow. McFarland et al. [10] investigated snowpack properties using SMMR brightness temperatures and were able to detect dry snow accumulation, and snow melting and refreezing processes.

9.1.3.6.1   Dry Snow

The normal dry snow microwave signature is the depression of brightness temperatures in the 37.0 GHz channels with respect to the 19.35 GHz channels due to volume scattering. At 37.0 GHz, scattering is the main component of the total extinction loss of the medium [4].

Channel combination data for footprints containing dry snow over the northern plains of the U.S. during a few days in February, 1988 are shown in Table 9.11. The ground truth snow depth values were obtained from the NOAA cooperative network of weather stations in the central plains states of the U.S. Average daily snow depth values and corresponding SSM/I brightness temperature data were gridded to 0.25 degree latitude/longitude cells for analysis. The developed rule for dry snow can be written as:

$$22V - 19V <= 4.0 \qquad\qquad [a]$$
$$((19V + 37V)/2.0) - ((19H + 37H)/2.0 > 4.0 \qquad\qquad [b]$$
$$37V - 19V < -6.5 \qquad\qquad [c]$$
$$19V - 19H >= 5.0 \qquad\qquad [i]$$
$$225 < 37V <= 257.0 \qquad\qquad [h]$$

Conditions [c], [i] and [h], together differentiate snow from cold bare soil situations as well as large precipitating thunderstorm clouds. The threshold for combination [c] is approximately the lower limit for this brightness temperature difference over dry arable soil with no snow. When snow is present, brightness temperatures in the 19.35 GHz channels also decrease, partly due to scattering and partly due to the decreased physical temperature of the snow and underlying soil.

Results in Table 9.11 also show the greater variability in brightness temperatures (reflected by the larger standard deviations) caused by the spatial distribution of snow at different depths.

TABLE 9.11 SELECTED BRIGHTNESS TEMPERATURE COMBINATIONS OVER DRY SNOW IN THE CENTRAL PLAINS OF THE UNITED STATES

Approximate box size:
NW corner: 49°N 105°W   SE corner: 45°N 100°W

Combination		SNOW Mean (K)	SD (K)	NO SNOW Mean (K)	SD (K)	Calendar Date of Overpass
22V - 19V	[a]	-2.62	1.04	-0.95	1.06	
$\frac{(19V+37V)}{2} - \frac{(19H+37H)}{2}$	[b]	12.02	2.33	10.09	1.41	51A 55A
37V - 19V	[c]	-11.55	4.10	-2.24	1.21	
85V - 37V	[d]	-8.36	5.92	-0.86	1.34	55D (1988)
85H - 37H	[e]	-4.48	5.72	1.42	1.37	

9.1.3.6.2   Wet Snow and Refrozen Snow

The classification of wet snow or melted snow containing water in liquid form as well as refrozen snowpacks requires the use of the dynamic database scheme, as their microwave signature could be confused with other surfaces. A small amount of liquid water (volumetric water content of 0.01) will increase the volume absorption coefficient of the medium to a value greater than the scattering coefficient, thus reducing the scattering albedo to a very small value [4]. For higher volumetric water contents, scattering is practically non-existent and the

snowpack begins to behave like a blackbody radiator. Such conditions are normally encountered in the spring, when the snowpack undergoes successive cycles of thawing and refreezing. During the day, when temperatures are above the freezing point, the top layers of the snowpack will partially melt, increasing the volumetric water content of the snow. As a result, the microwave brightness temperatures at 37.0 H GHz will increase with respect to the T_B at 19.35 H GHz. This is shown in Figure 9.4 for a 10 day sequence of SSM/I data over north-central Nebraska during February, 1988. The last snowfall over that region occurred on Day 50-51 with 50 to 75 mm of new snow being reported by the weather stations in the area. Average snow depths changed throughout the period from 254 mm on day 51, to 55 mm on day 59 (Figure 9.5). Maximum and minimum air temperatures are also shown in Figure 9.4, corresponding to descending and ascending overpasses respectively. The first available SSM/I data after the snowfall is for the ascending overpass on day 51. The T_B difference of approximately 20 K between the 19.35 H and 37.0 H GHz channels as well as the low brightness temperatures in both channels are an indication of dry snow and were classified as such with the dry snow rule. The signature for the ascending overpass on day 53 is similar to that of day 51 with slightly cooler temperatures. On both days, the minimum air temperature, which probably occurred a few hours prior to the overpass, was below the freezing point.

The descending overpass on day 53, showed a marked increase in the 37.0 H GHz brightness temperature to within 1 K of the 19.35 H GHz channel as a result of a wet snow surface layer. Thawing at the snow surface occurred during the day due to warm air temperatures (the maximum air temperature was 7° C). Data from the ascending overpass on day 55 shows a drastic decrease in brightness temperatures in both horizontally polarized channels. The minimum temperature reached approximately -14 °C, sufficient to re-freeze any liquid water in the snowpack. The thawing and refreezing process increases the size and changes the shape of the ice crystals, which tend to become spherical as the snowpack ripens and undergoes several of these cycles. The increase in particle size will increase the scattering albedo and decrease the polarization dependence causing additional scattering at longer wavelengths and lowering the T_B's.

Figure 9.6 shows brightness temperatures in the horizontal and vertical polarizations for the 19.35 and 37.0 GHz channels for the same period. Brightness temperatures for the ascending overpass on day 55 were lower than those of day 53 for both frequencies and polarizations while the polarization difference was smaller, indicating a refrozen snow surface layer. Subsequent overpasses beginning with the descending overpass on day 56 indicate a cycle of thawing during the day with re-freezing overnight. During this period, the variability in the 19.35 H - 37.0 H GHz T_B difference for the descending overpasses was probably due to different volumetric water content in the snowpack. In addition, the snow depth continuously decreased throughout the period (Figure 9.5), with the snowpack depth on day 59 being less than half its depth on day 53.

These variable microwave snow signatures are difficult to accurately interpret with stand alone independent rules. A dynamic database scheme should be implemented, if accuracy is desired, in order to differentiate dry snow from re-frozen snow and wet snow signatures as the

Dry Snow–Wet Snow–Refrozen Snow Cycle
Nebraska (Lat:42.5N Long:99.0W)
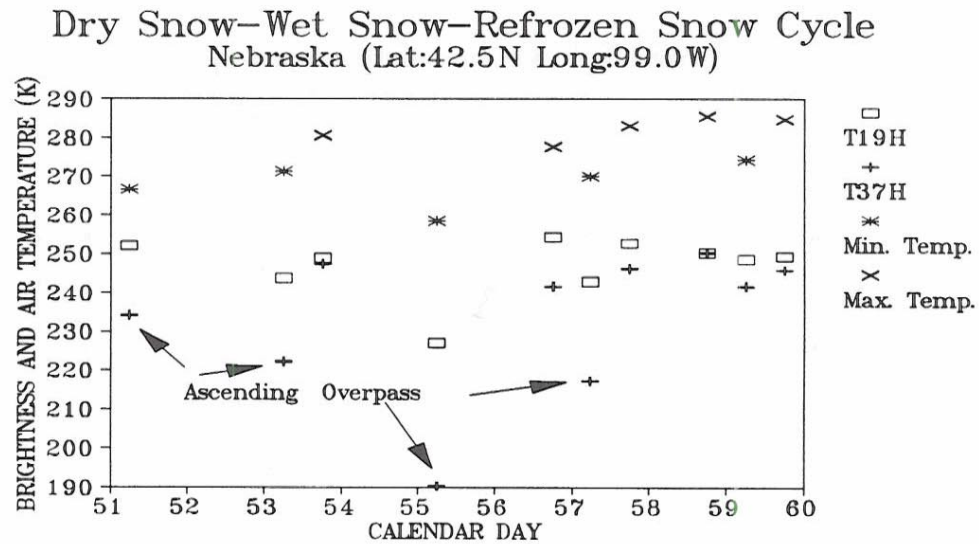
Figure 9.4 Variation in horizontally polarized brightness temperatures in the 19.35 GHz and 37.0 GHz channels throughout a ten day period in February, 1988.

Snow Depth Variation
Nebraska (Lat:42.5N Long:99.0W)
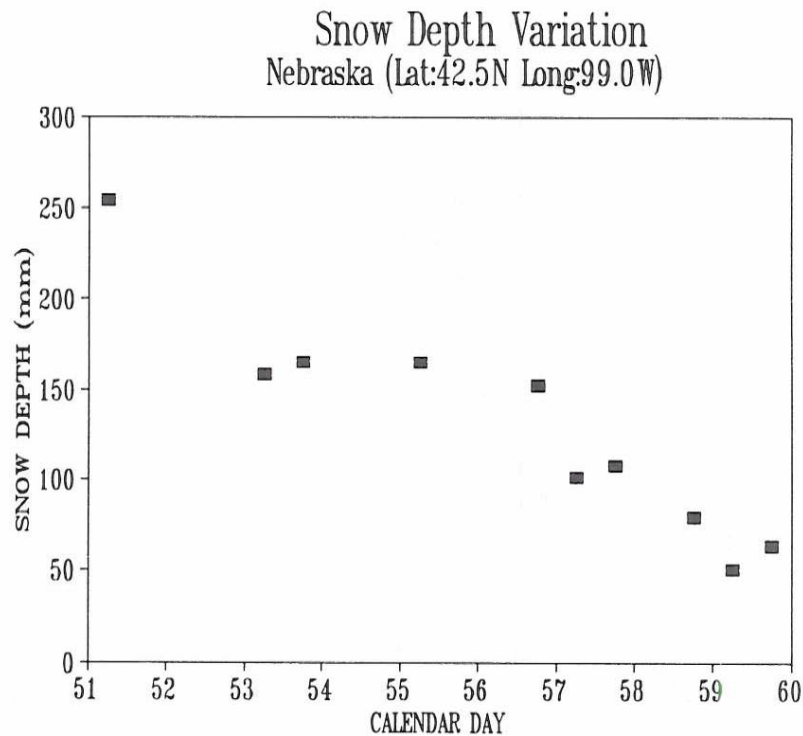
Figure 9.5 Snow depth changes over a ten day period in February 1988, for a region in central Nebraska.
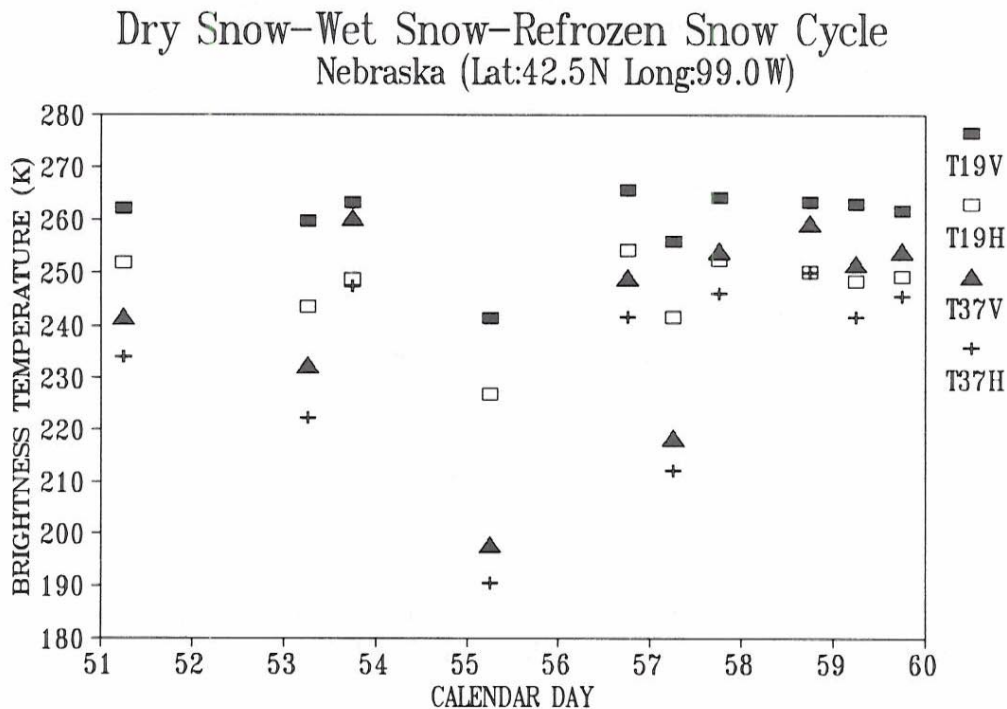
Figure 9.6 Variation in both polarizations of the 19.35 GHz and 37.0 GHz channels throughout a ten day period in February 1988.

snow depth retrieval algorithm should only apply to dry snow conditions. A dynamic database scheme would check for the accumulation of snow during the winter period as well as the onset of the thawing and re-freezing processes, and allow for the consideration of the previous history of the snowpack in the classification logic.

The following additional snow classification rules attempt to classify some of the changing snowpack conditions without the use of a dynamic database. It must be noted that due to the variability in snow microwave signatures, the rules are not perfectly accurate. Footprints which contain a mixture of snow with different degrees of liquid water content, wet soil and vegetation result in complex microwave signatures that cannot be interpreted by stand alone rules, requiring the knowledge of previous history for accurate classifications.

Based on the analysis of several time series of SSM/I data along with snow cover ground truth data as shown in Figures 9.4, 9.5 and 9.6, the wet snow rule can be written as:

$$22V - 19V <= 4.0 \qquad\qquad [a]$$
$$((19V + 37V)/2.0) - ((19H + 37H)/2.0) > 9.8 \qquad\qquad [b]$$
$$-6.5 <= 37V - 19V <= -0.8 \qquad\qquad [c]$$
$$85V - 37V < 0.5 \qquad\qquad [d]$$
$$-1.8 <= 37H - 19H <= 6.5 \qquad\qquad [j]$$
$$253 < 37V <= 268 \qquad\qquad [h]$$

where condition [b] ensures that a high polarization exists due to the presence of liquid water, condition [c] sets the range of scattering in the 37V caused by snow or a snow/soil mixture. It differentiates the wet snow pack from cold semi-arid surfaces. Condition [d] is the flag used to differentiate between dry and moist soil, condition [j] identifies the liquid water in the snowpack and condition [h] allows a range of brightness temperatures within which wet snowpacks usually occur, based on observations of SSM/I data. The rule is complex as a result of the complexity of the surface being classified. Cold semi-arid surfaces with moisture in the top layer, could be confused with wet snowpacks. Also, frozen ground signatures can add to the misclassifications.

Re-frozen snowpacks have a distinct signature from dry and wet snow. Brightness temperatures decrease with increasing frequency in both polarization channels, and additional scattering at 37.0 GHz and 85.5 GHz results in very low brightness temperatures. Thus:

$$22V - 19V <= 4.0 \qquad\qquad [a]$$
$$((19V + 37V)/2.0) - ((19H + 37H)/2.0) > 4 \qquad\qquad [b]$$
$$37V - 19V < -6.5 \qquad\qquad [c]$$
$$37V <= 225 \qquad\qquad [h]$$
$$19V < 37V < 85V$$
$$19H < 37H < 85H$$

9.1.3.6.3   Snow Over Lake Ice and Composite Snow Over Soil and Lake Ice

Some additional interesting microwave signatures involving snow include snow over ice in freshwater lakes during the winter and footprints which contain a mixture of snow over ice and surrounding soil surfaces.

Snow over lake ice can be identified by:

$$22V - 19V <= 6 \qquad\qquad [a]$$
$$((19V + 37V)/2.0) - ((19H + 37H)/2.0) > 4 \qquad\qquad [b]$$
$$37V - 19V >= 0 \qquad\qquad [c]$$
$$37V < 250 \qquad\qquad [h]$$
$$85V - 37V < 37V - 19V$$

and a snow-soil-lake ice mixture is detected by:

22V - 19V < = 4	[a]
((19V + 37V)/2.0) - ((19H + 37H)/2.0) > 4	[b]
37V - 19V < 0	[c]
37V < 250	[h]
85V - 37V > = 37V - 19V	
85H - 37H > = 37H - 19H	

The above rules address anomalous cases with complex signatures and were based on observations of microwave brightness temperatures over lakes in Canada during winter. They need to be further tested with SSM/I data collected over lakes in other parts of the world.

9.1.3.6.4   Snow and Dense Vegetation

This is a fairly common naturally occurring surface as many mountain ranges in the temperate zones have evergreen forests and snow cover during the winter. The microwave signature is characterized by a small polarization difference due to the dense vegetation but with lower brightness temperatures as frequency increases due to surface scattering. The rule attempts to classify these cases, thus decreasing the number of "rain over vegetation" misclassifications which would result otherwise. The rule is:

22V - 19V < = 4.0	[a]
((19V + 37V)/2.0) - ((19H + 37H)/2.0 < = 4.0	[b]
37V - 19V < -4.0	[c]
19V < = 268	[g]

In the winter, under dry snow cover conditions, brightness temperatures in the 19.35 V are usually well below 268 K. However in the spring, the snowpack at higher elevations under trees is usually the last to melt and contributes to surface scattering in the footprint scene. Physical surface temperatures are much higher and the overall effect is a higher brightness temperature in the 19.35 V GHz channel. Snow cover and vegetation can still occur with 19.35 V brightness temperatures greater than 268 K, as it is also theoretically possible for heavy rainfall to occur over a cool vegetated surface, resulting in similar microwave signatures. Most of the time it is possible to differentiate between both surfaces as scattering in the 85.5 GHz channels is greater for atmospheric phenomena such as thunderstorm clouds while for snow covered surfaces, the scattering occurs in both the 37.0 GHz and 85.5 GHz channels.

9.1.3.6.5   Snow Edge

No particular classification rule was developed to detect snow edge due to lack of precise ground truth data. However under the present scheme, footprints would be classified as dry or wet snow, refrozen snow, moist soil or dry soil. Thus, the position of the snow edge would be determined geographically by the classification of congruent footprints as one of these surface types. However, as mentioned above, microwave signatures from footprints that contain a mixture of snow at different liquid water contents, along with wet soil and/or vegetation are

complex and difficult to classify with stand alone rules. In many instances, these situations occur at the edge of the snowpack and will result in an indeterminate classification (a default classification when no other classification rule applies). The use of additional satellite instrument datasets, such as visible data from the OLS or AVHRR under clear conditions would be useful in identifying the exact position of the snowpack edge and serve as "ground truth" for the SSM/I data.

9.1.4 Summary and Conclusions

A list of the developed classification rules is shown in Table 9.12. All temperature threshold values are in degrees Kelvin, based on SDR brightness temperatures. Unless otherwise stated, all conditions within a rule must be true for the rule to apply.

It must be noted that no surface type classification scheme based solely on microwave brightness temperatures will be perfectly accurate. Over land, the large SSM/I footprints integrate emissions from highly heterogeneous surfaces with different microwave properties (soils, vegetation, water). Thus, the rules will classify a given footprint according to the surface type which is most prevalent within it. However misclassifications can still occur, as composite microwave signatures from a mixture of surfaces with different microwave emission properties can be misleading. Misclassifications could possibly occur between:

(1) heavy rainfall over cold, wet soil and snow covered soil

(2) snow and dense vegetation and heavy rain over cool dense vegetation

(3) wet snow with cold wet soil surface

(4) cold wet semi-arid surface and ripe snow covered soil surface.

(5) snow edge or snow-soil mixtures classified as indeterminate

The rules presented in Table 9.12 were designed to be used in combination with the developed overland parameter retrieval algorithms defined in sections 9.2, 9.3 and 9.4. Due to the complex mixture of surfaces which can naturally occur, there will be instances of indeterminate classifications. It is expected that these will be kept to a minimum, not affecting the retrieval of the parameters. Additional rules to deal with these anomalous cases could be developed in the future if necessary.

A listing of the parameter retrieval algorithms which apply to each surface type classification rule are shown in Table 9.13.

TABLE 9.12 SUMMARY OF SURFACE TYPE CLASSIFICATION RULES USING THE
SEVEN CHANNELS OF THE SSM/I

LAND SURF. TYPE	BRIGHTNESS TEMPERATURE COMBINATION THRESHOLD VALUES							
	[a] (K)	[b] (K)	[c] (K)	[d] (K)	[e] (K)	[g] (K)	[h] (K)	[j] (K)
Flooded Cond.	> 4							
Dense Veg.	≤ 4	≤ 1.9		≥ -1	< 4.5	> 262		
Dense Agric./ Range Veg.	≤ 4	> 1.9 ≤ 4		≥ -1	< 4.5	> 262		
Dry Arable Soil	≤ 4	> 4 ≤ 9.8	≥ -6.5	< 0.5 ≥ -5	< 4.2			
Moist Soil	≤ 4	> 4 < 19,7	≥ -6.5	≥ 0.5 < 4	< 4.2			
Semi-Arid Surface	≤ 4	> 9.8 < 19.7		< 0.5	< 6			< -1.8
Desert	≤ 2	≥ 19.7			> -1	> 268		
Precipitation Over Veg.	≤ 4	≤ 4		< -1		> 268		
Precipitation Over Soil	≤ 4	> 4	< -3	< -5	< -4.1	> 268		
Comp. Veg. and Water	≤ 4	< 6.4		≥ -1	≥ 4.5		> 257	
Comp. Soil & Water/Wet Soil	≤ 4	≥ 6.4	≥ -6.5	≥ 0.5	≥ 4.2			
Dry Snow[1]	≤ 4	> 4	< -6.5				> 225 ≤ 257	
Wet Snow	≤ 4	> 9.8	≤ -0.8 ≥ -6.5	< 0.5			≤ 268 > 253	≥ -1.8 ≤ 6.5
Refrozen Snow[2]	≤ 4	> 4	< -6.5				≤ 225	

[a] 22V - 19V [b] (19V+37V)/2 - (19H+37H)/2 [c] 37V - 19V
[d] 85V - 37V [e] 85H - 37H [f] 37V - 37H
[g] 19V [h] 37V [j] 37H - 19H
Additional conditions: [1] 19V - 19H ≥ 5 [2] 19V > 37V > 85V , 19H > 37H > 85H

9-32

TABLE 9.13 CLASSIFICATION RULES AND APPLICABLE ALGORITHMS

Surface Type		Algorithm
Flooded conditions	(7)	No EDR
Dense vegetation	(1)	Surface temperature over vegetation (STV)
Dense agriculture crops and range vegetation	(3)	Surface temperature over land (STL)
Dry arable soil	(9)	Surface temperature over desert (STD)
Moist soil	(18)	Surface temperature over moist soil (STML), Surface moisture (SM)
Semi-arid surface	(15)	STD
Desert	(10)	STD
Precipitation over vegetation	(4)	Precipitation over land (RL)
Precipitation over soil	(8)	RL
Composite vegetation and water	(2)	STV
Composite soil and water/wet soil	(6)	STML, SM
Dry snow	(14)	Snow depth (SD)
Wet Snow	(19)	No EDR
Re-frozen snow	(13)	No EDR
Indeterminate Classification	(0)	No EDR

STV, STL, STD, STML, SM, RL, and SD denote retrieval algorithm codes. Numbers in parenthesis are the proposed EDR surface type codes.

9.1.5   Alternative Rules

In mid-March 1988, after the SSM/I was turned back on, an increase in the noise level of the 85.5 GHz vertical polarization channel was observed. This channel continued to deteriorate until the data was rendered useless by the middle of that year. Later, similar problems with the 85.5 GHz horizontal polarization channel occurred.

The failure of both channels posed a problem for the use of some retrieval algorithms, including the land surface type classification scheme which depends on these brightness temperatures for the accurate classification of water in the footprint scenes as well as precipitation events over land. To address the unavailability of data from these channels, two alternative schemes were developed: (1) classification rules to be used when only the 85.5 V GHz is not available and (2) rules to be used when both 85.5 GHz channels are unusable.

9.1.5.1. Rules for the Loss of the 85.5 V GHz Channel

The methodology used in the development of these rules was the same as described in section 9.1.2. The difference being that the 85.5 V GHz was not included in the analysis. Channel brightness temperature differences and combinations in the original rules which were based on the 85.5 V GHz channel were substituted, for most part, by combinations using the 85.5 H GHz channel. An analysis of the changes to the original rules is conducted in the following Sections.

9.1.5.1.1   Dense Vegetation, Dense Agricultural and Rangeland Vegetation, Composite Dense Vegetation and Water, Rain Over Vegetation

In the original scheme, the 85.5 V - 37.0 V channel combination is used as an indicator of rainfall for the above listed rules. The modified rules are based on the 85.5 H - 37.0 H combination instead. Microwave radiation in the 85.5 H GHz channel will be scattered by hydrometeors in the atmosphere in a similar manner as the 85.5 V Ghz channel, due to its small wavelength. Based on the analysis of numerous storms and, considering a lower limit for this combination of approximately three standard deviations from the mean in the case of no precipitation (Table 9.3), a threshold value of -0.8 K was determined for the flagging of rainfall over vegetation. Therefore, if the combination 85.5 H - 37.0 H $<$ -0.8, rainfall is present within the footprint. Table 9.14 summarizes the new rules.

9.1.5.1.2   Dry Arable Soil, Semi-Arid Conditions, Desert, Composite Soil and Water or Wet Soil, Moist Soil

The 85.5 V - 37.0 V combination was used in these rules both as an indicator of rainfall and for the detection of surface moisture along with the 85.5 H - 37.0 H . The latter combination can be used on its own for both purposes with different brightness temperature threshold values. The classification results will be less accurate (more misclassifications) but overall, the rules perform satisfactorily.

To differentiate between dry arable soil and moist soil, the 85.5 H - 37.0 H combination must have a value less than 3 K, but greater than -4.1 K, the latter being the threshold value, below which rain is present within the footprint scene. For moist soil, the 85.5 H - 37.0 H is greater than 3 K but less than 4.3 K, while for wet soil or composite soil and water 85.5 H - 37.0 H > 4.3 K.   A summary of the rules is shown in Table 9.14.

9.1.5.1.3   Flooded Conditions, Snow Rules

These rules are unaffected by the loss of the 85.5 V GHz channel.

9.1.5.2   Rules for the Loss of Both 85.5 GHz Channels

With the unavailability of both 85.5 GHz channels, the classification of surface moisture is practically impossible with any acceptable degree of accuracy. To detect moisture on the surface due to precipitation it will be necessary to maintain a running average of the 19.35 H/37.0 V normalized brightness temperature and observe significant decreases in the value of this parameter due to moisture, as described in Section 9.3. Composite water and soil or vegetation scenes are harder to detect with just the lower frequency channels. Precipitation can be detected with the 37.0 V - 19.35 V channel combination, however, due to the lesser sensitivity to small hydrometeors of the longer wavelengths at 37.0 GHz, the classification of rainfall is less accurate. Thunderstorm events with large water and/or ice droplets will be classified, but smaller events with finer precipitation might not be detected without the 85.5 GHz channels.

The resulting rules to be used without the 85.5 V GHz channel are shown in Table 9.14 and the rules to be used if both of the 85.5 GHz channels are missing are shown in Table 9.15. The appropriate retrieval algorithms to be used when all channels are present or when the 85.5 V GHz channel is missing are given in Table 9.13. When both 85.5 GHz channels are missing the retrieval algorithms to be used are given in Table 9.16. The dry arable soil rule will be called arable soil as it will include surfaces that are moist or that contain water bodies. Thus, algorithm parameters such as land surface temperature retrieved over such areas could be less accurate.

Snow detection is limited to dry snow as the 85.5 GHz channels are required to categorize other snow surfaces such as wet snow or re-frozen snow.

TABLE 9.14 LAND SURFACE TYPE CLASSIFICATION RULES TO BE USED WHEN 85.5 V GHz CHANNEL IS MISSING

LAND SURF. TYPE	BRIGHTNESS TEMPERATURE COMBINATION THRESHOLD VALUES							
	[a] (K)	[b] (K)	[c] (K)	[d] (K)	[e] (K)	[g] (K)	[h] (K)	[j] (K)
Flooded Cond.	> 4							
Dense Veg.	≤ 4	≤ 1.9			< 4.5 ≥ -0.8	> 262		
Dense Agric./ Range Veg.	≤ 4	> 1.9 ≤ 4			≥ -0.8 < 4.5	> 262		
Dry Arable Soil	≤ 4	> 4 ≤ 9.8	≥ -6.5		≥ -4.1 < 3			
Moist Soil	≤ 4	> 4 < 19.7	≥ -6.5		≥ 3 < 4.3			
Semi-Arid Surface	≤ 4	> 9.8 < 19.7			> -4.1 < 6			< -1.8
Desert	≤ 2	≥ 19.7			> -1	> 268		
Precipitation Over Veg.	≤ 4	≤ 4			< -0.8	> 268		
Precipitation Over Soil	≤ 4	> 4	< -3		< -4.1	> 268		
Comp. Veg. and Water	≤ 4	< 6.4			≥ 4.5		> 257	
Comp. Soil & Water/Wet Soil	≤ 4	≥ 6.4	≥ -6.5		≥ 4.3			
Dry Snow[1]	≤ 4	> 4	< -6.5				> 225 ≤ 257	
Wet Snow	≤ 4	> 9.8	≤ -0.8 ≥ -6.5				≤ 268 > 253	≥ -1.8 ≤ 6.5
Refrozen Snow[2]	≤ 4	> 4	< -6.5				≤ 225	

[a] 22V - 19V   [b] (19V+37V)/2 - (19H+37H)/2   [c] 37V - 19V
[d] 85V - 37V   [e] 85H - 37H   [f] 37V - 37H
[g] 19V   [h] 37V   [j] 37H - 19H
Additional conditions: [1] 19V - 19H >= 5   [2] 19H > 37H > 85H

TABLE 9.15 LAND SURFACE CLASSIFICATION RULES TO BE USED WHEN BOTH 85.5 GHz CHANNELS ARE MISSING

LAND SURF. TYPE	BRIGHTNESS TEMPERATURE COMBINATION THRESHOLD VALUES							
	[a] (K)	[b] (K)	[c] (K)	[d] (K)	[e] (K)	[g] (K)	[h] (K)	[j] (K)
Flooded Cond.	> 4							
Dense Veg.	≤ 4	≤ 1.9	≥ -6.4			> 262		
Dense Agric./ Range Veg.	≤ 4	≤ 4 > 1.9	≥ -6.4			> 262		
Arable Soil[1]	≤ 4	≤ 9.8 > 4	≥ -6.5					
Semi-Arid Surface	≤ 4	< 19.7 > 9.8						< -1.8
Desert	≤ 2	≥ 19.7				> 268		
Precipitation Over Veg.	≤ 4	≤ 4	< -6.4			> 268		
Precipitation Over Soil	≤ 4	> 4	< -6.4			> 268		
Dry Snow[2]	≤ 4	> 4	< -6.5				> 225 ≤ 257	

[a] 22V - 19V   [b] (19V+37V)/2 - (19H+37H)/2   [c] 37V - 19V
[d] 85V - 37V   [e] 85H - 37H   [f] 37V - 37H
[g] 19V   [h] 37V   [j] 37H - 19H

[1]Arable soil type includes the types dry arable soil, moist soil, and composite soil and water/wet soil.
[2]Additional conditions: 19V - 19H >= 5

TABLE 9.16 CLASSIFICATION RULES AND APPLICABLE ALGORITHMS TO BE USED WHEN THE 85.5 GHz CHANNELS ARE MISSING

Surface Type		Algorithm
Flooded conditions	(7)	No EDR
Dense vegetation	(1)	Surface temperature over vegetation (STV)
Dense agriculture crops and range vegetation	(3)	Surface temperature over land (STL)
Arable soil[1]	(9)	Surface temperature all types (STA) Surface moisture (SM)
Semi-arid surface	(15)	Surface temperature over desert (STD)
Desert	(10)	STD
Precipitation over vegetation	(4)	Precipitation over land (RL)
Precipitation over soil	(8)	RL
Dry snow	(14)	Snow depth (SD)
Indeterminate Classification	(0)	No EDR

STV, STL, STA, SM, STD, RL, and SD denote retrieval algorithm codes. Numbers in parenthesis are the proposed EDR surface type codes.
[1]Arable soil type includes the types dry arable soil, moist soil, and composite soil and water/wet soil. The All Types land surface temperature algorithm is to be used.

9.1.7 References

[1] J. Hollinger, R. Lo, G. Poe, R. Savage and J. Pierce, <u>Special Sensor Microwave/Imager User's Guide</u>, Naval Research Laboratory, Washington, DC, 120pp., 1987.

[2] USDA-SCS, "Land Resource Regions and Major Land Resource Areas of the United States," Agricultural Handbook 296, Washington DC, 1981.

[3] F. T. Ulaby, R. K. Moore, and A. K. Fung, <u>Microwave Remote Sensing: Active and Passive. Volume II: Radar Remote Sensing and Surface Scattering and Emission Theory</u>, Reading, MA: Addison-Wesley, 1982.

[4] F. T. Ulaby, R. K. Moore, and A. K. Fung, <u>Microwave Remote Sensing: Active and Passive. Volume III: From theory to applications</u>, Norwood, MA, Artech House, 1986.

[5] W. H. Stiles and F. T. Ulaby, "The active and passive microwave response to snow parameters 1. Wetness," <u>J. Geophys. Res.</u>, vol. 85, No. C2, pp. 1037-1044, 1980.

[6] F. T. Ulaby and W. H. Stiles, "The active and passive microwave response to snow parameters 2. Water equivalent of dry snow," <u>J. Geophys. Res.</u>, vol. 85, No. C2, pp. 1045-1049, 1980.

[7] A. T. C. Chang and J. C. Shiue, " A comparitive study of microwave radiometer observations over snowfields with radiative transfer model calculations," <u>Remote Sens. Env.</u>, vol. 10, pp. 215-229, 1980.

[8] K. F. Kunzi, S. Patil, and H. Rott, "Snow-cover parameters retrieved from Nimbus-7 Scanning Multichannel Microwave Radiometer (SMMR) data," <u>IEEE Trans. Geosci. Remote Sensing,</u> vol. GE-20, pp. 452-467, 1982.

[9] E. Schanda, C. Matzler, and K. Kunzi, "Microwave Remote Sensing of Snow Cover," <u>Int. J. Remote Sensing</u>, vol. 4, pp. 149-158, 1983.

[10] M. J. McFarland, G. D. Wilke, and P H. Harder, "Nimbus 7 SSMR Investigation of snow pack properties in the northern Great Plains for the winter of 1978-79," <u>IEEE Trans. Geosci. Remote Sensing</u>, vol. GE-25, pp. 35-46, 1987.

[11] Cosmic Software Information Services, "CLIPS reference manual," Version 4.0, University of Georgia, Computer Services Annex, Athens, Georgia 30602, 1987.

9.2   LAND SURFACE TEMPERATURE ALGORITHMS

9.2.1   Algorithm Development Rationale

The retrieval of surface temperature over land has been an omission in the development of applications of passive microwave radiometry. Microwave radiometers on satellites have been designed to retrieve the atmospheric temperature profile and sea surface temperature. The radiometers for atmospheric profiles have several channels on the flank of an absorption band, such as the 50 to 60 GHz oxygen absorption band. In order to retrieve an atmospheric temperature profile, the temperature of the lowest atmospheric layer is needed. Sources of this temperature can be either surface temperature reports or a channel in a window adjacent to the absorption peak. While considerable research has been conducted in atmospheric temperature microwave sounders [1], the specification of the surface temperature field over land has not been a product of this research. A major complicating factor has been the variability of the land surface in the field of view of the radiometers. Water in any form in the atmosphere, on the land surface, or in soil (without significant vegetative cover) changes the emission, absorption, and scattering of the emitted radiation. These problems are generally viewed in terms of standardizing or normalizing the background temperature so that the water, in its various forms, may be quantified.

The potential exists for retrieval of land surface temperatures without *a priori* knowledge of the emissivity, absorption, or scattering. The temperatures of densely vegetated or dry land surfaces, each with a high emissivity, should be easily retrievable from vertically or horizontally polarized brightness temperatures. Lambert and McFarland [2] found excellent correlations between the Nimbus-7 Scanning Multichannel Microwave Radiometer (SMMR) in the 18 and 37 GHz vertical and horizontal channels, and air temperature for dry range and prairie areas in the northern Great Plains. The observed air temperatures were measured at screen height, 1.2 m, and reported as daily maximums and minimums in the NOAA climatological data. Incorporation of the 22 GHz vertical channel should aid in the correction for atmospheric water vapor absorption of the emitted radiation. The horizontally polarized brightness temperature at either 19 or 37 GHz should similarly correct for effects of surface or soil water on the emissivity. Land surface temperature retrieval from passive microwave may not be possible or meaningful in the presence of snow, ice, or water.

The original Hughes Aircraft Company (HAC) algorithm for the retrieval of land surface temperature had three forms. Temperature over cloudy land (TLC) was not investigated due to an inability to discriminate extensive cloud cover in the land surface classification module. Previous experience [3] indicates that temperature over snow (STS) and cloudy snow (TSC) would be extremely difficult, at best, to retrieve. The passive microwave radiation from a snow pack is a combination of attenuated radiation emitted from the underlying soil, the reflected sky radiation from the snow surface, and the radiation emitted from the snow. This radiation is strongly influenced by the crystalline structure of the snow, which changes slowly through hoar crystal development and rapidly through freezing and thawing cycles. Although snow is

regarded as a blackbody radiator, it does not function as a black body at the incidence angles of the SSM/I.

For surfaces with a high emissivity (dense vegetation, frozen soil, and glacial), the original HAC algorithm was:

$$ST = C_1 * 19V, \tag{1}$$

where:

ST = surface temperature (K)
C_1 = 1.09 for vegetation
       1.07 for frozen and glacial, and
$19V$ = 19.35 GHz vertically polarized brightness temperature.

These values for C_1 are the inverse of the modelled emissivities for these surfaces. The influences of the atmosphere on the emitted radiation were not considered in this algorithm. For surface temperature over arable land (agricultural and range land), desert, and snow, the original algorithm was:

$$ST = C_0 + C_1*37V - C_2*22V - C_3*19H + C_4*85H. \tag{2}$$

Here C_0 = -36.4 and C_N = the coefficients for the channels, as indicated. The physical explanation for this algorithm can be discerned by rewriting the equation as:

$$ST = A_0 + A_1*37V + A_2*(37V-22V) + A_3*(37V-19H) + A_4*85H. \tag{3}$$

In this form, the 37V channel is the primary channel to retrieve the land surface temperature. Three corrections were made to this estimate. The brightness temperature difference between 37V and 22V is a measure of the atmospheric water vapor which attenuates the emitted radiation. As the difference increases, the amount of the correction must also increase. The polarization difference between the 37 and 19 GHz brightness temperatures is a function of the water present in the land surface scene. As before, the greater the channel-polarization difference, the more the correction is required. The 85H correction is small, between 15 and 20 K, and is a function of the attenuation by atmospheric water. The actual correction for atmospheric water is less, but in this form of the equation, the constant is included in the overall regression constant (C_0). Rearrangement of these terms produces:

$$ST = C_0 + (A_1+A_2+A_3)*37V - A_2*22V - A_3*19H + A_4*85H \tag{4}$$

Here A_1 = the inverse of the emissivity of the dry scene in the 37V channel, A_2 = 0.127, the value of C_2, A_3 = 0.459, the value of C_3, and A_4 = 0.0636, the value of C_4. If the inverse of the emissivity is set at 1.024, the sum of the coefficients is equal to C_1, which is 1.610. If the contribution of the 85H channel were neglected, the inverse of the emissivity should be increased to 1.07.

The purpose of this investigation was to validate the basic rationale of the algorithms based on multiple linear regression, to select the optimal channels for various categories of land surface types, and to calibrate the coefficients of the regression.

9.2.2 Methodology

The management of SSM/I and climatological data has been discussed in Section 9.0.1. Files of SSM/I brightness temperatures and air temperatures (24 hour minimum, maximum and temperature at time of observation) were created for the Western Desert and Central Plains test areas. Air temperatures at screen height (1.2 m) were used for the calibration and validation of the coefficients, as opposed to estimates of the temperature of the emitting layer. Air temperature at screen height is the standard for incorporation of temperature into numerical meteorological and agricultural meteorological models. Under cloudy or high humidity conditions, the surface and screen temperatures should be very close. For clear, dry conditions the deviations between the surface and screen temperatures will be the greatest, especially if the radiating surface is characterized by a high emissivity and a low density. These conditions promote strong radiational inversions in the early morning hours and superadiabatic lapse rates near the surface in the early afternoon hours. A coefficient of determination of 0.94 for 974 pairs of screen temperatures and the radiometric temperature of the earth surface from thermal infrared or microwave sounder measurements from NOAA 6 has been reported [4] with a standard deviation usually less than 2.0 K during the summer months, but in the 3 to 4 K range in the winter months.

Temperatures from the climatological network were used, as opposed to hourly temperatures from first order weather stations, in order to achieve the required density of surface temperature observations. The operation of climatological stations requires volunteer observers to record temperatures and other climatological elements each day. The temperatures are the maximum and minimum during the past 24 hours and the temperature at the time of observation. The time of observation is either during the early morning or the late afternoon, normally at the convenience of the observer. Federal stations (National Weather Service and Flight Service) record the climatological elements at midnight, local time. Consequently, the temperatures recorded for a given day may have occurred the previous day. The actual time of the satellite overpass was about 0615 local standard time, which corresponds fairly closely with the early morning observing time for the temperatures. The late afternoon observing times are generally in the 1700 to 1900 time range, ordinarily several hours after the time of the occurrence of the maximum temperatures. An attempt was made using curve fitting techniques to estimate the screen air temperature at the time of the satellite overpass (near 1800 local standard time). A combination of sine and exponential terms incorporating the times of sunrise, sunset, and normal occurrence of the air maxima and minima were used to determine the 1800 temperature. A large variance was noted when the estimate was compared with the temperature at an observation time of 1800. A decision was made to confine the data set for algorithm development to the ascending, or early morning, overpass.

The means and standard deviations of the differences between the reported minimum temperature and the temperature at time of observation was calculated for about 600 climatological stations in the Central Plains for days 231, 234, 235, and 240, 1987. The comparisons are presented in Table 9.17. For climatological stations with 0500 and 0600 observing times, the reported 24 hour minimum temperature did not agree with the temperature at observation time. This is probably due to an occurrence of the minimum temperature on the previous morning. For this reason, stations with 0500 and 0600 reporting times were excluded from the ground truth data set. For stations with 0700 and 0800 reporting times, the average differences were generally in the 2 to 4 C range, with standard deviations of 2.6 to 3.4 C. In general, the minimum temperatures that constituted the ground truth were about 2 C less than the air temperatures at the satellite overpass time. The ground truth temperatures ranged from 1.1 to 26.7 C during the test period. Additional information on the variance within the ground truth is in Miller [5].

TABLE 9.17 COMPARISON OF REPORTED MINIMUM TEMPERATURES WITH TEMPERATURES AT TIME OF OBSERVATION FOR CENTRAL PLAINS TEST AREA, DAYS 231, 234, 235, AND 240

Means (C) of Differences, Temperature at Time of
Observation and 24 Hour Minimum Temperature

Time of	Calendar Day			
observation	231	234	235	240
0500	19.5	12.5	12.2	16.5
0600	6.2	3.5	4.7	4.8
0700	2.7	1.7	2.1	2.1
0800	4.2	3.3	3.3	3.5

Standard Deviations (C) of Differences Temperature at Time of
Observation and 24 Hour Minimum Temperature

Time of	Calendar Day			
observation	231	234	235	240
0500	4.95	9.62	4.63	7.74
0600	9.02	6.15	5.48	6.89
0700	2.97	2.67	2.95	2.69
0800	3.40	3.35	3.13	3.24

A multiple linear regression analysis was performed for each surface type identified in the land classification module (EXTLND). Initially, all seven channels were used in the regression. The best four channels were identified, based on the coefficient of determination (R^2) and the root mean square error (RMSE). Similar categories of surface type were aggregated, primarily to prevent gradients in the temperature field across surface classifications [6]. Pixels classified as rain, flooded, and snow were excluded from the regression analyses. No stratification was made for cloudiness, due to an inability to classify different cloud types and amounts in the land classification module. The algorithms that were identified were then tested against independent data for both the Central Plains and the Western Desert test areas. The surface temperature files from the Climatological Data contained temperatures from single stations.

9.2.2.1   Multiple Linear Regression With Brightness Temperatures

Four major assumptions are inherent in multiple linear regression analysis. The basic assumption is that the regression model is linear. The other assumptions are that the values of the dependent variable (the retrieved variables, or EDRs, in the SSM/I analyses) are independent of each other and are normally distributed and that the variance of the independent variable is the same for all values of the independent variables. Violation of any of these four assumptions leads to problems with the analysis [7].

In the land surface temperature investigations, the basic form of the D-matrix algorithm was linear. The validity of this assumption was examined by plotting the predicted values against the observed values of land surface temperature and curvilinearity was not apparent. The relationships between the surface temperature and the brightness temperatures were also expected to be linear from a theoretical basis, primarily for single channel regression models. The data were autocorrelated both spatially and temporally within a specific time frame and a test area and the regression equations were tested against independent data for other locations and seasons, so the implications of autocorrelation are not expected to be significant.

Multicollinearity is a problem when two or more of the independent variables are highly correlated with each other. In this event, the regression model will not be able to separate out the effect of each brightness temperature on the surface temperature. In the presence of pronounced multicollinearity, the estimates of the coefficients will have large standard errors and will tend to be unreliable. Multicollinearity is present when a high coefficient of determination is accompanied by statistically insignificant estimates of the regression coefficients [7]. This degree of multicollinearity in the D-matrix approach is evident from the correlation matrix of SSM/I brightness temperatures for the test area and period as shown in Table 9.18. Every channel was highly correlated with every other channel. The highest correlation coefficient was 0.99 between 19V and 37V, the channels least influenced by atmosphere (after rain and standing water pixels were removed from the analysis). The lowest channel to channel correlation was 0.84 between 85V and both 19H and 37H, the channels least sensitive and most sensitive to surface moisture, respectively. The 85V is also the most sensitive to clouds. The correlation between the horizonal and vertical components at a given frequency, or the within channel

correlation, was very high for the 19, 37, and 85 GHz channels. In data not presented, the within channel correlation in vegetated terrain was highest for the 85 GHz channels.

TABLE 9.18 CORRELATION MATRIX OF SSM/I BRIGHTNESS TEMPERATURES FOR LAND SURFACE TYPES USED IN THE LAND SURFACE TEMPERATURE RETRIEVAL

SSM/I CHANNEL

	19V	19H	22V	37V	37H	85V	85H
19V	1.00	0.96	0.98	0.99	0.95	0.91	0.93
19H		1.00	0.93	0.94	0.98	0.84	0.90
22V			1.00	0.98	0.93	0.95	0.96
37V				1.00	0.96	0.93	0.94
37H					1.00	0.84	0.91
85V						1.00	0.97
85H							1.00

A recommended procedure when multicollinearity is present is to drop the correlated variables from the equation, depending on the test of significance of the regression coefficient and the judgement of the researcher [7]. If a highly correlated variable is dropped from the regression equation, the coefficient of determination will not change. This was apparent in the multiple linear regression analyses, as will be shown in the discussion.

Another recommended approach to remove the effects of multicollinearity is to change the form of the independent variables. Normalization or differencing techniques may be employed. A principal component analysis was performed to account for the effects of multicollinearity of the SSM/I brightness temperatures. Principal components is a multivariate analysis technique used to describe relationships between independent variables. A set of linear transformations is used to create a new set of independent variables that are jointly uncorrelated [8]. The first principal component has the largest variance of any linear function of the original brightness temperatures. The second component has the second largest variance, and so forth.

Principal component (or factor) analysis was used to determine the most significant physical factors that relate the SSM/I brightness temperatures to the land surface temperature. The principal components were determined for each land surface category of the aggregated set and used as independent variables in a linear regression analysis. The importance of each coefficient was determined based on the value of the probability level (p) and the t statistic. The

p value is a two-tailed significance probability that the coefficient (and correlation) is zero. A low value of p indicates a high probability that the correlation is significant.

Ideally, the variance of the independent variable is not a function of the values of the independent variables - a condition known as homoscedasicity. If the variances are not equal, then heteroscedasticity will be a problem. This condition was not rigorously tested in the SSM/I data sets, but is not believed to be a problem. A visual examination of the scatter plots of predicted versus observed land surface temperatures did not reveal any pronounced change in variance distribution as a function of the value of the observed land surface temperature. A simple mean and standard deviation analysis of the SSM/I brightness temperatures in the analysis also did not indicate a problem with heteroscedasticity. Because the principal components are standardized and uncorrelated, the coefficient estimates have standard errors, thus avoiding heteroscedasticity. The standard deviations of the brightness temperatures were of the same order of magnitude as the land surface temperatures, as shown in Table 9.19. The standard deviations were slightly higher in the horizontal channels, as expected from influences of surface moisture. The standard deviations of the SSM/I brightness temperatures were also higher in the lower frequencies, as expected from the decreased influence of scattering at the lower frequencies.

In the linear regression analysis, the C_p statistic was used to determine the optimum multiple linear regression models for each surface type aggregate. The C_p is a measure of the total squared error for a model with n independent variables [8]. The C_p provides a measure of the error variance plus the bias introduced by failing to include significant variables in a model. The smallest value of the C_p statistic indicates the optimum model, but the subsets that show a wide divergence between the C_p values are indicative of useful subset sizes. The C_p values are in Miller [5].

TABLE 9.19 MEAN AND STANDARD DEVIATION OF THE SSM/I BRIGHTNESS
TEMPERATURES FOR LAND SURFACE TYPES USED IN THE LAND
SURFACE TEMPERATURE RETRIEVAL

Variable	Mean(K)	Std.Dev(K)	Range(K)
19V	275.6	7.47	258.0-290.6
19H	269.0	8.06	243.3-286.0
22V	275.5	6.52	260.2-289.2
37V	272.7	6.82	257.3-286.7
37H	267.6	7.56	240.1-284.5
85V	275.9	5.56	265.5-288.4
85H	272.9	6.13	255.0-286.0
TEMP	285.9	5.09	272.2-299.9

9.2.3 Results and Discussion

Two sets of results will be presented; with and without the 85 GHz channels as a consequence of the degradation of the 85 GHz channels on the SSM/I on DMSP F-8. The primary data set used for the analysis was from days 231, 234, 235, and 240 in August, 1987 when the 85 GHz channels were not as yet degraded. Consequently, two sets of algorithms are developed. In the event that the loss of the 85 GHz channels degrades the capability to discriminate between land surface types, four categories are used: agricultural/range, dry soil, moist soil, dense vegetation, and all categories. The number of data points in each category for the Western Desert and Central Plains test areas is shown in Table 9.20.

TABLE 9.20 NUMBER OF DATA POINTS, BY CATEGORY, FOR PRINCIPAL COMPONENT AND MULTIPLE LINEAR REGRESSION ANALYSIS OF LAND SURFACE TEMPERATURE

Surface Type	Western Desert	Central Plains	Total
Dense agric./range	317	122	439
All moist soils	955	900	1855
All dry soils	399	107	506
Dense vegetation	133	6	139
All types	1804	1135	2939

9.2.3.1   Algorithms Without the 85 GHz Channels

The results of the regression analysis of principal components for all land surface types in the Western Desert and Central Plains test areas are contained in Table 9.21. The independent variable was land surface temperature, as inferred by screen temperature, at time of overpass. The five eigenvalues which sum to 5.0 correspond to the five SSM/I channels used in this investigation. A set of eigenvalues of relatively small and equal magnitude indicates that the multicollinearity is small, which is not the case with the SSM/I brightness temperatures. The cumulative sum of the variance explained is 1.00. The first factor explains nearly 90 percent of the variance and the second factor explains eight percent. The p level is very low for all five factors, as expected due to the intercorrelations.

TABLE 9.21 PRINCIPAL COMPONENTS ANALYSIS FOR ALL LAND SURFACE TYPES
FOR 19, 22, AND 37 GHz CHANNELS, WESTERN DESERT AND CENTRAL
PLAINS TEST AREAS

| | Factor | | | | |
	1	2	3	4	5
Eigenvalue	4.47	0.40	0.0878	0.0319	0.0115
Cumulative	0.894	0.974	0.991	0.998	1.000
t statistic	46.24	-48.54	-8.05	15.04	-6.07
p level	0.0001	0.0001	0.0001	0.0001	0.0001

The significance or factor loading of each channel within each factor is shown in Table
9.22. These correlations within a factor help with the physical explanation of the factor. Factor
1 represents radiative emission. Although all correlations are high (over 0.92), the 19V and 37V
GHz channels have the highest correlations. This is expected because vertically polarized
radiation is affected less by surface moisture and reflections from bare, dry soil than horizontally
polarized radiation.

TABLE 9.22 CORRELATION COEFFICIENTS OF THE PRINCIPAL COMPONENTS FOR
ALL LAND SURFACE TYPES FOR 19, 22, AND 37 GHz CHANNELS,
WESTERN DESERT AND CENTRAL PLAINS TEST AREAS

| Channel | Factor | | | | |
	1	2	3	4	5
19V	0.966	-0.215	0.091	-0.100	0.055
19H	0.920	0.346	0.181	0.003	-0.040
22V	0.943	-0.303	0.034	0.133	0.004
37V	0.974	-0.148	-0.144	-0.058	-0.067
37H	0.923	0.347	-0.157	0.027	0.048

Factor 2 is a polarization difference term, as indicated by the opposite signs of the
correlations in the vertical and horizontal polarization channels. Two sources of polarization

difference in radiation emitted from land surfaces are surface moisture and reflection from smooth, dry surfaces such as deserts. As surface moisture increases, the emissivity decreases due to a higher dielectric constant. This effect is frequency dependent, with a greater effect at the low frequencies. The horizontally polarized brightness temperatures will decrease more than the vertically polarized brightness temperatures. The net effect of increased soil moisture is to decrease the brightness temperatures and to increase the polarization differences. This is consistent with the principal component analysis with land surface temperature as the dependent variable.

A second source of polarization difference is sky and cloud reflection from bare, dry soils. Deserts are identified by the large polarization differences, which reached 30 C at 19 GHz for deserts in Africa and North America. As the land surface becomes rougher and more vegetated, the polarization difference decreases due to decreased reflection. A principal components analysis for the Central Plains data set, without the bare dry soil influence, is presented in Tables 9.23 and 9.24. The value of the eigenvalue decreased from 0.40 to 0.11 when the Western Desert was excluded from the data set. The correlations of factor 2 with all channels also are lower; attributed to the decreased influence of strong polarization differences from bare, dry soils.

Factor three is characterized by positive, but low, correlations with the 19 and 22 GHz channels and negative correlations for the 37 GHz channels. Factors 4 and 5 are characterized by very low correlations with all channels.

The effects of atmospheric water vapor and cloud and precipitation particles do not appear to have a major influence on land surface temperature retrieval from the SSM/I brightness temperatures. The percent variance explained by factors 3, 4, and 5 is very low when compared to the variance explained by the emission and polarization difference factors. This indicates that regression coefficients to correct for these influences for land surface temperature retrieval may not be statistically significant.

TABLE 9.23 PRINCIPAL COMPONENTS ANALYSIS FOR ALL LAND SURFACE TYPES FOR 19, 22, AND 37 GHz CHANNELS, CENTRAL PLAINS TEST AREA

	Factor				
	1	2	3	4	5
Eigenvalue	4.83	0.11	0.03	0.02	0.005
Cumulative	0.967	0.989	0.995	0.999	1.000
t statistic	51.96	-28.53	12.25	12.99	7.06
p level	0.0001	0.0001	0.0001	0.0001	0.0001

TABLE 9.24 CORRELATION COEFFICIENTS OF THE PRINCIPAL COMPONENTS FOR
ALL LAND SURFACE TYPES FOR 19, 22, AND 37 GHz CHANNELS,
CENTRAL PLAINS TEST AREA

Channel	Factor				
	1	2	3	4	5
19V	0.993	-0.076	0.001	-0.085	-0.040
19H	0.977	0.081	0.016	-0.028	0.024
22V	0.979	-0.175	0.075	0.079	-0.005
37V	0.989	-0.100	-0.093	-0.024	0.044
37H	0.979	0.172	-0.089	0.059	-0.022

The principal component analysis for the agricultural/range land surface type is shown in Tables 9.25 and 9.26. Emission, the first factor, accounts for 98.7 percent of the variance and all channels are very highly correlated with this factor (over 0.98). Because the agricultural/range land surface type was characterized by low polarization difference, the physical interpretation of the second factor will change. The 19 GHz channels are positively correlated and the other channels are negatively correlated with this factor.

A principal components analysis of the Central Plains data set shows only the emission factor to be significant at the 0.05 level, as shown in Tables 9.27 and 9.28. The correlations are over 0.99 for all channels. This suggests that a single channel algorithm could be used for land surface temperature for this land surface category.

TABLE 9.25 PRINCIPAL COMPONENTS ANALYSIS FOR AGRICULTURAL/RANGE
LAND SURFACE TYPES FOR 19, 22, AND 37 GHz CHANNELS, WESTERN
DESERT AND CENTRAL PLAINS TEST AREAS

	Factor				
	1	2	3	4	5
Eigenvalue	4.88	0.054	0.044	0.014	0.010
Cumulative	0.975	0.986	0.995	0.998	1.000
t statistic	27.80	-8.18	-0.19	-2.05	2.20
p level	0.0001	0.0001	0.8464	0.0407	0.0284

TABLE 9.26 CORRELATION COEFFICIENTS OF THE PRINCIPAL COMPONENTS FOR AGRICULTURAL/RANGE LAND SURFACE TYPES FOR 19, 22, AND 37 GHz CHANNELS, WESTERN DESERT AND CENTRAL PLAINS TEST AREAS

Channel	Factor				
	1	2	3	4	5
19V	0.993	0.062	0.033	-0.074	-0.058
19H	0.983	0.174	0.012	0.046	0.031
22V	0.982	-0.110	0.151	0.029	0.007
37V	0.992	-0.061	-0.075	-0.052	0.067
37H	0.988	-0.066	-0.120	0.053	-0.046

TABLE 9.27 PRINCIPAL COMPONENTS ANALYSIS FOR AGRICULTURAL/RANGE LAND SURFACE TYPES FOR 19, 22, AND 37 GHz CHANNELS, CENTRAL PLAINS TEST AREA

	Factor				
	1	2	3	4	5
Eigenvalue	4.94	0.028	0.022	0.008	0.005
Cumulative	0.987	0.993	0.997	0.999	1.000
t statistic	22.36	-1.84	1.91	-1.59	1.03
p level	0.0001	0.0681	0.0584	0.1145	0.3039

TABLE 9.28 CORRELATION COEFFICIENTS OF THE PRINCIPAL COMPONENTS FOR AGRICULTURAL/RANGE LAND SURFACE TYPES FOR 19, 22, AND 37 GHz CHANNELS, CENTRAL PLAINS TEST AREA

Channel	Factor				
	1	2	3	4	5
19V	0.997	0.051	-0.009	-0.041	-0.048
19H	0.992	0.121	0.001	0.034	0.025
22V	0.991	-0.053	0.125	0.002	0.004
37V	0.996	-0.044	-0.055	-0.050	0.037
37H	0.994	-0.075	-0.061	0.055	-0.017

The principal component analysis for the moist soils land surface type is presented in Tables 9.29 and 9.30. As expected, the emission term was factor 1 and the polarization difference term was factor 2. These terms together accounted for 97.5 percent of the variance. All factors were significant, however. This is perhaps a result of correlations between land surface types and atmospheric conditions. For example, when the land surface is moist, the atmospheric water vapor and cloud water content may have more of an effect on microwave emission than when the surface is dry.

TABLE 9.29 PRINCIPAL COMPONENTS ANALYSIS FOR MOIST SOILS LAND SURFACE TYPES FOR 19, 22, AND 37 GHz CHANNELS, WESTERN DESERT AND CENTRAL PLAINS TEST AREAS

	Factor				
	1	2	3	4	5
Eigenvalue	4.59	0.283	0.085	0.028	0.010
Cumulative	0.919	0.975	0.992	0.998	1.000
t statistic	44.00	-32.05	-3.00	13.31	-5.52
p level	0.0001	0.0001	0.0027	0.0001	0.0001

TABLE 9.30 CORRELATION COEFFICIENTS OF THE PRINCIPAL COMPONENTS FOR MOIST SOILS LAND SURFACE TYPES FOR 19, 22, AND 37 GHz CHANNELS, WESTERN DESERT AND CENTRAL PLAINS TEST AREAS

| Channel | Factor | | | | |
	1	2	3	4	5
19V	0.974	-0.185	0.073	-0.092	0.055
19H	0.942	0.273	0.191	0.000	-0.034
22V	0.959	-0.254	0.028	0.125	0.004
37V	0.979	-0.124	-0.134	-0.055	-0.065
37H	0.937	0.307	-0.157	0.025	-0.041

The principal components and correlations for the dry soils land surface type are presented in Tables 9.31 and 9.32 Four factors were significant, including emission and polarization difference, but the first three factors accounted for 98.9 percent of the variance. The emission and polarization difference factors were apparent. Factor 3, characterized by positive correlations between the factor and the 19 and 22 GHz channels and negative correlations with the 37 GHz channels, was also present.

TABLE 9.31 PRINCIPAL COMPONENTS ANALYSIS FOR DRY SOILS LAND SURFACE TYPES FOR 19, 22, AND 37 GHz CHANNELS, WESTERN DESERT AND CENTRAL PLAINS TEST AREAS

| | Factor | | | | |
	1	2	3	4	5
Eigenvalue	4.39	0.402	0.155	0.041	0.015
Cumulative	0.877	0.958	0.989	0.997	1.000
t statistic	12.32	-2.76	3.36	2.75	-1.73
p level	0.0001	0.0068	0.0011	0.0071	0.0859

TABLE 9.32 CORRELATION COEFFICIENTS OF THE PRINCIPAL COMPONENTS FOR DRY SOILS LAND SURFACE TYPES FOR 19, 22, AND 37 GHz CHANNELS, WESTERN DESERT AND CENTRAL PLAINS TEST AREAS

Channel	Factor				
	1	2	3	4	5
19V	0.954	-0.243	0.122	-0.110	0.061
19H	0.920	0.303	0.244	-0.004	-0.054
22V	0.933	-0.327	0.028	0.146	-0.006
37V	0.962	-0.084	-0.240	-0.071	-0.065
37H	0.913	0.372	-0.149	0.044	0.064

The dense vegetation principal component analysis and the correlations with the factors are shown in Tables 9.33 and 9.34. Three factors were significant at the 0.05 level. As with the agricultural/range land surface type, the dense vegetation category is characterized by a very low polarization difference. Consequently, polarization difference did not emerge as an obvious factor in the analysis. Factor 2 was characterized by negative, but small correlation coefficients with the 19 and 37 GHz channels and a higher, positive correlation with the 22V channel. The coefficient of determination, however, was very low. Theoretically, a single channel, vertical polarization, should be sufficient to retrieve the emitting layer temperature. With the best single channel, the 22V, the coefficient was only 0.21. With all five channels, the coefficient of determination increased to 0.29. Addition of the 85 GHz channels in the regression did not

TABLE 9.33 PRINCIPAL COMPONENTS ANALYSIS FOR DENSE VEGETATION LAND SURFACE TYPES FOR 19, 22, AND 37 GHz CHANNELS, WESTERN DESERT AND CENTRAL PLAINS TEST AREAS

	Factor				
	1	2	3	4	5
Eigenvalue	4.65	0.156	0.097	0.053	0.039
Cumulative	0.931	0.962	0.982	0.992	1.000
t statistic	5.92	2.71	0.14	-3.25	0.67
p level	0.0001	0.0075	0.0829	0.0014	0.5059

TABLE 9.34 CORRELATION COEFFICIENTS OF THE PRINCIPAL COMPONENTS FOR DENSE VEGETATION LAND SURFACE TYPES FOR 19, 22, AND 37 GHz CHANNELS, WESTERN DESERT AND CENTRAL PLAINS TEST AREAS

Channel	Factor				
	1	2	3	4	5
19V	0.967	-0.038	0.212	-0.125	0.035
19H	0.966	-0.180	0.072	0.169	-0.007
22V	0.939	0.339	0.008	0.061	0.014
37V	0.981	-0.030	-0.098	-0.060	-0.155
37H	0.970	-0.080	-0.192	-0.041	0.116

improve the performance statistics. The poor performance of the retrieval algorithms is most likely due to the non-representative ground truth. Virtually all of the dense vegetation land surface types were in the mountainous areas of the Western Desert test area. The temperatures of the emitting surfaces, the coniferous tree canopies in the mountains, are not represented by the nearest climatological station. These stations tend to be in lower elevations along river valleys.

Factor analysis and regression of the principal components indicated that a four channel linear regression model should include 19V, 19H, 22V, and 37H. The recommended four channel land surface temperature retrieval algorithm, without the 85 GHz channels, based on the C_p statistic for each of the land surface types is given in Table 9.35. It is interesting to note that 22V was the single channel with the highest correlation with the surface temperature ground truth. In the factor analysis, 22V did not have the highest correlations with the individual factors.

As discussed previously, the estimated variance in the ground truth in the Western Desert was about twice that of the Central Plains. Consequently, the algorithm development is based on the Central Plains data set, with the addition of the dense vegetation land surface type from the Western Desert. The performance statistics of the recommended algorithm are given in Table 9.36. The coefficients of determination range from 0.64 for dry soil to 0.81 for agricultural/range land surface types. The root mean square errors are around 2.5 C.

9.2.3.2   Algorithms With the 85 GHz Channels

The results of the regression analysis of principal components for all land surface types for the Western Desert and the Central Plains is contained in Table 9.37. The factors will not

TABLE 9.35 RECOMMENDED LAND SURFACE TEMPERATURE RETRIEVAL
ALGORITHMS WITHOUT THE 85 GHz CHANNELS

$$\text{Temp (K)} = C_0 + C_1 {*} T19V + C_2 {*} T19H + C_3 {*} T22V + C_4 {*} T37H$$

Surface Type	C_0	C_1	C_2	C_3	C_4
Dense veg. (STV)	32.4	0.31	-0.26	0.82	0.04
Ag/range (STL)	32.4	0.31	-0.26	0.82	0.04
Moist soils (STML)[1]	89.6	-0.47	0.01	1.49	-0.32
Dry soils (STD)	76.7	-0.39	0.31	1.24	-0.42
All types (STA)	83.7	-0.49	-0.02	1.58	-0.34

[1] Included in case future dry/moist soils differentiation is developed.

TABLE 9.36 PERFORMANCE OF RECOMMENDED LAND SURFACE TEMPERATURE
RETRIEVAL ALGORITHMS WITHOUT THE 85 GHz CHANNELS

Surface Type	RMSE (C)	R_2
Dense vegetation	3.03	0.27
Agricultural/range	2.61	0.81
Moist soils	2.32	0.79
Dry soils	2.43	0.64
All types	2.45	0.77

correspond identically to the factors without the 85 GHz channels, due to the nature of principal components analysis. However, the dominant factors will continue to have physical explanations. The magnitude of the eigenvalues and the cumulative variance explained by each factor indicates that there is relatively little variance explained by the third through seventh factors (less than two percent). The p level, however, remains very low for all factors, as expected from the intercorrelations of the channels.

TABLE 9.37 PRINCIPAL COMPONENTS ANALYSIS FOR ALL LAND SURFACE TYPES
FOR 19, 22, 37, AND 85 GHz CHANNELS, WESTERN DESERT AND
CENTRAL PLAINS TEST AREAS

	Factor				
	1	2	3	4	5
Eigenvalue	6.07	0.48	0.32	0.07	0.03
Cumulative	0.867	0.936	0.982	0.992	0.997
t statistic	59.59	-61.16	7.37	7.55	4.96
p level	0.0001	0.0001	0.0001	0.0001	0.0001

Factor 1 has a high positive correlation with all channels, as shown in Table 9.38, and is therefore interpreted as the emission factor. The correlations are the highest for the 19V, 22V, and 37V channels, as expected from physical considerations.

TABLE 9.38 CORRELATION COEFFICIENTS OF THE PRINCIPAL COMPONENTS FOR
ALL LAND SURFACE TYPES FOR 19, 22, 37, AND 85 GHz CHANNELS,
WESTERN DESERT AND CENTRAL PLAINS TEST AREAS

Channel	Factor				
	1	2	3	4	5
19V	0.948	-0.098	-0.289	0.009	-0.049
19H	0.892	0.411	-0.072	0.166	-0.029
22V	0.946	-0.225	-0.191	0.038	0.128
37V	0.977	-0.091	-0.096	-0.144	-0.060
37H	0.913	0.372	0.079	-0.134	0.033
85V	0.928	-0.022	0.361	0.008	-0.077
85H	0.914	-0.318	0.227	0.069	0.059

The correlations between factor 2 and the 19 and 37 GHz channels indicate that factor 2 is due primarily to polarization difference. The 85 GHz channels are less polarized than the other channels as a result of increased surface and atmospheric scattering and a lower response

to surface moisture. Factor 3 may represent a cloud factor, due to the relatively higher positive correlations with the 85 GHz channels.

Principal component and correlation analyses for the Central Plains are contained in Figures 9.39 and 9.40 and show more variance explained by factor 1 and a higher correlation of all channels with factor 1. As with the analysis without the 85 GHz channels, when the analysis is performed on the Central Plains data set, there is less of an influence of the strongly

TABLE 9.39 PRINCIPAL COMPONENTS ANALYSIS FOR ALL LAND SURFACE TYPES FOR 19, 22, 37 AND 85 GHz CHANNELS, CENTRAL PLAINS TEST AREA

	Factor				
	1	2	3	4	5
Eigenvalue	6.62	0.25	0.03	0.02	0.01
Cumulative	0.946	0.981	0.986	0.988	0.989
t statistic	65.55	37.95	11.16	-3.13	3.71
p level	0.0001	0.0001	0.0001	0.0018	0.0002

TABLE 9.40 CORRELATION COEFFICIENTS OF THE PRINCIPAL COMPONENTS FOR ALL LAND SURFACE TYPES FOR 19, 22, 37, AND 85 GHz CHANNELS, CENTRAL PLAINS TEST AREA

Channel	Factor				
	1	2	3	4	5
19V	0.987	-0.068	0.007	0.035	-0.037
19H	0.961	-0.237	0.118	0.031	0.001
22V	0.987	0.069	0.057	-0.096	0.012
37V	0.989	-0.025	-0.093	0.023	-0.001
37H	0.965	-0.227	-0.085	-0.021	0.034
85V	0.948	0.311	0.025	0.051	0.044
85H	0.970	0.183	-0.027	-0.020	-0.050

polarized emission from dry soils. Consequently, factor 2 appears to shift from polarization difference due to dry soils and surface moisture to an overall moisture term. This could include

surface moisture, cloud water, and atmospheric water vapor. Factors 3 through 7, although significant, account for less than two percent of the variance. The very low correlations of the SSM/I channels with these factors renders physical interpretation more precarious. The interpretation of the physical rationale for each of the factors in the principal component analysis is similar to that of the analysis without the 85 GHz channel. The additional data is contained in Miller [5].

Table 9.41 contains a comparison of the performance of the land surface retrieval algorithms with and without the 85 GHz channels. In general, the incorporation of the 85 GHz channels improved the algorithm performance. The improvement in root mean square error was about 0.5 C, with a corresponding increase in coefficient of determination of 0.1. The principal

TABLE 9.41 COMPARISON OF STATISTICS FROM PRINCIPAL COMPONENT ANALYSIS FOR SELECTED LAND SURFACE TYPES AND TEST AREAS WITH AND WITHOUT INCLUSION OF THE 85 GHz CHANNELS

| | RMSE (C) | | R_2 | |
	Without 85 GHz	With 85 GHz	Without 85 GHz	With 85 GHz
All surfaces				
WD	3.79	3.17	0.64	0.72
CP	2.42	2.05	0.77	0.70
All	3.67	3.18	0.62	0.72
Agric./range				
WD		2.82		0.50
CP		2.37		0.85
All	3.00	2.77	0.66	0.71
Moist soils				
WD	3.87	3.45	0.64	0.71
CP	2.30	1.80	0.80	0.88
All	3.53	2.93	0.63	0.74
Dry soils				
WD	3.95	3.89	0.54	0.55
CP	2.44	2.47	0.64	0.64
All		3.47		0.53
Dense vegetation				
WD	2.92	2.93	0.32	0.33

component statistics will not be in exact agreement with the statistics of the multiple linear regression, but will be sufficiently similar for a conclusion. The 85 GHz channels improve the accuracy of the land surface temperature retrieval algorithm. If the 85 GHz channels are not available for accurate classification of land surface type, an aggregation of all surface types leads to an RMSE of about 3.7 C. Scatter plots of the land surface temperature calculated with the recommended algorithm with the minimum air temperature show a linear dependence [6].

Optimum model selection based on the C_p statistic for all land surface types is shown in Table 9.42. The 85V is the single channel model with the highest coefficient of determination

TABLE 9.42 MULTIPLE LINEAR REGRESSION MODEL FOR ALL CHANNELS FOR RETRIEVAL OF LAND SURFACE TEMPERATURE FOR ALL LAND SURFACE TYPES

Channels in Regression Model							C_p	R^2
22V							1213.	0.666
85H							978.	0.700
85V							380.	0.784
85V	19V						181.	0.813
85V	37V						138.	0.819
85V	37H						116.	0.822
85V	19V	22V					89.1	0.827
85V	37H	22V					70.5	0.829
85V	37V	22V					58.3	0.831
85V	37V	22V	19H				28.6	0.835
85V	37H	22V	19V				17.0	0.837
85V	37H	22V	37V				15.7	0.837
85V	37H	22V	37V	85H			13.9	0.838
85V	37H	22V	19V	19H			8.1	0.839
85V	37H	22V	37V	19V			7.2	0.839
85V	37H	22V	19V	19H	85H		8.9	0.839
85V	37H	22V	37V	19V	85H		8.4	0.839
85V	37H	22V	37V	19V	19H		6.9	0.839
85V	37H	22V	37V	19V	19H	85H	8.0	0.839

and lowest value of the C_p statistic, in contrast to the 22V of the analysis without the 85 GHz channels. The coefficients of determination indicate that only two or three terms in the regression equation are sufficient to produce essentially the same value as the full seven term model. This is due to the multicollinearity of the channels.

A larger ground truth data set was assembled, but the areas and dates remained the same as previously, to compare algorithm performance with and without the 85 GHz channels. Additional data from days 50, 51, 53, 56, 57, and 59, 1988 were included as an independent data set. A multiple linear regression analysis on the Central Plains and Western Desert data for days 231, 234, 235, and 240, 1987 was conducted for all land surface types except snow, rain, and standing water. The coefficients of determination ranged from 0.15 for dense vegetation to 0.86 for vegetation with some water present. The RMSE's ranged from 1.87 C for vegetation with some water present to 3.58 C for dense vegetation. The lower statistics for the dense vegetation may be more of a function of the variance between the surface temperature observations and the temperature of the emitting canopy than of the site to site or day to day variance of the brightness temperatures. The same land surface type categories were constructed as previously. A multiple linear regression analysis was performed for each category with the Central Plains data set. The channels that are optimal, based on the C_p statistic and the statistical significance of the regression, are presented in Table 9.43 for the Central Plains data set only. The 85V and 37V channels are dominant, followed by the 22V and 19V channels. However the 85V, 37V, 22V, and 19H channels were selected for a four channel retrieval algorithm. The coefficients of determination showed essentially no change from the optimal channels. The recommended four channel land surface retrieval algorithm, with the 85 GHz channels, for each of the land surface types is given in Table 9.44. The performance statistics are given in Table 9.45.

TABLE 9.43 OPTIMUM MODELS FOR LAND SURFACE TEMPERATURE RETRIEVAL FROM THE 19, 22, 37, AND 85 GHz CHANNELS, BASED ON THE C(p) STATISTIC, CENTRAL PLAINS TEST AREA

Land surface type	C_p	Channels						R^2	R^2 all ch.
Agricultural/range	0.90	85V	37H					0.812	0.817
Moist soils	6.23	85V	37V	22V	37H	19V	19H	0.851	0.851
Dry soils	3.55	22V	85V	37H				0.622	0.634
All types	5.55	85V	37H	22V	19V	19H		0.791	0.791

TABLE 9.44 RECOMMENDED LAND SURFACE TEMPERATURE RETRIEVAL ALGORITHMS WITH THE 85 GHz CHANNELS

$$Temp\ (K) = C_0 + C_1*T19H + C_2*T22V + C_3*T37V + C_4*T85V$$

Surface Type	C_0	C_1	C_2	C_3	C_4
Dense veg. (STV)	24.94	-1.2784	0.8800	0.5933	0.7299
Ag/range (STL)	6.97	-0.6266	0.2716	-0.1297	1.4820
Moist soils (STML)	23.16	-0.1873	0.5221	-0.6271	1.2320
Dry soils (STD)	72.68	-0.4598	0.5984	-0.8828	-0.2623
All types (STA)[1]	26.46	-0.3133	0.7327	-0.4469	0.9540

[1] Can be used if above four types cannot be differentiated.

TABLE 9.45 PERFORMANCE OF RECOMMENDED LAND SURFACE TEMPERATURE RETRIEVAL ALGORITHMS WITH THE 85 GHz CHANNELS

Surface Type	N	RMSE (C)	R_2
Dense Veg.	68	3.45	0.21
Ag/range	237	2.69	0.77
Moist soils	1230	2.78	0.76
Dry soils	229	3.60	0.46
All types	1764	3.14	0.71

9.2.4 Recommendations

1. Algorithms should be implemented, with or without the 85 Ghz channels, for land surface temperature retrieval.

2. Dense vegetation will have the same coefficients as agricultural and range lands for the algorithms to be used when the 85 GHz channels are not available.

3. If the land surface classification is degraded due to the loss of the 85 Ghz channels, an all surface types land surface temperature retrieval algorithm should be implemented.

4. Algorithms should be deferred for surface temperatures for cloudy land, snow, cloudy snow, and glacial.

Additional research should include digital thermal infrared surface temperatures under clear, relatively dry sky conditions as the ground truth. The logical source of this information is from the Operational Line Scanner (OLS) on the DMSP satellite. The major difficulty to date with the use of OLS thermal data as ground truth for the SSM/I surface temperature retrieval is the difficulty in assigning a latitude and longitude for each OLS pixel (the operational uses of the OLS data are based on visual interpretation of the images) and merging this information with the SSM/I information. The use of OLS data under near-ideal conditions will facilitate the calibration of algorithms for surface conditions of dense vegetation and forest areas, mountainous areas, and areas with a low density of surface weather stations. Particularly with areas of dense vegetation and forests, the thermal infrared channel will provide a source of ground truth that is representative of the emitting surface. The variance of the ground truth may also be determined for all land surface types and locations. The algorithms developed are intended for use on both ascending and descending passes with the knowledge that a bias will be inherent with the descending pass (late afternoon) retrieval. This bias can be determined with the use of OLS thermal infrared data.

An additional recommendation is to use SSM/I land surface temperature data in the retrieval of atmospheric profiles with other sensors on the DMSP satellite series.

9.2.5 References

[1] F. T. Ulaby, R. K. Moore, and A. K. Fung. Microwave Remote Sensing: Active and Passive. Volume III: From Theory to Applications, Norwood, MA, Artech House, 1986.

[2] V. M. Lambert and M. J. McFarland, "Land surface temperature estimation over the northern Great Plains using dual polarized passive microwave data from the Nimbus 7," ASAE Paper 87-4041, Baltimore MD, 1987.

[3] M. J. McFarland, G. D. Wilke, and P. H. Harder II, "Nimbus 7 SSMR Investigation of snow pack properties in the northern Great Plains for the winter of 1978-79," IEEE Trans. Geosci. Remote Sensing, vol. GE-25, pp. 35-46, 1987.

[4] H. W. Yates, J. D. Tarpley, S. R. Schneider, D. F. McGinnis, and R. A. Scofield, "The role of meteorological satellites in agricultural remote sensing," Remote Sens. Env., vol. 14, pp. 219-233, 1984.

[5] R. L. Miller, "Land surface temperature estimation using special sensor microwave imager data," Ph.D. dissertation, Texas A&M University, College Station, TX, 1990.

[6] M. J. McFarland, R. L. Miller, and C. M. U. Neale. "Land surface temperature derived from the SSM/I passive microwave brightness temperatures," <u>IEEE Trans. Geosci. Remote Sensing</u>, vol. 28, pp. 839-845, 1990.

[7] A. R. Hoshmand, <u>Statistical Methods for Agricultural Sciences</u>, Portland, OR: Timber Press, 1989.

[8] R. J. Freund and R. C. Littell, <u>SAS system for regression</u>, Cary NC: SAS Inst. Inc., 1986.

9.3   SURFACE MOISTURE ALGORITHMS

9.3.1   Algorithm Development Rationale and Background

Free water in pores of soil surface layers will lower the emissivity of those layers due to the increase in the soil permittivity. The results of several field experiments have generally shown a linear relationship between normalized brightness temperature and soil moisture expressed either on a gravimetric or percent of field capacity basis [1], [2], [3]. Wang et al. [3] also showed a linear relationship between normalized brightness temperatures and the volumetric water content in the top 10 cm layer of soils at different frequencies (1.4, 5 and 10.7 GHz). The linear variation of brightness temperatures at a particular wavelength with the volumetric moisture content of the soil will be approximately the same for most soil texture types.

Emitted microwave brightness temperatures have also been correlated to estimates of surface moisture such as the antecedent precipitation index (API) [4], [5], [6]. High correlations were found at several frequencies including 19 GHz and 37 GHz. Recent studies by Choudhury et al. [7] and Owe et al. [8] have also used the API as a measure of soil surface wetness and incorporated soil evaporation in the estimation of the recession coefficient.

Several physical factors affect the sensing of soil moisture at different microwave frequencies. At short wavelengths, most of the brightness temperature contributions from a soil are emitted by a shallow layer at the soil surface. For a wet soil, this moisture sensing depth is on the order of ten percent of the wavelength. This would represent an emitting layer of only a couple of millimeters at the 19.35 GHZ (1.55 cm) channel of the SSM/I. Soil surface roughness and texture also affect the measured brightness temperatures by decreasing the sensitivity to soil moisture. This was shown to be the case by Wang et al. [3] and Newton and Rouse [9] for several microwave frequencies.

Vegetation cover will also decrease the sensitivity to soil moisture due to self emission as well as scattering and de-polarization of microwave radiation emitted by the soil. Several studies have indicated that longer wavelengths can better penetrate vegetation cover and therefore are better suited for soil moisture sensing. Vegetation effects on microwave sensitivity to soil moisture have been studied and discussed by Wang et al. [10], Burke and Schmugge [11], Theis and Blanchard [12] and Ulaby et al. [13].

The short wavelengths of the SSM/I will result in a small soil penetration depth as well as a reduced sensitivity to surface moisture if any vegetation is present above the soil surface. In addition, the large SSM/I footprint sizes will lead to the introduction of noise due to surface type variability as well as the random nature of precipitation occurrences and spatial patterns at that scale. For these reasons, the correlation of SSM/I variables based on brightness temperatures with an antecedent precipitation index was deemed the best approach for moisture retrievals at the soil surface.

9.3.2 Methodology

9.3.2.1 Passive Microwave Data

SSM/I overpasses over the Central Plains and Western areas of the United States were selected based on the potential presence of surface moisture. This was determined by locating storm systems with significant precipitation on daily and weekly weather maps published by NOAA. Only large frontal systems were selected as small convective storms can be a source of error due to the spatial resolution of the sensor and the relatively sparse distribution of weather stations at that scale. The SSM/I data were ordered to cover a time period ranging from 1 or 2 days before the storm to several days after it. In this way, it was possible to detect abrupt changes in surface moisture on the day of the storm and the subsequent dry-down period. SDR brightness temperatures over the area of interest were downloaded to disk using software supplied by NRL for the VAX VMS operating system. The data were then submitted to a set of programs developed at Texas A&M University which removed header records and prepared the data for the surface-type classification expert system program [14]. The classification scheme as described in section 9.1 determined the major surface types, i.e., water, snow, and dry and wet soil surfaces as well as vegetation densities based on average polarizations in the 19.35 GHz and 37.0 GHz channels. Classified footprints were gridded to 0.25 degree latitude/longitude cells which contained the seven brightness temperatures and a surface type classification code. Because the distance between concentric A-scan footprints of the SSM/I is on the order of 0.25 degrees at mid-latitudes, most of the time only one footprint was placed in each latitude/longitude 0.25 grid cell.

9.3.2.2 Ground Truth Data

Climatic data used as "ground truth" in this study covered a period from July to October 1987 and January to December 1988 and consisted of daily maximum and minimum temperatures and precipitation amounts from the cooperative network of weather stations operated by NOAA. The data tapes were ordered from NOAA and were downloaded to disk using a VAX mainframe computer with special software developed for this purpose. The climatic variables for each weather station were gridded to 0.25 degree latitude/longitude cells for the entire USA and for each calendar day of the year. If more than one weather station were present in a particular grid cell, the values for each climatic variable were averaged.

Daily antecedent precipitation index (API) values were calculated for each latitude/longitude cell based on the available temperature and precipitation data. The API was calculated as:

$$API_i = (API_{i-1} + P_i) * K \tag{1}$$

where K is the recession coefficient, P is the effective precipitation, and $_i$ and $_{i-1}$ represent the current and preceding days respectively.

Two methods of estimating the recession coefficient (K) were used: (1) a method proposed by Wilke and McFarland [6] where the K factor was allowed to vary between a minimum of 0.70 in the summer and a maximum of 0.92 in the winter to account for seasonal changes in evaporation potential, and (2), a K factor described by Choudhury et al. [7]:

$$K_i = \exp(-E_i/W_m) \tag{2}$$

where E_i is the evaporation on day i and W_m is the maximum depth of soil water available for evaporation.

As stated by Choudhury et al. [7], the magnitude of W_m is uncertain because of the small sampling depth at microwave frequencies. They concluded based on a sensitivity analysis that the correlations between brightness temperature and API consistently increased as W_m is decreased. Due to the small moisture sensing depth at 19.35 GHz and the empirical nature of the above mentioned equation, five recession coefficients were computed for values of W_m equal to 5, 7.5, 10, 15, and 20 mm which resulted in five API values with notation API_1, API_2, API_3, API_4, and API_5. The notation API_m was used for the API estimated using the recession coefficient proposed by Wilke and McFarland [6]. Because the API value for a given day at a given grid cell location depends on the API of the previous day, missing records in weather data files were checked and reported in the output file as number of days since last missing record. If for a given grid cell location on a given day, the precipitation data were available but no temperature data were reported, an estimate of the maximum and minimum temperatures for that grid cell was obtained by averaging data from surrounding cells. This was acceptable because temperature is a fairly spatially conservative variable across similar elevations. For these cases, a flag was set in the output file in order to allow further screening of those data if their reliability was questioned during the analysis. An additional variable calculated for each grid cell was the number of days since the last rainfall event. In this way its significance in the algorithm development could be evaluated.

The potential evaporation or evapotranspiration was computed using the Hargreaves equation [15]. This equation was selected because it required input data which were readily available such as the day of the year, the latitude, and daily minimum and maximum temperatures. The Hargreaves equation is in good agreement with the Penman equation in relatively dry climates with no or moderate wind. The Hargreaves equation can be written as:

$$ETP = 0.0023 \, Ra \, (TC + 17.8) \, TD^{0.5} \tag{3}$$

where ETP is the potential evapotranspiration for grass in mm/day, Ra is the extraterrestrial solar radiation in mm/day, TC is given by (Tmax + Tmin)/2 (Average Daily Temperature), TD is given by (Tmax - Tmin), Tmax is the maximum daily temperature, and Tmin is the minimum daily temperature.

The extraterrestrial radiation can be expressed as a function of the latitude and the day of the year:

$$Ra = \frac{9167.32 \; ES \; [OM \sin(Lat) \sin(DEC) + \cos(Lat) \cos(DEC) \sin(OM)]}{596 - 0.55 TC} \quad (4)$$

where:

Lat = latitude of the location in radians,
DEC = 0.40876 cos(0.0172142 (J + 192)),
J = day of year,
ES = 1.00028 + 0.03269 cos(0.0172142 (J + 192)),
OM = Arc.cos (-tan(Lat) tan(DEC)).

The latitude of the lower right hand corner of each grid cell was used as the latitude value in this equation. Other ground truth data used consisted of the Major Resource Regions and Major Land Resource Areas of the United States (MLRA) by the Soil Conservation Service [16]. This land classification scheme groups areas with similar soils, natural vegetation, climate, and topography and assigns a code number to each area. These code numbers were also gridded to 0.25 degree cells.

9.3.2.3   Regression Data File Preparation

Gridded classified SSM/I files were merged with gridded API files for the appropriate dates and overpasses with a computer program, resulting in output files which contained primary data for statistical regression. The implied assumptions with the form of the API equation used (Equation 1) were that no evapotranspiration occurred during the night and that the precipitation events occurred between 12 am and 6 pm. Therefore, the morning overpass gridded SSM/I files were merged with the gridded API file for the previous day while afternoon overpasses were merged with API files for the same day. If no API data were available for a particular 0.25 degree cell, data from that cell were omitted in the output daily regression file. The merged daily regression files contained the latitude/longitude location of each grid cell, the corresponding seven measured brightness temperatures, the classified surface type code, the surface temperature, the precipitation amount for the day, the number of days since the last rainfall event for that location, the MLRA region code, missing data flags and 6 estimated API values. All gridded daily files were then transferred from the VAX to a microcomputer via KERMIT for further preparation.

A computer program was written for the PC to read these files and search for locations with high API values in order to confirm on independent weather maps that storms actually had occurred. A second computer program screened the regression files with the purpose of building time series for a given MLRA region (MLRA region files) which included data prior to the storm as well as for the dry-down period after the storm. In this procedure, grid cells with at least one high API value were detected and tagged during the successive reading of all input files. The average polarization in the 19 GHz and 37 GHz channels was calculated for dry

surface conditions just prior to the storm in order to further categorize vegetation cover density for the tagged grid cells. During a second reading of the input files, time series for grid cells with at least one high API value were created and the data placed in different MLRA sub-region output files according to vegetation of the dry cell. A maximum of six possible output files (MLRA sub-region files) were created for each MLRA region according to vegetation cover density for the period of analysis. The six initial classes corresponded to average polarizations in the 19.35 GHz and 37.0 GHz of: (6) less than 4 K, (5) between 4 and 6 K, (4) between 6 and 8 K, (3) between 8 and 10 K, (2) between 10 and 12 K and, (1) greater than 12 K.

9.3.2.4 Data Analysis

The regression analysis was conducted on MLRA sub-region files grouped according to vegetation density class. The ground truth variables consisted of the 6 API values while the SSM/I variables consisted of several forms of normalized brightness temperatures and apparent emissivities. Apparent emissivities were obtained for each channel by dividing the brightness temperatures by the surface physical temperature. As simultaneous OLS surface temperature data were not available, the measured air temperature at the time of the overpass was used. Air temperatures can be fairly good substitutes under stable climatic conditions. For ascending overpasses, the minimum air temperature was used as it usually occurs close to 6 am in a semi-arid environment. For descending overpasses, the temperature at 6 pm was predicted using a sinusoidal curve adjustment based on maximum and minimum temperature for the day [17].

For best operational use, the surface moisture algorithms ideally should be based solely on SSM/I brightness temperatures without requiring any additional ground information. Because both vertically and horizontally polarized brightness temperatures are influenced by the physical soil and vegetation temperatures in the same way, but vary in magnitude with moisture (at the 53° incidence angle of the SSM/I), normalized brightness temperature ratios consisting of horizontal channels divided by the vertical channels were tested as well. As it will be seen later in the analysis section, the T19H GHz channel was found to be the most sensitive to surface moisture. The following SSM/I variables were tested :

1) T_{19h}
2) T_{19h} / T_{phys}
3) T_{19h} / T_{19v}
4) T_{19h} / T_{37v}
5) $T_{19h} / [0.5*(T_{19v} + T_{37v})]$.

As previously mentioned in section 9.3.1, most of the correlations between microwave emission versus surface moisture have been explained by linear relationships. This was examined in the data analysis by using the following models:

$$X = \beta_0 + \beta_1 \text{ API}$$

where:

$$API = \text{Antecedent Precipitation Index } (API_m, API_1, ..., \text{ or } API_5),$$
$$X = \text{brightness temperature, apparent emissivity,}$$
$$\text{or normalized temperature ratio,}$$
$$\beta_0, \beta_1 = \text{regression coefficients.}$$

The second model tested was a logarithmic transformation of the API values versus the radiometric data:

$$X = \beta_0 + \beta_1 \ln (API)$$

The third model correlated the SSM/I variables with estimated reflectivity coefficients obtained by transforming API values using a simple radiative transfer model:

$$X = \beta_0 + \beta_1 \Gamma_{estim}$$

The radiative transfer equation was defined in section 9.1 and the assumptions involved are described by Ulaby et al. [18]. In this procedure, volumetric soil moisture values were estimated from API values using $\theta_v = C_{trans} API$. The value of the C_{trans} coefficient was determined by looking at its effect on the goodness of the fit of the linear model. The soil dielectric constant can be estimated from the volumetric soil moisture content using an empirical equation proposed by Hallikainen et al. [19]:

$$\varepsilon_r = (a_0 + a_1 S + a_2 C) + (b_0 + b_1 S + b_2 C)\, \theta_v + (c_0 + c_1 S + c_2 C)\, \theta_v^2 \quad (5)$$

where:

$$\theta_v = \text{volumetric soil moisture content,}$$
$$S = \text{sand content in \%,}$$
$$C = \text{clay content in \%,}$$
$$a,b,c = \text{empirical coefficients.}$$

Hallikainen et al. [19] provide values for the empirical coefficients to determine the real and imaginary part of the dielectric constant for frequencies between 1.4 GHz and 18 GHz. They showed that as the frequency increases, the soil dielectric constant versus soil moisture relationship is less sensitive to soil texture. At 18 GHz the influence is minimum. For this study, the empirical dielectric behavior at 18 GHz was assumed to be applicable to the 19.35 GHz which was the SSM/I frequency selected for the soil moisture algorithm development.

The real and imaginary parts of the dielectric constant were then used to estimate the specular reflectivity coefficient after the trigonometric transformation of the following equations [18]:

$$\Gamma^{sp}(\theta_1, h) = \left| \frac{\cos\theta_1 - \sqrt{\varepsilon_r - \sin^2\theta_1}}{\cos\theta_1 + \sqrt{\varepsilon_r - \sin^2\theta_1}} \right|^2 \quad (6)$$

$$\Gamma^{sp}(\theta_1, v) = \left| \frac{\varepsilon_r \cos\theta_1 - \sqrt{\varepsilon_r - \sin^2\theta_1}}{\varepsilon_r \cos\theta_1 + \sqrt{\varepsilon_r - \sin^2\theta_1}} \right|^2 \qquad (7)$$

Finally, the specular reflectivity was then used to estimate the soil surface reflectivity coefficient (Γ_{estim}). The use of this regression model for surface moisture prediction however would require an iterative method to solve for API from Γ_{estim} because the equation is implicit for unknown soil moisture.

The regression analysis was conducted on a VAX using the SAS statistical package and on a microcomputer using a simple regression program written in Turbo Pascal 5.0. The latter program was tailored to handle the above mentioned models, allowing for the selection of different SSM/I and ground truth variables. Figure 9.7 represents a general flow-chart of the data analysis methodology.

9.3.3 Selection of the SSM/I Channels Most Sensitive to Surface Moisture

9.3.3.1 SSM/I Channel Selection

Many studies in passive microwave remote sensing have shown a decrease in sensitivity to surface moisture as wavelengths decrease, due to smaller penetration depths. To test this fact, stepwise regression using the logarithmic model was conducted with SAS on 21 MLRA sub-region files for the 1987 data set. Regressions of apparent emissivities (SSM/I brightness temperatures divided by the surface physical temperature) versus API were carried out and in all cases, the channel resulting in the highest correlation was the 19.35 GHz horizontal polarization. The decrease in correlation was quite drastic when other channels with shorter wavelengths were used in the regression. Based on this analysis and due to the larger available penetration depth and sensitivity at the 53° incidence angle, the 19.35 H channel was considered best suited for surface moisture retrievals.

9.3.3.2 Vegetation Cover Effects

Vegetation overlying the soil surface will decrease the sensitivity of moisture detection at the short wavelengths of the SSM/I. At high vegetation densities, retrievals of surface moisture are physically impossible. To illustrate this fact, several time series of SSM/I signatures, API values, and precipitation values were plotted for single latitude/longitude cells in regions with different vegetation covers. The SSM/I variables or signatures represented in the following graphs consist of the apparent emissivity for the 19.35 GHz horizontal channel (e19H = T19H/Ts) and the normalized temperature ratio 19.35 H GHz divided by the 19.35 V GHz channel (T19H/T19V). The API values were obtained using Equation 1 and 2 with a W_m of 10 mm.

SSM/I Data Def Tape

NOAA
Weather Data

Do for each satellite overpass of interest

Create Daily files for minimum, maximum temperatures and precipitation for each quarter of a degre box where at least one weather station is present.

- Read Data Portion of Interest
- Clean Data, Remove Headers & Prepare Data for Surface Type Algorithms
- Surface Type Algorithms.

API computation

$$API_i = (API_{i-1} + P_i) k_i$$

$$k_i = \exp(-E_i / W_m)$$

E_i : ETP by Hargreaves eq.
W_m = 5, 7.5, 10, 15 or 20 mm

Daily files
- location in 1/4 deg.
- SSM/I brightness temp.
- Classification type.
- Surface Temperature.
- MLRA
- API values

Build time series files for MLRA sub-regions based on vegetation density (polarization over dry soil).

Merge Files of Similar Vegetation Cover

Regression Analysis:
SSM/I Variable:
t19h/ts, t19h/t19v, t19h/t37v, t19h/[0.5*(t19v+t37v)], or t19h
Ground Truth:
API1, API2, API3, API4, API5, or APIm
Models:
1) Linear
2) Logarithmic
3) Transformation of API using Radiative Transfer Model

Model Selection and Coefficients Determination for Various Vegetation Covers

Test Model with Independent Set of Data

Figure 9.7 Global flow chart of data analysis methodology.

In Figure 9.8, for an agricultural area in West Texas, the normalized temperature (T19H/T19V) at the beginning of the season was relatively low, gradually increasing and peaking around day 220. The increase in normalized temperature as the season progressed was a result of a decrease in polarization caused by increased vegetation density. Precipitation events resulting in considerable surface moisture caused an abrupt lowering of the SSM/I variables due to a decrease in soil emissivity. The storm on day 141 resulted in a much greater microwave response than the larger event around day 190. This could have occurred due to a combination of two factors: a denser vegetation cover on day 190 and/or a localized storm which did not thoroughly wet the entire footprint. Figure 9.9 shows a similar pattern with peak vegetation occurring around day 180. A well vegetated footprint from a location further east is shown in Figure 9.10. The normalized temperature had a value closer to one indicating small polarizations and the sensitivity to surface moisture resulting from precipitation was lower. In all the series examined, the apparent emissivity carried more unexplained variability than the normalized temperature.

To further study vegetation effects on surface moisture retrievals, an analysis was also conducted on the MLRA sub-region files for the central plains of the United States for the year 1988. In this way, differences due to vegetation types and seasonal effects could be considered. Table 9.46 shows the resulting vegetation density classifications of the sub-region files for the three seasons analyzed. A change in vegetation density over the time period studied (from spring to fall) occurred for most of the areas under investigation. For cropland areas in the central plains, spring time is characterized by relatively bare soils followed by a rapid increase in vegetation density at the end of spring, to full cover during summer and low vegetation cover after harvest in late summer or fall. Agriculture in the central plains relies mostly on natural precipitation with the exception of the south eastern plain regions (West Texas, Oklahoma, East Colorado, and Nebraska) which have a significant area under irrigation. Other major vegetation types consist of rangeland, and pasture. Natural vegetation is mostly comprised of short, medium, or tall grasses with peak vegetation density occurring in late spring and early summer depending on the latitude. Changes in vegetation density for this type of land cover are not as extreme as the case of cropland regions. The area covered in this study ranged from a latitude of 30 to 49 degrees north latitude, which implied a spectrum of vegetative calendars according to location and elevation.

The MLRA sub-region files were clipped to contain data for a period of not more than 25 days in a particular season and included data from the dry surface prior to the storm, the passage of the storm and increase in surface moisture, and the subsequent recession period as the surface dried. For most data sets shown in Table 9.46, the average polarization over dry soil was highest in the spring, lowest in the summer, and showed an intermediate value in the fall. Polarizations in spring and fall were similar for cropland areas, with a decrease of several Kelvins in the summer. Some regions did not show any significant changes across seasons and were either 1) semi-arid regions or lakes if the average polarization was large, or 2) dense natural vegetation if the average polarization was low. Vegetation density had a major influence on the sensitivity to surface moisture. Figure 9.11 shows distinct differences between the three seasons for MLRA 106 (Nebraska and Kansas Loess-Drift Hills) in the Central Feed Grains and
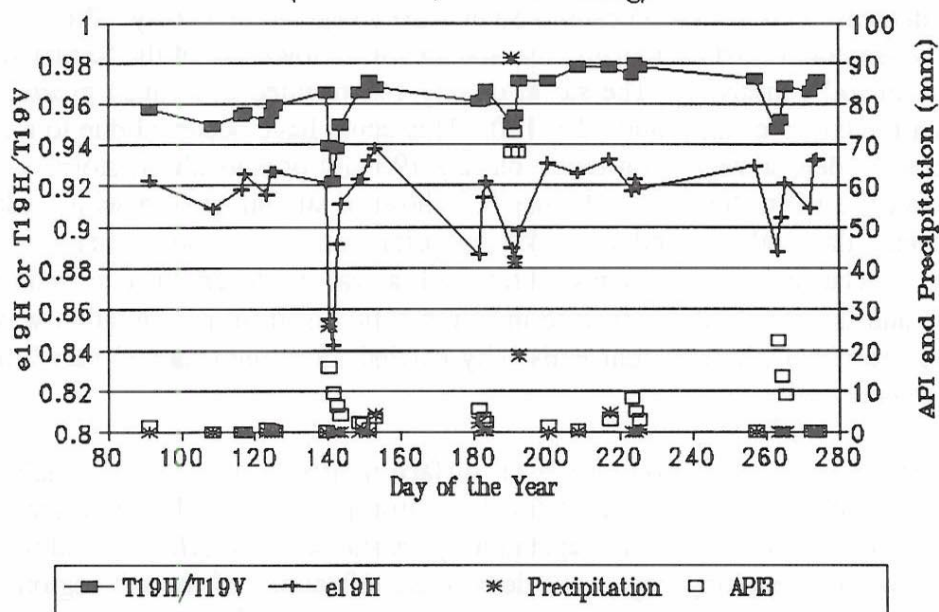
Figure 9.8 SSM/I signature response to surface moisture over a grid cell with medium density vegetation.
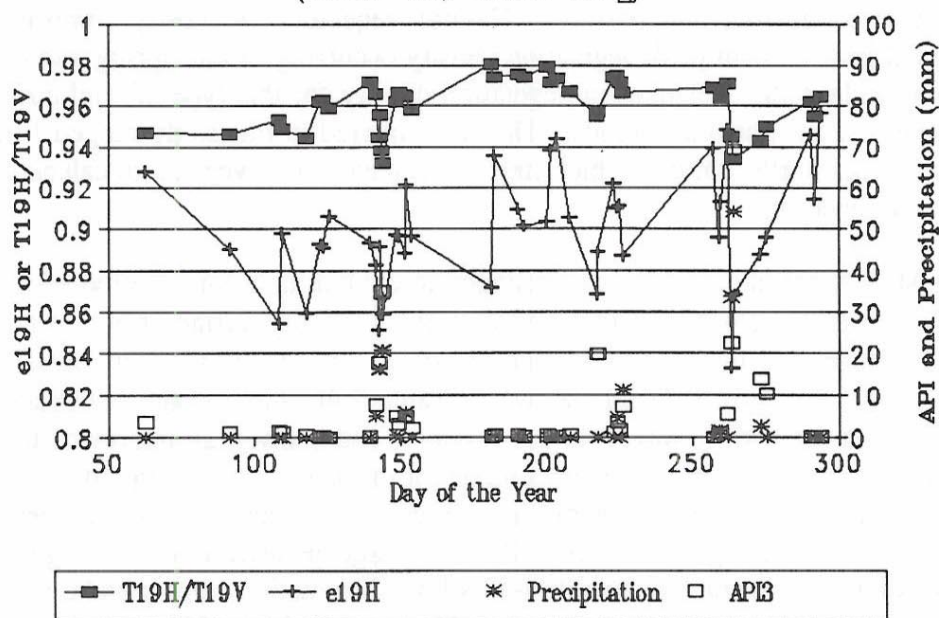
Figure 9.9 SSM/I signature response to surface moisture over a grid cell with low density vegetation at the beginning of the season.

Signature Response to Surface Moisture
(41.25 Lat, 94.75 Long)
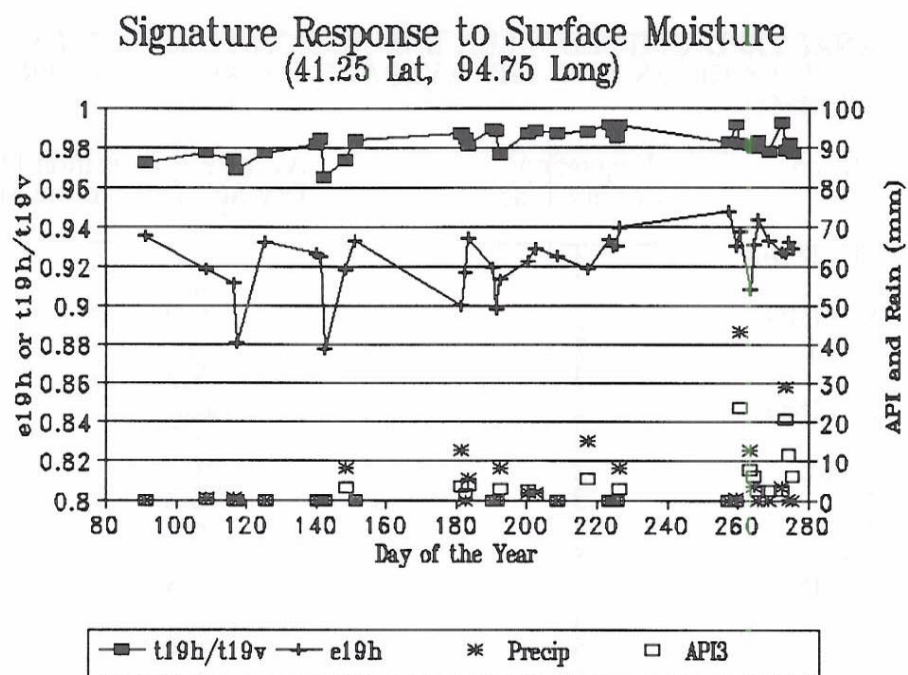
Figure 9.10 SSM/I signature response to surface moisture over a grid cell with dense vegetation.

SSM/I Signature Response to Surface Moisture for MLRA 106
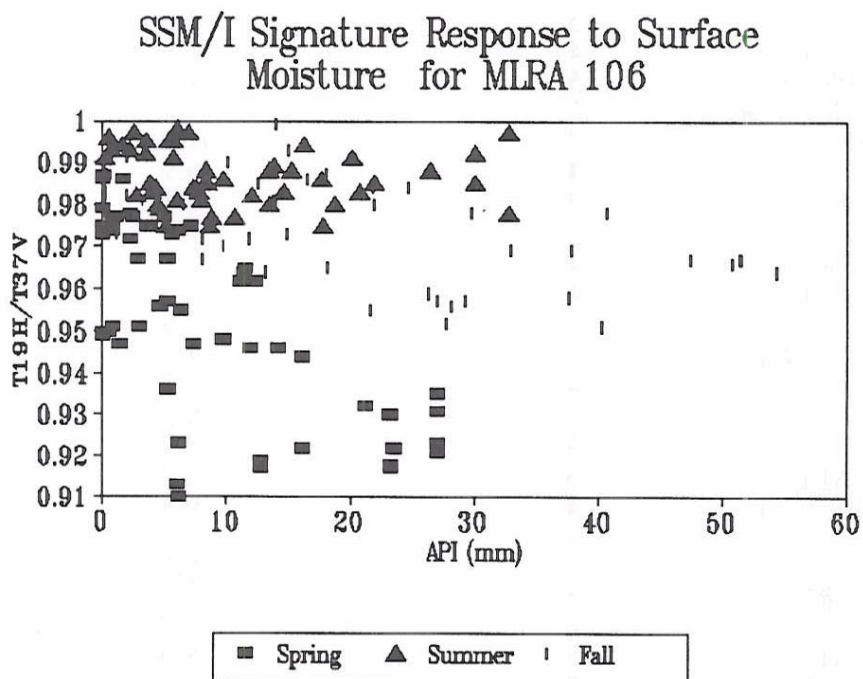
Figure 9.11 The influence of vegetation cover on the sensitivity to surface moisture for the MLRA 106 region.

9-75

TABLE 9.46 ANALYZED CLIPPED MLRA SUB-REGION FILES WITH VEGETATION DENSITY CLASS AND THE AVERAGE POLARIZATION PRIOR TO THE STORM

MLRA	Season	Vegetation Density Class	Average Polarization Over Dry Soil Prior to the Storm (K)
54	SUMMER	3	8.8
"	"	5	4.6
55B	SPRING	3	9.0
"	FALL	2	-
"	"	3	8.3
56	SPRING	2	10.6
"	"	3	9.7
"	FALL	2	10.7
"	"	3	8.2
"	"	4	6.2
57	FALL	5	5.6
67	SPRING	3	8.5
"	"	4	6.5
"	"	5	5.4
69	SPRING	5	5.9
70	SPRING	4	7.6
71	SPRING	3	8.0
"	"	5	4.0
"	FALL	6	3.2
72	SPRING	4	6.6
"	SUMMER	3	8.9
"	"	4	6.4
"	"	5	4.5
"	FALL	4	6.8
73	SPRING	5	4.4
75	SUMMER	5	5.6
77	SPRING	2	10.1
"	"	4	6.9
"	FALL	3	8.1
"	"	4	7.7
78	SPRING	4	7.9
"	FALL	4	6.3
"	"	5	4.9
80A	SPRING	4	7.1
81	FALL	4	6.5
"	"	5	5.5
"	"	3	8.7
"	FALL	3	8.3
102B	SPRING	4	6.4
"	"	4	6.0
"	"	5	6.8
106	SPRING	3	9.5
"	"	4	8.0
"	FALL	5	5.3
107	SPRING	3	8.4

Livestock Region. Corn and wheat are the main crops in the area. The average polarization in the 19 and 37 GHz channels over dry grid cells was 8.2 K for spring, 4.6 K for summer, and 7.0 K for fall. The greater sensitivity to soil moisture in the spring resulted in lower normalized temperatures as API values increased. Figure 9.12 corresponds to MLRA 103 (Central Iowa and Minnesota Till Prairies), a corn and soybean region with overall denser vegetation in the summer and fall seasons. The average polarization over dry grid cells was 7.5 K in the spring, 4.5 K in the summer, and 5.4 K in the fall.

SSM/I Signature Response to Surface Moisture for MLRA 103
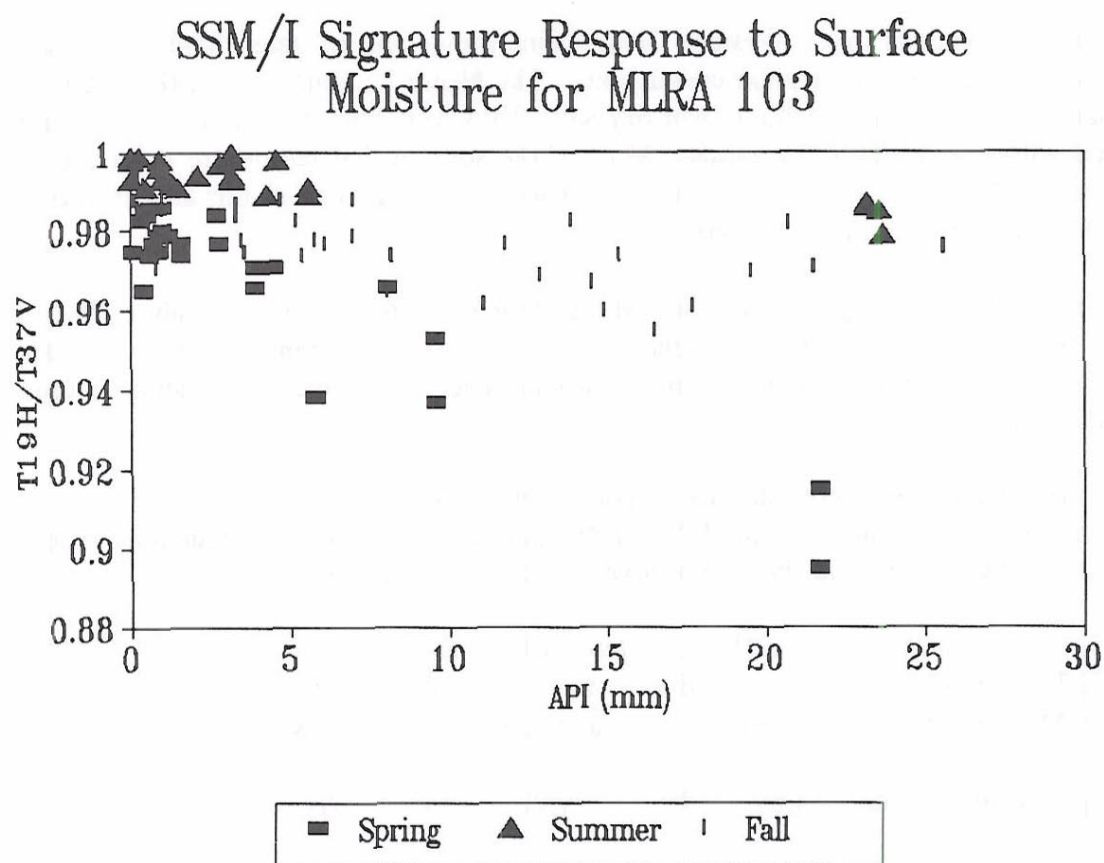
Figure 9.12 The influence of vegetation cover on the sensitivity to surface moisture for the MLRA 103 region.

Analysis of the T19H/T37V versus API relationship for the MLRA sub-region files lead to the following conclusions:

1) The strongest correlation between T19H/T37V and API occurred in the Northern Great Plains Spring Wheat Region (MLRA 54, 55B, 56), in the Central Great Plains Winter Wheat and Range Region (MLRA 70, 71, 73, 77, 78), and in the Central Feed Grains and Livestock Region (MLRA 102A, 102B, 106). The calculated polarization difference over dry

soil for those regions was usually relatively high in spring (7-10 K), low in summer (4-6 K), and high again in fall (6-10 K).

2) The Desertic Basins, Plains, and Mountains (MLRA 42) located in New Mexico and Texas had high polarizations over dry soils suggesting low vegetation densities. However, precipitation over that area was mainly due to local convective storm systems. At the SSM/I resolution scale, the correlation between T19H/T37V and API was low even over footprints with low density vegetation.

3) Observations in the Western Great Plains Range and Irrigated Region (MLRA 65, 67, 69, 70) resulted in usually good correlations. The Nebraska Sand Hills (MLRA 65) in the spring had high average polarization over dry soil. However, high API values weren't always associated with low T19H/T37V ratios. Most of the soils in that region are deep and sandy which result in fast drainage and low moisture retention. A large part of this area is also under irrigation (sprinkler and sub-irrigation).

4) The screening program classified the MLRA 119 region (which is about 76 percent forested) as high density vegetation (class 6) for the spring, summer, and fall seasons. Observations over time did not show any significant decrease in brightness temperatures for large API values.

Three main vegetation density classes were selected from the original six using regression analysis on the data of Table 9.46 and moisture retrieval equations would be developed. These classes and their respective threshold values were:

1) Low density vegetation: for avg. pol. diff. > 8 K
2) Medium density vegetation: 6 K < avg. pol. diff. <= 8 K
3) Medium high density veg.: 4 K < avg. pol. diff. <= 6 K

where avg. pol. diff. is defined by: (19V + 37V)/2 - (19H + 37H)/2.

A fourth class (dense vegetation) would encompass average polarizations of less than 4 K. However, a moisture retrieval equation was not developed for this class due to very small sensitivities.

The clipped MLRA sub-region files were randomly grouped into the three above mentioned classes according to average polarization prior to the storm. Two independent data sets were created for each class: one for algorithm development and one for verification.

9.3.4 Algorithm Development

The analysis was conducted on MLRA sub-region files stratified according to vegetation density and generated by the screening program described in 9.3.2. The objectives and general procedure of algorithm development were to:

1) Test the use of different SSM/I variables for model development and select the most sensitive to surface moisture.

2) Select the most appropriate API values according to moisture sensing depth.

3) Select the statistical model and test against independent data.

4) Develop and test the surface moisture retrieval algorithm logic.

9.3.4.1   Selection of the Best SSM/I Variable For Algorithm Development

Microwave brightness temperatures are influenced by the physical temperature of the emitting surface; therefore, the apparent emissivity (brightness temperature divided by the physical temperature) should be a more accurate indicator of surface moisture because it removes some variability in the observations due to changes in the surface physical temperature. For this reason, the apparent emissivity along with normalized 19.35 H GHz brightness temperatures using the 19.35 GHz and 37.0 GHz vertical polarization channels were compared in order to select the most significant SSM/I variable for algorithm development.

Table 9.47 shows the correlation coefficients obtained for the linear model between the different SSM/I variables tested and the API_3. The best correlation was obtained for the normalized brightness temperature T19H/T37V. Physically, this can be explained by the fact that the 37 GHz channel is closer to the skin temperature due to its smaller penetration depth. The apparent emissivity using air temperature (e19H) resulted in the worst correlation for most cases. Unfortunately, surface skin temperatures were not available for this research so air temperatures recorded by the weather station network were used instead. This introduced some additional unexplained variance to the data.

TABLE 9.47 CORRELATION COEFFICIENTS OBTAINED FOR THE LINEAR MODEL BETWEEN SSM/I VARIABLES AND API_3 FOR DIFFERENT VEGETATION CLASSES

Vegetation Density Class	Correlation Coefficient				
	T19H	e19H	T19H/T19V	T19H/T37V	2 * T19H/(T37V + T19V)
1	-0.7622	-0.6845	-0.7451	-0.7745	-0.7668
1	-0.7467	-0.6532	-0.7507	-0.7726	-0.7667
2	-0.6825	-0.5291	-0.6644	-0.6921	-0.6857
2	-0.5903	-0.4784	-0.6742	-0.6952	-0.6914
3	-0.6058	-0.4060	-0.6984	-0.7074	-0.7162
3	-0.6381	-0.5260	-0.6402	-0.6912	-0.6756

Models: 1) T19H = a + b API_3 2) e19H = a + b API_3
           3) T19H/T19V = a + b API_3 4) T19H/T37V = a + b API_3
           5) 2*T19H/(T19V+T37V) = a + b API_3.

The relationship between API and normalized brightness temperature was non linear for large API values. Therefore, API_3 values greater than 70 mm were not included in this analysis.

9.3.4.2   Selection of the API Ground Truth Values

The linear model was used to determine the best correlation between API and the normalized brightness temperature (T19H/T19V). As described in section 9.3.2, each file prepared for statistical analysis contained a set of 6 API values, five of which were computed using a recession coefficient estimated from local potential evapotranspiration with a soil water depth available for evaporation (W_m) of 5, 7.5, 10, 15, and 20 mm (designated as API_1 to API_5, respectively). Correlation coefficients obtained through this analysis are shown in Table 9.48. The best correlation among the three vegetation density classes resulted from the API_4 (W_m = 15 mm). Except for one case, the API_m resulted in poorer correlation coefficients than the API values derived from daily evapotranspiration.

TABLE 9.48 CORRELATION COEFFICIENTS FOR DIFFERENT API ESTIMATES
             USING THE LINEAR MODEL

Vegetation Density Class	Correlation Coefficient					
	API_1	API_2	API_3	API_4	API_5	API_m
1	-0.7391	-0.7629	-0.7745	**-0.7835**	-0.7769	-0.7014
1	-0.7430	-0.7637	-0.7726	**-0.7746**	-0.7603	-0.6753
2	-0.6790	**-0.6925**	-0.6921	-0.6785	-0.6573	-0.6855
2	-0.5797	-0.6682	-0.6952	-0.6943	-0.6774	**-0.7048**
3	-0.6808	-0.6983	-0.7074	**-0.7137**	-0.7058	-0.6348
3	-0.5137	-0.6389	-0.6912	-0.7263	**-0.7282**	-0.6946

Model: T19H/T37V = a + b API.

Vegetation Density Class:

CLASS 1: (T19V + T37V)/2 - (T19H + T37H)/2 > 8 K
CLASS 2: 6 K < (T19V + T37V)/2 - (T19H + T37H)/2 <= 8 K
CLASS 3: 4 K < (T19V + T37V)/2 - (T19H + T37H)/2 <= 6 K

9.3.4.3   Model Selection

The relationship between volumetric moisture content and normalized temperature is non-linear according to the Radiative Transfer Model (RTM). Such a trend was observed in the data for large API values when normalized temperatures (T19H/T19V) were plotted against API for footprints grouped by MLRA class (Figure 9.13). It appeared from the observation and regression analysis of several such cases that the relationship became non-linear for API_4 values greater than 70 mm.

T19V/T37V vs API4 for
MLRA 78 in the Fall Season
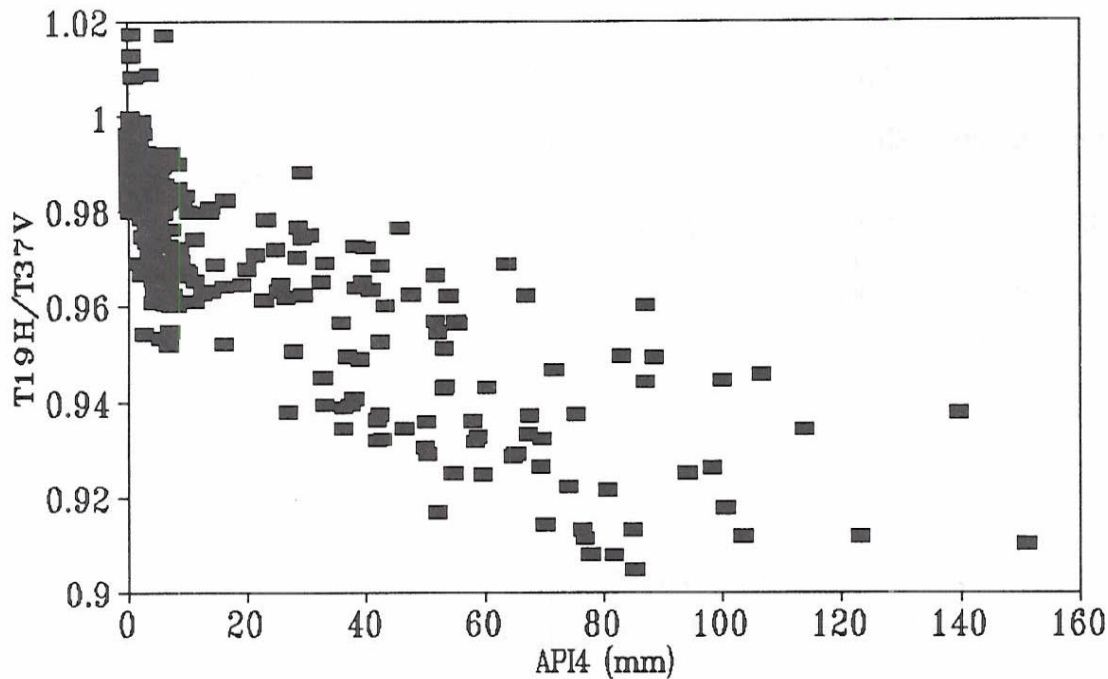
Figure 9.13 Normalized temperature versus API relationship showing the decrease in sensitivity to surface moisture for large API values.

The third model tested, which used an estimated soil surface reflectivity coefficient based on the API and the RTM, resulted in a better correlation for those data sets with very large API values, due to linearization of the data. However, that model was deemed impractical to be used for the final surface moisture retrieval equations because the surface reflectivity coefficient is implicit for API when the normalized temperature is known. In addition, and an abnormally low normalized temperature could lead to very large and unrealistic API estimates.

The linear model was determined to be the most appropriate and simple for algorithm development providing that vegetation density was taken into consideration. Therefore, observations with API_4 values greater than 70 mm were not included in the analysis so that the linear model would apply.

Curve fitting was conducted on the three regression data sets representing the three vegetation densities. The normalized temperature (T19H/T37V) was expressed as a linear combination of API_4. Table 9.49 shows the slope and intercept as well as the regression coefficients obtained. The resulting regression equations are plotted together in Figure 9.14. The absolute value of the slope was directly proportional to the average polarization in the 19.35

GHz and 37.0 GHz channels prior to the storm and thus inversely proportional to the vegetation density. The intercept (T19H/T37V for $API_4 = 0$) increased as the vegetation density increased due to the decrease in polarization. As expected, the standard error of estimate for the API increased as the vegetation density increased.

TABLE 9.49 REGRESSION COEFFICIENTS FOR THREE VEGETATION DENSITIES

Vegetation Density Class	Slope	Intercept	R	Standard Error Est. Norm. Temp.	API_4 (mm)
Low Density	-0.001481	0.9765	-0.7835	0.0167	8.8
Medium Density	-0.000873	0.9835	-0.6785	0.0151	11.7
Med High Density	-0.000580	0.9902	-0.7137	0.0098	12.0

Model: $T19H/T37V = a + b \, API_4$.

Figure 9.14 Normalized Temperature versus API_4 regression lines for three vegetation densities.

The model scatter plots and corresponding residual plots for each vegetation density class can be seen in Figures 9.15 through 9.20. The residual plots indicate that the relationship between T19H/T19V and API can be assumed linear for API_4 values less than 70 mm. The T19H/T19V residuals were larger for observations over low density vegetation (maximum of 0.05) and medium density vegetation. For relatively high density vegetation, the residuals were less than 0.03. This is in agreement with passive microwave theory as the variability in normalized temperature should be smaller for observations over vegetated areas. The largest residuals found in the data sets always corresponded to an overestimation of T19H/T37V for small API values. In other words, relatively small API values were sometimes associated with low normalized temperature values. This can be explained by the fact that small precipitation depths uniformly spread over a SSM/I footprint area just prior to the satellite overpass could result in a low normalized temperature. Contamination by water bodies not detected by the classification scheme would also produce the same effect.

Other sources of noise resulted from the methodology used in estimating the API and merging those gridded files with the SSM/I gridded files. The assumption that precipitation events occurred between 12 am and 6 pm might not have held for all cases. If rainfall occurred at night (after 6 pm), the morning SSM/I overpass would record low brightness temperatures but the computed API values would have included the recession coefficient for the previous day. In addition, some weather stations report on an evening schedule (5 or 6 pm) and a precipitation event occurring after 6 pm would be considered the next day. These two facts could lead to observations in the data where abnormally low normalized temperatures were associated with API values of zero. The alternative however, would have been to group the morning SSM/I overpasses with API values based on precipitation of the same day. This would have lead to high API values associated to high normalized temperatures for precipitation occurring after 6 am. These cases would have been more common, producing numerous leverage points in the data which would have artificially reduced the slope of the regression line. The heterogeneity of precipitation over a 30 km grid cell even for the large frontal systems used in this analysis resulted in some observations with high API values being paired with normalized temperatures higher than expected. Such problems are unavoidable at the spatial resolution of the SSM/I.

9.3.4.4 Model Testing

The regression equations for each vegetation density class were tested with independent data sets described in section 9.3.2. These equations are shown in Table 9.50 in their operational form, as inversions of the equations developed in Table 9.49. If API values predicted by the regression equations were negative, they were set to zero, and if they were larger than 70 mm, they were set to 70 mm. Figures 9.21 to 9.23 show plots of the predicted API versus actual "ground truth API" values. The correlation between predicted API and actual "ground truth" API was satisfactory for the three vegetation density classes. The best correlation occurred for the low density vegetation class (R = 0.7686).

Scatter Plot and Regression Line
Low Density Vegetation (Class 1)
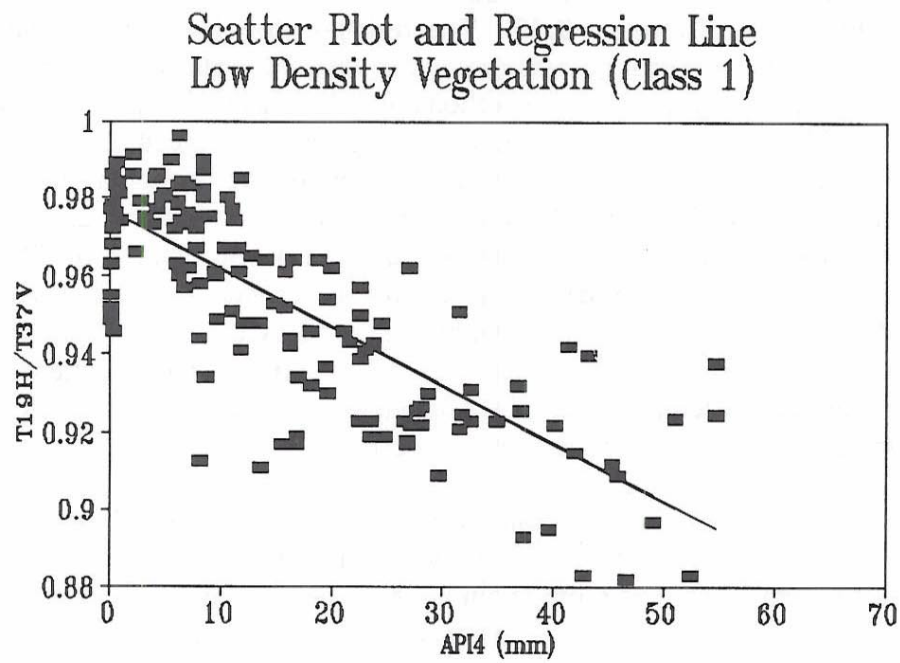
Figure 9.15 Scatter plot and regression line for the low density vegetation class (R = -0.7835).

Residual Plot
Low Density Vegetation (Class 1)
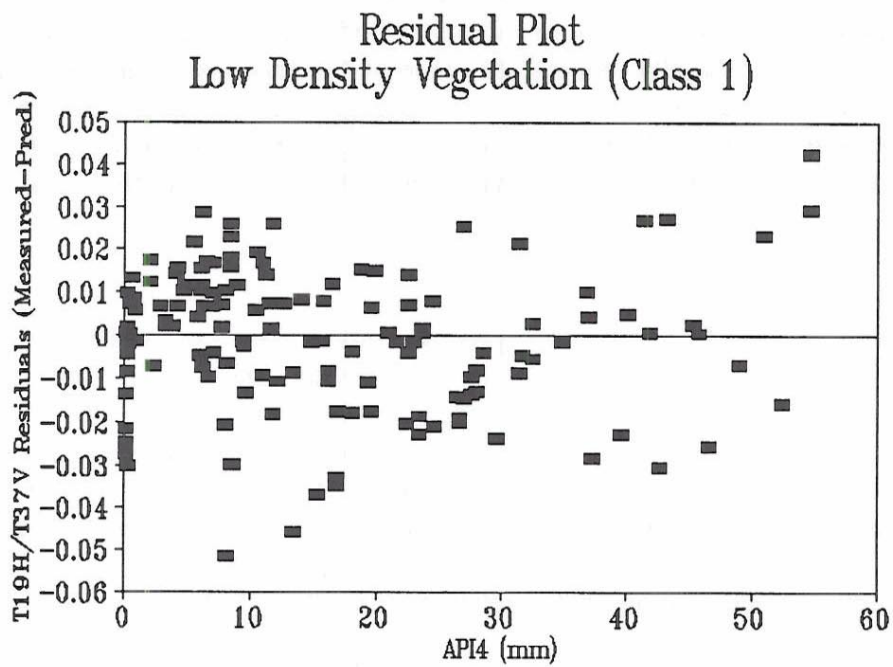
Figure 9.16 Residual plot for the low density vegetation class.
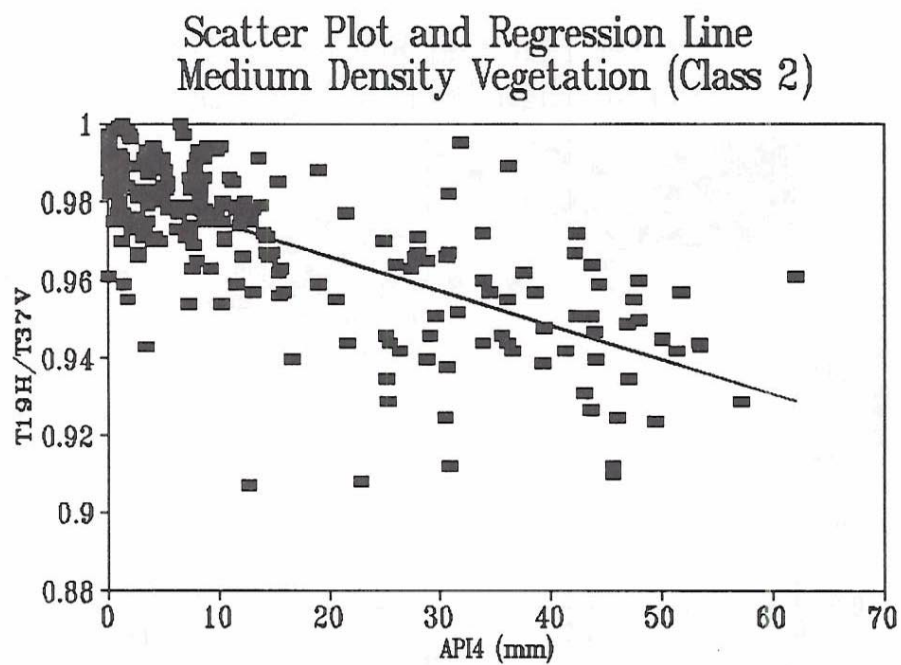
Figure 9.17 Scatter plot and regression line for the medium density vegetation (R = -0.6785).
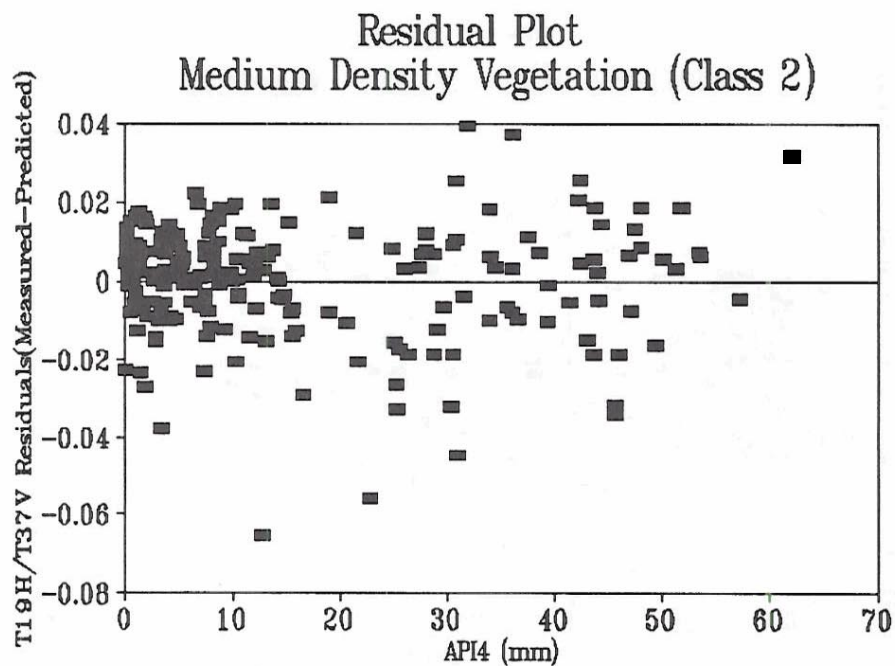
Figure 9.18 Residual plot for the medium density vegetation class.
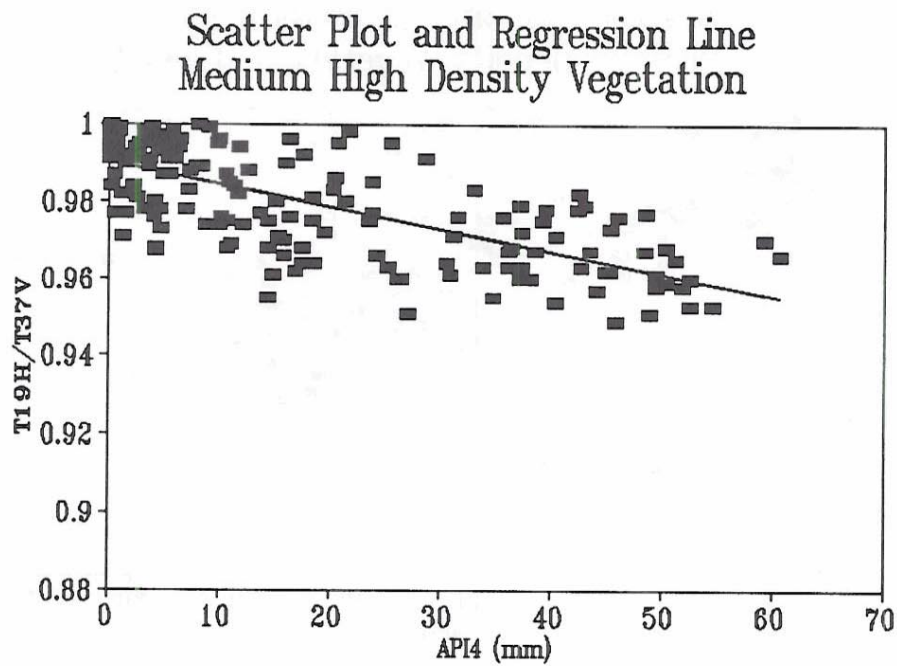
Figure 9.19 Scatter plot and regression line for the medium high density vegetation class (R = -0.7137).
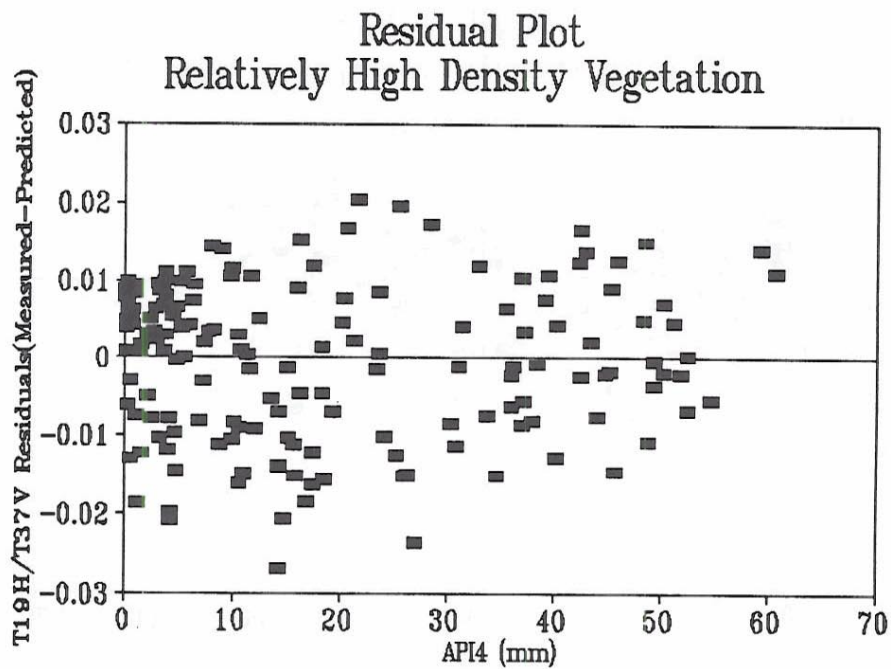
Figure 9.20 Residual plot for the medium high density vegetation class.

TABLE 9.50 RECOMMENDED SURFACE MOISTURE RETRIEVAL ALGORITHMS FOR THREE VEGETATION DENSITY CLASSES

Vegetation Density Class	B1	A1	TEST
Low Density Vegetation	-675.22	659.35	[b] > 8 K
Medium Density Vegetation	-1145.48	1126.58	6 < [b] ≤ 8 K
Med High Density Vegetation	-1724.14	1707.24	4 < [b] ≤ 6 K

$$API = A1 + B1 \ T19H/T37V \qquad [b] \ \frac{19V + 37V}{2} - \frac{19H + 37H}{2}$$
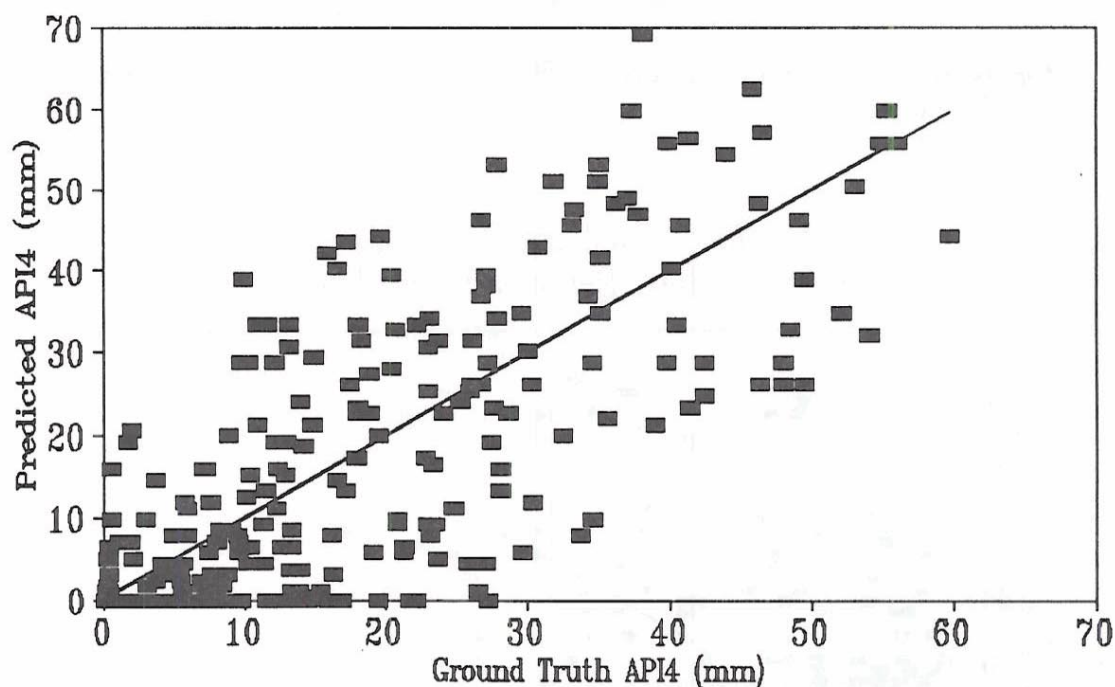
Figure 9.21 Plot of predicted versus actual API_4 surface moisture values for the independent data set (R = 0.7686).

Surface Moisture Retrieval
Medium Density Vegetation
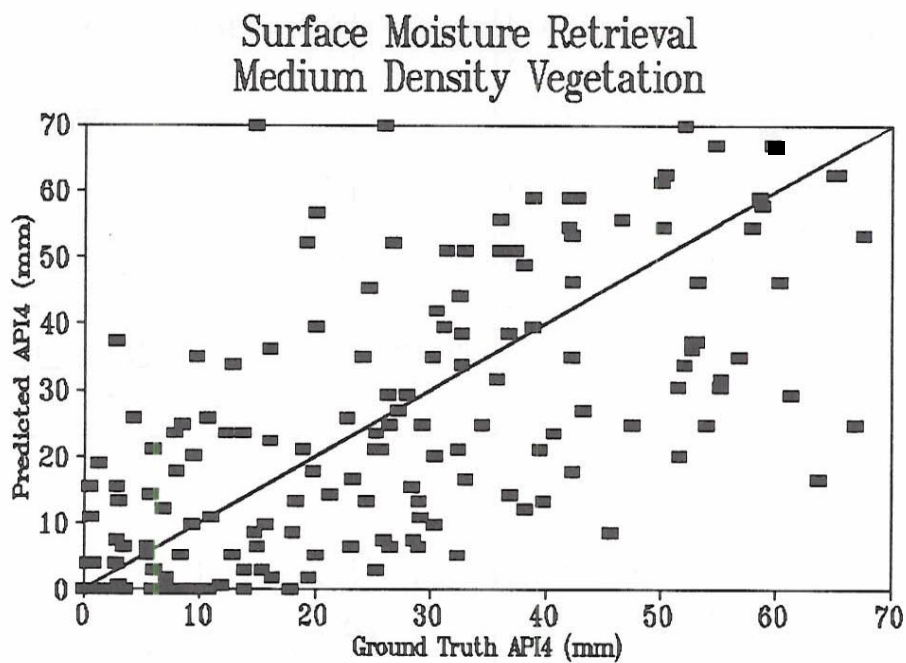
Figure 9.22 Predicted versus actual API_4 values for the independent data set (R = 0.6871).

Surface Moisture Retrieval
Medium High Density Vegetation
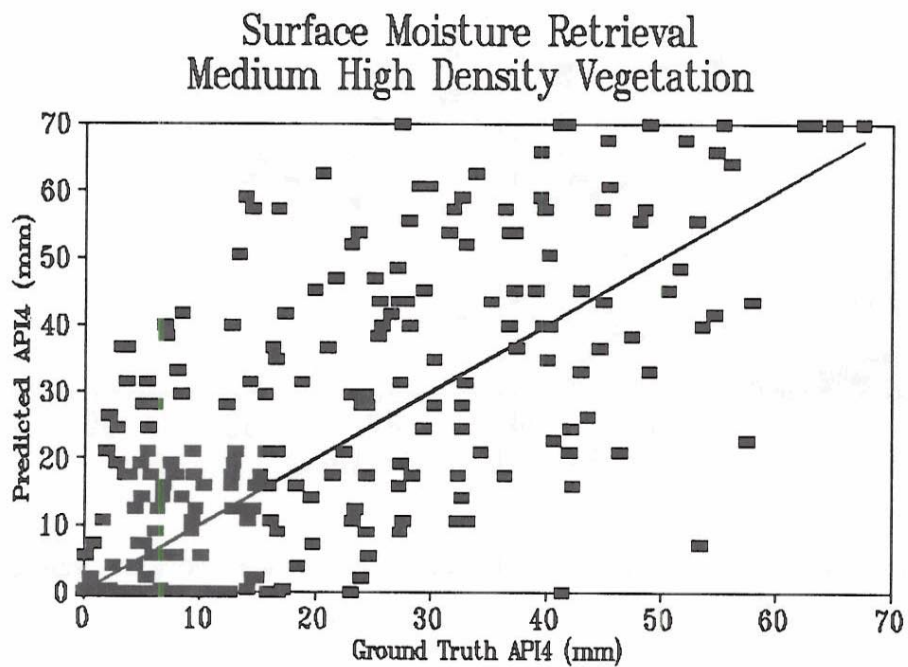
Figure 9.23 Predicted versus actual API_4 values for the independent data set (R = 0.7226).

9.3.4.5 Surface Moisture Prediction Algorithm

The normalized temperature ratio does not vary considerably from day to day over dry soil conditions. In regions with similar climatic conditions to the Central Plains of the U.S., it will slowly increase and decrease over the growing season following the growth and senescence of natural vegetation or the seasonal variations in agricultural vegetation densities. However, if the vegetation is not too dense, a precipitation event will cause a sharp decrease in the ratio which will gradually, over a period of time, return to its value prior to the event, assuming that the vegetation density has not changed considerably during the period.

Thus, for the best use of the developed models, we recommend that the algorithm be implemented in the dynamic database framework described in the land surface type classification section of this report (section 9.1). This implies calculating and storing a running average of certain SSM/I variables for grid cell locations of interest, which are updated at each available overpass of the instrument.

The following steps are recommended for the use of the algorithm:

1) Compute a running average of T19H/T37V and of the average polarization in the 19 and 37 GHz channels for each overpass and grid cells in the area of interest. The average polarization is used as a vegetation density index while the T19H/T37V normalized temperature is the indicator of surface moisture. The running averages would include brightness temperatures for the five last overpasses.

2) Before including the SSM/I variables from the latest overpass in the running averages, compare T19H/T37V to its running average. If T19H/T37V is not significantly different and the surface type classification code does not indicate moisture, the soil is considered dry. If a significant reduction in T19H/T37V has occurred and the surface type classification code indicated moisture, the surface is considered moist or wet.

3) If the soil surface is determined to be dry, include the latest values for the SSM/I variables in the running averages.

4) If the soil is classified as moist, the latest values for the SSM/I variables should not be included in the running averages. The value presently in the database for the average polarization in the 19 GHz and 37 GHz channels is used to select the appropriate surface moisture retrieval equations for that vegetation density class.

5) For subsequent overpasses, estimate surface moisture using the selected equation until the predicted API reaches zero or until the T19H/T37V normalized temperature is close to the running average value prior to the storm.

The algorithm was applied to many grid cells representing different MLRA regions in the central plains over the snow-free period in 1988. Examples are shown in Figures 9.24 and

Surface Moisture Retrieval Algorithm
(MLRA 55B, Lat: 45.3, Long: 98.3)

Legend: —□— Predicted API   —— Ground Truth API
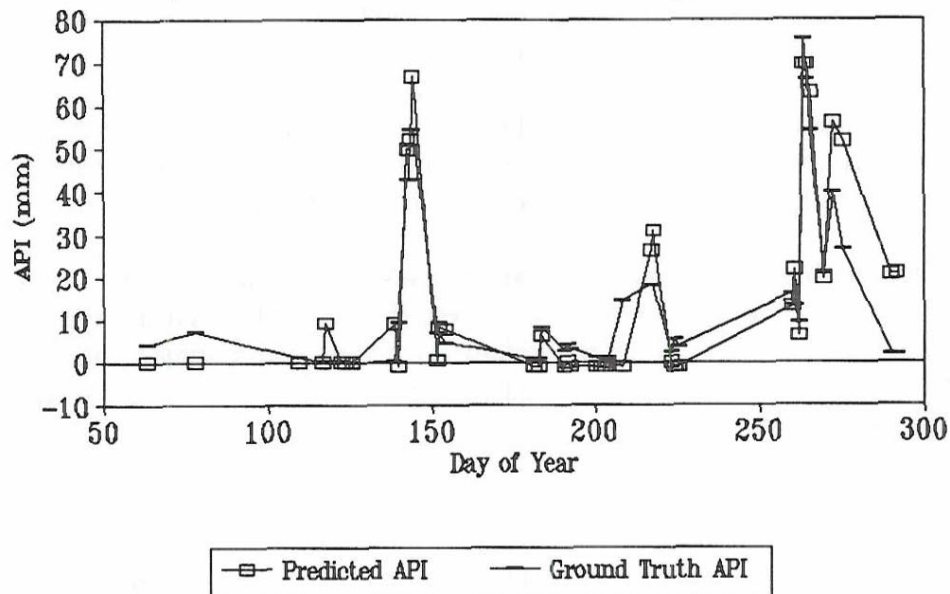
Figure 9.24 Application of the surface moisture retrieval algorithm to a grid cell in the Central Black Glaciated Plains Region in North Dakota during the snow-free period in 1988.

Surface Moisture Retrieval Algorithm
(MLRA 102, Lat: 44.8, Long: 97.0)

Legend: —□— Predicted API   —— Ground Truth API
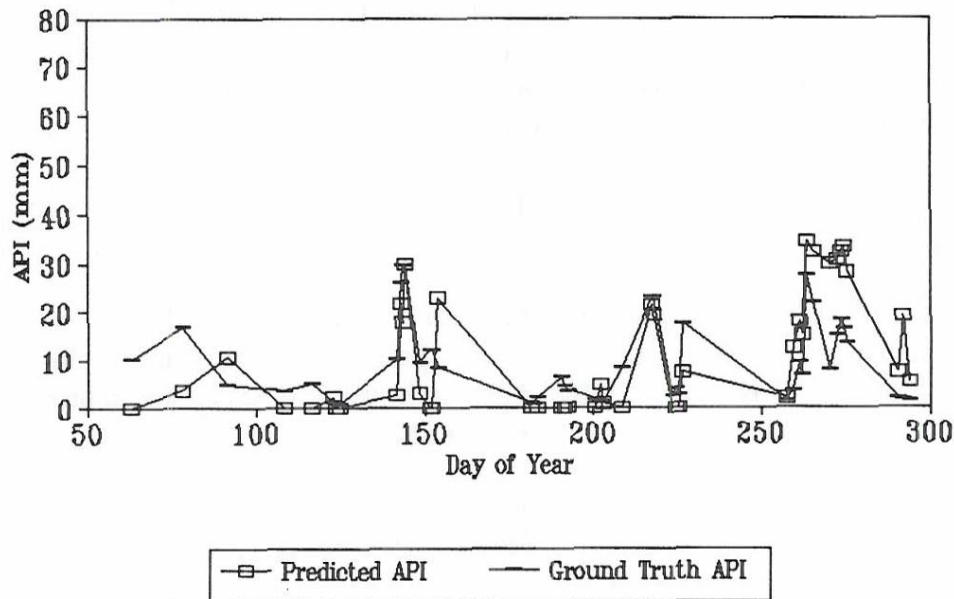
Figure 9.25 Application of the surface moisture retrieval algorithm to a grid cell in the Rolling Till Prairie Region of Eastern South Dakota, during the snow-free period of 1988.

9.25 for two different MLRA regions. The algorithm reasonably predicted surface moisture for both large and small precipitation events.

9.3.5 Observations and Recommendations

The developed surface moisture retrieval algorithms are satisfactory considering the physical limitations of the SSM/I instrument for this purpose. The short wavelengths of the SSM/I channels only permit very small soil moisture sensing depths. Vegetation is an additional complication which causes further decreases in moisture sensitivity. The large footprint size at 19 GHz introduces unavoidable noise due to the spatial variability in surface types as well as the random nature of precipitation and consequently soil moisture at that scale. Therefore, these algorithms are a compromise, retrieving surface moisture with the API surrogate while taking into consideration vegetation density effects. Surface type variability effects are partially removed in the classification scheme as the retrieval equations only apply to certain surface types i.e. moist soil and arable soil. Under flooded conditions, the API will usually be greater than 70 mm rendering accurate retrievals almost impossible due to the non-linear nature of the response.

The most accurate operational use of these algorithms will require the maintenance of running averages for the appropriate SSM/I variables within a dynamic database continuously updated with each overpass as described in section 9.3.4.5. The running average of SSM/I variables are necessary for establishing the vegetation cover density and selecting the appropriate retrieval equation.

If the use of the dynamic database is not possible, the retrieval equations can be used over certain surface type classifications providing the predictions are limited to zero for the lower limit and 70 mm for the upper limit. It should be understood that the errors in surface moisture retrievals could be greater than the standard errors described in the analysis section. The applicable surface type classifications would be moist soils and wet soils. The equations would not be applicable to the other surface types.

Theoretically, volumetric soil moisture to a certain depth can be estimated from a time series of soil surface temperature and moisture, both retrievable with the SSM/I. This would require the knowledge of certain soil physical characteristics such as water holding capacity, infiltration and hydraulic conductivity. At the spatial resolution of the SSM/I however, any such modelling attempt would be questionable. As an alternative, a gross value of soil moisture for a grid cell location could be estimated during a limited period after a storm by assuming an average value for the soil water holding capacity of the grid cell and distributing the retrieved surface moisture down to a certain depth of the soil profile. The average value of soil water holding capacity would be stored in the database as well.

9.3.6 References

[1]   T. P. Schmugge, P. Gloersen, T. Wilheit, and F. Geiger, "Remote sensing of soil moisture with microwave radiometers," J. Geophys. Res., vol. 79, pp. 317-323, 1974.

[2]   T. J. Schmugge, "Effect of texture on microwave emission from soils," IEEE Trans. Geosci. Remote Sensing, vol GE-18, pp. 353-361, 1980.

[3]   J. R. Wang, P. E. O'Neill, T. J. Jackson, and E.T. Engman, "Multifrequency measurements of the effects of soil moisture, soil texture and surface roughness," IEEE Trans. Geosci. Remote Sensing, vol. GE-21, pp. 44-51, 1983.

[4]   M. J. McFarland, "The correlation of Skylab L-Band brightness temperatures with antecedent precipitation," Conference on Hydrometeorology, AMS, Boston, MA, 1976.

[5]   M. J. McFarland and B.J. Blanchard, "Temporal correlations of antecedent precipitation with Nimbus 5 ESMR brightness temperatures," Second conference on hydrometeorology, AMS, Boston, MA, 1977.

[6]   G. D. Wilke and M. J. McFarland, "Correlations between Nimbus 7 scanning multichannel microwave radiometer (SMMR) data and an antecedent precipitation index," JCAM, vol. 25, pp. 227-238, 1986.

[7]   B. J. Choudhury, M. Owe, S. N. Goward, R. E. Golus, J. P. Ormsby, A. T. C. Chang, and J. R. Wang, "Quantifying spatial and temporal variabilities of microwave brightness temperature over the U.S. Southern Great Plains," Int. J. Remote Sens., vol. 8, no. 2, pp. 177-191, 1987.

[8]   M. Owe, A. Chang, and R. E. Golus, "Estimating surface moisture from satellite microwave measurements and a satellite derived vegetation index," Remote Sens. Environ., vol. 24, pp. 331-345, 1988.

[9]   R. W. Newton and J.W. Rouse, Jr., "Microwave radiometer measurements of soil moisture content," IEEE Trans. Antennas and Propag., vol. AP-28, pp. 680-686, 1980.

[10]   J. R. Wang, J. C. Shiue, and J. E. McMurtrey, "Microwave remote sensing of soil moisture content over bare and vegetated fields," Geophys. Res. Letters, vol. 7, no. 10, pp. 801-804, 1980.

[11]   H. K. Burke and T. J. Schmugge, "Effects of varying soil moisture contents and vegetation canopies on microwave emissions," IEEE Trans. Geosci. Remote Sensing, vol. 20, no. 3, pp. 268-274, 1982.

[12] S. W. Theis and A. L. Blanchard, "The effect of measurement error and confusion from vegetation on passive microwave estimates of soil moisture," Int. J. Remote Sens., vol. 9, no. 2, pp. 333-340, 1988.

[13] F. T. Ulaby, M. Razani, and M. C. Dobson, "Effects of vegetation cover on the microwave sensitivity to soil moisture," IEEE Trans. Geosci. Remote Sensing, vol. 21, no. 1, pp. 51-61, 1983.

[14] C. M. U. Neale, M. J. McFarland, and K. Chang, "Land surface type classification using microwave brightness temperatures from the Special Sensor Microwave/Imager," IEEE Trans. Geosci. Remote Sensing, vol. 28, no. 5, pp. 829-838, 1990.

[15] G. H. Hargreaves, and Z. A. Samani, "Reference crop evapotranspiration from temperature," Trans. Am. Soc. Agric. Eng. vol. 1, no. 2, pp. 96-99, 1985.

[16] Soil Conservation Service, "Land resource regions and major land areas of The United States," United State Department of Agriculture, SCS, Washington, DC, 1981.

[17] M. J. McFarland, I. R. McCann, and K. S. Kline, "Synthesis and measurement of temperature for insect models," ASAE Monograph, St. Josephs, MI, 1991.

[18] F. T. Ulaby, R. K. Moore, and A. K. Fung, Microwave Remote Sensing: Active and Passive, Volumes I, II, and III, Reading, MA: Addison-Wesley, 1981, 1982, and 1986.

[19] M. T. Hallikainen, F. T. Ulaby, M. C. Dobson, M. A. El-Rayes, and L. K. Wu, "Microwave dielectric behavior of wet soil - part 1: empirical models and experimental observations," IEEE Trans. Geosci. Remote Sensing, vol. 23, no. 1, pp. 25-34, 1985.

9.4   SNOW PARAMETER ALGORITHMS

9.4.1   Algorithm Development Rationale

Passive microwave radiometry has significant promise for the remote sensing of snowpack properties. Snow particles behave as volume scatterers of the radiation emitted from the underlying surface [1], [2]. The scattering is a strong function of wavelength. At wavelengths less than 1 cm, Mie scattering by individual snow particles such as crystals and grains is pronounced [3]. The longer wavelengths, such as the 1.66 cm (18 GHz) of the Nimbus 7 SMMR and the 1.55 cm (19 GHz) of the SSM/I, are scattered less by a typical snowpack. Consequently, radiation at these wavelengths emitted from a snow covered surface will be a function of the state of the surface beneath the snowpack. Frozen ground has a high emissivity, greater than 0.90, due to the low permittivity of ice [4]. Moist soil with free water present has an emissivity as low as 0.70 at these wavelengths due to the high permittivity of liquid water [5]. The physical temperature of a dry snowpack is not a major influence due to the low contribution of emitted radiation from the snow [6]. In reality, the passive microwave radiation received from a snowpack is a function of the frequency distribution of snow crystal and grain sizes. This frequency distribution is highly correlated with snow depth and with snow water content for typical snow densities.

Frozen ground prior to a snowfall event will have fairly similar brightness temperatures, with low polarization differences, at the SSM/I frequencies. Vegetation and roughness elements will decrease the polarization differences, while bare and moist soil will increase the polarization differences. The primary effect of a new, dry snowfall will be to depress the brightness temperatures in the shorter wavelengths (higher frequencies). If the snow crystals are very small, the 85 GHz channels will show a marked drop in brightness temperature. Pronounced decreases in the 37 GHz channels are more typical of a new snow. The polarization differences will also increase dramatically. For new, dry snow of the order of tens of centimeters depth, the brightness temperatures in the longer wavelengths are essentially unchanged from those prior to the snowfall.

The crystalline structure of a new snowpack initially is influenced by snow crystal size, dryness, temperature, and wind. The crystalline structure will change on a day to day basis as a result of thermal gradients in the snowpack and the energy balance of the snow surface layer. A net effect of both processes - hoar crystal formation in the lower layer of the snowpack [7], [8] and larger grain sizes and layer formation from thaw and freeze cycles - is to progressively increase the mean crystal size. This will produce a further decrease in brightness temperatures at 37 GHz (see, for example, [9]). Large crystals will decrease the radiation at the 22 and 19 GHz frequencies, as will be shown.

Snow edge is relatively easy to detect with passive microwave. The radiation from the underlying surface is scattered more at 0.81 cm than at 1.55 or 1.66 cm. The polarization difference increases markedly at 0.81 cm. Grody [10], Kunzi, et al. [11], and McFarland, et al. [9] used comparisons of 0.81 cm with 1.66 cm brightness temperatures or polarization

difference at 0.81 cm to detect snow edge or discriminate snow covered areas from areas without snow with SMMR data. The threshold was of the order of 2 cm snow depth, although others have noted a threshold of 5 cm [12].

Investigations on determination of snow depth or snow water equivalent have focused on the 37 GHz channels [9], [11], [12], [13]. Foster, et al. [3] reported coefficients of determination approaching 0.9 at the 0.81 cm wavelength (37 GHz) of the Nimbus 6 ESMR. These correlations were obtained with the vertically polarized brightness temperatures for one degree latitude-longitude cells in North Dakota and Montana. Chang [14] used the difference between the 1.66 (18 GHz) and 0.81 cm horizontally polarized brightness temperatures of SMMR to retrieve snow depths over several large open land areas in Canada, the U. S. Great Plains, and central Russia. Correlation coefficients of 0.85 were obtained. Gloersen, et al. [12] retrieved snow water equivalent as a linear function of the brightness temperature difference between the 1.66 and 0.81 horizontally polarized brightness temperatures from SMMR. Goodison, et al. [15] found excellent correlations between snow depth and the 37V channel of an airborne radiometer, with the ground truth from airborne gamma and surface snow surveys. The coefficient of determination was 0.86 and the slope of the linear regression line was 1.83 mm/K.

Kunzi, et al. [11] noted that microwave signature of a snowpack was independent of depth for dry snow depths greater than 50 cm. McFarland, et al. [9], in their study of snow depths in the northern Great Plains, found the upper threshold to be somewhat lower, around 40 cm. This upper threshold apparently marks the depth where all emitted radiation from the underlying surface is scattered or absorbed. For the snow depths above the threshold, the radiation is a function of the crystal morphology in the pack and reflected radiation from the crystals and internal layers.

Ideally, the brightness temperatures before the first snowfall would be incorporated into the algorithm to retrieve the snow depth or water equivalent [6]. The decrease in microwave emission due to snow would be a result of the scattering. The algorithms that combine the 37 GHz channels with lower frequency channels represent an attempt to incorporate the pre-snow passive microwave signature into the algorithm. This procedure would be especially useful with varying vegetation, soil, and soil moisture within a region. This procedure is not feasible in the D-matrix algorithm approach, however. Incorporation of a longer wavelength (lower frequency) in the snow depth or water equivalent algorithm could provide information on the state of the ground underlying the snowpack [14].

Any liquid water in the snowpack increases the microwave brightness temperatures [1], [14]. A change of one percent in liquid water results in a change of 70 K in the 0.81 cm horizontal polarization brightness temperature [4].

McFarland, et al. [9] separated the snow season into two phases; the accumulation phase and the ripening and melting phase. Schanda, et al. [16] had essentially the same classification scheme with winter snow (no melting metamorphism), wet spring snow (with a layer of wet

snow crystals at the surface), and dry, refrozen spring snow. The brightness temperatures and the polarization differences do not have the same patterns after the onset of crystal metamorphism produced by melting. Daytime melting produces marked increases in brightness temperature [9], [17]. Nightime refreezing after a daytime melt does not return the brightness temperatures to pre-melt values. McFarland, et al. [9] noted a gradual increase in the nighttime brightness temperatures from the onset of the ripening and melting phase to full ripeness and melting. The obvious implication is that an algorithm to retrieve the snowpack parameters has to initially discriminate between these phases or classes. Different algorithms are needed: one to determine snow depth or water equivalent before the melt phase and another to determine the stage of ripening.

9.4.2 Methodology

The data sets of SSM/I brightness temperatures and climatological data were assembled as 0.25 or 0.5 degree grid files previously described. The climatological data consisted of snowfall in the preceding 24 hours, total snow depth, and water equivalent of the new snow. Daily air maximum and minimum temperatures were also available.

Several separate data sets were analyzed. The full data set consisted of the SSM/I brightness temperatures from day 343, 1989 to day 60, 1990 for the Central Plains test areas for quarter degree grid boxes. The cases analyzed included those overpasses, both ascending and descending, when the test area was largely covered by the overpass. The SSM/I data set consisted of 344a, 344d, 346d, 353d, 363d, 007a, 008a, 024a, 047a, 049a, 050a, 055a, 056a, 057a, and 058a, where each overpass is identified by the calendar day number and a for ascending or d for descending. Only those grid boxes with a climatological reporting station were included in the analysis. Additional data sets were processed from February 1988.

9.4.3 Results and Discussion

The snow depths and the microwave brightness temperatures are highly correlated, as shown in Figures 9.26 through 9.29. Figure 9.26 shows the reported snow depths and Figures 9.27, 9.28, and 9.29 the SSM/I brightness temperatures at 19V, 37V, and 85V GHz respectively over the Central Plains for day 51, 1988, ascending pass. Figure 9.30 shows the minimum surface air temperature for this same day. A visual correlation of the snow depth with the brightness temperatures appears to show excellent agreement. However, when multiple linear regression was performed, the best R squared was in the vicinity of 0.20 with an RMSE of 11 cm. The snow depths for this case generally match the observed depressions in the SSM/I channels, but the localized nature of the heavier amounts may be a source of variance. A geolocation correction was not applied to the SSM/I data.

The full data set was analyzed spatially and temporally for two separate sections of the Central Plains, as shown in Figure 9.31. The eastern area was defined by 41 to 47 degrees north latitude and 88 to 96 degrees west longitude. This area covers Iowa, Minnesota, and the

SNWD51A
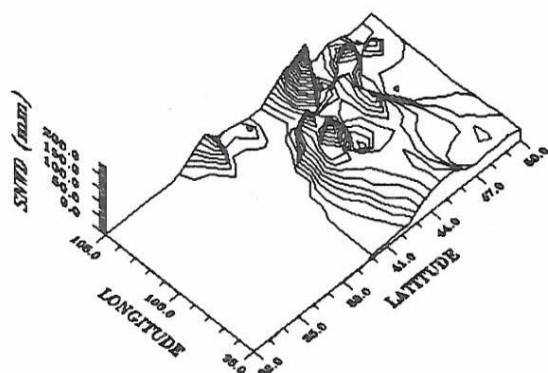
T19V51A

Figure 9.26 Accumulated snow depth for the Central Plains on day 51, 1988.
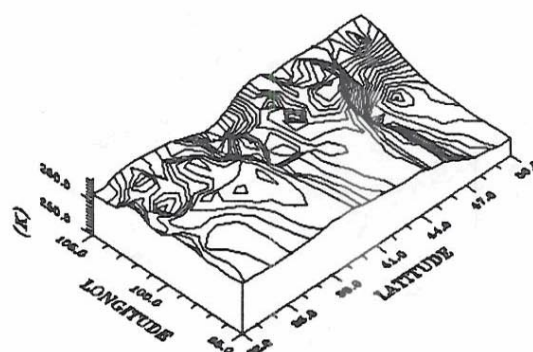
Figure 9.27 SSM/I brightness temperatures at 19V over the Central Plains on day 51, 1988.

T37V51A
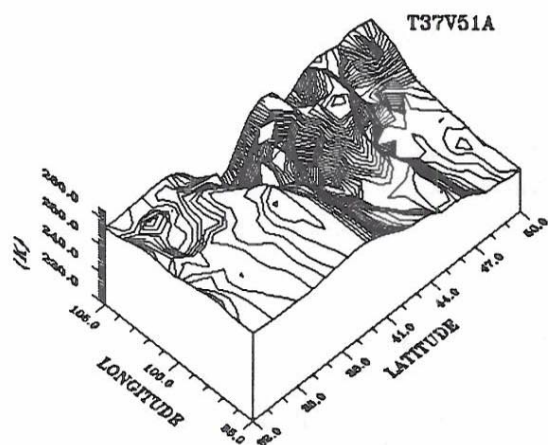
T85V51

Figure 9.28 SSM/I brightness temperatures at 37V over the Central Plains on day 51, 1988.
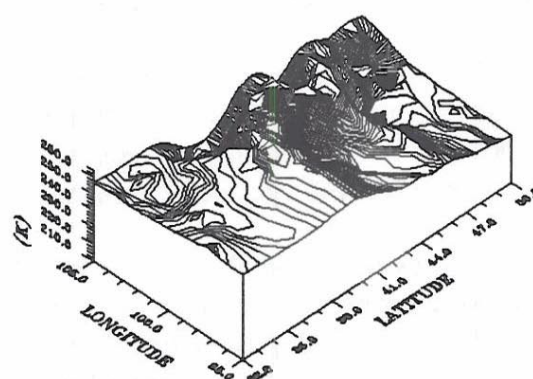
Figure 9.29 SSM/I brightness temperatures at 85V over the Central Plains on day 51, 1988.
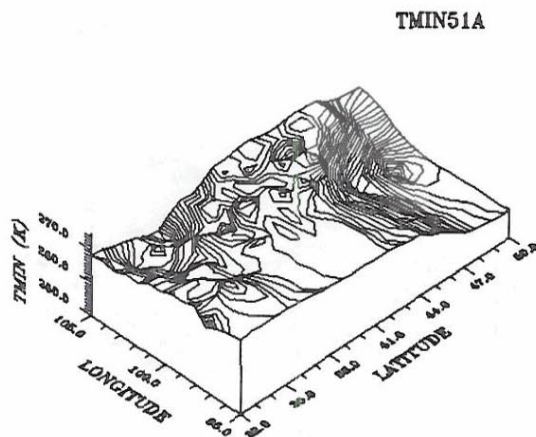
Figure 9.30 Observed surface minimum air temperature for the Central Plains for day 51, 1988.

Figure 9.31 The test area for the analysis of snow.

western half of Wisconsin. The western area was defined by 41 to 47 degrees north latitude and 96 to 104 degrees west longitude. This area includes North Dakota, South Dakota, and the northern two-thirds of Nebraska. Spatial data sets were organized in a spread sheet by day, with all seven brightness temperatures, the major land resource area [18], the classification by the EXTLND surface type classification module, and the climatological data. The first step in the analysis was to calculate the correlation coefficient matrix for all variables. This correlation analysis included several derived variables from the brightness temperatures, such as channel and polarization differences. Separate correlations were calculated for various categories of snow depth, major land resource area, and location.

The snow season for the winter of 1989 in the test area was characterized by a few major snowfalls that melted/sublimed significantly in the several week periods between the snowfalls. The snow cover reported in the Weekly Weather and Crop Bulletin showed very little snow on January 16, 1989 and again on February 12, 1990.

Figures 9.32 and 9.33 show the correlation coefficients between the 19V and 37V brightness temperatures and snow depth for days between 344, 1989 and day 58, 1990. All grid cells with snow depths greater than 0 mm and less than 400 mm were used in the correlation analysis. No stratification was done for land surface type and no points were removed based on obvious outlier locations, such as the Black Hills of South Dakota. Several patterns are readily apparent. The 37V channel has the higher correlation coefficient than the 19V channel, but the 19V channel shows a marked response to snow depth (actually grain size characteristics
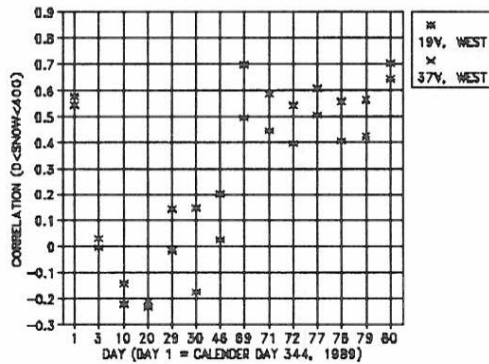
Figure 9.32 Correlation coeff-
icients between snow depth and
37V brightness temperatures,
western test area.
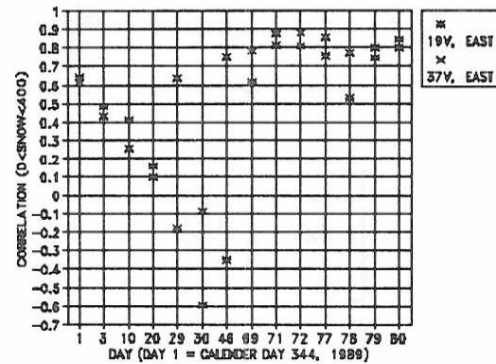
Figure 9.33 Correlation coeff-
icients between snow depth and
37V brightness temperatures,
eastern test area.

within the snowpack). This indicates that the 19V channel will not necessarily provide information from the surface beneath the snowpack. The 19V channel will provide information on the aging of the snowpack and the development of larger crystals in response to surface thawing and refreezing and to hoar crystal formation in response to thermal gradients.

Another pattern is that the correlation coefficients are highly erratic with conditions of light and decreasing snow amounts. Correlation coefficients that remained fairly stable from day to day did not occur until late February. For days 047, 050, 055, 056, 057, and 058 which were all ascending overpasses, correlation coefficients were calculated between snow depth in mm and the 37V brightness temperature in K for all grid cells with the land surface category of snow present (EXTLND). These correlation coefficients ranged from -0.50 to -0.84. The correlation coefficients were higher for the eastern test area, although the regression coefficients were similar. Selected scatter plots and descriptive statistics are presented in Figures 9.34 through 9.37. The 85V channel is considered unreliable in these data sets due to the high standard deviations and the means less than those of the 85H channel. In Figures 9.34 and 9.35 for the western test area, the maximum snow depths in the data set were less than 400 mm and the relationships between snow depth and the 37V brightness temperature were fairly linear. In contrast, note the relationship between snow depths greater than about 400 mm and the 37V brightness temperatures in Figures 9.36 and 9.37 for the eastern test area. As noted in previous investigations, the passive microwave response is significantly decreased with snow depths greater than about 400 mm.

As shown by the statistics in Tables 9.51 through 9.54, which correspond to Figures 9.34 through 9.37, the correlation coefficients were high for all channels, which is expected due to the high intercorrelations between the channels (excluding the 85V). The 37V channel was consistently a better predictor of snow depth, which is consistent with several other investigations. Several combinations of channels were also examined. These included the

9-99
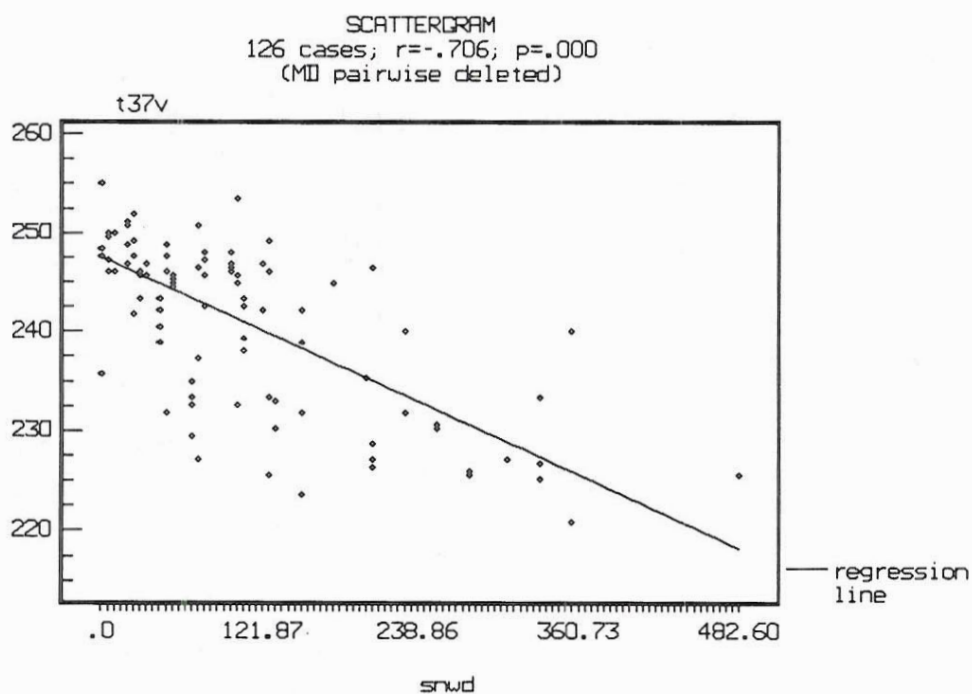
Figure 9.34   37V GHz brightness temperature and reported snow depth for the western test area, day 47, 1990, ascending pass.
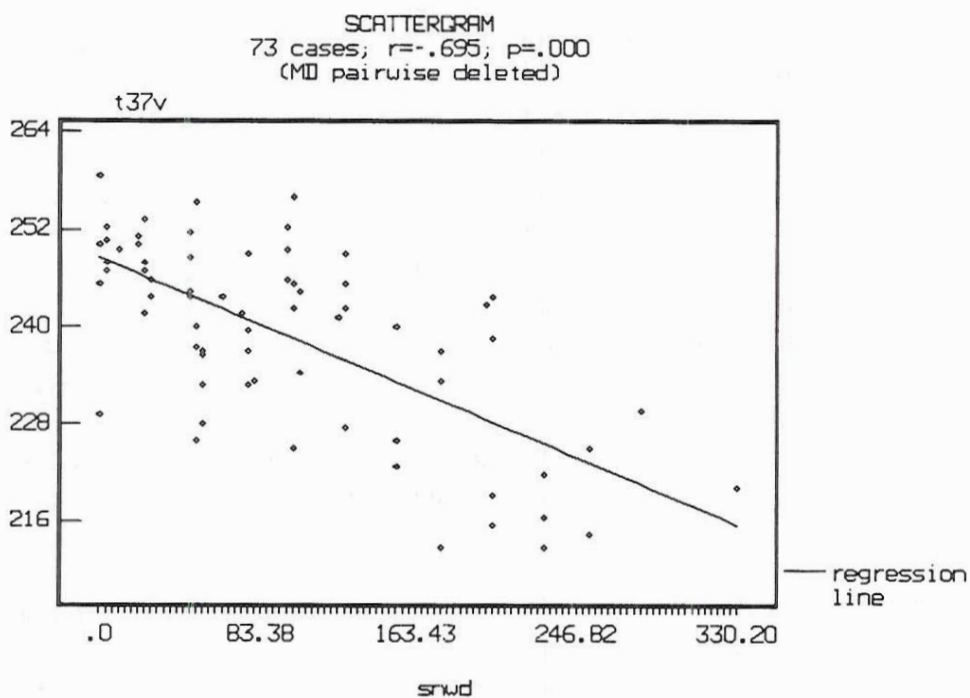
Figure 9.35   37V GHz brightness temperature and reported snow depth for the western test area, day 58, 1990, ascending pass.

SCATTERGRAM
131 cases; r=-.830; p=.000
(MD pairwise deleted)
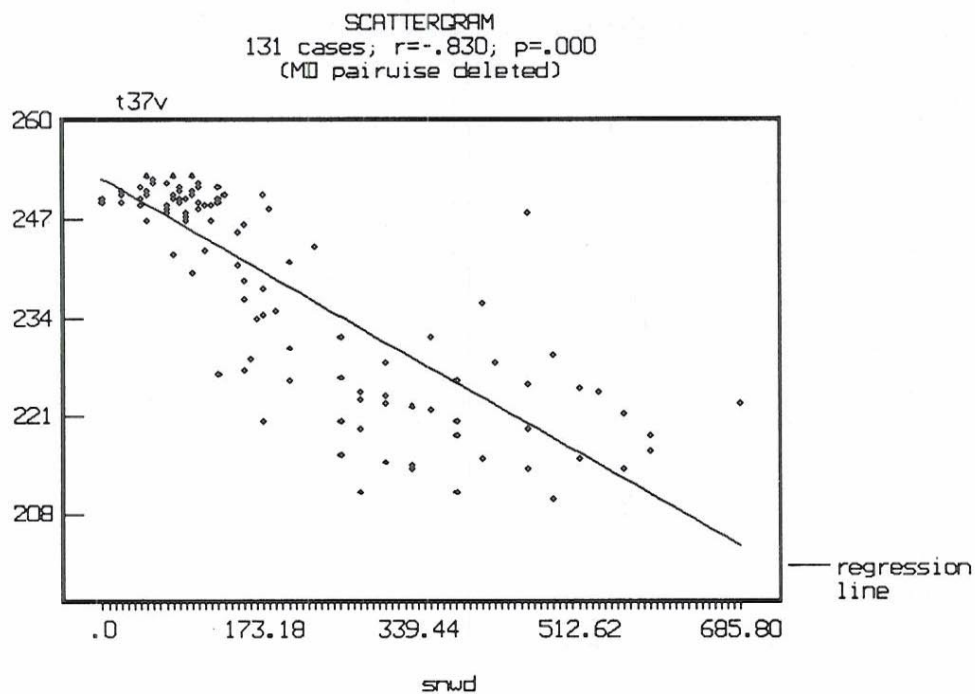
Figure 9.36 37V GHz brightness temperature and reported snow depth for the eastern test area, day 50, 1990, ascending pass.

SCATTERGRAM
130 cases; r=-.726; p=.000
(MD pairwise deleted)
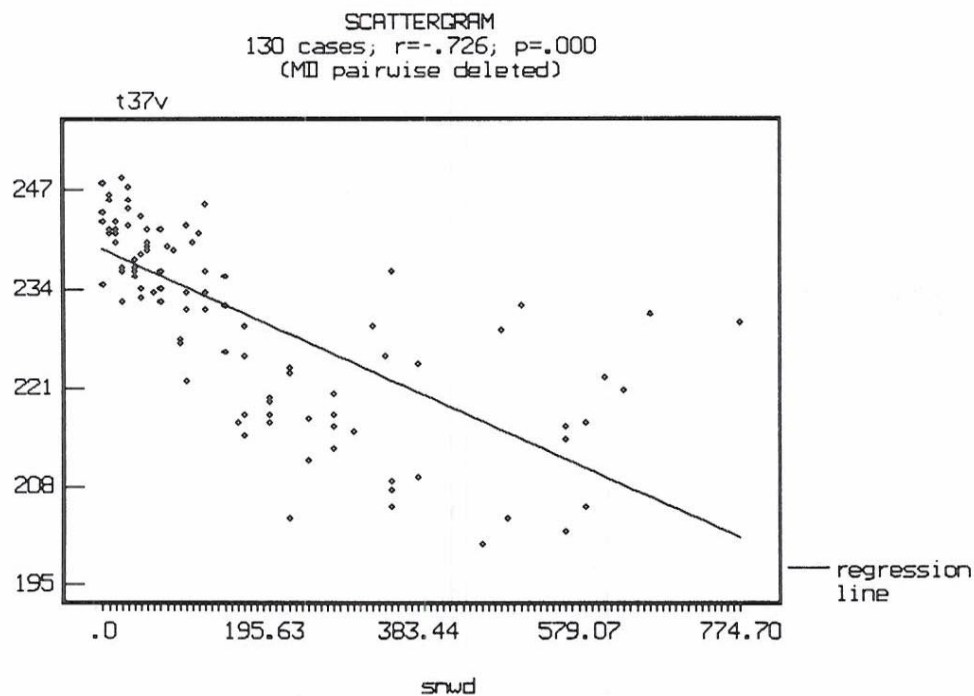
Figure 9.37 37V GHz brightness temperature and reported snow depth for the eastern test area, day 58, 1990, ascending pass.

TABLE 9.51   RELEVANT STATISTICS BETWEEN SSM/I BRIGHTNESS TEMPERATURES AND REPORTED SNOW DEPTH FOR THE WESTERN TEST AREA, DAY 47, 1990, ASCENDING PASS

css/pc: basic stats	Descriptive statistics in dbl precision Nunmber of Cases = 126 (MD pairwise deleted)				
	N	Min	Max	Mean	Std.Err	Std.Dev.
lat	126	41.0000	47.0000	44.5349	.156786	1.75992
lon	126	96.0000	104.0000	99.7492	.215347	2.41726
t19v	126	246.4000	261.9000	254.0595	.325702	3.65599
t19h	126	225.5000	252.0000	239.4095	.621787	6.97954
t22v	126	242.9000	261.0000	251.4151	.373291	4.19018
t37v	126	220.6000	255.2000	242.3556	.733083	8.22884
t37h	126	204.8000	249.0000	231.5802	.914381	10.26390
t85v	126	104.6000	335.9000	209.5564	3.724459	41.80695
t85h	126	172.3000	277.5000	223.0738	1.588973	17.83617
surft	126	14.0000	14.0000	14.0000	.000000	.00000
mlra	126	53.2000	106.0000	65.6389	1.470974	16.51164
snwd	126	.0000	482.6000	86.2460	8.446839	94.81554
snow	126	.0000	50.8000	6.0175	1.205538	13.53213

css/pc: basic stats	Correlations r(x,y) Number of Cases = 126 (MD pairwise deleted)				
std.mode	lat	lon	t19v	t19h	t22v	t37v
lat	1.0000	-.0909	-.7957	-.7492	-.7946	-.6969
lon	-.0909	1.0000	.1624	.2784	.1559	.2714
t19v	-.7957	.1624	1.0000	.9397	.9780	.8796
t19h	-.7492	.2784	.9397	1.0000	.9387	.9124
t22v	-.7946	.1559	.9780	.9387	1.0000	.9058
t37v	-.6969	.2714	.8796	.9124	.9058	1.0000
t37h	-.6618	.3610	.8637	.9375	.8825	.9799
t85v	.1127	.1565	-.0309	-.0615	-.0372	.0198
t85h	-.3936	.2602	.5342	.6043	.5432	.6585
surft	--	--	--	--	--	--
mlra	-.3210	-.4363	.2417	.2119	.2569	.3058
snwd	.5189	-.3768	-.5183	-.5937	-.5509	-.7064
snow	.3040	-.1293	-.2870	-.2729	-.2594	-.1763

css/pc: basic stats	Correlations r(x,y) N. of CASES = 126 (MD pairwise deleted)				
std.mode	t37h	t85v	t85h	surft	mlra	snwd
lat	-.6618	.1127	-.3936	--	-.3210	.5189
lon	.3610	.1565	.2602	--	-.4363	-.3768
t19v	.8637	-.0309	.5342	--	.2417	-.5183
t19h	.9375	-.0615	.6043	--	.2119	-.5937
t22v	.8825	-.0372	.5432	--	.2569	-.5509
t37v	.9799	.0198	.6585	--	.3058	-.7064
t37h	1.0000	-.0081	.6982	--	.2376	-.7072
t85v	-.0081	1.0000	-.1719	--	-.1016	-.0051
t85h	.6982	-.1719	1.0000	--	.0768	-.5412
surft	--	--	--	--	--	--
mlra	.2376	-.1016	.0768	--	1.0000	-.1151
snwd	-.7072	-.0051	-.5412	--	-.1151	1.0000
snow	-.1865	-.0273	-.0587	--	.1479	.1655

TABLE 9.52   RELEVANT STATISTICS BETWEEN SSM/I BRIGHTNESS
TEMPERATURES AND REPORTED SNOW DEPTH FOR THE WESTERN
TEST AREA, DAY 58, 1990, ASCENDING PASS

css/pc: basic stats	Descriptive statistics in dbl precision Nunmber of Cases = 73 (MD pairwise deleted)					
	N	Min	Max	Mean	Std.Err	Std.Dev.
lat	73	41.0000	47.0000	45.4370	.153368	1.31038
lon	73	96.0000	104.0000	99.2110	.264342	2.25854
t19v	73	247.1000	265.7000	254.1069	.554792	4.74015
t19h	73	225.9000	253.4000	236.8781	.887327	7.58132
t22v	73	242.9000	263.3000	251.6438	.613062	5.23800
t37v	73	212.5000	258.7000	238.9534	1.395477	11.92296
t37h	73	201.3000	251.4000	227.3671	1.573632	13.44512
t85v	73	143.3000	299.2000	211.6479	4.293872	36.68686
t85h	73	170.5000	273.1000	220.4137	2.842392	24.28541
surft	73	14.0000	14.0000	14.0000	.000000	.00000
mlra	73	53.2000	106.0000	64.4945	2.236502	19.10868
snwd	73	.0000	330.2000	96.7288	9.600029	82.02268
snow	73	.0000	63.5000	10.7904	1.787244	15.27022

css/pc: basic stats	Correlations r(x,y) Number of Cases = 73 (MD pairwise deleted)					
std.mode	lat	lon	t19v	t19h	t22v	t37v
lat	1.0000	.3627	-.2450	-.3137	-.2645	-.3265
lon	.3627	1.0000	.6542	.5039	.6252	.4815
t19v	-.2450	.6542	1.0000	.9583	.9780	.8796
t19h	-.3137	.5039	.9583	1.0000	.9548	.8721
t22v	-.2645	.6252	.9780	.9548	1.0000	.9216
t37v	-.3265	.4815	.8796	.8721	.9216	1.0000
t37h	-.3546	.4291	.8954	.9229	.9290	.9848
t85v	.1123	.1713	.0858	.0900	.1065	.1228
t85h	-.2033	.3745	.6451	.6589	.6578	.7865
surft	--	--	--	--	--	--
mlra	-.5898	-.5641	-.0668	.0764	-.0056	.1982
snwd	.3503	-.3523	-.6445	-.6413	-.6503	-.6948
snow	.2005	-.1106	-.1865	-.1752	-.1708	-.0630

css/pc: basic stats	Correlations r(x,y) Number of Cases = 73 (MD pairwise deleted)					
std.mode	t37h	t85v	t85h	surft	mlra	snwd
lat	-.3546	.1123	-.2033	--	-.5898	.3503
lon	.4291	.1713	.3745	--	-.5641	-.3523
t19v	.8954	.0858	.6451	--	-.0668	-.6445
t19h	.9229	.0900	.6589	--	.0764	-.6413
t22v	.9290	.1065	.6578	--	-.0056	-.6503
t37v	.9848	.1228	.7865	--	.1982	-.6948
t37h	1.0000	.1202	.7850	--	.2220	-.6755
t85v	.1202	1.0000	-.0280	--	-.0966	-.0077
t85h	.7850	-.0280	1.0000	--	.2088	-.5684
surft	--	--	--	--	--	--
mlra	.2220	-.0966	.2088	--	1.0000	-.1044
snwd	-.6755	-.0077	-.5684	--	-.1044	1.0000
snow	-.0968	-.0239	-.0195	--	.2103	.2677

TABLE 9.53   RELEVANT STATISTICS BETWEEN SSM/I BRIGHTNESS TEMPERATURES AND REPORTED SNOW DEPTH FOR THE EASTERN TEST AREA, DAY 50, 1990, ASCENDING PASS

css/pc: basic stats	Descriptive statistics in dbl precision Nunmber of Cases = 131 (MD pairwise deleted)					
	N	Min	Max	Mean	Std.Err	Std.Dev.
lat	131	41.0000	47.0000	43.7519	.16533	1.8923
lon	131	90.0000	96.0000	92.6664	.15851	1.8143
t19v	131	238.8000	260.7000	253.8069	.51849	5.9344
t19h	131	221.7000	251.1000	239.9221	.69832	7.9927
t22v	131	235.5000	259.1000	251.1534	.57253	6.5529
t37v	131	209.7000	253.0000	239.2802	1.17946	13.4995
t37h	131	198.5000	246.7000	229.3076	1.24080	14.2016
t85v	131	117.7000	303.4000	202.4687	3.53612	40.4818
t85h	131	173.4000	280.2000	218.1198	1.48938	17.0468
surft	131	14.0000	14.0000	14.0000	.00000	.0000
mlra	131	-99.0000	107.0000	-66.5916	6.36067	72.8012
snwd	131	.0000	685.8000	187.5916	13.68397	156.6202
snow	131	.0000	88.9000	9.2863	1.49697	17.1336

css/pc: basic stats	Correlations r(x,y) Number of Cases = 131 (MD pairwise deleted)					
std.mode	lat	lon	t19v	t19h	t22v	t37v
lat	1.0000	.2429	-.8964	-.8426	-.8974	-.8318
lon	.2429	1.0000	-.1739	-.3031	-.1808	-.2383
t19v	-.8964	-.1739	1.0000	.9313	.9855	.9082
t19h	-.8426	-.3031	.9313	1.0000	.9481	.9451
t22v	-.8974	-.1808	.9855	.9481	1.0000	.9372
t37v	-.8318	-.2383	.9082	.9451	.9372	1.0000
t37h	-.7994	-.3102	.8685	.9577	.9012	.9851
t85v	.0044	-.2065	.0029	.0278	-.0095	.0063
t85h	-.5691	-.2205	.5954	.7271	.6265	.7731
surft	--	--	--	--	--	--
mlra	.1116	.7143	.0021	-.0675	.0287	.0161
snwd	.8054	.2609	-.8145	-.8212	-.8196	-.8301
snow	-.1985	-.2361	.2774	.3907	.2844	.3269

css/pc: basic stats	Correlations r(x,y) Number of Cases = 131 (MD pairwise deleted)					
std.mode	t37h	t85v	t85h	surft	mlra	snwd
lat	-.7994	.0044	-.5691	--	.1116	.8054
lon	-.3102	-.2065	-.2205	--	.7143	.2609
t19v	.8685	.0029	.5954	--	.0021	-.8145
t19h	.9577	.0278	.7271	--	-.0675	-.8212
t22v	.9012	-.0095	.6265	--	.0287	-.8196
t37v	.9851	.0063	.7731	--	.0161	-.8301
t37h	1.0000	.0160	.8092	--	-.0449	-.8245
t85v	.0160	1.0000	.0228	--	-.2140	.0562
t85h	.8092	.0228	1.0000	--	-.0036	-.6537
surft	--	--	--	--	--	--
mlra	-.0449	-.2140	-.0036	--	1.0000	.0947
snwd	-.8245	.0562	-.6537	--	.0947	1.0000
snow	.3836	.0968	.3638	--	-.0554	-.2971

TABLE 9.54   RELEVANT STATISTICS BETWEEN SSM/I BRIGHTNESS
TEMPERATURES AND REPORTED SNOW DEPTH FOR THE EASTERN
TEST AREA, DAY 58, 1990, ASCENDING PASS

css/pc: basic stats	Descriptive statistics in dbl precision Nunmber of Cases = 130 (MD pairwise deleted)					
	N	Min	Max	Mean	Std.Err	Std.Dev.
lat	130	41.0000	47.0000	43.6331	.16737	1.9083
lon	130	90.0000	96.0000	92.4831	.15089	1.7204
t19v	130	239.8000	258.6000	250.0223	.37497	4.2753
t19h	130	217.1000	243.6000	234.2392	.57502	6.5563
t22v	130	235.2000	258.0000	246.8938	.47530	5.4193
t37v	130	200.1000	248.7000	231.8923	1.04472	11.9116
t37h	130	186.9000	240.3000	222.2762	1.13770	12.9718
t85v	130	74.5000	302.3000	208.4746	3.75692	42.8355
t85h	130	140.3000	270.9000	203.1923	2.00235	22.8303
surft	130	14.0000	14.0000	14.0000	.00000	.0000
mlra	130	-99.0000	107.0000	-74.8892	5.68800	64.8532
snwd	130	.0000	774.7000	152.8885	15.33124	174.8030
snow	130	.0000	184.1000	7.5685	1.94649	22.1934

css/pc: basic stats	Correlations r(x,y) Number of Cases = 130 (MD pairwise deleted)					
std.mode	lat	lon	t19v	t19h	t22v	t37v
lat	1.0000	.1066	-.9082	-.7156	-.8786	-.6995
lon	.1066	1.0000	.0194	.1620	.0097	.0413
t19v	-.9082	.0194	1.0000	.8701	.9724	.8483
t19h	-.7156	-.1620	.8701	1.0000	.8915	.9015
t22v	-.8786	.0097	.9724	.8915	1.0000	.8994
t37v	-.6995	-.0413	.8483	.9015	.8994	1.0000
t37h	-.6393	-.0967	.8090	.9258	.8653	.9880
t85v	-.0667	.0469	.0441	.0462	.0719	.0450
t85h	-.3198	.0074	.4823	.6262	.5358	.7335
surft	--	--	--	--	--	--
mlra	.0437	.6871	.1146	-.0028	.1245	.1001
snwd	.7985	.0784	-.8436	-.6990	-.8192	-.7259
snow	.3009	.1036	-.2414	-.1834	-.2229	-.1393

css/pc: basic stats	Correlations r(x,y) Number of Cases = 130 (MD pairwise deleted)					
std.mode	t37h	t85v	t85h	surft	mlra	snwd
lat	-.6393	-.0667	-.3198	--	.0437	.7985
lon	-.0967	.0469	.0074	--	.6871	.0784
t19v	.8090	.0441	.4823	--	.1146	-.8436
t19h	.9258	.0462	.6262	--	-.0028	-.6990
t22v	.8653	.0719	.5358	--	.1245	-.8192
t37v	.9880	.0450	.7335	--	.1001	-.7259
t37h	1.0000	.0416	.7578	--	.0510	-.6870
t85v	.0416	1.0000	.0087	--	-.0281	-.0486
t85h	.7578	.0087	1.0000	--	.1073	-.3911
surft	--	--	--	--	--	--
mlra	.0510	-.0281	.1073	--	1.0000	-.0135
snwd	-.6870	-.0486	-.3911	--	-.0135	1.0000
snow	-.1280	.0792	.0273	--	.1567	.3407

polarization difference in the 37 GHz channels and the difference between the 19V and 37V channels. The 37 GHz polarization difference was a poor predictor, with correlation coefficients for all data sets less than 0.50. The correlation coefficients between the 19V and 37V channel difference term and snow depth were generally in the range 0.55 to 0.75, which were about 0.15 lower than those of the 37V channel. On the few occasions when the correlation coefficients were higher, the interchannel correlation coefficient was also high. This indicated that no new information was available from the 19V channel. Based on these complete analyses, with analyses of partial data sets from February, 1989, the conclusion is that the use of a single channel, the 37V brightness temperature, provides the highest correlation coefficient of any SSM/I channel or channel combination.

The results for the 37V regression with snow depth for all grid cells were:

Test Area	n	Intercept	Slope
East	614	246.8	-0.0488
West	609	248.4	-0.0625

The regression equation for this combined data set is:

$$37V \ (K) = 247.6 - 0.557*SD \ (mm)$$

9.4.4 Recommendations

It was not possible to construct an algorithm which is suited for automatic determination of snow depth or snow water equivalent under all snow conditions. The interpretation of the algorithm predictions should be conducted, with previous data, other sources of data, and a knowledge of the areas of concern. However if the snowpack is known to be dry, that is classified as dry snow by the surface type identifier (see Section 9.1), the snow depth may be extracted with a high degree of accuracy with a single channel algorithm based on the 37V GHz brightness temperature. The snow depth (SD) algorithm, in millimeters, is the inverse of the regression equation determined in Section 9.4.3 and is given in Table 9.55.

TABLE 9.55	RECOMMENDED SNOW DEPTH RETRIEVAL ALGORITHM FOR DRY SNOW

$$SD = 444.5 - 1.795*T37V$$

9.4.5 References

[1] C. Matzler, H. Aebischer, and E. Schanda, "Microwave dielectric properties of surface snow," IEEE J. of Oceanic Engr., vol. OE-9, no. 5, pp. 366-371, 1984.

[2] J. A. Kong, R. Shin, J. C. Shine, and L. Tsang, "Theory and experiment for passive microwave remote sensing of snowpacks," J. Geophys. Res., vol. 84, pp. 5669-5673, 1979.

[3] J. L. Foster, A. Rango, D. K. Hall, A. T. C. Chang, L. J. Allison, and B. C. Diesen III, "Snowpack monitoring in North America and Eurasia using passive microwave satellite data," Remote Sens. Env., vol. 10, pp. 285-298, 1980.

[4] F. T. Ulaby, R. K. Moore, and A. K. Fung, Microwave Remote Sensing: Active and Passive, Volume III: From Theory to Applications, Dedham, MA: Artech House, 1986.

[5] M. J. McFarland and B. J. Blanchard, "Temporal correlations of antecedent precipitation with Nimbus 5 ESMR brightness temperatures," Amer. Meteor. Soc. 2nd Conf. Hydrometeor, Toronto, Canada, pp. 311-315, 1977.

[6] M. Hallikainen, and P. Jolma, "Retrieval of snow water equivalent from Nimbus-7 SMMR data," Snow Watch '85, Glaciological Data Report GD-18, Coop. Instit. for Res. in Environ. Sci., Boulder, CO, pp. 173-179, 1986.

[7] C. Matzler, "Can microwave signatures be used to retrieve the water equivalent of a dry snowpack?" Proc. 3rd Inter. Colloq. on Spectral Signat. of Objects in Remote Sensing, Les Arcs, France, pp. 277-284, 1985.

[8] D. K. Hall, "Influence of snow structure variability on global snow depth measurement using microwave radiometry," Snow Watch '85, Glaciological Data Report GD-18, Coop. Instit. for Res. in Environ. Sci, Boulder, CO, pp. 161-171, 1986.

[9] M. J. McFarland, G. D. Wilke, and P. H. Harder II, "Nimbus 7 SMMR investigation of snowpack properties in the northern Great Plains for the winter of 1978-79," IEEE Trans. Geosci. Remote Sensing, vol. GE-25, no. 1, pp. 35-46, 1987.

[10] N. C. Grody, "Snow cover monitoring using microwave radiometry," Snow Watch '85, Glaciological Data Report GD-18, Coop. Instit. for Res. in Environ. Sci., Boulder, CO, pp. 189-192, 1986.

[11] K. F. Kunzi, S. Patil, and H. Rott, "Snow cover parameters retrieved from Nimbus-7 scanning multichannel microwave radiometer (SMMR) data," IEEE Trans. Geosci. Remote Sensing, vol. GE-20, pp. 452-467, 1982.

[12]   P. Gloersen, D. J. Cavalieri, A. T. C. Chang, T. T. Wilheit, W. J. Campbell, O. M. Johannessen, K. B. Katasaros, K. F. Kunzi, D. B. Ross, D. Staelin, E. P. L. Windsor, F. T. Barath, P. Gudmansen, E. Langham, and R. O. Ramsier, "A summary of results from the first Nimbus-7 SMMR observations," <u>J. Geophys. Res.</u>, vol. 89, pp. 5335-5344, 1984.

[13]   T. J. Schmugge, "Microwave approaches in hydrology," <u>Photogram. Engr. Remote Sensing</u>, vol. 46, no. 4, pp. 495-507, 1980.

[14]   A. T. C. Chang, "Nimbus-7 SMMR snow cover data," Snow Watch '85, Glaciological Data Report GD-18, Coop. Instit. for Res. in Environ. Sci., Boulder, CO, pp. 181-187, 1986.

[15]   B. E. Goodison, I. Rubenstein, F. W. Thirkettle, and E. J. Langham, "Determination of snow water equivalent on the Canadian prairies using microwave radiometry," Proc. Modelling snow-melt-induced processes, IAHS Pub. 155, Budapest, pp. 63-173, 1986.

[16]   E. Schanda, C. Matzler, and K. Kunzi, "Microwave remote sensing of snow cover," <u>Int. J. Remote Sensing</u>, vol. 4, pp. 149-158, 1983.

[17]   H. K. Burke, C. J. Bowley, and J. C. Barnes, "Determination of snowpack properties from satellite passive microwave measurements," <u>Remote Sens. Env.</u>, vol. 15, pp. 1-20, 1984.

[18]   Soil Conservation Service, "Land resource regions and major land resource areas of the United States," <u>USDA-SCS Agriculture Handbook 296</u>, Washington, DC, 1981.

SECTION 10

SEA ICE VALIDATION

Atmospheric Environment Service, Environment Canada
North York, Ontario, Canada
Institute for Space and Terrestrial Science, York University
North York, Ontario, Canada
Norland Science and Engineering Ltd.
Ottawa, Ontario, Canada

Principal Investigator: Dr. R.O. Ramseier

PARTICIPATING ORGANIZATIONS

Scientific Authority:

Dr. R.O. Ramseier
Atmospheric Environment Service,
Centre for Research in Experimental Space Science, Microwave Group
York University
4700 Keele Street
North York, Ontario, Canada M3J 1P3

Contracting Authority:

Mr. S. Cooper
Science and Professional Services Directorate
Dept. of Supply and Services
12C1, Place du Portage, Phase III
11 Laurier Street
Hull, Quebec, Canada K1A 0S5

Prime Contractor:

Norland Science and Engineering, Ltd.
902-280 Albert Street
Ottawa, Ontario, Canada K1P 5G8

Algorithm Development[*]:

Institute for Space and Terrestrial Science
Centre for Research in Experimental Space Science, Microwave Group
York University
North York, Ontario, Canada M3J 1P3

Additional Support:
Ice Center, Environment Canada
365 Laurier Avenue
Ottawa, Ontario, Canada K1A 0H3

now at
Lasalle Academy E
373 Sussex Drive
Ottawa, Ontario, Canada K1A 0H3

[*] Initial algorithm development under contract to PhD Associates Inc., 4700 Keele Street, North York, Ontario, Canada.

PROJECT PARTICIPANTS

Atmospheric Environment Service, Ice Branch:

- Rene O. Ramseier, Senior Scientist
- Ken Asmus, AES Project Manager

Norland Science and Engineering Ltd.:

- Cathryn Bjerkelund, Validation Contract Project Manager, radar validation
- Tom Hirose, statistician
- Lyn D. Arsenault (Cold Regions Remote Sensing), radar validation

Institute for Space and Terrestrial Science, York University:

- Rene O. Ramseier, Principal Investigator for validation, Earth Observations
  Laboratory, Microwave Group
- Irene G. Rubinstein, Project Leader (Algorithm Development), Earth
  Observations Laboratory, Microwave Group

10.0 SEA ICE VALIDATION

10.1 INTRODUCTION

The validation of the Special Sensor Microwave/Imager (SSM/I) for sea-ice parameters was carried out by the Atmospheric Environment Service (AES), Environment Canada on data collected from June 1987 to September 1988. The objective of the validation project was to determine the accuracy of the Hughes Aircraft Company (HAC) sea ice algorithm* and the AES/YORK algorithm for total sea-ice concentration, ice age (i.e., first year or multi-year ice), ice type fractions, and the location of the ice edge. The aim was to see if these retrieved parameters could be predicted within the specifications given in Table 10.1 and, if retrieval parameters fail to meet the specifications, to determine, if possible, corrections needed to bring the parameters within specifications. A description of the HAC and AES/York algorithms is presented in Appendix 10A.

The performance was to be assessed for all seasons and in different geographic areas. In this project, four seasons were identified; ice formation (freeze-up), winter, initial melt, and advanced melt. The difference between the two stages of melt is the presence of snow cover during initial melt.

The validation also included the operational demonstration of the HAC and AES/YORK algorithms for ice reconnaissance and forecasting, which was carried out at the AES Ice Branch in Ottawa and at the U.S. Navy/NOAA Joint Ice Center in Washington, D.C.

The validation of the two algorithms for the three ice parameters involved the comparison of map products produced by the algorithms with airborne radar imagery flown over the same area as close in time as possible to the satellite overpass. This was no trivial task, because the radar imagery had to be obtained from AES ice reconnaissance aircraft which have operational constraints on the timing and location of flights. Therefore, the number of successful events, that is, where airborne radar imagery is collected within six hours of an SSM/I orbit and over a large enough area to match the SSM/I orbit, was only a fraction of the total planned events. Nevertheless, a sufficient number of events and numbers of validated SSM/I footprints were available to perform a statistical comparison for total ice concentration and ice edge location. There were insufficient ice fraction samples available to undertake any statistical analysis; only some trends in the data can be reported. Altogether 1.6 million sq km were validated for total ice concentration, and more than 6000 km were validated for ice edge position.

The sea ice validation program required that the overall accuracy of the ice parameters (regardless of geographic location, total ice concentration, or season) be determined. This

---

*Hughes Aircraft Company developed the sea ice algorithm and associated ground software used at FNOC and AFGWC to process the SSM/I data.

objective was achieved in this project by pooling all the data together as a combined areas data set.

TABLE 10.1				
ICE AND WIND PARAMETER REQUIREMENTS FOR SSM/I VALIDATION				
Parameter	Geometric Resolution (km)	Range of Values	Quantization Levels	Validation Criteria
(a) HAC Algorithm				
Ice				
Concentration	25	0 to 100%	5	±10%
Age	50	first-year	1 yr	none
		multi-year	>1 yr	none
Edge location	25	present/absent	N/A	±12.5 km
Ocean				
Surface Wind				
Speed	25	3 to 25	1	±2 m/s
(b) AES/YORK Algorithm				
Ice				
Concentration	25	0 to 100%	7	±10%
Age	25	Fractions of		
		0-100% first-year	1 yr	±10%
		old ice	> 1 yr	±10%
Edge location	25	present/absent	N/A	±12.5 km
Ocean				
Surface Wind				
Speed	25	3 - 40 m/s	1	±2 m/s

Because ice properties, ice concentrations and combinations of ice types differ between geographic areas and times of year, the performance of the algorithms were examined as a function of these parameters. Passive microwave sensors are sensitive to the amount of free water content in the overlying snow, a parameter that varies with season.

The validation project also examined algorithm performance over intervals of ice concentration as well as for different geographic areas and seasons. The statistical criteria for whether or not an algorithm met the originally defined criteria were more rigorously defined as follows:

1.   The algorithm is judged successful if the average difference in total ice concentration was within ±12% (HAC algorithm) or ±10% (AES/YORK algorithm) as well as at the 95% confidence interval.

2.   The algorithm is marginal if the average difference in total ice concentration was within ±12% (±10%), but was greater than ±12% (±10%) at the 95% confidence interval.

3.   The algorithm failed if the average difference in total ice concentration was greater than ±12% (±10%).

The same criteria were used to judge the algorithm performance for ice edge location, with a limit of ±12.5 km. Evaluation was not performed for ice fraction because of insufficient data.

The validation team felt that it was important to study the performance of each algorithm over intervals of ice concentrations as well as combining all the data. Sea-ice concentrations occur at 0-10% and 90-100% in many areas for lengthy periods of the year; however, during periods of break-up, movement and formation, ice concentrations vary widely and can change quickly. It is important to know how well the algorithms perform at intermediate concentrations, and to determine if the performance is consistent or varies as a function of concentration interval. The interval selected was 10%, which is the same division used by AES Ice Branch in reporting ice conditions.

The validation was carried out in two geographic regions, the Canadian Arctic and the Gulf of St. Lawrence where corroborating airborne radar data were available. In the Arctic, most of the validated SSM/I footprints were in the Beaufort Sea. A small percentage of the total sample (<10%) was in Northern Baffin Bay, Amundsen Gulf and M'Clure Strait. The Arctic data set for total ice concentration comprises slightly more than 80% of the validation samples where a sample is a validated SSM/I footprint. The validation results for these areas are discussed in detail below.

The performance of the two algorithms was estimated by performing statistical analysis of the data set. The statistics used include determining an average difference of all the samples combined. This provides an indication of algorithm performance in the real world, but it can be biased by the distribution of samples over the range of concentration. To overcome this sampling bias, a uniform sampling of the Arctic data set was undertaken. The resulting statistics, e.g. mean difference, standard deviation and 95% confidence interval between the algorithm and the radar total ice concentrations provides the overall estimate of accuracy.

The sampling for total ice concentration was biased towards the 0-10% and 90-100% concentration bins which made up a large proportion of the samples. Part of the problem was that the airborne radar imagery covered areas and time of year where an almost complete ice cover was present, e.g., Beaufort Sea in the fall and winter, or where the aircraft flew along or adjacent to the ice edge such that one of the two radar swaths imaged mostly open water

conditions. Figure 10.1 illustrates the distribution of samples for the SSM/I algorithm for all areas and seasons combined. Over 2/3 of the samples were either at 0-10% or 90-100% ice concentration intervals. The bias was removed using two techniques:
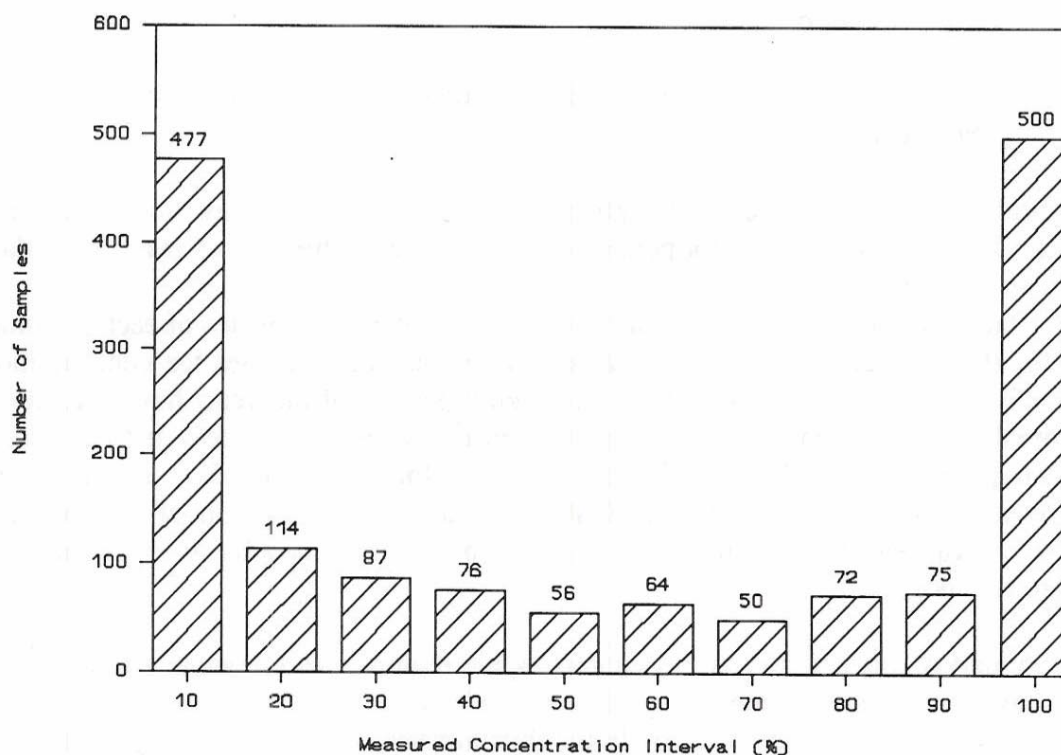
Figure 10.1 - Distribution of total ice concentration samples, HAC algorithm.

1. Examining the mean differences over 10% ice concentration intervals.

2. Extracting equal numbers of samples over the range of concentrations to produce statistics and distributions similar to those for the entire sample population.

10.2 TOTAL ICE CONCENTRATION RESULTS

10.2.1 Canadian Arctic

Using the acceptance criteria, the results of the two algorithms for all data pooled and by season are presented in Table 10.2.

TABLE 10.2					
TOTAL ICE CONCENTRATION EVALUATION CRITERIA FOR ARCTIC					
Algorithm	Pooled	Ice Formation	Winter	Initial Melt	Advanced Melt
HAC	Failed	Failed	Successful	Successful	Failed
AES/YORK	Successful	Marginal	Successful	Successful	Marginal

The results are for samples where there was less than 3 h between the radar imagery and SSM/I overpass. The average difference and standard deviation in concentration between algorithm and radar-based estimates for the SSM/I and AES/YORK algorithms are presented in Figures 10.2 and 10.3.

Figure 10.2 - Mean difference and standard deviation HAC vs radar for total ice concentration, Arctic, pooled.
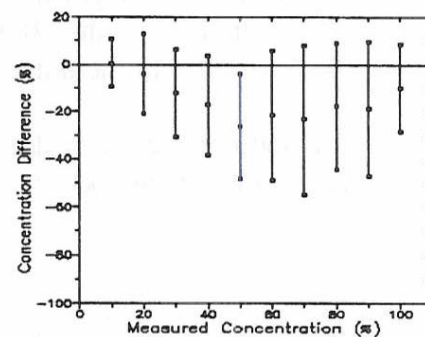
Figure 10.3 - Mean difference and standard deviation AES/YORK vs radar for total ice concentration, Arctic, pooled.

o   Both algorithms underpredict total ice concentration across the entire concentration range.

o   Both algorithms work best at low ice concentrations, less than or equal to 20% for HAC, less than or equal to 30% for AES/YORK.

o   The AES/YORK algorithm performs better at high concentrations (90-100%) where the average difference is underpredicted by less than 10%.

o   Both algorithms significantly underpredict in the middle range of ice concentrations. Both exhibit a characteristic "curve" where the underprediction increases with increasing ice concentration until it reaches a maximum at 70 to

90% concentrations for HAC, 50% for AES/YORK. Both algorithms improve for the 90-100% interval, but HAC still underpredicts by over 20%. The improved performance for 90-100% is still not as good as for the lower concentrations.

10.2.1 Ice Formation

For the ice formation phase the results reveal that:

o   Both algorithms perform successfully within the criteria for the 0-10% ice concentration bin but both underpredict for the higher intervals.

o   For concentrations 20-30% and above (including 90-100%), the HAC algorithm underpredicts well below the 12% acceptance criteria.

o   The AES/YORK algorithm shows the same trend, but begins to underpredict by more than 10% at the 40-50% concentration bin. There is improvement at 90-100%, similar to the pooled data results.

o   The underprediction is due to the presence of new ice and the refreezing of old ice freshwater meltponds.

10.2.2 Winter

For the winter ice phase the results show that:

o   Over 90% of the samples are in the 90-100% concentration bin, reflecting the typical ice conditions for the Arctic at this time of year.

o   Both algorithms perform well in cold conditions and at high ice concentrations. Their performance in winter is better than any other season.

o   There were insufficient samples at lower ice concentrations to fully test the accuracy of the algorithms.

10.2.3 Initial Melt

For the initial melt phase the results indicate that:

o   Between 65 and 70% of the samples were at ice concentrations less than or equal to 10%. This sampling is not typical of ice conditions in the Arctic at this time of year. The data set was limited by the lack of radar coverage of the area because the AES aircraft does not cover the area operationally at this time.

○ No inference can be made about the accuracy of the algorithms at higher ice concentrations. More samples are needed at the higher ice concentrations to confirm algorithm performance.

10.2.4 Advanced Melt

○ There were sufficient samples over the range of concentration to determine overall algorithm performance as well as between ice concentration bins.

○ Both algorithms underpredict total ice concentration for all the concentration bins.

○ The HAC predicts best for concentrations less than 20%. For all concentration bins above 10-20%, including the 90-100% bin, it significantly underpredicts ice concentration. This is probably the result of high water coverage on the ice at this time of year [1].

○ The AES/YORK algorithm shows a similar trend to HAC, except that the underprediction is less (by at least 10%) for all concentration bins.

10.2.5 Gulf of St. Lawrence Regional Results

The number of samples is considerably less than in the Arctic. Samples were available for only two seasons, ice formation and winter. The evaluation criteria applied to the two algorithms for this area is summarized in Table 10.3.

○ More than 80% of the samples are in the 90-100% concentration interval. The acceptance criteria can only be applied to this bin because of insufficient data in the other bins.

○ At high concentrations the HAC algorithm in particular has difficulty predicting the presence of new and thin ice types.

TABLE 10.3			
TOTAL ICE CONCENTRATION EVALUATION CRITERIA FOR GULF OF ST. LAWRENCE			
Algorithm	Pooled	Ice Formation	Winter
HAC	Failed	Failed	Failed
AES/YORK	Successful	Marginal	Marginal

o   Figures 10.4 and 10.5 present mean difference and standard deviation by concentration interval. The variability of the data reflects the low number of samples. The following observations may be made:
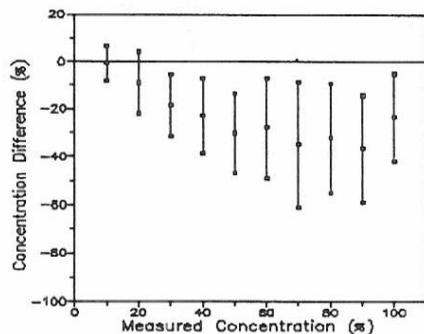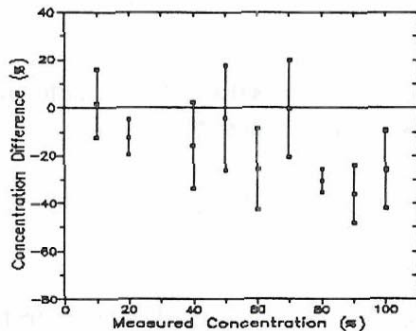

Figure 10.4 - Mean difference and standard deviation HAC vs radar for total ice concentration, Gulf of St Lawrence, pooled.
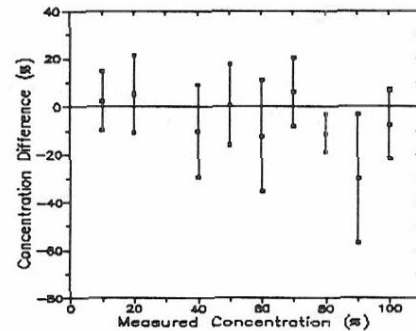
Figure 10.5 - Mean difference and standard deviation AES/YORK vs radar for total ice concentration, Gulf of St Lawrence, pooled.

10.2.2.1 Ice Formation and Winter

For the ice formation and winter phases the results indicate that:

o   The evaluation criteria results are greatly influenced by the high proportion of new and thin ice types which the algorithms appear to have difficulty predicting.

o   The first-year ice in the Gulf is not as thick as that in the Arctic in the winter months and, there is a higher percentage of new and thin ice in the matrix. These factors contribute to the differences in algorithm performance for the Gulf.

10.2.2   Uniform Sampling

To remove the bias of the total ice concentration pooled data set for the disproportionate number of samples in the 0-10% and 90-100% concentration intervals (where both algorithms perform better), an equal number of samples from each 10% interval was analyzed statistically. The resulting distributions of uniform sampling for the HAC and the AES/YORK algorithms are presented in Figures 10.6 and 10.7, respectively. Note that

o   Both algorithms show degraded performance, with increased mean differences and higher standard deviations.

o   The trend in the differences over the mid-range of concentrations is similar to the complete data set (Figures 10.2 to 10.5). The distributions confirm the tendency

for the algorithms to underpredict total ice concentration particularly in the middle concentration ranges.

o   Using the evaluation criteria for total concentration on the uniform sampling of the pooled data set, both algorithms failed.
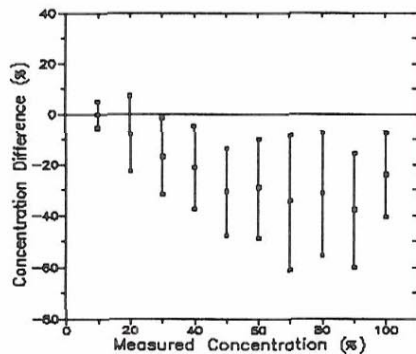
Figure 10.6 - Mean difference and standard deviation HAC vs radar for total ice concentration, Arctic, uniform sampling.
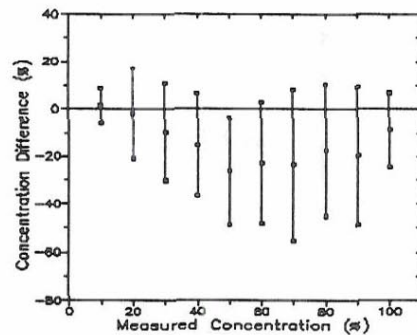
Figure 10.7 - Mean difference and standard deviation, AES/YORK vs radar for total ice concentration, Arctic, uniform sampling.

10.3   ICE EDGE LOCATION RESULTS

Validation of ice edge was only possible for relatively simple ice edges. Sections of the ice edge that were convoluted, consisted of plumes or embayments, or were otherwise complex could not be validated because there was no consistent way to make measurements between the radar and the algorithm ice edge locations. This difficulty reduced the number of samples available for subsequent statistical analysis.

During the ice formation and winter seasons ice edge comparisons were further complicated by the presence and formation of new ice. Ice edge measurements were made only where no ambiguity existed in interpreting new or thin ice in the radar imagery. However, the time between the radar coverage and the SSM/I overpass was critical because the two sensors may detect different distributions of the edge simply because of new ice growth. These factors reduced the length of ice edge available for comparison of the radar and SSM/I.

Almost 90% of ice edge measurements were made for the Beaufort Sea data. The distribution of ice edge displacements for the combined areas provide a representative and consistent measure of algorithm performance.

The distribution of ice edge measurements for the HAC algorithm (Figure 10.8) shows a tendency to underpredict the ice edge location. The samples are skewed into the ice with a mean difference between -11 and -20 km (bin no. -1).
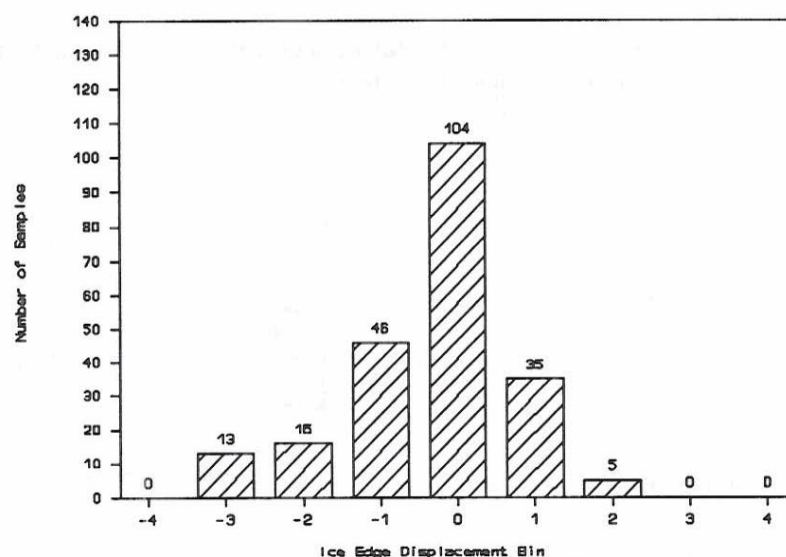
Figure 10.8 - Distribution of ice edge differences, HAC vs radar
for all areas combined.

The ice edge displacement results for the AES/YORK algorithm are presented in Figure 10.9. More than 90% of the samples fall within ±20 km of the ice edge as derived from airborne radar imagery. The samples are uniformly distributed about the ice edge location, with positive displacements representing edge locations beyond the ice edge and negative displacements indicating ice edge locations within the ice.

Tables 10.4 and 10.5 summarize the evaluation criteria for ice edge location for the two algorithms for the Arctic and Gulf of St. Lawrence respectively.

TABLE 10.4				
ICE EDGE EVALUATION CRITERIA FOR CANADIAN ARCTIC				
Algorithm	Pooled	Ice Formation	Initial Melt	Advanced Melt
HAC	Failed	Marginal	Failed	Failed
AES/YORK	Successful	Successful	Successful	Marginal
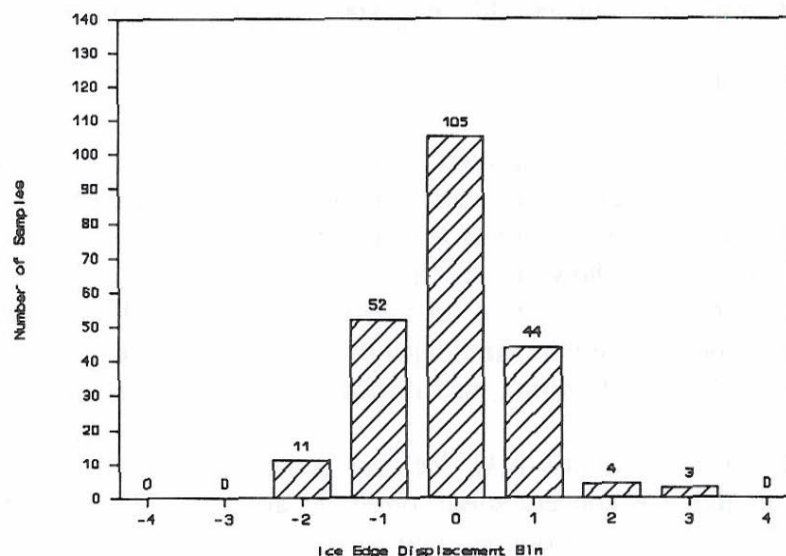
Figure 10.9 - Distribution of ice edge differences, AES/YORK vs radar
for all areas combined.

TABLE 10.5		
ICE EDGE EVALUATION CRITERIAL FOR GULF OF ST. LAWRENCE		
Algorithm	Pooled	Ice Formation
HAC	Marginal	Marginal
AES/YORK	Marginal	Marginal

10.4   ICE FRACTION RESULTS

The validation of ice fraction was not possible because of the low number of samples. Consequently no statistical analysis was undertaken. Of the total sample population, only 10% were at high old ice concentrations (81-100%) because of the scarcity of coincident airborne radar coverage for such areas.

The HAC algorithm does not produce an old ice concentration; it reports old ice if the concentration is above 35%. Therefore the validation of the HAC algorithm for ice fraction was really a question of whether or not it reliably reports old ice when its fraction is above 35%.

The only trends apparent were that for the AES/YORK algorithm, it was underestimating the old ice fraction for the limited number of samples available, and for old ice concentrations above 80%, the HAC algorithm flagged old ice in about half of the samples.

10.5 ADDITIONAL RESULTS

There were additional shortcomings with the HAC algorithm which are not apparent in the statistical results concerning adverse weather conditions and the ice edge contour. There were areas occasionally shown by the algorithm as ice covered where no ice should be present. An example in the Labrador Sea, is shown on the left in Figure 10.10, where ice along the coast was extended by the HAC algorithm into an apparent ice cover all the way to the west Greenland coast. The problem could be eliminated by a suitable weather filter algorithm, as is incorporated in the AES/YORK algorithm, and illustrated on the right in Figure 10.10.

The appropriateness of the 30% HAC ice edge contour as the one to define ice edge is questioned because of high ice concentrations observed along it. The ice edge (or 0% ice concentration contour) as determined on the radar imagery corresponded to a HAC algorithm ice concentration of between 25 and 50%, with an average of 35% ice concentration, depending on the ice types present. The 30% HAC algorithm ice contour was observed to correspond to an average ice concentration of 56%. In comparison, the AES/YORK algorithm at the 0% radar ice concentration contour corresponded to an ice concentration of between 0 and 25%, with an average ice concentration of 16%, depending on ice type and the 10% AES/YORK algorithm ice edge corresponded to an average ice concentration of 25%.

The HAC algorithm was designed to flag the presence of old ice only when concentrations reached 35% or more of the total ice concentration. Because it only flags, but does not determine the ice fraction concentration, its usefulness is reduced for operational purposes. The AES/YORK algorithm is designed to provide open water, first-year, and old ice fractions. It also allows retrieval of the ocean surface wind speeds, cloud cover, precipitation, and water vapor for ice-free areas.

10.6 CONCLUSIONS/RECOMMENDATIONS

o   The AES/YORK algorithm is recommended for operational use. It is superior to the HAC algorithm for total ice concentration estimates and ice edge location for the geographic areas and seasons validated in this project. AES/YORK also produces more specific estimates of old ice concentration.

It is recommended that a tailored or reduced version of the AES/YORK algorithm be implemented for operational use. See Appendix A for a description of the complete AES/YORK algorithm. This tailoring is necessary for two major reasons. First, the AES/YORK was constructed to retrieve not only the basic SSM/I parameter of sea ice concentration and identify first-year and multi-year ice types but also additional parameters such as the fractions of first-year, multi-year, and thin ice within the SSM/I footprint as well as ocean surface wind speed

SSM/I SEA ICE CONCENTRATION

BAFFIN BAY – DAVIS STRAIT – LABRADOR SEA

22 Nov 1987   Rev 2194
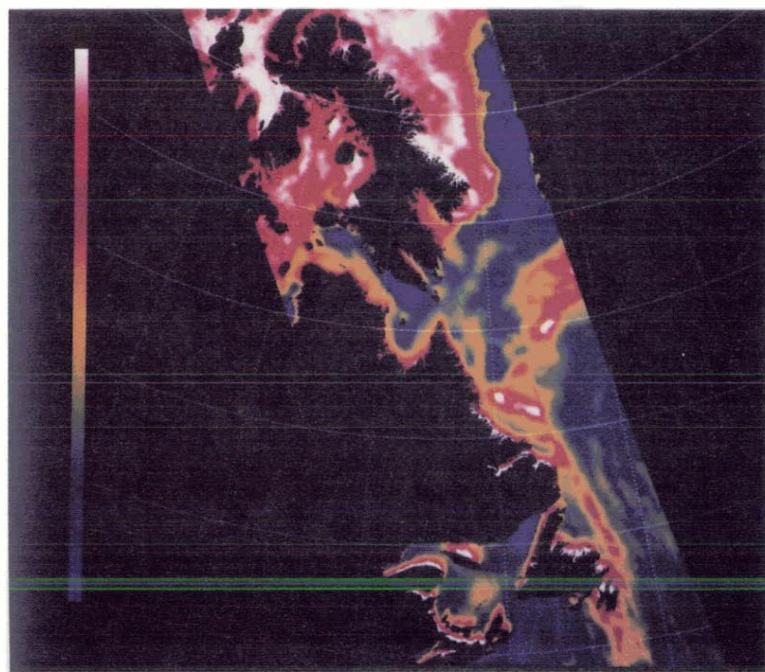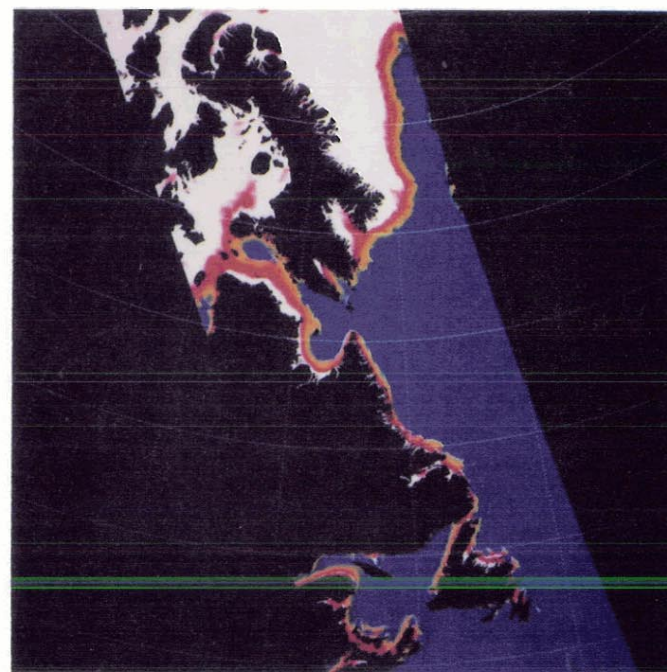
HAC Algorithm   AES/YORK Algorithm

Color Bar 0 – 100% Ice Concentration

Figure 10.10 SSM/I sea ice concentration retrieved from HAC and AES/YORK algorithms.

and vertical columns of water vapor and liquid cloud water. Second, the computer resources required to implement the complete AES/YORK algorithm are significantly greater than the proposed tailored version. The error in the retrieved sea ice concentration using the tailored AES/YORK algorithm is essentially that associated with the complete algorithm and as discussed earlier is typically less than 10-12%.

10.6.1 Recommended Sea Ice Algorithm

Figure 10.11 presents a flow chart of the reduced version of the AES/YORK sea ice algorithm. Specific equations and decision tests employed in the algorithm are presented in Table 10.6 and the ocean regions where sea ice is possible and the algorithm should be implemented are given in Table 10.7. The initial test identified in the flow chart, Test 1, checks for the reasonableness of the 19V, 19H, 37V, and 37H SDRs and polarization differences 19V-19H, 37V-37H. If any of the inequalities in Test 1 of Table 10.6 are true, no sea ice concentration or ice type identification is retrieved. If none of these inequalities are true, the SDRs are reasonable for open ocean or sea ice and total sea ice concentration, TOTICE, is computed either for winter/fall conditions or summer/spring conditions. Equation A in Table 10.6 is used to compute TOTICE and employs only the 19V and 37V SDRs. Depending on the value of TOTICE and several subsequent threshold tests, new values of TOTICE may be computed. As shown in Figure 10.11, a threshold TC is selected depending on the condition of winter/fall or summer/spring. TC is essentially an atmospheric offset threshold used later. The next step in the algorithm is to compute a discriminate D which is an estimate of the total ice concentration independent of Equation A and is expressed by Equation C in Table 10.6. Test 2 which follows the computation of D is a consistency check between TOTICE and D. If TOTICE and D are both less than or equal to 0.7, additional testing is necessary to determine the influence of clouds and/or ocean roughness. These tests are identified as tests, 3, 4, and 5 in the flow diagram. If TOTICE is greater than 0.7 or in the event the output of these tests results in TOTICE being less than or equal to 0.5 and D greater than 0.15 (test 6), then the effects of cloud and ocean roughness are unimportant and the algorithm recomputes TOTICE using only the 37V and 37H channels with Equation D of Table 10.6. This is done to take advantage of the higher resolution of the 37 GHz data and provides greater accuracy in determining sea ice edge. (The highest resolution 85 GHz channels are currently not employed in sea ice concentration retrievals. Under clear skies and calm ocean surface, the 85 GHz data offers the potential for determining sea ice edge to ±6 km). In the event clouds or ocean roughness is important, the previous value of TOTICE is used. Test 6 is followed by out-of-bounds checking of TOTICE and if TOTICE is less than 0.25 no ice type identification is made. If TOTICE is greater than or equal to 0.25, the ice type identifier TBI is computed with Equation E. If TBI is less than 238, the fraction of ice is predominately multi-year ice. Otherwise the fraction is predominately first-year.
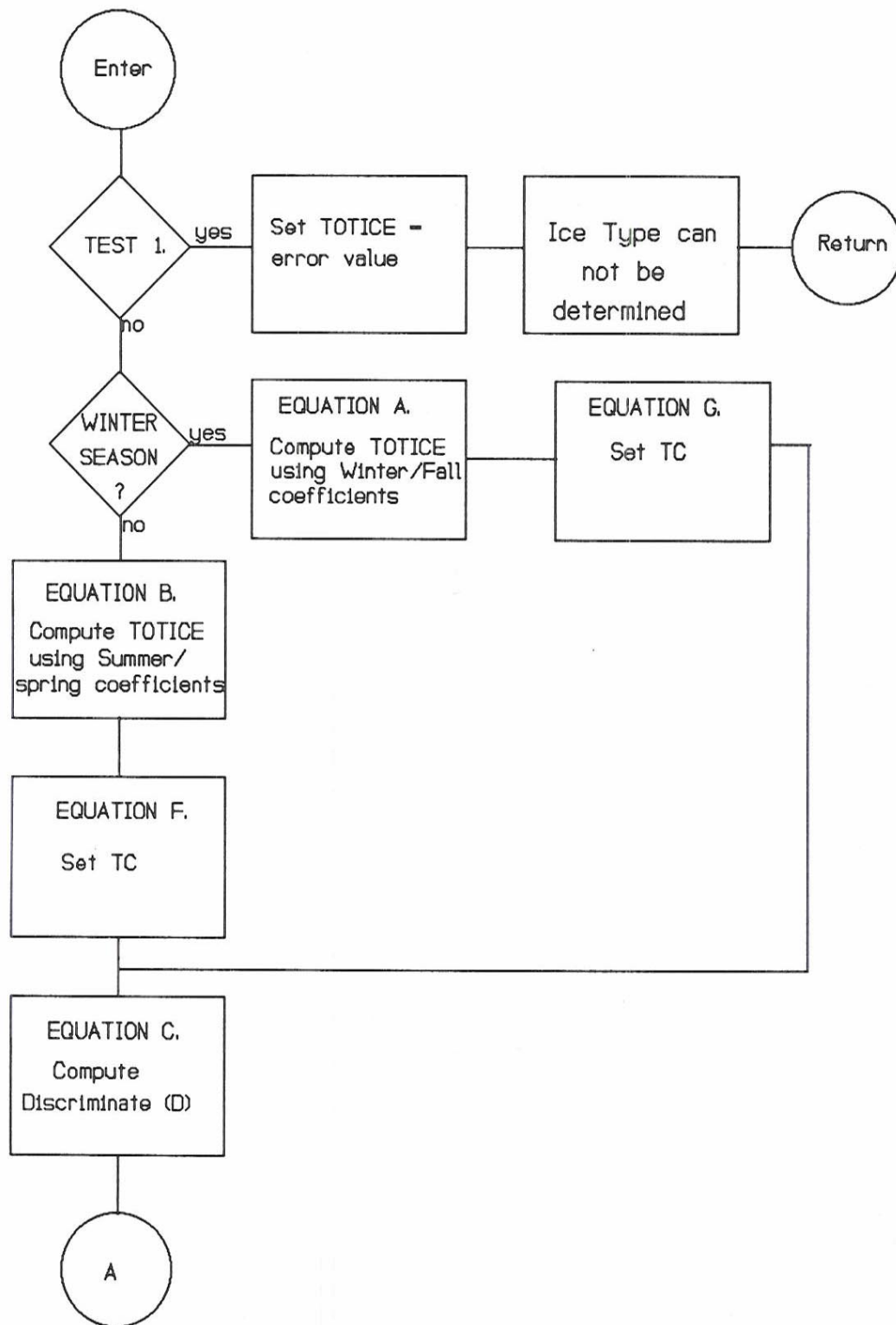
Figure 10.11 - Recommended sea ice algorithm flow chart.

10-15
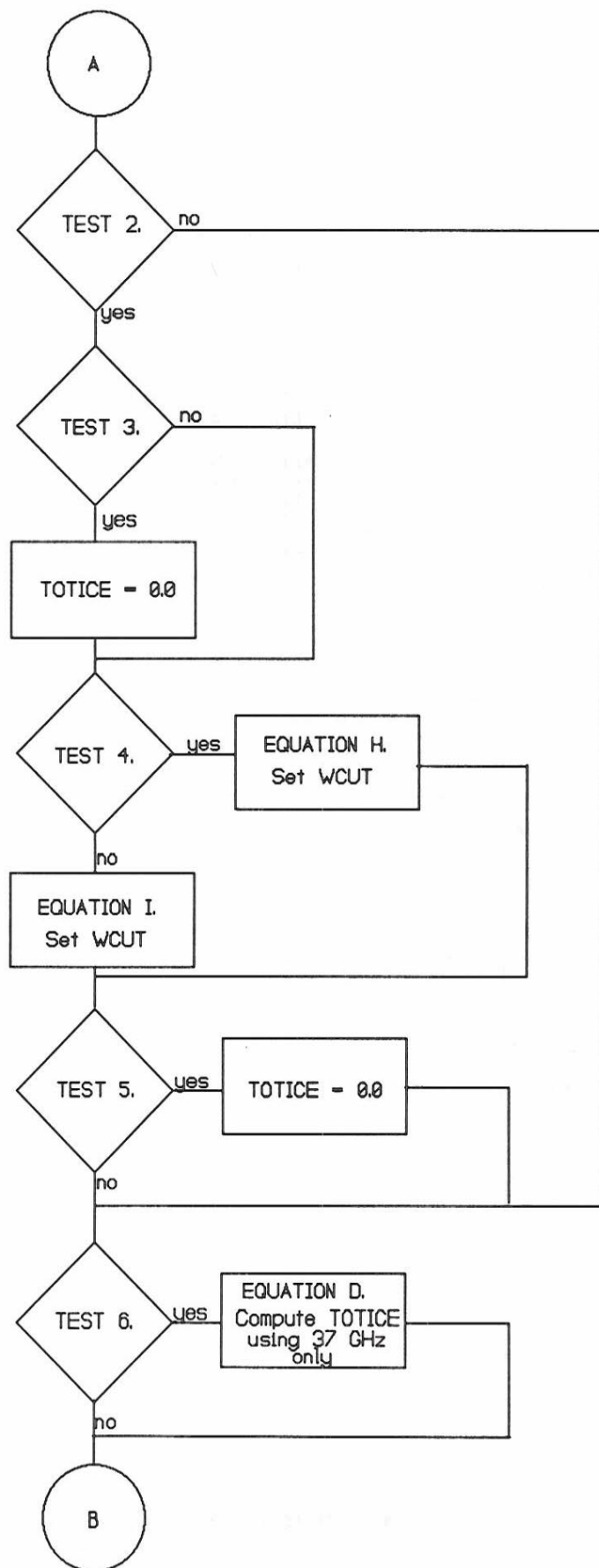
Figure 10.11 - continued.
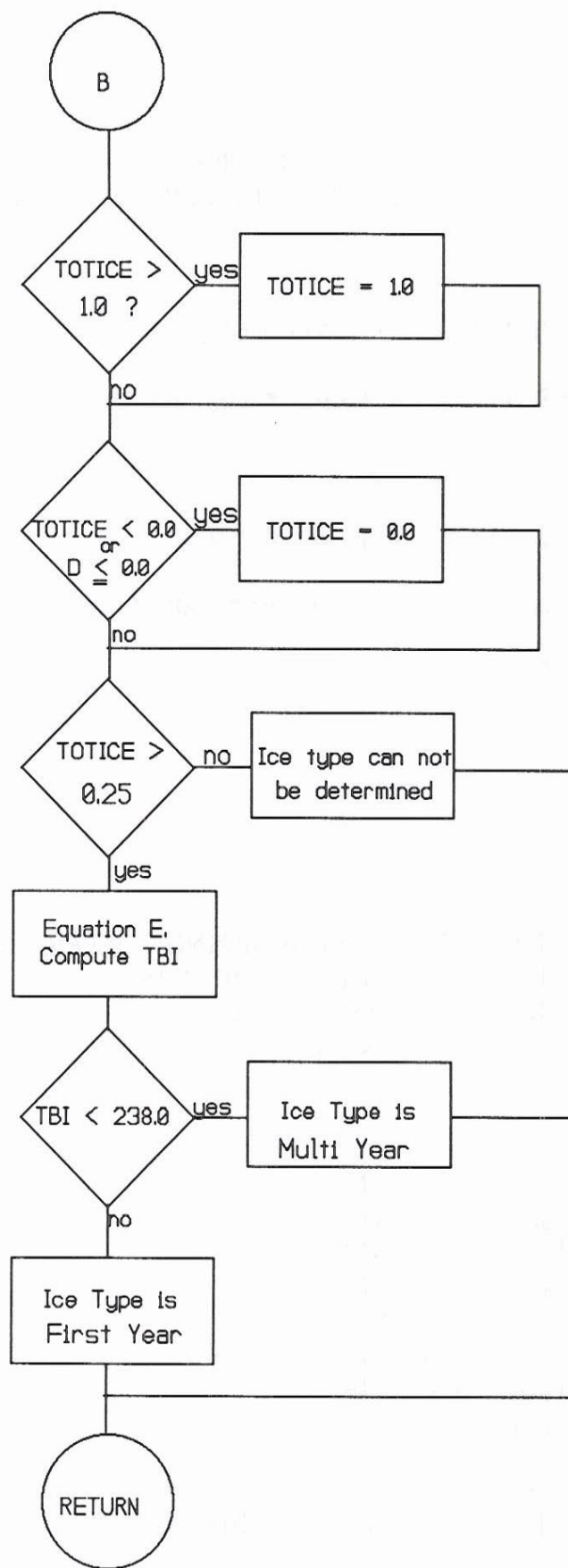
Figure 10.11 - continued.

TABLE 10.6
EQUATIONS AND TESTS USED IN RECOMMENDED SEA ICE ALGORITHM

EQUATIONS

A. $\text{TOTICE} = \text{CWF(1)} * T_{B37V} + \text{CWF(2)} * T_{B19V} + \text{CWF(3)}$

B. $\text{TOTICE} = \text{CSS(1)} * T_{B37V} + \text{CSS(2)} * T_{B19V} + \text{CSS(3)}$

C. $D = 1.0 - 0.0513 * (T_{B37V} - T_{B19V})$

D. $\text{TOTICE} = (T_{B37V} + 0.5 * T_{B37H} - 265.0) * 0.01$

E. $\text{TBI} = [T_{B19V} - TC - (1.0 - \text{TOTICE}) * 180]/\text{TOTICE}$

F. $TC = 14.0$

G $TC = 6.8$

H $\text{WCUT} = 6.0$

I $\text{WCUT} = 8.5$

WINTER/FALL COEFFICIENTS
CWF(1) = -0.013656219
CWF(2) = 0.024412842
CWF(3) = -1.677645

SUMMER/SPRING COEFFICIENTS
CSS(1) = -0.015231617
CSS(2) = 0.025911011
CSS(3) = -1.656920

TESTS

1.
$T_{B19V} \leq 151.0$ OR
$T_{B19H} \leq 92.0$ OR
$T_{B37V} \leq 171.0$ OR
$T_{B37H} \leq 125.0$ OR
$(T_{B19V} - T_{B19H}) \geq 80.0$ OR
$(T_{B37V} - T_{B37H}) \geq 80.0$ OR
$T_{B19H} > T_{B19V}$ OR
$T_{B37H} > T_{B37V}$

2. $\text{TOTICE} \leq 0.7$ AND $D \leq 0.7$
3. $D \leq 0.3$ AND $[T_{B37V} * 1.5 - T_{B19V}] > 120.0$
4. $T_{B37V} \leq 215.0$
5. $D \leq 0.15$ OR $[T_{B37H} - 2 * T_{B37V} + 270.0] \geq \text{WCUT}$
6. $\text{TOTICE} \leq 0.5$ AND $D > 0.15$

10-18

TABLE 10.7

REGIONS WHERE SEA ICE MAY EXIST AND THE RECOMMENDED
ALGORITHM SHOULD BE IMPLEMENTED

1. Southern Hemisphere: All ocean regions less than 50S latitude.

2. Northern Hemisphere: All ocean regions above 65N latitude.

    A. Alaska Area: longitude 165-200E and latitude 50-90N

    B. Gulf of St. Lawrence and Hudson Bay: longitude 240-315E and latitude 42-90N

    C. Sea of Japan and Sea of Okhotsk: longitude 130-155E and latitude 40-90N

    D. Baltic and North Sea: longitude 5-30E and latitude 53-90N

    E. Kamchatka Peninsula: longitude 155-165E and latitude 45- 90N

    F. Iceland: longitude 330-350E and latitude 60-90N

    G. Greenland: longitude 315-330E and latitude 55-90N

    H. White Sea: longitude 30-50E and latitude 63-90N

    I. Yellow Sea: longitude 115-130E and latitude 37-90N

    J. Kodiak Island: longitude 200-210E and latitude 55-90N

    K. Gulf of Alaska: longitude 210-240E and latitude 58-90N

10.7   REFERENCES

[1] D. A. Etkin and R. O. Ramseier, "Validation of a Passive Microwave Sea Ice Data Set for Hudson Bay," First Circumpolar Symposium on Remote Sensing of Arctic Environments, Yellowknife, N.W.T., Canada, 1990.

ACRONYMS

ACIF   Ice Forecasting Division of AES Ice Branch (Ottawa, Ontario, Canada)

AES   Atmospheric Environment Service (Canadian Department of the Environment)

AES/PhD   algorithm developed by AES and PhD Associates Inc.

AES/YORK   algorithm developed by AES and York University

AIMR   airborne imaging microwave radiometer

DEF   ephemeris data tapes

FNOC   Fleet Numeric Oceanographic Center (Monterey, California, U.S.A.)

ICEC   Ice Centre, Environment Canada (Ottawa, Ontario, Canada)

ISTS   Institute of Space and Terrestrial Science, York University

JIC   Joint (NOAA/NAVY) Ice Center (Washington, D.C., U.S.A.)

NOAA   U.S. National Oceanic and Atmospheric Administration

NPOC   U.S. Naval Polar Oceanographic Center (Washington D.C., U.S.A.)

NRL   Naval Research Laboratory (Washington, D.C., U.S.A.)

SAR   synthetic aperture radar

SLAR   side-looking airborne radar

SSM/I   special sensor microwave/imager

SSMR   scanning multichannel microwave radiometer

APPENDIX 10A

10A.0 DESCRIPTION OF HAC AND COMPLETE AES/YORK SEA ICE ALGORITHMS

10A.1 BACKGROUND

The sea ice algorithm used by the U.S. Navy for the SSM/I was developed during the 1970s by Environmental Research and Technology Inc. under a subcontract from Hughes Aircraft Corporation, and was tested during the NIMBUS satellite series of scanning multichannel microwave radiometers (SMMR). The HAC algorithm was tested extensively from 1982 to 1987 by the Ice Research and Development Division of Ice Branch, Atmospheric Environment Service (AES) which is part of the Canadian Department of the Environment, for both research and operational purposes. To improve the retrieval of ice information in all weather conditions and to optimize the use of SMMR channels another algorithm was developed (produced under contract to AES by PhD Associates Inc.). Known as the AES/PhD version, it also underwent research and operational testing from 1984 to 1987.

By the time of the SSM/I launch in June 1987, an updated version of the AES/PhD algorithm was introduced by AES and the Microwave Group at the Institute of Space and Terrestrial Science (ISTS), York University, which is now known as the AES/YORK algorithm. This algorithm has been evaluated with the HAC algorithm in this validation program. The AES/YORK algorithm incorporates weather and sea state corrections to aid in the retrieval of ice type (fraction), ice concentration, and ice edge position for operational ice reconnaissance. The Canadian validation program was based on the criteria listed in Table 10.1, and the more stringent Canadian criteria for resolution requirement of ice age and total ice concentration were applied to the AES/YORK algorithm while the U.S. criteria were applied to the HAC algorithm.

The Canadian validation program also involved an operational demonstration and evaluation project in which both the AES Ice Branch, Ice Forecasting Division (ACIF), and the U.S. Navy/NOAA Joint Ice Center participated. Both ice centres were given near real-time ice charts using the AES/YORK ice algorithm by pulling near real time SSM/I data from the Fleet Numeric Oceanographic Center (FNOC) in Monterey, California. This was made possible through support from the Naval Research Laboratory (NRL).

Because the AES/YORK ice algorithm uses weather and sea state corrections to enhance the retrieved ice parameters, a number of useful side products were obtained for the ice-free ocean area, such as wind speed, areas of precipitation, atmospheric water vapor, and cloud amount. Six Canadian weather centres participated in an operational demonstration and evaluation of these parameters from 20 January to 31 March 1988. The results of this exercise, which were very promising, have been published in a report by Ramseier et al. [1]

10A.2 THEORY

In the microwave region, the radiation intensity received by a radiometer is proportional to the absolute temperature of the medium. This apparent temperature is referred to as brightness temperature, T_B. The attenuation of the surface-emitted radiation and the transmittance of the atmosphere are both related to the optical thickness of the atmosphere (c).

In the absence of scattering, the brightness temperature sensed by a satellite radiometer can be represented by [2]:

$$T_B(z,\theta) = T_{B1}(z,\theta) = e^{-c}[\epsilon T_s + (1-\epsilon)T_{B2}] \tag{1}$$

where T_B = brightness temperature,
z = satellite location height,
θ = incidence angle,
ϵ = effective surface emissivity,
c = total opacity of the atmosphere along the line of sight,
T_s = surface temperature.

The quantities T_{B1} and T_{B2} are proportional to the upward and downward emission from the atmosphere, respectively, plus attenuated sky background radiation, and can be calculated from:

$$T_{B1} = \int_0^z T(z')g(z')e^{-\int_z^z g(z'')\sec\theta dz''} \sec\theta dz' \tag{2}$$

$$T_{B2} = T_{cosmic}e^{-c} + \int_0^\infty T(z')g(z') e^{-\int_0^{z'} g(z'')\sec\theta dz''} \sec\theta dz \tag{3}$$

where
$c(z) = \int_0^\infty g(z)\sec\theta dz$,
$g(z)$ = total opacity at height z, representing the sum of the contributions from water vapor, oxygen, and liquid water droplets in cloud.

As the mixing ratio of oxygen is essentially constant and the absorption coefficient is very weakly temperature dependent, they contribute to a constant offset in the 1 to 40 GHz region. The absorption caused by non-precipitating water droplets in the atmosphere has a linear dependence on the amount of liquid water and a quadratic variation with frequency [3].

The intensity of the atmospheric radiation can be calculated using results from Swift et al., [4]. For the frequencies used in the algorithms validated, some typical values for the opacity coefficient and the atmospheric contribution to the observed brightness temperature are presented in Table 10A.1

TABLE 10A.1				
TYPICAL VALUES FOR THE OPACITY COEFFICIENT (C) AND THE ATMOSPHERIC CONTRIBUTION (T_{Bl}) TO THE OBSERVED BRIGHTNESS TEMPERATURE				
Location	Typical Values at 19 GHz		Typical values at 37 GHz	
	c	$T_{Bl}CT_{Bl}$	c	$T_{Bl}CT_{Bl}$
Polar Regions	0.025	6.7	0.049	13.0
Midlatitudes	0.050	12.5	0.100	25.0

The emissivity of different targets is a consequence of their dielectric properties. As water is a polar molecule, it has a very large dielectric constant at microwave frequencies which results in a large reflectivity (low emissivity) for a liquid water surface such as the ocean. Most solid surfaces have emissivities in the range 0.80 to 0.95, so there is a significant contrast between liquid water surfaces, such as lakes, rivers, and the oceans, and solid surfaces such as land and sea-ice. The low emissivity of the open ocean makes it a good background for viewing the intervening atmosphere.

The higher salinity of first-year ice causes it to be optically opaque and, therefore, its microwave signature is almost frequency independent. The virtually de-salinated near surface portion of old ice makes it optically thin, i.e., radiation emanates from a thicker layer of old ice. A significant part of the radiation from old ice is suppressed by volume scattering within the ice because of air pockets formed during summer melt and brine drainage. The brightness temperature signature of old ice is, therefore, generally lower than that of first-year ice. As the sensitivity to volume scattering is inversely related to the wavelength of the radiation, at higher frequencies one would observe larger variability in the brightness temperature of old ice.

The upwelling brightness temperature of a scene containing open ocean and various amounts of sea-ice is a function of the ice concentration, ice emissivity (i.e., ice type), the physical temperature of the components, and the amount of water vapor and liquid water in the intervening atmosphere. Assuming that the ice cover within the field of view is a mixture of old ice and first-year ice, the brightness temperature sensed on the i-th channel of the radiometer can be expressed as:

10A.3

$$T_{Bi} = e^{ci} \{F[e_{Fi}T_F+(1-e_{Fi})T_{B2}] +M[e_{Mi}T_M +(1-e_{Mi})T_{B2} +(1-IC)[e_{wi}T_w+(1-e_{wi})T_{B2}]\} + T_{B1}$$

$$(4)$$

where

ci	=	total atmospheric opacity,
IC	=	total ice cover fraction,
F	=	fraction of first-year ice,
M	=	fraction of old ice,
ϵ_{Fi}, ϵ_{Mi}, ϵ_{Wi}	=	surface emissivities of first-year ice, old ice, and sea water,
T_F, T_M, T_W	=	surface temperatures of first-year ice, old ice, and open ocean,
T_{B2}	=	incident sky temperature at the surface caused by atmospheric downward emission,
T_{B1}	=	contribution from atmospheric upward emission.

From equation 4 it follows that by considering the difference of T_{BV} (from vertical channel) and T_{BH} (from horizontal channel) for 37 GHz, one minimizes the contribution from the atmosphere:

$$T_{BV}-T_{BH}=e^{-c}[Fd\epsilon_F(T_F-T_{B2})+Md\epsilon_M(T_M-T_{B2}) + (1-IC)d\epsilon_w(T_w-B2)] \qquad (5)$$

where $d\epsilon_F = \epsilon_{FV}-\epsilon_{FH}$,
$d\epsilon_M = \epsilon_{MV}-\epsilon_{MH}$,
$d\epsilon_w = \epsilon_{WV}-\epsilon_{WH}$.

This equation was used in developing the HAC algorithm.

An algorithm that calculates ice concentrations by solving equation 5 for IC, i.e., assuming that a possible solution can be of the form:

$$IC=A*DT_B+B \qquad (6)$$

where coefficients A and B are calculated from equation 5, after making reasonable assumptions about the physical temperatures of the various components and selecting appropriate atmospheric parameters. It can be demonstrated that $DT_B=T_{BV}-T_{BH}$ decreases with the increase in optical opacity (because of larger amounts of water vapor and cloud cover) and the increase in emissivities of open ocean (because of surface roughness), which implies that an algorithm of the type described above will yield erroneous ice concentration retrievals, particularly in weather where high levels of water vapor in the atmosphere, cloud cover, and wind-roughened seas are experienced. To improve on the ice information retrieval reliability for all weather conditions,

a dual frequency, dual polarization (19, 37 GHz) algorithm was developed and is described in Section 10A.3.

10A.3 HAC ALGORITHM

The HAC algorithm was derived by using equation 5. The following assumptions were used to be able to evaluate IC from equation 6: the surface temperatures $TF = T_M = T_W = T_S$, $d\epsilon_P = d\epsilon_M = d\epsilon_I$. Using simple algebra, the coefficients A and B can be calculated from the following:

$$A = e^{-c}/[(T_S - T_{B2})(d\epsilon_I - d\epsilon_W)] \tag{7}$$

$$B = d\epsilon_W/(d\epsilon_W - d\epsilon_I) \tag{8}$$

Climatological mean values of atmospheric water vapor, liquid water, ice surface temperature, and emissivities were used as inputs to evaluate parameters A and B for different climatic zones [5].

Determination of ice type is achieved by computing the effective average ice brightness temperature within the footprint and comparing it with a preselected value, T_C (e.g., brightness temperature of a sample of 35% old ice and 65% first-year ice cover). The equation for the calculation of effective brightness temperature, T_X, using the component from the 37 GHz vertical channel ($T_{B(37V)}$), is as follows:

$$T_X = [C_0 + C_1 * T_{B(37V)}] * IC + C_2 \tag{9}$$

The coefficients C_0, C_1, and C_2 are calculated using climatology. If $T_X > T_C$, the sea-ice fraction within the observed area is flagged as first-year ice. For $T_X < T_C$ the ice cover fraction is identified as old ice.

Weather correction criteria were imposed on sea-ice concentration retrieval after it was observed that false ice information was obtained because of the influence of wind and atmosphere. The correction procedure uses cut-off values for the 19 GHz horizontal component, $T_{B(19H)}$,, and $T_{B(37V)} - T_{B(37H)}$. The ice concentrations are calculated only if the following conditions are met:

$$IC > 10\%$$
$$T_{B(19H)} > 140°K$$
$$T_{B(37V)} - T_{B(37H)} > 5°K$$

If these conditions are not met, the footprint is declared to be ice free. Figures 10A.1 to 10A.4 illustrate the results of such weather corrections (Lo, personal communication, 1987).
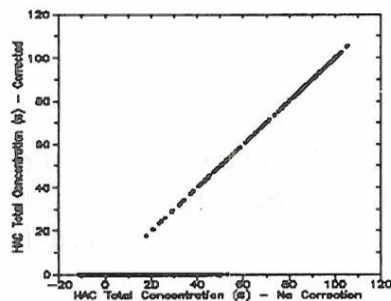
Figure 10A.1 - The effect of weather filtering on the HAC algorithm retrieval of total ice concentrations, orbit 3967, Labrador region.

Figure 10A.2 - Comparison between total ice concentrations retrieved with AES/YORK and HAC algorithms, orbit 3967, Labrador region.

Figure 10A.3 - The effect of weather filtering on the HAC algorithm retrieval of total ice concentrations, Gulf of St. Lawrence, 27 January 1988.

Figure 10A.4 - Comparison between total ice concentration retrieved with AES/YORK and HAC algorithms, orbit 3379, NE Newfoundland waters, 14 February 1988.

The HAC algorithm was first tested on SMMR data. During the evaluation it became obvious that although the retrieval of total ice concentration was within the specifications for ideal weather conditions, areas of rough seas and overcast sky were identified incorrectly as ice-covered ocean.

Prior to launch, a simple procedure for removing some of the weather effects on the ice retrieval was added to the algorithm, but as shown in Figures 10A.1 and 10A.2, the problem was not solved for severe weather conditions. In addition, the accuracy of the ice edge location was degraded by introducing a lower limit (10%) on the calculated ice concentrations. Using only the 37 GHz channel provides the highest available resolution, however, it could lead to errors in total ice concentration estimate and ice type flagging when the ice surface is wet or

under a heavy snow cover. At the onset of snow melt one would also observe large differences in retrieved ice concentrations, depending on the time of the observation.

10A.4 AES/YORK ALGORITHM

Equation 4 written for the 19 and 37 GHz channels can be solved for F and M, with seasonal/regional values for c (optical opacity) and T_{atm} (atmospheric component). Sample areas in the Arctic and the east coast of Canada were selected for establishing passive microwave signatures of first-year ice, old ice, and calm open ocean.

Equation 4 can be rewritten for each channel (1-4) in the following manner:

$$T_{B1} = A_1(B_1*F + C_1 *M + D_1 * W) + TO1 \tag{10}$$

$$T_{B2} = A_2(B_2*F + C_2 *M + D_2 * W) + TO2 \tag{11}$$

$$T_{B3} = A_3(B_3*F + C_3 *M + D_3 * W) + TO3 \tag{12}$$

$$T_{B4} = A_4(B_4*F + C_4 *M + D_4 * W) + TO4 \tag{13}$$

where F, M, and W are fractions of first-year ice, old ice, and open ocean, respectively. A_1, A_2, A_3, A_4, and TO1 to TO4 are atmospheric correction parameters for each channel. A_2, A_3, and A_4 can be expressed in terms of A_1 using frequency dependence of atmospheric absorption coefficients [3]. B_i, C_i, and D_i represent sensitivity coefficients to the presence of various ice types and open ocean. All these parameters were described in detail in PhD Associates Ltd., [6]. The set of equations 10 to 13 can be solved for A_1, F, M, and W. Prior to the calculations of ice concentration and ice type identification, the input brightness temperatures (at 37 GHz and 19 GHz for both polarizations) are subjected to a multi-step testing procedure. The results of this testing determine whether the sensed radiation was emitted from an ice-covered area or from open ocean, as shown in the flow chart given in Figure 10A.5.

The first test decision was made using the contrast between the brightness temperatures of open ocean under heavy cloud cover and ice cover radiance. A discriminating function (D) of 19 and 37 GHz vertical components was generated. Critical values for D were derived by simulating brightness temperatures for open ocean with heavy cloud cover and ice cover near the ice edge. Four ranges of critical values were selected to represent a cross section of atmospheric conditions. D_{MIN1} represents a critical value for discriminating the ice edge area from the open ocean with a surface roughness caused by wind speeds greater than 10 m/s. D_{MIN2} is used for differentiating the ice edge from partially overcast sky and wind-roughened open ocean surface. D_{MAX} is a criteria for distinguishing ice cover greater than 35% concentration. The open ocean, with fully overcast sky, will always result in $D < D_{MAX}$. When $D > 1$, the ice concentration will be more than 90% and old ice is present in the field of view.
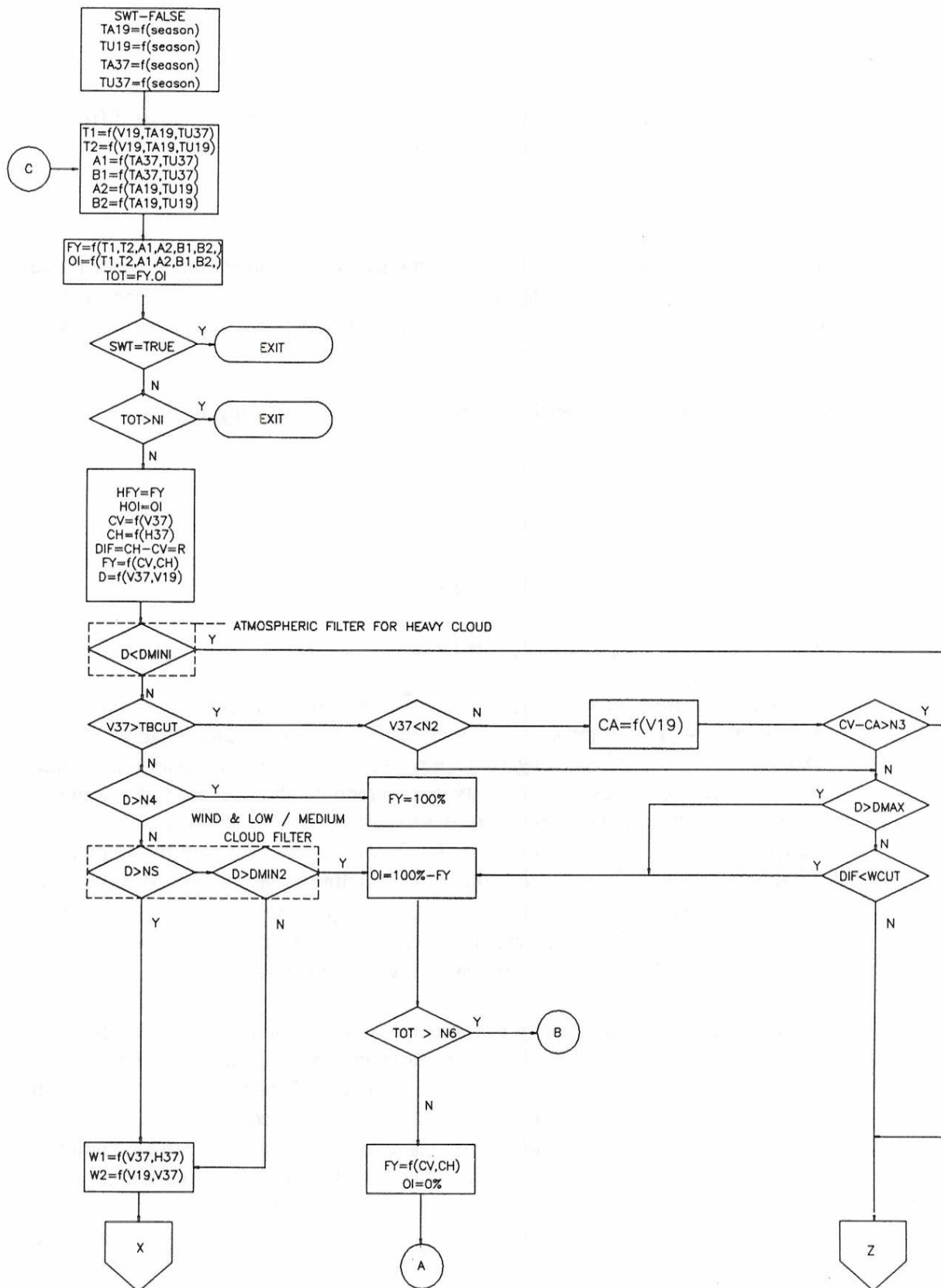
Figure 10A.5 - Complete AES/York temperature independent sea ice algorithm flow chart.
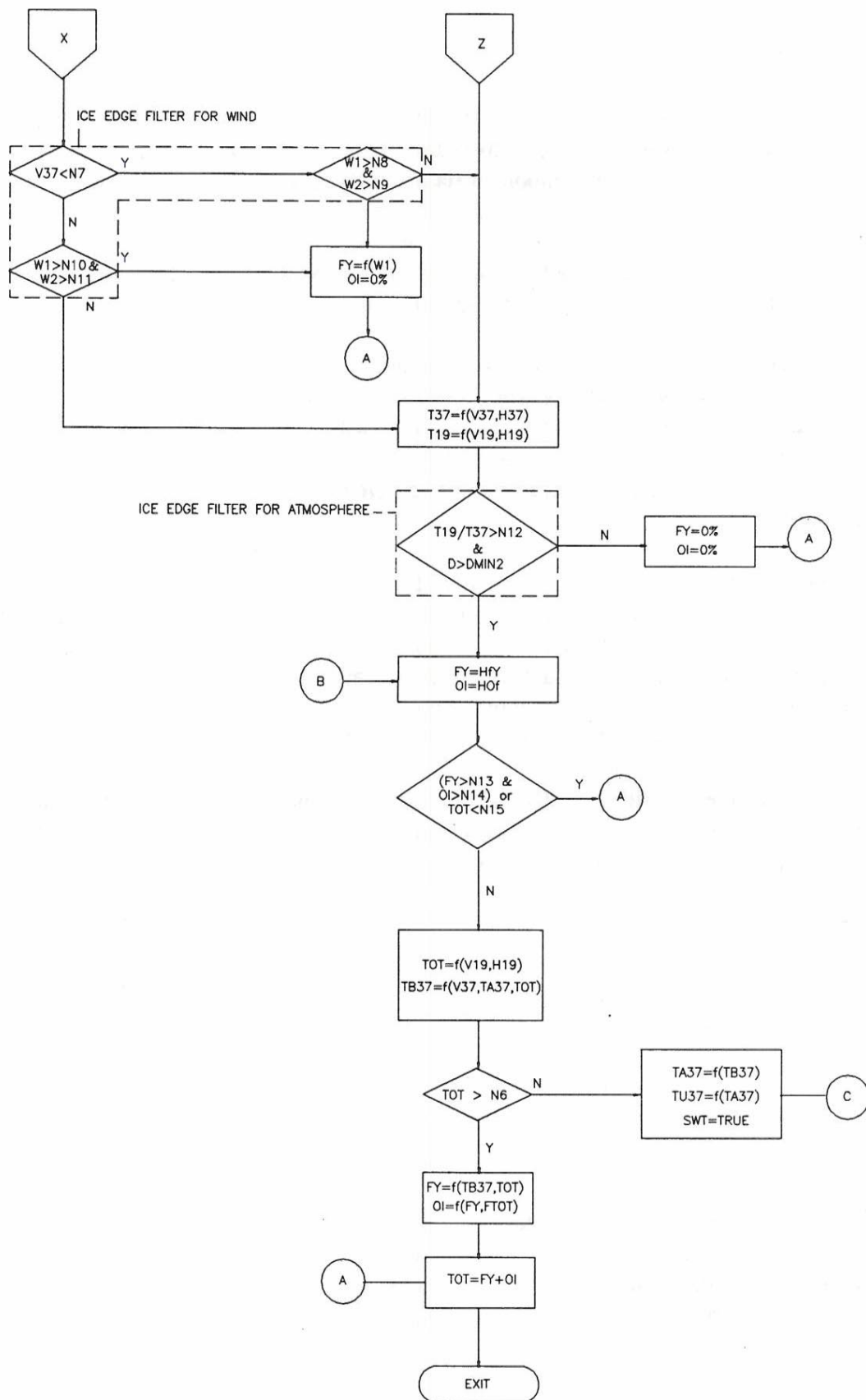
Figure 10A.5 - continued.

10A.9

The data points for which $D < D_{MAX}$ and $D > D_{MIN1}$ are given a second test. This test was designed to distinguish low ice concentration areas from open ocean with low winds and low to moderate cloud cover. This filter component (labelled R) relies on the contrast in sensitivities of the vertical and horizontal components of the 37 GHz channel to the presence of sea-ice.

The third step in the filtering algorithm was set up for the analysis of data points with a $T_{B(37V)}$ greater than that for open ocean, but less than that for a 50% ice-covered value.

The measured brightness temperatures at 19 and 37 GHz are assumed to originate from a partially ice-covered area and partially from open ocean roughened by wind. The possible ice concentrations and wind speed in the ice-free segment within the field of view are calculated using 37 GHz (vertical and horizontal components) and 19 GHz brightness temperatures. Only data points for which calculated ice concentrations are greater than 5% are sent to the data pool for ice chart plotting.

The fourth step is used on data points with $D_{MAX} > D > D_{MIN1}$, but with R values the same as for a rough ocean surface. Assuming that the field of view is an ice-free area, possible surface wind speed and atmospheric contribution to the observed 19 and 37 GHz are calculated. Comparison is then made between the atmospheric information from 37 GHz with the amount estimated for 19 GHz. If the ratio is outside the range (determined from theoretical simulations), the data points are assigned to be from the ice-covered areas.

To implement the algorithm correctly, the filtering components of the algorithm had to be tested. Figures 10A.6 to 10A.8 illustrate the testing of the various components of the filtering procedure in comparison with the HAC algorithm.
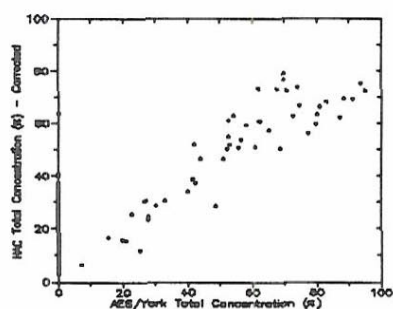
Figure 10A.6 - Comparison between total ice concentrations retrieved with AES/YORK and HAC algorithms, Gulf of St. Lawrence, 27 January 1988.
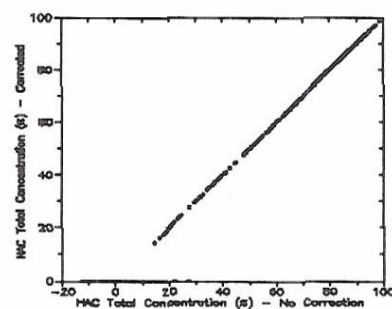
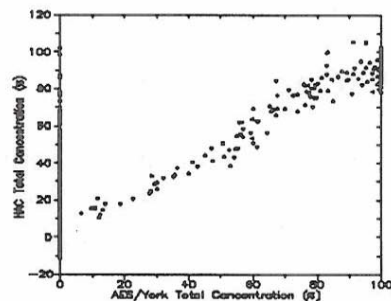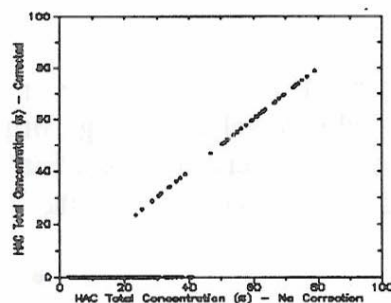Figure 10A.7 - The effect of weather filtering on the HAC algorithm retrieval of total ice concentrations, Beaufort Sea, 26 October 1987.

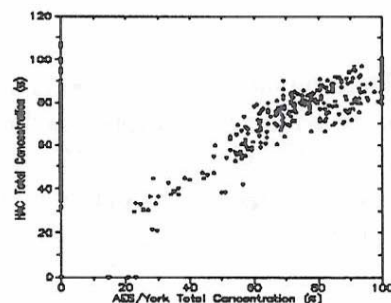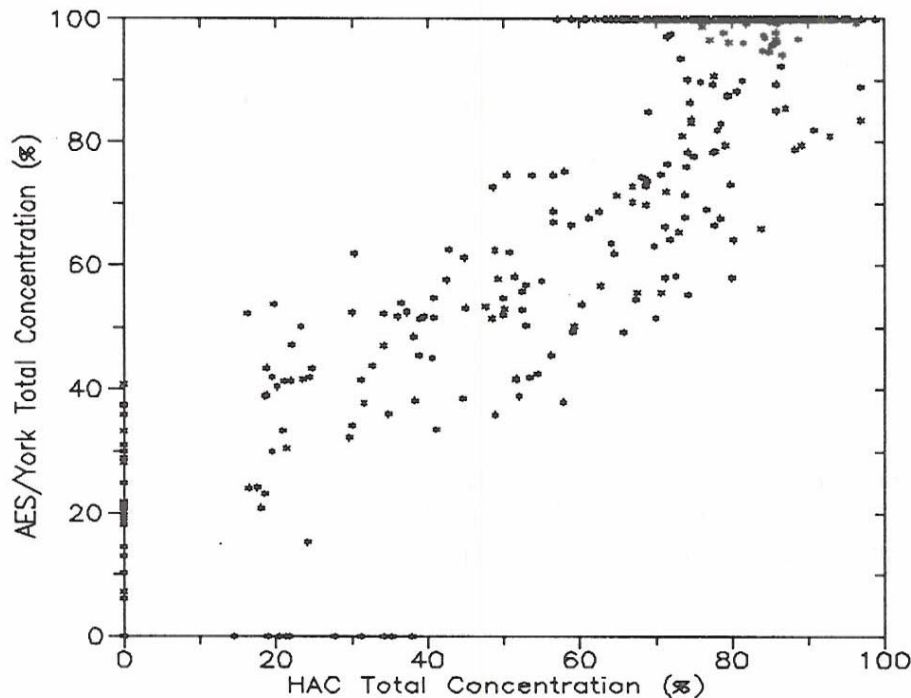Figure 10A.8 - Comparison between total ice concentrations retrieved with AES/YORK and HAC algorithms, Beaufort Sea, 26 October 1987.

After the filtering procedure is completed, ice concentration and ice fractions are calculated using equations 10 to 13, for data points with $D > D_{MAX}$. For data points with $D < D_{MAX}$, 37 GHz channels are used for the calculations of the ice concentrations, therefore using the best resolution in the proximity of an ice edge. The ice type fractions are checked for consistency with the brightness temperatures observed on all four channels used in the algorithm.

10.A5 REFERENCES

[1]   R. O. Ramseier, I. G. Rubinstein, and A. F. Davies, "Operational Evaluation of Special Sensor Microwave/Imager by the Atmospheric Environment Service," Centre for Research in Experimental Space Science, York University, North York, Ontario, Canada, 1988.

[2]   F. T. Ulaby, R. K. Moore, and A. K. Fung, <u>Microwave Remote Sensing: Active and Passive. Volume III: From Theory to Applications</u>, Dedham, MA: Artech House, pp. 1347-1348, 1986.

[3]   P. Gloersen and F. T. Barath, "A scanning multi-channel microwave radiometer for Nimbus-G and Seasat-A," <u>IEEE J. of Oceanic Eng.</u>, No. OE-2, pp. 172-178, 1977.

[4]   C. T. Swift, L. S. Fedor, and R. O. Ramseier, "An algorithm to measure sea-ice concentration with microwave radiometers," <u>J. of Geophys. Res.</u>, No. 90, pp. 1087-1099, 1984.

[5]   R. C. Lo, "A Comprehensive Description of the Mission Sensor Microwave Imager (SSM/I) Environmental Parameter Extraction Algorithm," Naval Research Laboratory, Memo. Rep. 5199, Washington, DC, 1983.

[6]   PhD Associates Inc., "Final Report on Sea-Ice Algorithm Development, AES/PhD Temperature Independent Ice Model," Under contract to Atmospheric Environment Service, Environment Canada.

SECTION 11

PRECIPITATION VALIDATION

by

William S. Olson, Frank J. LaFontaine, William L. Smith,
Robert T. Merrill, Barry A. Roth, and Thomas H. Achtor

Cooperative Institute for Meteorological Satellite Studies (CIMSS)
Space Science and Engineering Center
University of Wisconsin-Madison
1225 West Dayton Street
Madison, WI 53706

11.0 PRECIPITATION VALIDATION

11.1 INTRODUCTION

This section summarizes the results of the SSM/I rainfall rate retrieval algorithm validation effort, which has been completed for the midlatitude and tropical climate zones. The validation is presented, followed by an evaluation of the operational D-Matrix rainfall rate retrieval algorithm based upon available ground truth. This section concludes with recommendations for the improvement of the existing retrieval algorithm, and an example application of an alternate algorithm to tropical cyclone data.

11.2 VALIDATION PLAN

11.2.1 Navy Specifications

The validation plan was specifically designed to evaluate the performance of the Hughes "D-Matrix" algorithm for obtaining rainfall rates from SSM/I brightness temperature data. The SSM/I specifications called for an algorithm which would enable the retrieval of rainfall rates from the observations of the DMSP-F8 with 5 mm/hr accuracy over the range 0 - 25 mm/hr at 25 km spatial resolution.

11.2.2 Method

Raingages provide the most accurate standard for point estimates of surface rainfall. However, because of the high spatial variability of precipitation intensity and the requirement to validate 25 km space-scale estimates over ocean as well as land areas, area-averaged radar rain estimates were utilized as the primary source of validation data in this study.

In order to maximize the correlation between the radar rain rate estimates and surface rainfall amounts, only low antenna elevation angle ($\alpha \leq 1°$) plan-position indicator (PPI) scans were used. Also since the radar beam height increases with range, no radar measurements beyond a range of 220 km were considered. Ground clutter and obvious radar artifacts were also screened. The remaining bin reflectivities were converted to rainfall rate using a standard relationship between the reflectivity factor Z and the rainfall rate R ($Z = 200 R^{1.6}$) and then interpolated to a 5 km cartesian grid. In this report R is in units of mm/hr.

Typically three radar PPI sweeps bracketing the DMSP-F8 overpass time were processed. All gridded rainfall rates falling within a 625 km^2 circular area of a given SSM/I all-channel brightness temperature scene were time-interpolated to the SSM/I measurement time. The time-interpolated, gridded rain rates were subsequently area-averaged and then stored along with the corresponding seven sensor data record (SDR) brightness temperatures. In addition, the time-interpolated rainfall rates at 5 km resolution were recorded to allow for improved calibration.

Since individual radar-derived rainfall rate estimates can have a high uncertainty, simultaneous raingage measurements were also recorded for the purpose of calibration. Wilson and Brandes [1] demonstrated that errors in radar storm-total estimates of rainfall rate could be reduced from 63% to 24% using calibrating raingages.

11.2.3 Data Sources

SSM/I sensor data records (SDR) and environmental data records (EDR) coinciding with radar measurements of precipitation were obtained from the Naval Research Laboratory (NRL) archive.

Surface truth for midlatitude validations was obtained from seven radar sites in the United Kingdom operational network and the Patrick Air Force Base (PAFB) radar at Cape Canaveral, Florida (see Table 11.1 for specifications). Each of these radars provided significant coverage of both land and ocean areas, and were operating almost continuously since the launch of the DMSP-F8.

The United Kingdom data were obtained from archives maintained by the British Meteorological Office (BMO), and the PAFB data were retrieved from laser disk recordings compiled by personnel at the Severe Storms Laboratory at NASA/Goddard Space Flight Center.

Raingage observations from three to five telemetering raingages are recorded operationally for five of the seven United Kingdom sites. If sufficient raingage hourly totals are recorded in the same time frame as a given radar sweep, then a real-time correction is applied to the radar data using the scheme described by Collier, et al. [2]. Hourly raingage totals and corresponding hourly-integrated radar totals were provided along with the radar data on BMO archive tapes.

Hourly raingage data from National Weather Service (NWS) gages in the vicinity of the PAFB radar site were obtained from National Climatic Data Center archives at Asheville, North Carolina. Twenty-one raingages in the NWS network provided hourly rainfall rate totals within a 200 km radius of the PAFB site.

Radars operating continuously at Darwin, Australia and Kwajalein, Marshall Islands were utilized to validate rain rate retrieval algorithms at tropical latitudes. Located on the northwest coast of Australia, the Darwin radar provides reflectivity data both over the continent and over the Indian Ocean. The Kwajalein radar yields rain rate data exclusively over the Pacific Ocean. Networks of raingages were maintained in the vicinity of both radars to check the calibration of the rainrate estimates; however, no real-time correction was applied to the data from either site.

11.2.4 Satellite/Radar Data Geolocation

Since rainfall is highly variable in both space and time, accurate geolocation of both the satellite and radar data was essential to the validation effort.

	TABLE 11.1			
VALIDATION RADAR SITES AND NUMBER OF CALIBRATING RAINGAGES				
Radar Site	Specifications	Latitude	Longitude	Number of gages
(Midlatitudes)				
Patrick Air Force Base Cape Canaveral, Florida	5 cm, C-band, 1.1° beamwidth	28.255N	80.606W	21
Camborne, England, United Kingdom	10 cm, S-band, 2° beamwidth	50.218N	5.327W	3
Upavon, England, United Kingdom	10 cm, S-band, 2° beamwidth	51.299N	1.781W	3
Clee Hill, England, United Kingdom	5.6 cm, C-band, 1° beamwidth	52.297N	2.597W	3
Hameldon, England United Kingdom	5.6 cm, C-band, 1° beamwidth	53.756N	2.281W	5
Chenies, England United Kingdom	5.6 cm, C-band 1° beamwidth	51.688N	0.053W	5
Shannon, Ireland, United Kingdom	10 cm, S-band 2° beamwidth	52.791N	6.936W	0
Castor Bay, North Ireland, United Kingdom	5.6 cm, C-band 1° beamwidth	54.503N	6.341W	0
(Tropics)				
Darwin, Northern Territory, Australia	5.3 cm, C-band, 1.7° beamwidth	12.457S	130.925E	26
Kwajalein, Marshall Islands	10.7 cm, S-band 2.2° beamwidth	8.72N	167.73E	9

Errors as great as 30 km were observed in the position of land-ocean boundaries in the DMSP-F8 SSM/I imagery. Since precipitation fields can vary greatly on a spatial scale of 10 km or less, correlations between brightness temperature features in the SSM/I imagery and radar echoes were degraded in many situations. Through a cooperative effort between scientists at University of Wisconsin, University of Massachusetts (UMASS), and the Naval Research Laboratory (NRL), a method was developed to automatically relocate the SSM/I data.

The method consisted of an optimization routine which searched for corrections in the spacecraft pitch and yaw angles that maximized the correlation between discontinuities in the 85.5 GHz horizontally-polarized SSM/I brightness temperature imagery and the known location of coastal boundaries as specified in the World Data Base II (WDB II) coastline map. The transformation between pitch and yaw perturbations and perturbations in the earth coordinates of SSM/I measurements was provided by Mark Goodberlet of UMASS. The World Data Base II coastline map was provided by Gene Poe and Pete Conway of NRL.

A digital edge detector was applied to the 85.5 GHz horizontally-polarized brightness temperature data in original scan format to locate coastal discontinuities in the imagery. If the edge detector identified a brightness temperature discontinuity between adjacent footprints of at least 30 K in the United Kingdom or Kwajalein imagery, or a discontinuity of 15 K in the Florida or Darwin imagery, then the location of the discontinuity was recorded on a 4 km resolution grid using a standard map projection. The World Data Base II coastlines were referred to the same grid. A smaller edge detector threshold was utilized at tropical and subtropical latitudes to account for the smaller land/ocean contrast at those latitudes. Swath data from Alaska or the U.S.S.R. was used to geolocate Kwajalein data from the same orbit, due to the paucity of large land masses in the vicinity of Kwajalein Island.

A simplex algorithm was invoked to iteratively search for the spacecraft pitch and yaw perturbations which maximized the number of grid-point "matches" between the edge-detected coastline and the World Data Base II coastline over a 2000 km section of the SSM/I swath centered on the region of interest. The effect of adding a roll perturbation to the optimization scheme did not significantly improve image registration.

Upon review of 10 to 20 relocated SSM/I images, the automated procedure appeared to locate the satellite data to within about 6 km of the World Data Base II coastline. The WDB II coastline is reported to be accurate to within 3 km.

The validation radar data were earth-located using the recorded range of the radar bin and the elevation and azimuth angle of the radar antenna. The range of the radar bin along the earth's surface and the bin altitude were computed using the standard formulae presented in Battan [3]. Given the earth range of the bin, the azimuth of the radar antenna, and the known latitude and longitude of the radar site, the earth location of the radar bin was determined using the geodetic formulae of Sodano [4]. The uncertainty in the earth location of any radar bin is estimated to be on the order of 1-2 km.

11-4

11.2.5 Radar Calibration

With the exception of the Shannon and Castor Bay radars, for which no raingage data were available, an attempt was made to calibrate the midlatitude radar-derived precipitation intensities using coincident hourly raingage recordings. Approximately 50% of all the United Kingdom radar data corresponding to DMSP-F8 overpasses had been pre-calibrated using the scheme described in Collier, et al. [2]. Their scheme relies upon a time-series analysis of radar-to-gage ratios, determination of radar "bright-bands", and adjustments for orographically-forced precipitation. Rainfall rates obtained from the uncalibrated United Kingdom radars, which had been assessed using $Z = 200 R^{1.6}$ (Marshall and Palmer [5]), were left unaltered.

For the remaining DMSP-F8 overpass times, there were generally insufficient raingage data to perform a radar calibration, unless gage data covering a period of one day or more were incorporated into the analysis. Despite the relatively large number of raingages recording in the vicinity of the PAFB radar site, there were again insufficient data to perform instantaneous radar calibrations for most DMSP-F8 overpass times. The inadequacy of the gage networks for instantaneous calibrations is due to the high space- and time-variability of precipitation.

The radar data available from Darwin and Kwajalein were insufficient for performing raingage calibrations. The Marshall and Palmer [5] relationship was utilized to interpret the reflectivity data from these radars.

All radar-derived rainfall rates within a 625 km² circular area centered on a given SSM/I all-channel scene were time-interpolated to the SSM/I measurement time and subsequently area-averaged to yield a ground truth rainfall rate product.

11.3 VALIDATION ERROR

The total validation error can be divided into two general categories: (1) errors in the gridded radar estimates of rainfall rate (at 5 km resolution), and (2) errors arising from atmospheric variability linked to discrepancies in the space and time collocation of the 625 km² area-average rainfall rates and the SSM/I all-channel measurements.

These error categories can be further subdivided into contributing error sources. It will be assumed in this analysis that raingages provide an accurate standard for surface rain totals over a period of one hour. If it is also assumed that the errors from contributing sources are uncorrelated, then the error variance of a gridded and time-interpolated radar rainfall rate with respect to a gage estimate can be expressed as

$$\sigma_{rgint}^2 = \sigma_{rg}^2 + \sigma_{int}^2, \tag{1}$$

where σ_{rg} is the error of an instantaneous gridded radar rain rate, and σ_{int} is the error arising from time-interpolation of the gridded radar measurement to the SSM/I measurement time.

In practice it is only feasible to estimate the error of hourly-integrated radar rain rates with respect to gage totals over the same period. In the British Meteorological Office radar calibration scheme, rain rates from approximately 12 radar sweeps are averaged to obtain an hourly total. The error variance of hourly-integrated radar rain rates may be approximated by

$$\sigma_{rgh}^2 = \{\sigma_{rg}^2/n\} + \sigma_f^2. \tag{2}$$

Here, n is the number of radar sweeps utilized in the hourly integration and σ_f is the error introduced by integrating a finite number of sweeps to form an hourly total. Combining Eqs. (1) and (2):

$$\sigma_{rgint}^2 = n\{\sigma_{rgh}^2 - \sigma_f^2\} + \sigma_{int}^2 \tag{3}$$

A value for σ_{rgh} of six tenths of the radar-derived rainfall rate (i.e. 0.6 R, where R is the rain rate) was taken from a study by Wilson and Brandes [1]. Harrold, et al. [6] showed that σ_f was on the order of 0.1 R. It is assumed that R, which is based on hourly averages in the preceding estimates, can be approximated by the instantaneous rain rate for the purpose of making an error estimate. This assumption may lead to an overestimate of the error for high instantaneous rain rates, since the average rain rate over an hour period which includes a high rain rate event is likely to be lower than the instantaneous rain rate.

The radar rain rate time interpolation error is almost negligible, since the data are interpolated from radar measurements separated by 15 minutes at most. A value of 0.05 R is estimated for σ_{int}.

The errors due to area-averaging of gridded radar data and co-registration with the D-Matrix estimates can be expressed as

$$\sigma_{rgt}^2 = \{\sigma_{rgint}^2/m\} + \sigma_{rgs}^2 + \sigma_d^2. \tag{4}$$

where m is the number of time-interpolated, radar grid boxes averaged over a 625 km² area, σ_{rgs} is the error due to misregistration of SSM/I and radar measurements, and σ_d is the error introduced by the discretization of the D-Matrix rain rate estimates.

Typically 25 radar grid boxes are averaged per 625 km² area. After relocation of the SSM/I measurements according to the method described in Section 11.2.4, the total misregistration between SSM/I and radar measurements is approximately 7 km, based upon comparisons to the World Data Base II coastlines and an estimated 1 - 2 km error in the radar measurements. The validation error due to misregistration is estimated by considering the error incurred by estimating a "reference" area-averaging rain rate using area-averaged rain rates at different displacements from the reference location. Based upon this approach, σ_{rgs} for a displacement of 7 km is found to vary as a logarithmic function of the rainfall rate.

The rounding of D-Matrix rain rate estimates to integral values leads to a constant $\sigma_d = .29$ mm/hr.

Utilizing the individual errors estimated above, the total validation error is evaluated and plotted in Figure 11.1. Also plotted is the validation error that would result if the geolocation of the SSM/I data was not corrected, assuming an average 25 km misregistration error for uncorrected data. It may be noted that from the figure that a 35% to 60% reduction in the validation error is achieved by relocating the SSM/I data using the automated procedure. The validation error of the relocated data varies almost linearly with rainfall rate, with about a 45% error at 24 mm/hr rain rate.

11.4 D-MATRIX ALGORITHM EVALUATION

11.4.1 Data Samples

DMSP-F8 overpasses of the United Kingdom and PAFB validation sites were separated by season into summer, spring-fall, and winter overpasses. Collocated SSM/I and radar measurements from nine overpasses of the United Kingdom sites and three overpasses of the PAFB site during August of 1987 composed the summer validation data set. Seven United Kingdom overpasses during September of 1987 and six overpasses of PAFB during September and November of 1987 and March of 1988 contributed to the spring-fall validation data set. The winter validation data set was derived from twenty-five overpasses of United Kingdom sites during January and February of 1988.

Only radar data from the tropical warm season were available from the Darwin and Kwajalein sites. Data from eleven SSM/I overpasses of Darwin and nine overpasses of Kwajalein were collocated with the averaged radar rain data to produce the tropical validation data set. The Darwin overpasses occurred during February and March of 1988, while the Kwajalein overpasses spanned the months of August - November of 1988.

The total number of collocated area-averaged radar rain rate estimates and SSM/I all-channel scenes are listed by season in Table 11.2. Listed separately are the number of collocated measurements over land and ocean backgrounds. Also included are the number of collocated measurements for which the area-average radar rain rate was at least 1 mm/hr.

Despite prescreening of overpasses to identify those in which significant rain events were observed by both the SSM/I and radar, only a small fraction of the total number of collocated data exceeded the 1 mm/hr threshold. The skewed distribution of rain data is further illustrated by the seasonal histograms in Figure 11.2. The histograms indicate that a large percentage of rain events at 25 km resolution have intensities less than 1 mm/hr. At midlatitudes, the winter data are more highly skewed towards light precipitation than the summer and spring-fall data. The highly skewed winter distribution results in a relatively small number of collocated winter
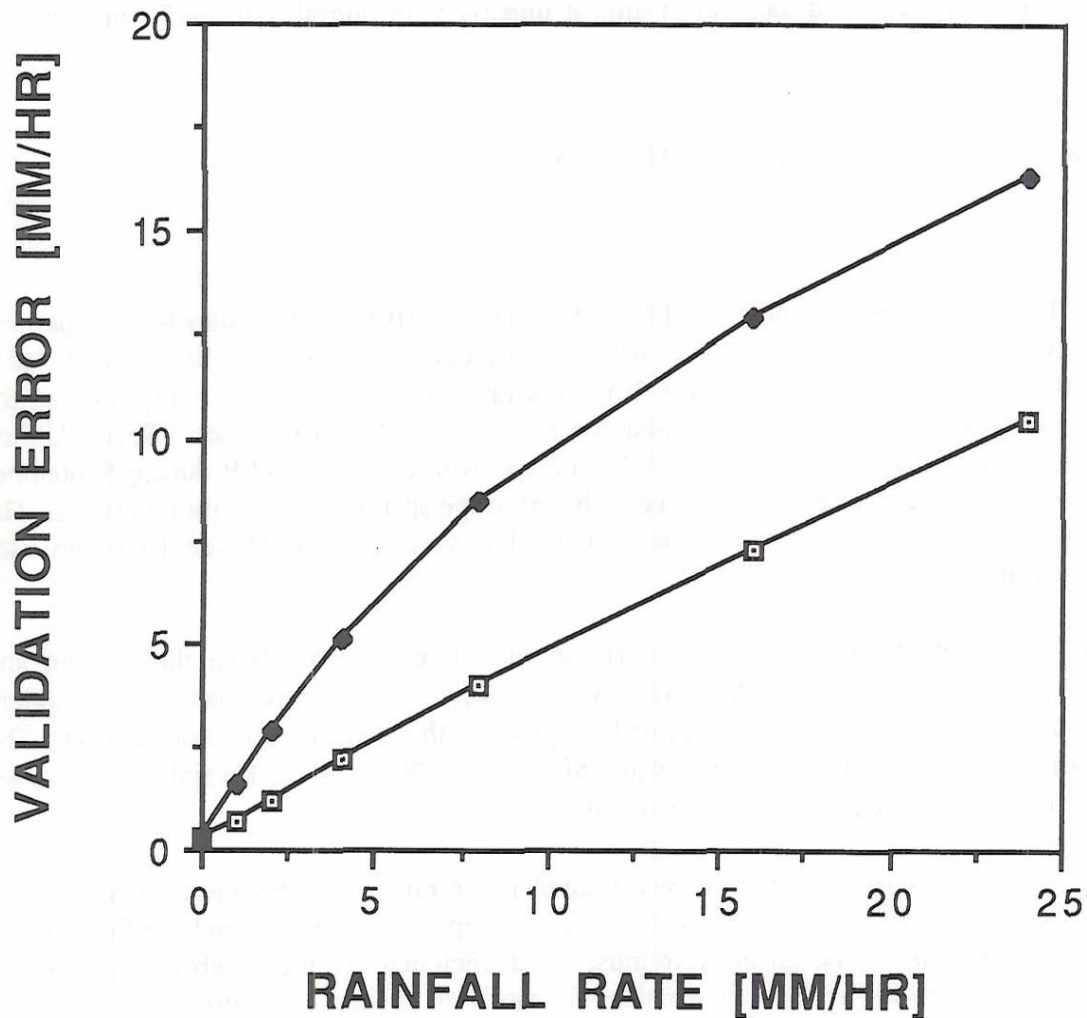
Figure 11.1 - Validation error as a function of rainfall rate. Squares indicate the error standard deviation for an average SSM/I-radar measurement misregistration of 7 km. An average misregistration error of 7 km is expected after the geolocation method described in Section 11.2.4 is applied to the SSM/I measurements. If the geolocation of SSM/I measurements is not corrected, then the average misregistration of SSM/I and radar measurements is approximately 25 km. The validation error for uncorrected data is indicated by the solid dots in the figure.

TABLE 11.2

NUMBERS OF COLLOCATED SSM/I ALL-CHANNEL SCENES AND RADAR DERIVED RAINFALL RATES

	LAND	OCEAN	TOTAL
(Midlatitudes)			
Summer	1155 (85)	551 (28)	1706 (113)
Spring-Fall	1794 (217)	1045 (181)	2839 (398)
Winter	4067 (205)	2236 (9)	6303 (214)
All Seasons	7016 (507)	3832 (218)	10,848 (725)
(Tropics)			
Warm Seasons:	342 (41)	1365 (180)	1707 (221)

Numbers of collocated SSM/I all-channel scenes and radar-derived rainfall rates for both the midlatitude summer, spring-fall, and winter seasons, and for the tropical warm season. The numbers of collocated data over land and ocean regions are also individually tabulated. The number in parentheses is the subset of the total sample for which the radar-derived rainfall rate was at least 1 mm/hr.
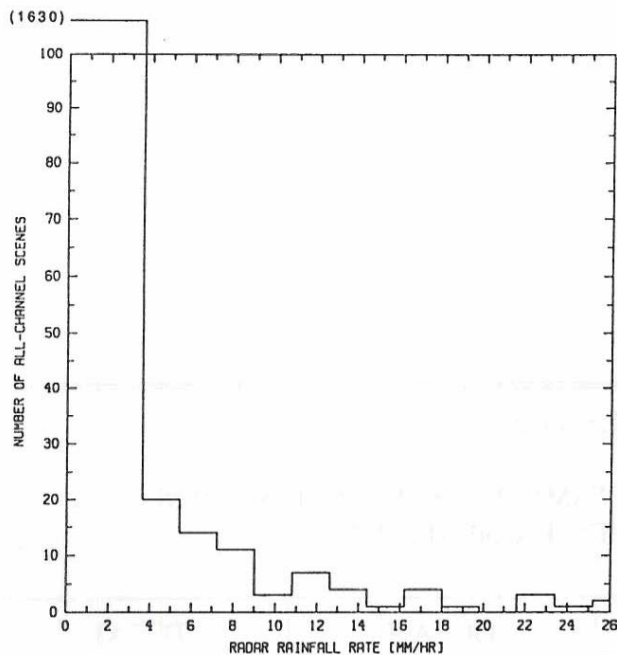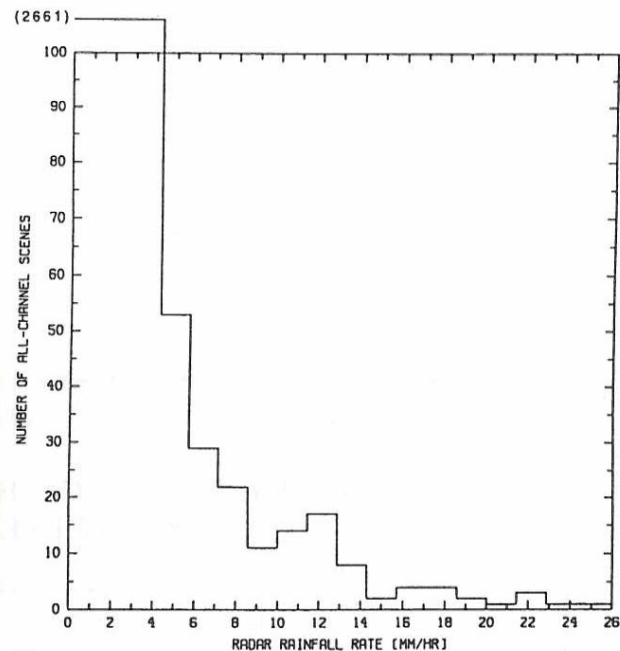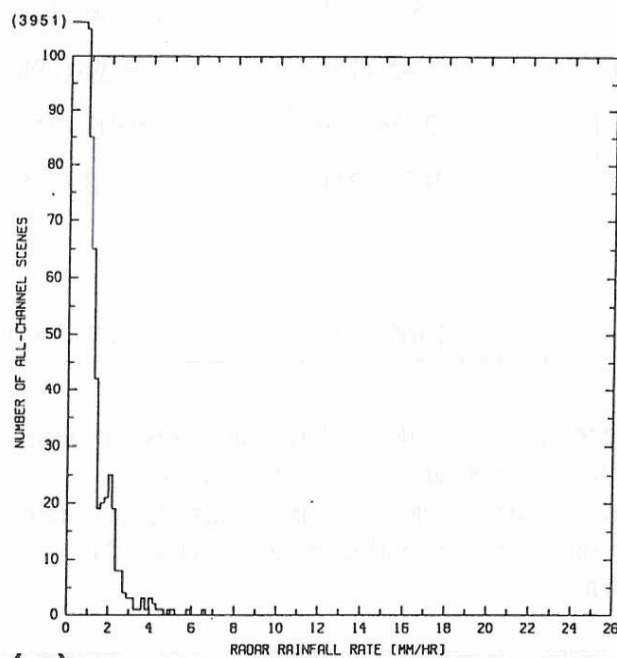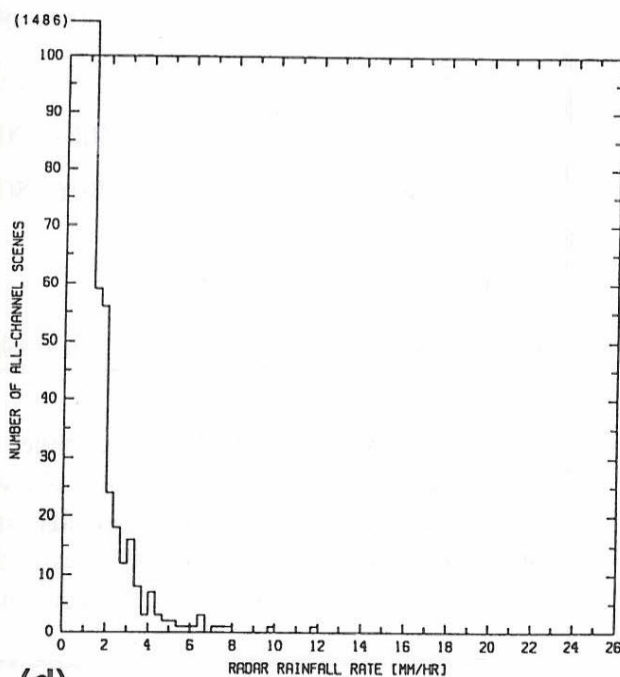
Figure 11.2 - Number of collocated SSM/I all-channel scenes and area-average radar measurements versus radar-derived rainfall rate for (a) the midlatitude summer climate zone, (b) the midlatitude spring-fall climate zone, (c) the midlatitude winter climatic zone, and (d) the tropical warm climatic zone. Data over land and ocean are included in the histograms. The number of data in the first rain rate interval is given in parentheses.

11-10

data over the ocean with rainfall rates greater than or equal to 1 mm/hr (See Table 11.2). In the tropics, the rain rate distribution is also skewed towards light rain events (see Figure 11.2d).

The skewed rain distributions have an important bearing on the statistical analyses to be presented in the following sections.

11.4.2 Midlatitude D-Matrix Algorithm Error Statistics

Error statistics of the D-Matrix algorithm rain rate retrievals are presented for the six midlatitude climatic zones in Tables 11.3 through 11.8. To compensate for the naturally skewed distribution of rainfall rates, the retrieval error statistics are stratified. Statistics are computed for different subsamples of the collocated data, such that only D-Matrix estimates and radar-derived rainfall rates exceeding specified minimum thresholds are included. As the minimum rain rate threshold defining a subsample is increased from 0.0 to 0.5, 1.0, 1.5, and 2.0 mm/hr, more emphasis is placed upon the performance of the D-Matrix algorithm at higher rainfall rates. An increase in the minimum rainfall rate threshold is reflected in an increase in the subsample mean rainfall rate and standard deviation (see Tables 11.3 - 11.8). The "error" standard deviation (σ_e) is the standard deviation of the difference between the SSM/I estimated rainfall rate and the radar "ground truth" estimate. In addition to the traditional statistical quantities, the success ratio S, which is the ratio of the rain rate estimate "error" variance, σ_e^2, to the sum of the variances of the validation error and 5 mm/hr retrieval tolerance, is listed for each subsample. An S-ratio greater than 1 indicates that the rain rate estimate falls outside the Navy specifications. An S-ratio less than or equal to 1 implies that the algorithm performance is within the validation accuracy permitted by the data.

Scatterplots of the D-Matrix rain rate estimates versus radar-derived rainfall rates for each of the six midlatitude climatic zones are presented in Figure 11.3. The solid diagonal lines drawn on each of the plots define the ±5 mm/hr retrieval error range.

The D-Matrix error statistics can be compared to the error statistics of the best possible linear model estimate which are included in the second section of each table. The best linear model of the radar rain rates is obtained by regressing the SSM/I corrected brightness temperatures (SDR data) against the radar rainfall rates using a stepwise procedure. Regressions are performed on the same subsamples of collocated data from which the D-Matrix error statistics were derived. Since a lower bound of 0 mm/hr is imposed on the D-Matrix rain retrievals, the same lower bound is imposed on the regression estimates. Each regression relationship represents the best possible model of the radar rain rates which is linear in the brightness temperature data, and therefore it defines an upper limit on the performance that can be expected from any linear retrieval algorithm for the data sample in question.

This section will conclude with a general discussion of the D-Matrix retrieval error statistics. Because the channels selected for rain retrievals over land are the same for each season and the brightness temperature weightings are similar, the discussion will first focus on land

retrievals for the three specified seasons and then move to a discussion of rain retrievals over the ocean.

TABLE 11.3							
STATISTICS FOR THE MIDLATITUDE SUMMER LAND CLIMATIC ZONE							
CASE	n	R_m	σ_R	b	σ_e	r	S
D-Matrix:							
$0.0 \leq R$, $RL \leq 25$ mm/hr	885	.535	2.31	-.086	1.50	.761	.089
$0.5 \leq R$, $RL \leq 25$ mm/hr	48	6.78	6.99	-1.05	5.18	.694	.735
$1.0 \leq R$, $RL \leq 25$ mm/hr	37	8.59	7.00	-1.64	5.87	.596	.809
$1.5 \leq R$, $RL \leq 25$ mm/hr	31	9.74	7.05	-2.00	6.17	.564	.814
$2.0 \leq R$, $RL \leq 25$ mm/hr	27	10.9	6.79	-2.62	6.55	.476	.834
All channel regressions:							
$0.0 \leq R$, $RL \leq 25$ mm/hr	885	.535	2.31	.215	1.47	.784	.086
$0.5 \leq R$, $RL \leq 25$ mm/hr	48	6.78	6.99	.037	4.57	.756	.572
$1.0 \leq R$, $RL \leq 25$ mm/hr	37	8.59	7.00	.012	5.13	.681	.618
$1.5 \leq R$, $RL \leq 25$ mm/hr	31	9.74	7.05	-.001	5.72	.585	.700
$2.0 \leq R$, $RL \leq 25$ mm/hr	27	10.9	6.79	-.001	5.67	.551	.625

The number of collocated SSM/I and radar observations in the sample (n), the mean radar-derived rainfall rate (R_m) and standard deviation (σ_R) of the sample, the bias (b) and error standard deviation (σ_e) of the rain rate estimate, the correlation coefficient (r) between the radar and SSM/I-derived rain rates, and the success ratio (S) for each case are listed above. R_m, σ_R, b, and σ_e are given in units of mm/hr. The first five rows are the statistics of the D-Matrix rain rate estimates (RL) for the indicated subsets of the full data sample. The statistics are stratified because the full data sample is dominated by low rainfall rates. In the second section the statistics of linear regression fits to the same subsets of points are listed. To maintain consistency with the D-Matrix estimates, a lower bound of 0 mm/hr was imposed upon the regression estimates.

TABLE 11.4

STATISTICS FOR THE MIDLATITUDE SPRING/FALL LAND CLIMATIC ZONE

CASE	n	R_m	σ_R	b	σ_e	r	S
D-Matrix:							
$0.0 \leq R,\ RL \leq 25$ mm/hr	1386	.549	1.50	1.56	3.99	.438	.632
$0.5 \leq R,\ RL \leq 25$ mm/hr	205	2.56	2.62	5.82	7.68	.279	2.18
$1.0 \leq R,\ RL \leq 25$ mm/hr	132	3.54	2.82	5.87	8.02	.193	2.24
$1.5 \leq R,\ RL \leq 25$ mm/hr	99	4.12	2.85	6.39	8.35	.163	2.34
$2.0 \leq R,\ RL \leq 25$ mm/hr	79	4.72	2.89	5.97	8.09	.157	2.11
All channel regressions:							
$0.0 \leq R,\ RL \leq 25$ mm/hr	1386	.549	1.50	.043	1.20	.602	.057
$0.5 \leq R,\ RL \leq 25$ mm/hr	205	2.56	2.62	.007	2.31	.472	.197
$1.0 \leq R,\ RL \leq 25$ mm/hr	132	3.54	2.82	.000	2.55	.423	.227
$1.5 \leq R,\ RL \leq 25$ mm/hr	99	4.12	2.85	.002	2.61	.400	.229
$2.0 \leq R,\ RL \leq 25$ mm/hr	79	4.72	2.89	.000	2.65	.399	.226

The number of collocated SSM/I and radar observations in the sample (n), the mean radar-derived rainfall rate (R_m) and standard deviation (σ_R) of the sample, the bias (b) and error standard deviation (σ_e) of the rain rate estimate, the correlation coefficient (r) between the radar and SSM/I-derived rain rates, and the success ratio (S) for each case are listed above. R_m, σ_R, b, and σ_e are given in units of mm/hr. The first five rows are the statistics of the D-Matrix rain rate estimates (RL) for the indicated subsets of the full data sample. The statistics are stratified because the full data sample is dominated by low rainfall rates. In the second section the statistics of linear regression fits to the same subsets of points are listed. To maintain consistency with the D-Matrix estimates, a lower bound of 0 mm/hr was imposed upon the regression estimates.

TABLE 11.5

STATISTICS FOR THE MIDLATITUDE WINTER LAND CLIMATIC ZONE

CASE	n	R_m	σ_R	b	σ_e	r	S
D-Matrix:							
$0.0 \leq R$, RL ≤ 25 mm/hr	3797	.171	.462	1.90	5.22	.387	1.09
$0.5 \leq R$, RL ≤ 25 mm/hr	199	1.39	.898	13.6	14.6	.037	8.29
$1.0 \leq R$, RL ≤ 25 mm/hr	110	1.90	.926	13.3	14.4	.018	7.90
$1.5 \leq R$, RL ≤ 25 mm/hr	66	2.38	.920	12.3	13.5	.144	6.79
$2.0 \leq R$, RL ≤ 25 mm/hr	36	2.90	.973	12.4	13.6	.123	6.71
All channel regressions:							
$0.0 \leq R$, RL ≤ 25 mm/hr	3797	.171	.462	.009	.409	.465	.007
$0.5 \leq R$, RL ≤ 25 mm/hr*							
$1.0 \leq R$, RL ≤ 25 mm/hr*							
$1.5 \leq R$, RL ≤ 25 mm/hr*							
$2.0 \leq R$, RL ≤ 25 mm/hr*							

*None of the channels could explain a significant portion of the variance; therefore no regression fit was attempted.

The number of collocated SSM/I and radar observations in the sample (n), the mean radar-derived rainfall rate (R_m) and standard deviation (σ_R) of the sample, the bias (b) and error standard deviation (σ_e) of the rain rate estimate, the correlation coefficient (r) between the radar and SSM/I-derived rain rates, and the success ratio (S) for each case are listed above. R_m, σ_R, b, and σ_e are given in units of mm/hr. The first five rows are the statistics of the D-Matrix rain rate estimates (RL) for the indicated subsets of the full data sample. The statistics are stratified because the full data sample is dominated by low rainfall rates. In the second section the statistics of linear regression fits to the same subsets of points are listed. To maintain consistency with the D-Matrix estimates, a lower bound of 0 mm/hr was imposed upon the regression estimates.

TABLE 11.6

STATISTICS FOR THE MIDLATITUDE SUMMER OCEAN CLIMATIC ZONE

CASE	n	R_m	σ_R	b	σ_e	r	S
D-Matrix:							
$0.0 \leq R$, $RO \leq 25$ mm/hr	551	.293	1.31	-.094	1.26	.446	.063
$0.5 \leq R$, $RO \leq 25$ mm/hr	14	6.63	2.90	-2.71	4.27	.012	.506
$1.0 \leq R$, $RO \leq 25$ mm/hr	14	6.63	2.90	-2.71	4.27	.012	.506
$1.5 \leq R$, $RO \leq 25$ mm/hr	12	6.56	2.91	-2.14	3.66	.149	.374
$2.0 \leq R$, $RO \leq 25$ mm/hr	11	6.98	2.64	-2.62	3.71	.317	.371
All channel regressions:							
$0.0 \leq R$, $RO \leq 25$ mm/hr	551	.293	1.31	.098	.909	.728	.033
$0.5 \leq R$, $RO \leq 25$ mm/hr	14	6.63	2.90	.000	2.61	.435	.189
$1.0 \leq R$, $RO \leq 25$ mm/hr	14	6.63	2.90	.000	2.61	.435	.189
$1.5 \leq R$, $RO \leq 25$ mm/hr	12	6.56	2.91	.000	2.62	.437	.191
$2.0 \leq R$, $RO \leq 25$ mm/hr	11	6.98	2.64	.000	2.27	.508	.139

The number of collocated SSM/I and radar observations in the sample (n), the mean radar-derived rainfall rate (R_m) and standard deviation (σ_R) of the sample, the bias (b) and error standard deviation (σ_e) of the rain rate estimate, the correlation coefficient (r) between the radar and SSM/I-derived rain rates, and the success ratio (S) for each case are listed above. R_m, σ_R, b, and σ_e are given in units of mm/hr. The first five rows are the statistics of the D-Matrix rain rate estimates (RO) for the indicated subsets of the full data sample. The statistics are stratified because the full data sample is dominated by low rainfall rates. In the second section the statistics of linear regression fits to the same subsets of points are listed. To maintain consistency with the D-Matrix estimates, a lower bound of 0 mm/hr was imposed upon the regression estimates.

TABLE 11.7

STATISTICS FOR THE MIDLATITUDE SPRING/FALL OCEAN CLIMATIC ZONE

CASE	n	R_m	σ_R	b	σ_e	r	S
D-Matrix:							
$0.0 \leq R, RO \leq 25$ mm/hr	1034	.961	2.77	-.551	2.91	.160	.333
$0.5 \leq R, RO \leq 25$ mm/hr	72	4.66	4.98	-1.23	5.98	-.229	1.16
$1.0 \leq R, RO \leq 25$ mm/hr	53	6.07	5.11	-2.99	6.44	-.146	1.20
$1.5 \leq R, RO \leq 25$ mm/hr	33	6.20	4.22	-2.35	5.17	-.036	.768
$2.0 \leq R, RO \leq 25$ mm/hr	30	6.64	4.18	-2.77	5.37	-.054	.799
All channel regressions:							
$0.0 \leq R, RO \leq 25$ mm/hr	1034	.961	2.77	.173	1.96	.711	.151
$0.5 \leq R, RO \leq 25$ mm/hr	72	4.66	4.98	.006	3.74	.660	.452
$1.0 \leq R, RO \leq 25$ mm/hr	53	6.07	5.11	.001	4.24	.560	.522
$1.5 \leq R, RO \leq 25$ mm/hr	33	6.20	4.22	.000	3.44	.581	.340
$2.0 \leq R, RO \leq 25$ mm/hr	30	6.64	4.18	-.001	3.49	.550	.337

The number of collocated SSM/I and radar observations in the sample (n), the mean radar-derived rainfall rate (R_m) and standard deviation (σ_R) of the sample, the bias (b) and error standard deviation (σ_e) of the rain rate estimate, the correlation coefficient (r) between the radar and SSM/I-derived rain rates, and the success ratio (S) for each case are listed above. R_m, σ_R, b, and σ_e are given in units of mm/hr. The first five rows are the statistics of the D-Matrix rain rate estimates (RO) for the indicated subsets of the full data sample. The statistics are stratified because the full data sample is dominated by low rainfall rates. In the second section the statistics of linear regression fits to the same subsets of points are listed. To maintain consistency with the D-Matrix estimates, a lower bound of 0 mm/hr was imposed upon the regression estimates.

TABLE 11.8

STATISTICS FOR THE MIDLATITUDE WINTER OCEAN CLIMATIC ZONE

CASE	n	R_m	σ_R	b	σ_e	r	S
D-Matrix:							
$0.0 \leq R$, RO ≤ 25 mm/hr	2236	.049	.147	.093	.920	.088	.034
$0.5 \leq R$, RO ≤ 25 mm/hr*	2	1.73	.261				
$1.0 \leq R$, RO ≤ 25 mm/hr*	2	1.73	.261				
$1.5 \leq R$, RO ≤ 25 mm/hr*	2	1.73	.261				
$2.0 \leq R$, RO ≤ 25 mm/hr*	0						
All channel regressions:							
$0.0 \leq R$, RO ≤ 25 mm/hr	2236	.049	.147	.005	.132	.441	.001
$0.5 \leq R$, RO ≤ 25 mm/hr*							
$1.0 \leq R$, RO ≤ 25 mm/hr*							
$1.5 \leq R$, RO ≤ 25 mm/hr*							
$2.0 \leq R$, RO ≤ 25 mm/hr*							

*Sample size insufficient for analysis to be performed.

The number of collocated SSM/I and radar observations in the sample (n), the mean radar-derived rainfall rate (R_m) and standard deviation (σ_R) of the sample, the bias (b) and error standard deviation (σ_e) of the rain rate estimate, the correlation coefficient (r) between the radar and SSM/I-derived rain rates, and the success ratio (S) for each case are listed above. R_m, σ_R, b, and σ_e are given in units of mm/hr. The first five rows are the statistics of the D-Matrix rain rate estimates (RO) for the indicated subsets of the full data sample. The statistics are stratified because the full data sample is dominated by low rainfall rates. In the second section the statistics of linear regression fits to the same subsets of points are listed. To maintain consistency with the D-Matrix estimates, a lower bound of 0 mm/hr was imposed upon the regression estimates.

Figure 11.3 - D-Matrix retrievals of rainfall rate versus radar-derived rainfall rate at midlatitudes for (a) summer over land, (b) spring-fall over land, (c) winter over land, (d) summer over ocean, (e) spring-fall over ocean, and (f) winter over ocean. Solid lines define the ±5 mm/hr retrieval error limits.
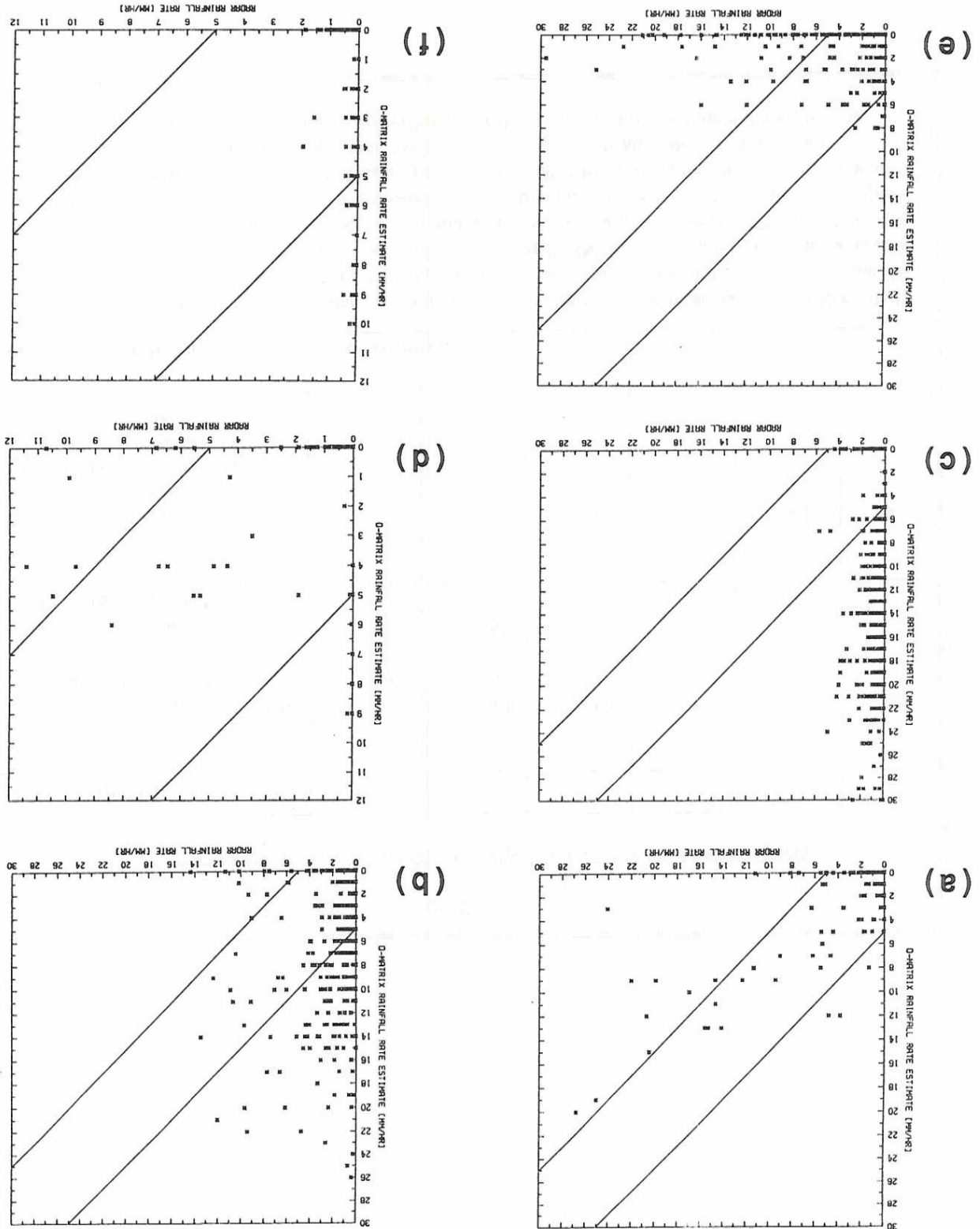
The D-Matrix algorithm for the midlatitude summer land climatic zone shows the best overall ability to estimate surface rainfall rates. Although the errors in the rain rate estimates are somewhat greater than the specified ±5 mm/hr at high rainfall rates, the correlation between the D-Matrix retrievals and the radar rain rates (maximum of .761) is relatively high and comparable to correlations obtained using other microwave sensors over land. For example, Spencer [7] found a .795 correlation between regressed Scanning Multichannel Microwave Radiometer (SMMR) brightness temperatures and radar-derived rainfall rates in summer rainfall over the midwest United States. The D-Matrix estimates are somewhat low-biased (∼ 1-2 mm/hr) with respect to radar; see Figure 11.3a. The linear regression estimates based upon the same samples of data yield slightly better estimates, with a maximum correlation of .784 over the entire sample of data.

The D-Matrix spring-fall land algorithm performs poorly in relation to the summer algorithm. In general the D-Matrix algorithm greatly overestimates light rain rates, which leads to positive biases of approximately 6 mm/hr and random errors of 8 mm/hr, and very low correlations to radar rain rates; see also Figure 11.3b. The success ratio (∼ 2 for rainfall rates ≥ .5 mm/hr) indicates that the D-Matrix retrievals fall outside the Navy specifications. Regression-based estimates of the data are superior, with lower mean errors (∼ 2 - 3 mm/hr) and modest but somewhat higher correlations with radar rain rates (maximum r = .602).

D-Matrix winter rain rate estimates are extremely high-biased, with very large mean errors (∼ 14 mm/hr) and almost no correlation to radar. The regression models are not much better (see Table 11.5 and Figure 11.3c).

A comparison of D-Matrix rain imagery and brightness temperature imagery indicated that lower land background brightness temperatures during the fall and winter seasons may have been interpreted as signatures of rainfall, leading to extreme positive biases in retrievals.

The midlatitude summer ocean D-Matrix algorithm shows less skill in estimating surface rainfall rate in comparison to the land algorithm for the same season, although the number of collocated D-matrix estimates and rainfall rates greater than 1 mm/hr (14) is admittedly small. The D-Matrix retrieval errors (∼ 4 mm/hr) are within the Navy specifications, but the correlation between retrieved and radar rain rates is low (maximum r = .446). The D-Matrix estimates are also low-biased on the order of 2 - 3 mm/hr for rainfall rates greater than .5 mm/hr. Regression estimates based upon the summer ocean data also yield low correlations with radar except in the range of very low rainfall rates (Table 11.6 and Figure 11.3d).

The spring-fall D-Matrix rain rate estimates over ocean are essentially uncorrelated with radar-derived rainfall rates (see Table 11.7 and Fig. 11.3e). Radar-derived rainfall rates are typically underestimated, with mean errors on the order of 5 - 6 mm/hr. D-Matrix success ratios exceed 1 for two of the subsamples, which indicate a performance outside of the prescribed ±5 mm/hr tolerance. Regression estimates based upon the same data yield a much greater correlation with radar rain rates, and errors are within specifications. The regression results

suggest that significant improvements can be made in the retrieval of rainfall rates over the ocean in the spring-fall season.

Although only a small number of collocated radar rain rates greater than 1 mm/hr were obtained during the winter season over the ocean, the plot in Figure 11.3f reveals a large positive bias in D-Matrix retrievals. Errors are within specifications only because the mean of the rain rate sample is extremely small ($\sim$ .05 mm/hr). Stratification of the winter ocean sample by a minimum threshold of 1 mm/hr eliminates all but two collocated measurements. Regression estimates based upon the entire sample of data yield low correlations with radar (Table 11.8).

<div style="border:1px solid black; padding:10px">

TABLE 11.9

STATISTICS FOR THE TROPICAL WARM SEASON LAND CLIMATIC ZONE

CASE	n	R_m	σ_R	b	σ_e	r	S
D-Matrix:							
$0.0 \leq R$, RL ≤ 25 mm/hr	120	.916	1.19	.301	1.88	.526	.139
$0.5 \leq R$, RL ≤ 25 mm/hr	37	1.74	1.46	2.07	3.06	.400	.359
$1.0 \leq R$, RL ≤ 25 mm/hr	23	2.28	1.63	2.54	3.67	.186	.504
$1.5 \leq R$, RL ≤ 25 mm/hr	14	2.94	1.82	3.20	4.52	-.302	.739
$2.0 \leq R$, RL ≤ 25 mm/hr	8	3.90	1.92	2.22	4.20	-.547	.601
All channel regressions:							
$0.0 \leq R$, RL ≤ 25 mm/hr	120	.916	1.19	.332	1.04	.576	.043
$0.5 \leq R$, RL ≤ 25 mm/hr	37	1.74	1.46	.214	1.13	.662	.049
$1.0 \leq R$, RL ≤ 25 mm/hr	23	2.28	1.63	.282	1.37	.577	.070
$1.5 \leq R$, RL ≤ 25 mm/hr	14	2.94	1.82	.245	1.43	.641	.074
$2.0 \leq R$, RL ≤ 25 mm/hr*							

*Sample size insufficient for analysis to be performed.

The number of collocated SSM/I and radar observations in the sample (n), the mean radar-derived rainfall rate (R_m) and standard deviation (σ_R) of the sample, the bias (b) and error standard deviation (σ_e) of the rain rate estimate, the correlation coefficient (r) between the radar and SSM/I-derived rain rates, and the success ratio (S) for each case are listed above. R_m, σ_R, b, and σ_e are given in units of mm/hr. The first five rows are the statistics of the D-Matrix rain rate estimates (RL) for the indicated subsets of the full data sample. The statistics are stratified because the full data sample is dominated by low rainfall rates. In the second section the statistics of linear regression fits to the same subsets of points are listed. To maintain consistency with the D-Matrix estimates, a lower bound of 0 mm/hr was imposed upon the regression estimates.

</div>

TABLE 11.10

STATISTICS FOR THE TROPICAL WARM SEASON OCEAN CLIMATIC ZONE

CASE	n	R_m	σ_R	b	σ_e	r	S
D-Matrix:							
$0.0 \leq R,\ RO \leq 25$ mm/hr	1361	.428	.892	.642	1.78	.630	.126
$0.5 \leq R,\ RO \leq 25$ mm/hr	241	1.73	1.42	3.17	3.56	.224	.486
$1.0 \leq R,\ RO \leq 25$ mm/hr	157	2.26	1.51	2.79	3.28	.185	.403
$1.5 \leq R,\ RO \leq 25$ mm/hr	100	2.83	1.64	2.47	3.14	.052	.359
$2.0 \leq R,\ RO \leq 25$ mm/hr	65	3.43	1.75	1.86	2.73	.079	.262
All channel regressions:							
$0.0 \leq R,\ RO \leq 25$ mm/hr	1361	.428	.892	.161	.618	.760	.015
$0.5 \leq R,\ RO \leq 25$ mm/hr	241	1.73	1.42	.956	1.69	.407	.110
$1.0 \leq R,\ RO \leq 25$ mm/hr	157	2.26	1.51	.111	1.40	.429	.073
$1.5 \leq R,\ RO \leq 25$ mm/hr	100	2.83	1.64	-.087	1.48	.441	.079
$2.0 \leq R,\ RO \leq 25$ mm/hr	65	3.43	1.75	.309	1.69	.357	.100

The number of collocated SSM/I and radar observations in the sample (n), the mean radar-derived rainfall rate (R_m) and standard deviation (σ_R) of the sample, the bias (b) and error standard deviation (σ_e) of the rain rate estimate, the correlation coefficient (r) between the radar and SSM/I-derived rain rates, and the success ratio (S) for each case are listed above. R_m, σ_R, b, and σ_e are given in units of mm/hr. The first five rows are the statistics of the D-Matrix rain rate estimates (RO) for the indicated subsets of the full data sample. The statistics are stratified because the full data sample is dominated by low rainfall rates. In the second section the statistics of linear regression fits to the same subsets of points are listed. To maintain consistency with the D-Matrix estimates, a lower bound of 0 mm/hr was imposed upon the regression estimates.

11.4.3 Tropical D-Matrix Algorithm Error Statistics

Statistics of the tropical algorithm rain rate estimates for land and ocean environments are presented in Tables 11.9 and 11.10, respectively. The statistical analyses are identical to those performed on the midlatitude data. Over either land or ocean, it is evident from the tables that although the D-Matrix algorithms may perform within specifications (S-factors all <1), the correlation of rain estimates with ground truth estimates is relatively low. Over either surface, correlations are all less than .2 for radar derived rainfall rates greater than 1 mm/hr; the bias and error standard deviation of the estimates are on the order of 2 to 3 mm/hr. These results are reflected in the D-Matrix retrieval plots in Figure 11-4a and 11-4b. Large positive biases in the D-Matrix estimates are noted at rainfall rates less than about 4 mm/hr, while there is a trend of negative biases at higher rainfall rates.

The linear regression estimates yield consistently higher correlations with the surface radar data in comparison to the D-Matrix estimates. Although the data sample is admittedly small over land in the tropics, the correlation coefficients of the regression estimates are close to .6, while the error standard deviations range from 1.0 to 1.5 mm/hr. Bias in the regression estimates is positive, but about an order of magnitude smaller than the bias in the D-Matrix estimates ($\sim$ .2 to .3 mm/hr). Over the ocean, the correlation coefficients of the regression estimates are somewhat smaller than those over land ($\sim$ .4 to .7), but again the error standard deviations and bias figures are significantly reduced in comparison to those of the D-Matrix estimates. Error standard deviations are on the order of 1.5 mm/hr, and the bias figures are all less than 1.0 mm/hr in absolute value. The regression estimates which were based upon the full land and ocean data samples are plotted versus the radar derived rainfall rates in Figures 11.4c and 11.4d, respectively.

Although lower rainfall rates tend to be overestimated and higher rainfall rates tend to be underestimated by the regression formulae, the overall bias and scatter in the regression estimates is significantly smaller than the bias and scatter of the D-Matrix estimates. Most of the regression estimates fall within approximately ± 2 mm/hr of the radar derived rainfall rates. These statistics and plots suggest that regression-based algorithms may be constructed which yield rain rate estimates which are superior to the D-Matrix estimates.

11.5 ALTERNATE ALGORITHMS

The plots of the D-Matrix rain rate estimates versus radar-derived rainfall rates in Figures 11.3 and 11.4 all show a common trend: at low radar-derived rainfall rates, the D-Matrix algorithm tends to overestimate rainfall rate, while at high rainfall rates the D-matrix algorithm tends to underestimate rain intensity. Linear regression models, in general, have this feature since scatter due to errors and nonlinearities in the relationship between variables is minimized with respect to the mean value of the independent data (i.e., in this case the mean rainfall rate).
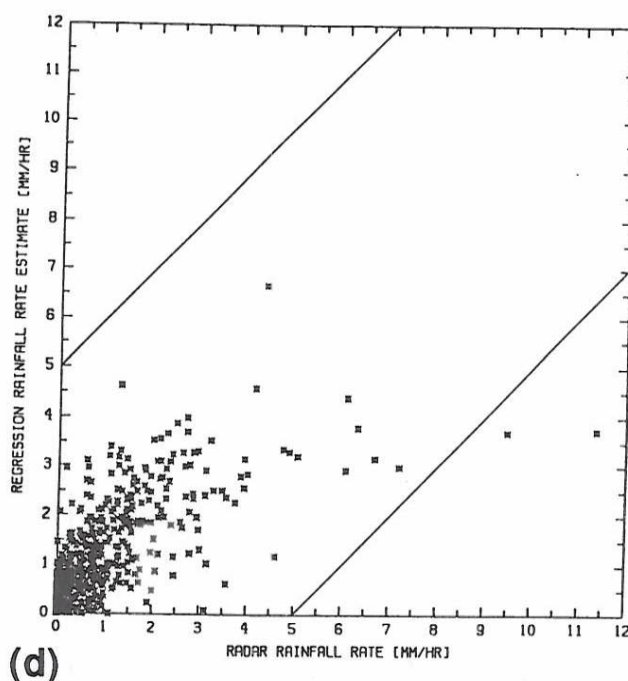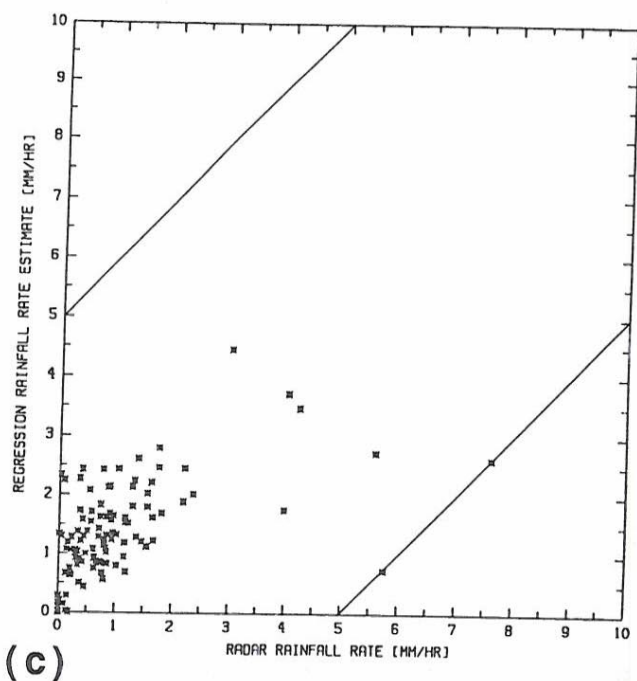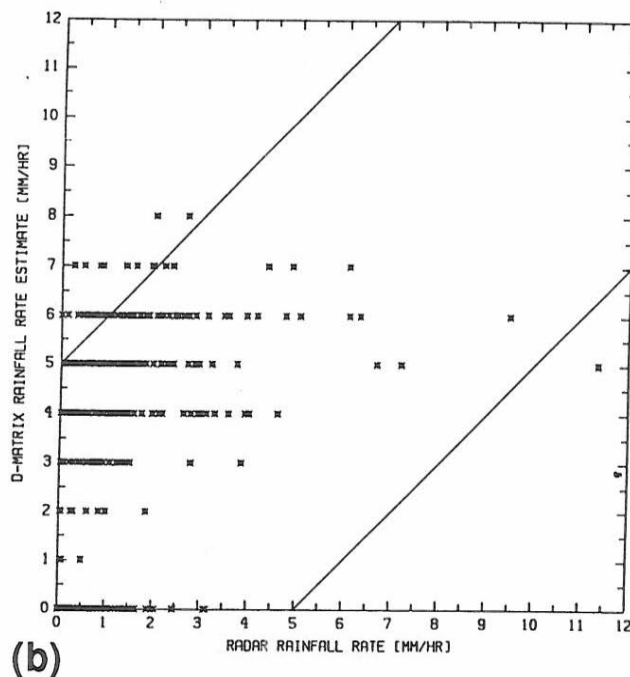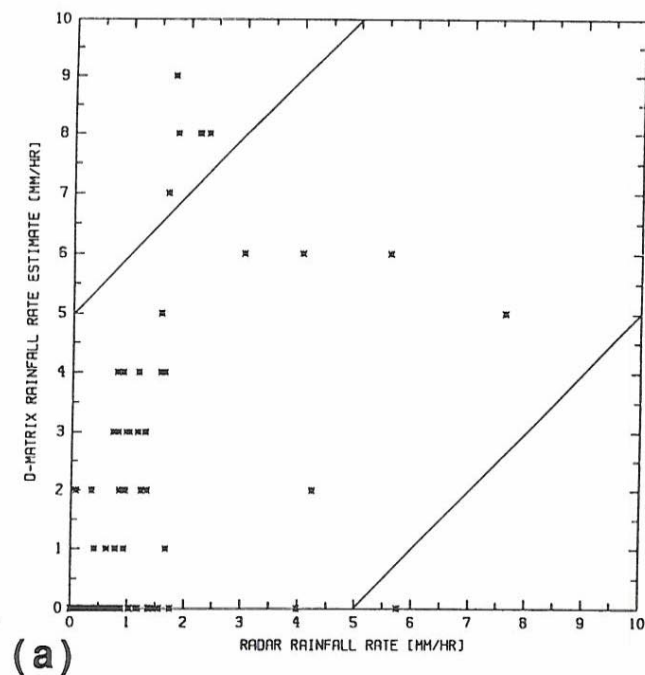
Figure 11.4 - D-Matrix retrievals of rainfall rate versus radar-derived rainfall rate for (a) the tropical warm season over land, and (b) the tropical warm season over ocean. Also shown are the linear regression estimates of rainfall rate based upon the collocated SSM/I brightness temperatures and radar derived rainfall rates between 0 and 25 mm/hr for (c) the tropical warm season over land, and (d) the tropical warm season over ocean. Solid lines define the ±5 mm/hr retrieval error limits.

11-23

(a)

(b)
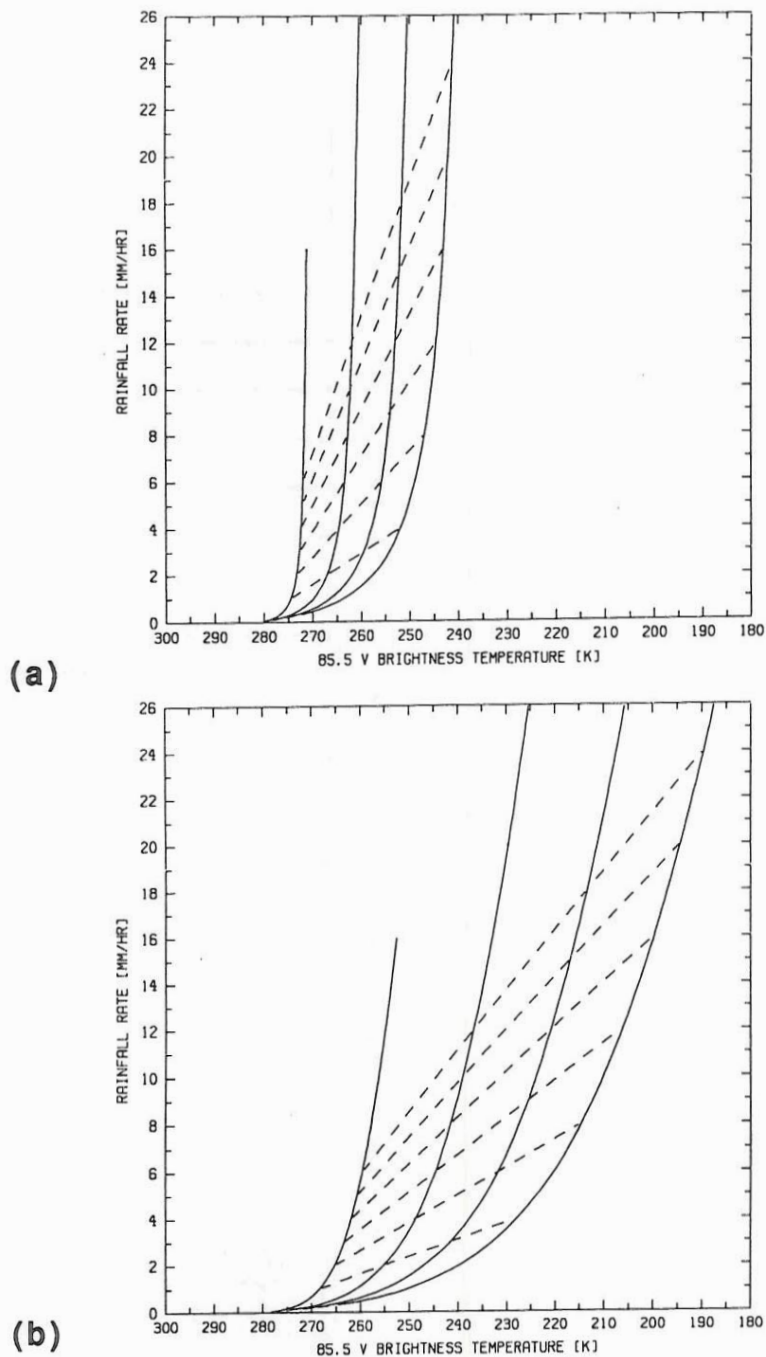
Figure 11.5 - Radiative transfer model computations of the 85.5 GHz vertical polarization brightness temperature upwelling from a cloud over land containing (a) only liquid hydrometeors, and (b) both liquid and ice hydrometeors. The cloud vertical structure in (a) is designed to simulate stratiform precipitation, whereas in (b) a convective cloud is modeled. The footprint-average rainfall rate is plotted as a function of the footprint-average upwelling brightness temperature. Solid lines are isolines of cloud fraction within the radiometer footprint, which run in the sequence .25, .50, .75, and 1.0 from left to right in the plots. Dashed lines are isolines of in-cloud rainfall rate which runs in the sequence 4, 8, 12, 16, 20, and 24 mm/hr from bottom to top in the plots. Model computations are provided by Kummerow [8].

11-24

A physical reason for the rainfall rate-dependent bias can be understood using model simulations of the brightness temperature upwelling from precipitating clouds. In Figure 11.5 are plotted model simulations of the 85.5 GHz vertically-polarized brightness temperature upwelling from precipitating clouds over land. Separate simulations are performed for clouds which fill different fractions of the radiometer footprint. The clouds in Figure 11.5a contain only liquid precipitation, in an attempt to simulate stratiform precipitation in which cloud updraft speeds are relatively low. In the figure, the footprint-average rainfall rate is plotted as a function of the footprint-average upwelling brightness temperature. Isolines of cloud fraction (solid) and in-cloud rainfall rate (dashed) are also indicated. The clouds in Figure 11.5b contain both liquid and ice precipitation-sized particles in a vertical distribution consistent with the structure of strongly convective clouds or thunderstorms. Figure 11.5 clearly indicates a nearly exponential relationship between rainfall rate and upwelling brightness temperature. Model simulations of upwelling brightness temperatures over land at the other SSM/I channel frequencies and polarizations show a similar nonlinear relationship. Over ocean backgrounds, model simulations indicate more complicated brightness temperature-rainfall rate relationships due to the generally lower ocean emissivity and the effects of raindrop emission (see Kummerow [9]).

These simulations suggest that linear models are in most cases inadequate to describe the relationship between brightness temperature and rain rate. An exception was shown by Spencer [10] to exist for convective precipitation over the ocean, where a linear combination of the brightness temperatures in the vertically and horizontally polarized 37 GHz channels of the SMMR was found to be linearly related to area-average rainfall rate. However, in tropical cyclones, where a mixture of convective and stratiform precipitation is present, Olson [11, 12] demonstrated that rain retrievals using the 37 GHz data alone tended to overestimate the intensity of lighter rainfall. Superior rain rate estimates were obtained when data from channels at the lower SMMR frequencies (e.g. 18 GHz) were incorporated into a physical retrieval method.

In addition to the nonlinear relationship between rainfall rate and brightness temperature, a comparison of the model curves in Figure 11.5 indicates that the type of precipitation (stratiform or convective) also has a bearing on the microwave signature of rainfall. The lack of ice in the stratiform cloud simulation causes the brightness temperature to become relatively insensitive to changes in the rainfall rate at rainfall rates greater than a few mm/hr. On the other hand, scattering from increasing numbers of ice hydrometeors in the convective cloud causes the brightness temperature to decrease with rain rate at higher rainfall rates. Clearly a mixture of the two precipitation regimes could lead to difficulties in making rain estimates, unless information from other channels is incorporated.

Two approaches are undertaken in an attempt to obtain improved regression models for rainfall rate. First, the residuals in the linear regression analyses are weighted to emphasize errors at the higher rainfall rates. Weighting by an increasing function of the rainfall rate helps to compensate for the skewed distribution of rainfall rates, which is dominated by low rain rates (see Figure 11.2). Although the errors in regression estimates of rainfall rate in the range of low rainfall rate tend to increase by this approach, they are more likely to remain within the specified

±5 mm/hr error limits because their magnitudes are initially small. Alternatively, improved rain rate estimates are obtained at higher rainfall rates due to the weighting.

The problems experienced with linear regression models can be partly overcome by utilizing nonlinear predictors/predictands which more closely match the physical relationship between brightness temperature and rainfall rate. The simplest nonlinear algorithm to implement operationally is

$$R = \exp \left(a_o + \sum_{i=1}^{7} a_i T_{B_i} \right) - c, \tag{5}$$

where the rainfall rate is expressed as an exponential function of the seven SSM/I-measured brightness temperatures T_{Bi}, with fitted constants a_i and c. The coefficients a_i are determined by regressing $ln(R+c)$ against the SSM/I brightness temperatures for different values of the constant c. Experimentation with the sets of collocated SSM/I and radar data indicate that values of c between 1 and 16 may be adequate for most climatic zones and seasons. The value of formula (5) is that the nearly exponential dependence of rainfall rate on brightness temperature is established. The fitted coefficients a_i allow for the variations in curvature of the brightness temperature to rainfall rate relationship which may be induced by varying cloud ice contents or fractional footprint coverage.

Regressions of both R and $ln(R+c)$ against SSM/I brightness temperature data were performed in an attempt to find a general retrieval formula for rainfall rate. Because the calibration of the radars at Darwin and Kwajalein was checked frequently as part of the Tropical Rainfall Measuring Mission (TRMM) program (see Simpson, et al. [13]), only data from these radars were utilized in the rain retrieval algorithm development. Collocated SSM/I and radar data were separated into land and ocean samples using a bitmap, and all data within 69 km (approximately one 19.35 GHz footprint width) of coastlines were filtered. In addition, flooded soil regions over land, as determined by the McFarland and Neale (personal communications) brightness temperature discriminant function, were filtered from the analysis. Residuals in the regressions were weighted by a factor of $R^{0.5}$ to compensate for the naturally skewed rainfall distribution.

Statistics of the rainfall rate regression estimates over land and ocean are presented in Tables 11.11 and 11.12, respectively. It may be noted from Table 11.11 that over land, either the linear or exponential model estimates are substantially better than the D-Matrix estimates. The mean error of the linear regression estimates for rainfall rates greater than or equal to .5 mm/hr is 1.14 mm/hr, which is significantly less than the D-Matrix error standard deviation (3.06 mm/hr). It would appear that the exponential models do not perform quite as well as the linear regression models, based upon the statistics in Table 11.11. Correlations to the radar rain rates are slightly lower, and error standard deviations are roughly the same. However, the overall bias of the exponential model estimates is somewhat lower, and an application of the exponential regression formulae to diverse rain systems over the tropics and midlatitudes indicated generally superior performance with respect to linear models. The exponential models

TABLE 11.11

REGRESSION MODEL STATISTICS BASED UPON THE TROPICAL WARM SEASON DATA OVER LAND

CASE	n	R_m	σ_R	b	σ_e	r	S
D-Matrix:							
$0.0 \leq R$, $RL \leq 25$ mm/hr	120	.916	1.19	.301	1.88	.526	.139
$0.5 \leq R$, $RL \leq 25$ mm/hr	37	1.74	1.46	2.07	3.06	.400	.359
Linear Regression:							
$0.0 \leq R$, $RL \leq 25$ mm/hr	120	.916	1.19	.373	1.04	.594	.043
$0.5 \leq R$, $RL \leq 25$ mm/hr	37	1.74	1.46	.185	1.14	.674	.050
Log regression (c=16.0):							
$0.0 \leq R$, $RL \leq 25$ mm/hr	120	.916	1.19	.339	1.03	.575	.042
$0.5 \leq R$, $RL \leq 25$ mm/hr	37	1.74	1.46	.063	1.16	.630	.052
Log regression (c=8.0):							
$0.0 \leq R$, $RL \leq 25$ mm/hr	120	.916	1.19	.304	1.02	.578	.041
$0.5 \leq R$, $RL \leq 25$ mm/hr	37	1.74	1.46	.016	1.16	.630	.052
Log regression (c=4.0):							
$0.0 \leq R$, $RL \leq 25$ mm/hr	120	.916	1.19	.246	.996	.581	.039
$0.5 \leq R$, $RL \leq 25$ mm/hr	37	1.74	1.46	-.058	1.16	.629	.052
Log regression (c=2.0):							
$0.0 \leq R$, $RL \leq 25$ mm/hr	120	.916	1.19	.205	.983	.586	.038
$0.5 \leq R$, $RL \leq 25$ mm/hr	37	1.74	1.46	-.112	1.16	.627	.052

The residuals in all regressions were weighted by the square root of the rainfall rate. The number of collocated SSM/I and radar observations in the sample (n), the mean radar-derived rainfall rate (R_m) and standard deviation (σ_R) of the sample, the bias (b) and error standard deviation (σ_e) of the rain rate estimate, the correlation coefficient (r) between the radar and SSM/I-derived rain rates, and the success ratio (S) for each case are listed above. R_m, σ_R, b, and σ_e are given in units of mm/hr. The models are categorized as linear models, which include the D-Matrix algorithm, and exponential models, in which $\ell n(R+c)$ is regressed against the SSM/I brightness temperatures. Statistics are stratified by an imposed minimum on the D-Matrix (RL) and radar-derived (R) rainfall rates. A sample minimum of 0.5 mm/hr emphasizes the errors at higher rainfall rates. To maintain consistency with the D-Matrix estimates, a lower bound of 0 mm/hr was imposed upon the regression estimates.

TABLE 11.12

REGRESSION MODEL STATISTICS BASED UPON THE TROPICAL WARM SEASON DATA OVER OCEAN

CASE	n	R_m	σ_R	b	σ_e	r	S
D-Matrix:							
$0.0 \leq R, RO \leq 25$ mm/hr	1361	.428	.892	.642	1.78	.630	.126
$0.5 \leq R, RO \leq 25$ mm/hr	241	1.73	1.42	3.17	3.56	.224	.015
Linear Regression:							
$0.0 \leq R, RO \leq 25$ mm/hr	1361	.428	.892	.161	.618	.761	.015
$0.5 \leq R, RO \leq 25$ mm/hr	241	1.73	1.42	.258	1.21	.558	.056
Log regression (c=16.0):							
$0.0 \leq R, RO \leq 25$ mm/hr	1361	.428	.892	.155	.602	.768	.014
$0.5 \leq R, RO \leq 25$ mm/hr	241	1.73	1.42	.214	1.20	.559	.055
Log regression (c=8.0):							
$0.0 \leq R, RO \leq 25$ mm/hr	1361	.428	.892	.126	.586	.772	.014
$0.5 \leq R, RO \leq 25$ mm/hr	241	1.73	1.42	.149	1.19	.557	.054
Log regression (c=4.0):							
$0.0 \leq R, RO \leq 25$ mm/hr	1361	.428	.892	.123	.580	.775	.013
$0.5 \leq R, RO \leq 25$ mm/hr	241	1.73	1.42	.124	1.19	.551	.054
Log regression (c=2.0):							
$0.0 \leq R, RO \leq 25$ mm/hr	1361	.428	.892	.106	.572	.778	.013
$0.5 \leq R, RO \leq 25$ mm/hr	241	1.73	1.42	.073	1.20	.541	.055

The residuals in all regressions were weighted by the square root of the rainfall rate. The number of collocated SSM/I and radar observations in the sample (n), the mean radar-derived rainfall rate (R_m) and standard deviation (σ_R) of the sample, the bias (b) and error standard deviation (σ_e) of the rain rate estimate, the correlation coefficient (r) between the radar and SSM/I-derived rain rates, and the success ratio (S) for each case are listed above. R_m, σ_R, b, and σ_e are given in units of mm/hr. The models are categorized as linear models, which include the D-Matrix algorithm, and exponential models, in which $\ell n(R+c)$ is regressed against the SSM/I brightness temperatures. Statistics are stratified by an imposed minimum on the D-Matrix (RO) and radar-derived (R) rainfall rates. A sample minimum of 0.5 mm/hr emphasizes the errors at higher rainfall rates. To maintain consistency with the D-Matrix estimates, a lower bound of 0 mm/hr was imposed upon the regression estimates.

TABLE 11.13

REGRESSION MODEL STATISTICS BASED UPON THE TROPICAL WARM SEASON DATA OVER LAND, WITHOUT THE 85.5 GHz DATA

CASE	n	R_m	σ_R	b	σ_e	r	S
D-Matrix:							
$0.0 \leq R$, $RL \leq 25$ mm/hr	120	.916	1.19	.301	1.88	.526	.139
$0.5 \leq R$, $RL \leq 25$ mm/hr	37	1.74	1.46	2.07	3.06	.400	.359
Linear Regression							
w/o 85 GHz data:							
$0.0 \leq R$, $RL \leq 25$ mm/hr	120	.916	1.19	.373	1.04	.594	.043
$0.5 \leq R$, $RL \leq 25$ mm/hr	37	1.74	1.46	.185	1.14	.674	.050
Log regression (c=1.0)							
w/o 85 GHz data:							
$0.0 \leq R$, $RL \leq 25$ mm/hr	120	.916	1.19	.149	.922	.643	.034
$0.5 \leq R$, $RL \leq 25$ mm/hr	37	1.74	1.46	-.062	1.13	.687	.049

The residuals in all regressions were weighted by the square root of the rainfall rate. The number of collocated SSM/I and radar observations in the sample (n), the mean radar-derived rainfall rate (R_m) and standard deviation (σ_R) of the sample, the bias (b) and error standard deviation (σ_e) of the rain rate estimate, the correlation coefficient (r) between the radar and SSM/I-derived rain rates, and the success ratio (S) for each case are listed above. R_m, σ_R, b, and σ_e are given in units of mm/hr. The models are categorized as linear models, which include the D-Matrix algorithm, and exponential models, in which $\ell n(R+c)$ is regressed against the SSM/I brightness temperatures. Statistics are stratified by an imposed minimum on the D-Matrix (RL) and radar-derived (R) rainfall rates. A sample minimum of 0.5 mm/hr emphasizes the errors at higher rainfall rates. To maintain consistency with the D-Matrix estimates, a lower bound of 0 mm/hr was imposed upon the regression estimates.

TABLE 11.14

REGRESSION MODEL STATISTICS BASED UPON THE TROPICAL WARM SEASON DATA OVER OCEAN, WITHOUT THE 85.5 GHz DATA

CASE	n	R_m	σ_R	b	σ_e	r	S
D-Matrix:							
$0.0 \leq R,\ RO \leq 25$ mm/hr	1361	.428	.892	.642	1.78	.630	.126
$0.5 \leq R,\ RO \leq 25$ mm/hr	241	1.73	1.42	3.17	3.56	.224	.486
Linear Regression w/o 85 GHz data:							
$0.0 \leq R,\ RO \leq 25$ mm/hr	1361	.428	.892	.206	.658	.735	.017
$0.5 \leq R,\ RO \leq 25$ mm/hr	241	1.73	1.42	.278	1.21	.558	.056
Log regression (c=2.0) w/o 85 GHz data:							
$0.0 \leq R,\ RO \leq 25$ mm/hr	1361	.428	.892	.134	.593	.762	.014
$0.5 \leq R,\ RO \leq 25$ mm/hr	241	1.73	1.42	.057	1.17	.569	.053

The residuals in all regressions were weighted by the square root of the rainfall rate. The number of collocated SSM/I and radar observations in the sample (n), the mean radar-derived rainfall rate (R_m) and standard deviation (σ_R) of the sample, the bias (b) and error standard deviation (σ_e) of the rain rate estimate, the correlation coefficient (r) between the radar and SSM/I-derived rain rates, and the success ratio (S) for each case are listed above. R_m, σ_R, b, and σ_e are given in units of mm/hr. The models are categorized as linear models, which include the D-Matrix algorithm, and exponential models, in which $\ell n(R+c)$ is regressed against the SSM/I brightness temperatures. Statistics are stratified by an imposed minimum on the D-Matrix (RO) and radar-derived (R) rainfall rates. A sample minimum of 0.5 mm/hr emphasizes the errors at higher rainfall rates. To maintain consistency with the D-Matrix estimates, a lower bound of 0 mm/hr was imposed upon the regression estimates.

worked better because the 85.5 GHz SSM/I brightness temperature data, which provide greater signal at lower rainfall rates, were selected in the stepwise regression procedure. The 85.5 GHz data were not selected in the linear regression over land.

Plots of the D-Matrix estimates and the exponential model estimates (c=8.0 mm/hr) versus radar derived rainfall rate are presented in Figures 11.6a and 11.6c, respectively. Although rainfall rates greater than 5 mm/hr tend to be underestimated by the exponential model, the majority of rain estimates are within ±2 mm/hr of the radar rainfall rates. The exponential model estimates compare favorably with the D-Matrix estimates, which are generally high biased.

Linear regression estimates of rain rates over the ocean also show an improvement over the D-Matrix estimates (Table 11.12). Errors with respect to radar rainfall rates are reduced significantly (3.56 mm/hr to 1.21 mm/hr for rainfall rates $\geq$ .5 mm/hr), and correlations increase dramatically. A maximum correlation of .761 is achieved over the entire data sample.
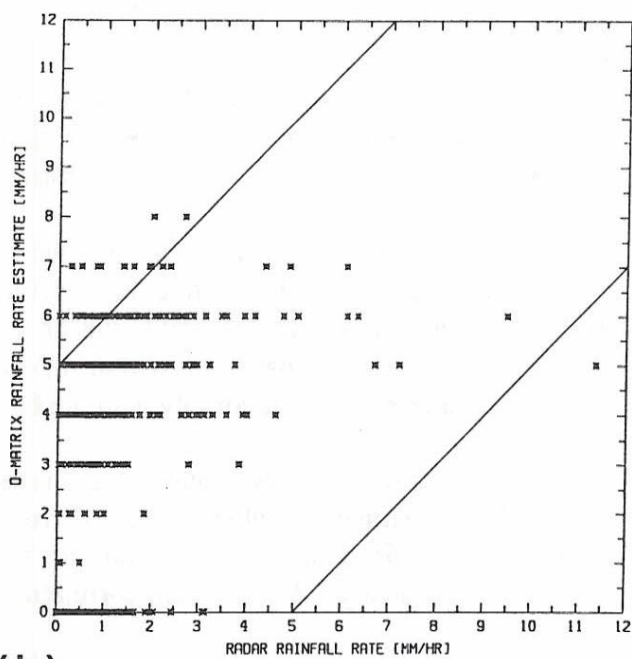
Rainfall rate estimates obtained from the exponential models are slightly more accurate than the linear regression estimates, and the bias in the exponential model estimates is generally reduced. A minimum error standard deviation of 1.19 mm/hr is achieved by the exponential model with c=8.0 mm/hr for radar rainfall rates greater than or equal to .5 mm/hr. Plots of the D-Matrix and exponential model (c=8.0 mm/hr) rain rate estimates versus the radar derived rain rates over ocean are presented in Figures 11.6b and 11.6d. As noted earlier in the regression analyses over land (Figure 11.6c), the ocean regression estimates tend to be low at rainfall rates greater than 5 mm/hr, but the majority of estimates fall within ±2 mm/hr of the radar rain rates. In contrast, the D-Matrix rain rate estimates are generally high biased and show much greater deviation from the radar rain rates.

Due to the recent degradation of the SSM/I 85.5 GHz channels on the DMSP-F8, the regression analyses were repeated with the 85.5 GHz brightness temperature data in both polarizations removed. Selected statistics from these analyses for land and ocean backgrounds are presented in Tables 11.13 and 11.14, respectively. One may recall that over land, the 85.5 GHz data were not selected in the linear regression analysis by the stepwise procedure; therefore the statistics of the linear models in Tables 11.11 and 11.13 are identical.

It is curious to note that the most accurate exponential model (c=1.0 mm/hr) appears to outperform all other models when the 85.5 GHz data are removed. This result is an artifact of the stepwise procedure. Since the 85.5 GHz data are generally most highly correlated with rainfall rate, these data were selected first in the all-channel regressions; the partial correlations of data from the remaining channels did not warrant substitution of the 85.5 GHz data with data from the lower-frequency channels. However, with the 85.5 GHz brightness temperature data removed, a different combination of channels was selected which yielded regression estimates with a somewhat higher correlation to the radar rain rates. The relatively small sample of validation data over land may have contributed to some ambiguity in the selection of an optimal regression model.
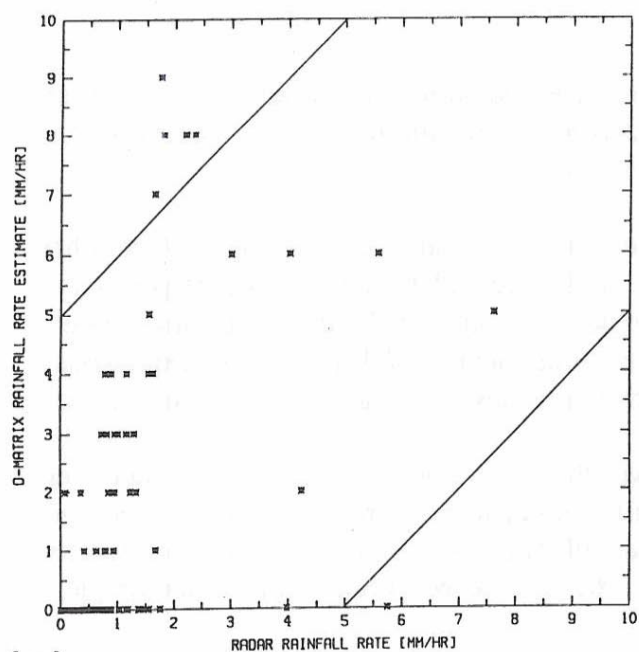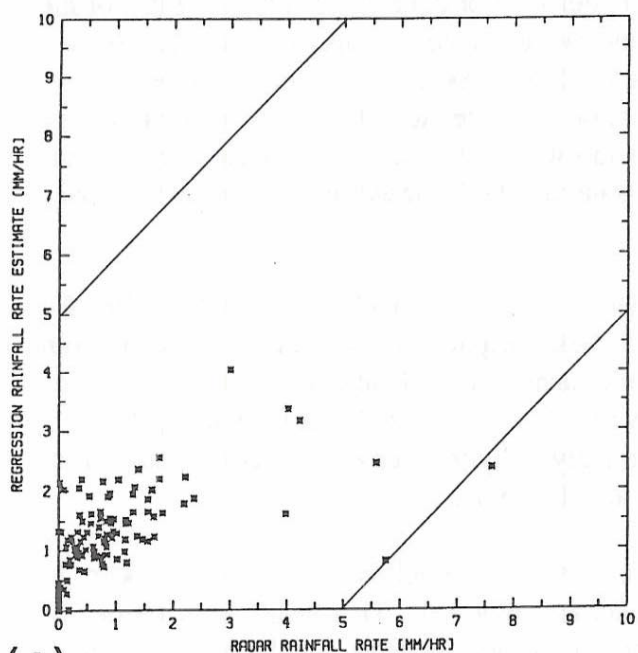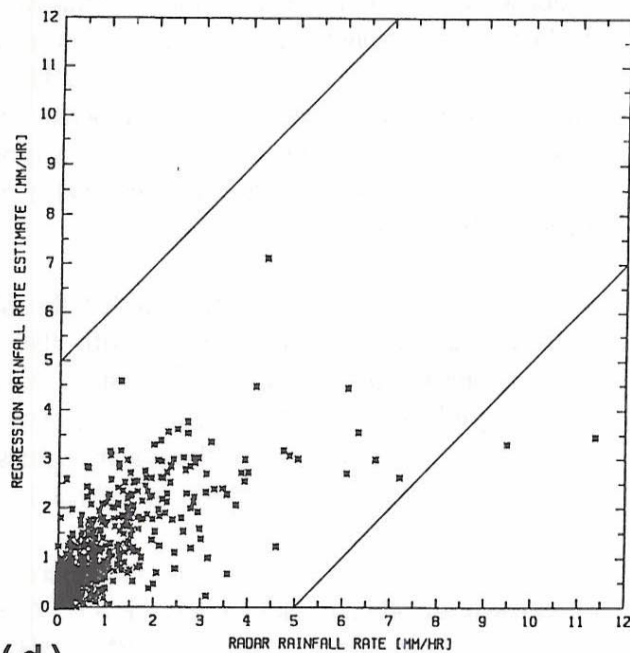
Figure 11.6 - D-Matrix and regression estimates of rainfall rate versus radar-derived rainfall rate from the tropics. D-Matrix rain rate estimates over land and ocean are plotted in (a) and (b), respectively. Logarithmic regression estimates over land and ocean (c=8.0 mm/hr) are plotted in panels (c) and (d), respectively. The regression formulae were based upon collocated SSM/I brightness temperatures and radar-derived rainfall rates obtained from the Darwin and Kwajalein validation sites. Solid lines define the ±5 mm/hr retrieval error limits.

Application of the regression models to SSM/I observations of diverse rain systems in the tropics and midlatitudes revealed that the formulae which did not incorporate the 85.5 GHz data tended to underestimate light rains over land. This is because the signal from light rainfall is relatively small in the lower-frequency SSM/I data, and this smaller signal is obscured by variations in surface emission.

Regressions over the ocean which did not include the 85.5 GHz SSM/I brightness temperatures yielded rain rate estimates which were about as accurate as those which included the 85.5 GHz data (see Tables 11.12 and 11.14). Only small differences in retrieved rain distributions were noted upon application of both formulae to SSM/I observations of several storms. Of the regression models which did not include the 85.5 GHz data, the exponential model with c=2.0 mm/hr produced optimal rain rate estimates over the ocean.

11.6 RECOMMENDATIONS

The regression analyses performed in the last section provided simple formulae which may be utilized to improve the retrieval of rainfall rates over land and ocean within the framework of the SSM/I operational retrieval software. It should be noted that although the regression formulae determined in Section 11.5 were based upon tropical radar data, application of these formulae to midlatitude rain systems yielded rain rate estimates which were climatologically realistic and consistent with available radar.

The statistics and independent application of the regression formulae suggest that if the 85.5 GHz SSM/I data are available, then the exponential models with c=8.0 mm/hr provide the best estimates of rainfall rate over land and ocean. Similar testing revealed that the exponential models with c=1.0 mm/hr over land and c=2.0 mm/hr over ocean yielded optimal results if the 85.5 GHz data were not available. These formulae would be applied if the screening logic described below is satisfied.

The screening logic utilizes the Hughes' negative polarization test for bad data. After passing this test, if the all-channel SSM/I brightness temperature scene is over land, then the McFarland and Neale screening logic is applied. If the brightness temperature scene is over the ocean, then a discriminant function developed by the authors is applied to eliminate false rain signatures near coasts. Coastal pixels are not processed.

SUMMARY:

The following is the recommended rain retrieval algorithm, including screening logic to test for the presence of rain.

SCREENING LOGIC:

If the 85 GHz channels are available then

> If $T_{B85V} - T_{B85H} <$ -2 K or
> $T_{B37V} - T_{B37H} <$ -2 K or
> $T_{B19V} - T_{B19H} <$ -2 K, then flag as indeterminate

Else if SSM/I measurement is over land, then

> If $T_{B22V} - T_{B19V} \leq$ 4 K and
> $(T_{B19V} + T_{B37V})/2 - (T_{B19H} + T_{B37H})/2 \leq$ 4 K and
> $T_{B85V} - T_{B37V} <$ -1 K and
> $T_{B19V} >$ 268 K

> or

> If $T_{B22V} - T_{B19V} \leq$ 4 K and
> $(T_{B19V} + T_{B37V})/2 - (T_{B19H} + T_{B37H})/2 >$ 4 K and
> $T_{B37V} - T_{B19V} <$ -3 K and
> $T_{B85V} - T_{B37V} <$ -5 K and
> $T_{B85H} - T_{B37H} <$ -4.1 K and
> $T_{B19V} >$ 268 K, then compute rain rate over land,

Else rain rate = 0 mm/hr.

Else if SSM/I measurement is over the ocean, then

> If -11.7939 - .02727 T_{B37V} + .09920 $T_{B37H} >$ 0 K, then
> compute rain rate over ocean,

Else rain rate = 0 mm/hr.

Else SSM/I measurement is coastal; flag as indeterminate.

Else the 85 GHz channels are not available then

> If $T_{B37V} - T_{B37H} <$ -2 K or
> $T_{B19V} - T_{B19H} <$ -2 K, then flag as indeterminate

Else if SSM/I measurement is over land, then

> If $T_{B22V} - T_{B19V} \leq$ 4 K and
> $(T_{B19V} + T_{B37V})/2 - (T_{B19H} + T_{B37H})/2 \leq$ 4 K and

$$T_{B37V} - T_{B19V} < -6.4 \text{ K and}$$
$$T_{B19V} > 268 \text{ K}$$

or

If $T_{B22V} - T_{B19V} \leq 4$ K and
$(T_{B19V} + T_{B37V})/2 - (T_{B19H} + T_{B37H})/2 > 4$ K and
$T_{B37V} - T_{B19V} < -6.4$ K and
$T_{B19V} > 268$ K, then compute rain rate over land,

Else rain rate $= 0$ mm/hr.

Else if SSM/I measurement is over the ocean, then

If $-11.7939 - .02727 \ T_{B37V} + .09920 \ T_{B37H} > 0$ K, then
compute rain rate over ocean,

Else rain rate $= 0$ mm/hr.

Else SSM/I measurement is coastal; flag as indeterminate.

RECOMMENDED ALGORITHMS:

If a rainfall rate over land is to be computed, then use

$$R = \exp(3.29716 - .01290 \ T_{B85V} + .00877 \ T_{B85H}) - 8.0 \text{ mm/hr.}$$

If a rainfall rate over the ocean is to be computed, then use

$$R = \exp(3.06231 - .0056036 \ T_{B85V} + .0029478 \ T_{B85H} - .0018119 \ T_{B37V}$$
$$- .00750 \ T_{B22V} + .0097550 \ T_{B19V}) - 8.0 \text{ mm/hr.}$$

Alternatively, if the 85.5 GHz channel data are unusable, then over land apply

$$R = \exp(-17.76849 - .09612 \ T_{B37V} + .15678 \ T_{B19V}) - 1.0 \text{ mm/hr,}$$

and over the ocean use

$$R = \exp(5.10196 - .05378 \ T_{B37V} + .02766 \ T_{B37H} + .01373 \ T_{B19V}) - 2.0 \text{ mm/hr.}$$

If any of these formulae yield a rainfall rate less than zero, then set the rain rate equal to 0 mm/hr.

11.7 APPLICATION OF THE ALTERNATE ALGORITHM TO TROPICAL CYCLONE DATA

The suggested retrieval formulae presented in Section 11.6 are applied to SSM/I data from an overpass of Hurricane Florence at 00:21 GMT on September 10, 1988.

Hurricane Florence originated in a stagnant frontal zone over the south central Gulf of Mexico and began to move northward and strengthen on September 9th. A middle-tropospheric trough to the west interacted with the vortex to stimulate strong convection over the center and an area of midlevel subsidence and drying to the west.

Just prior to the SSM/I overpass (00:01 GMT) Florence reached its peak intensity, with a minimum pressure of 982 mb and maximum winds of 35 m/s. The low level center was located just off of the Mississippi delta. The 6.7 micron water vapor imagery from GOES (not shown), indicated an influx of dry air into the circulation from the southwest. As a result, the convection indicated by the imagery of Figure 11.7 was weak and poorly organized, and had been decaying even before landfall.

Florence's disrupted convection and steady forward motion at 6 m/s kept rainfall totals relatively small. Twenty-four hour amounts along the track ranged from 35 to 105 mm, and similar amounts fell in a secondary convective area over the Florida panhandle, well to the east of the center. The secondary circulation also spawned 9 tornadoes, and the rains, although not extraordinary for a tropical cyclone, added to the already swollen rivers to produce the worst floods in ten years on the two Florida panhandle rivers. Damage in Louisiana was confined to beach erosion and wind damage to trees and power lines.

The 85.5 GHz horizontal and 19.35 GHz vertical channel SSM/I data, which are utilized in the alternate retrieval algorithms, are presented in Figure 11.7a and b, respectively. Warm colors indicate areas of high microwave brightness temperature, whereas cooler colors correspond to areas of lower brightness temperature. Signatures of precipitation are identified as depressions in the 85.5 GHz horizontally-polarized brightness temperatures (Figure 11.7a).

Over land the signal from precipitation is much smaller in the 19.35 GHz vertically-polarized channel. This is partly due to the fact that microwave scattering by raindrops is much weaker at 19.35 GHz, while the absorption/re-emission signature of rain does not contrast greatly with emission by the land background. The relatively low spatial resolution of the 19.35 GHz channels also contributes to reduced rain response. However, since the ocean emits at a relatively low brightness compared to emission by rain at 19.35 GHz, the 19.35 vertical channel provides rain information for oceanic rain retrievals. The small band of precipitation about 500 km southeast of New Orleans is identified as a region of increased microwave brightness in relation to the low emissivity ocean background in Figure 11.7b.

The alternate algorithm retrieval of rainfall rates in Florence is presented in Figure 11.7c. Retrievals within about 25 km of the coast were filtered because the radiometer measurements
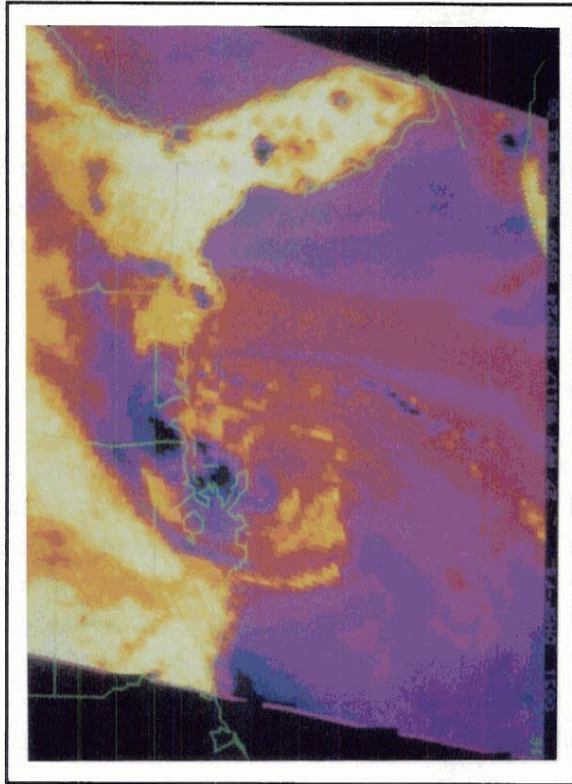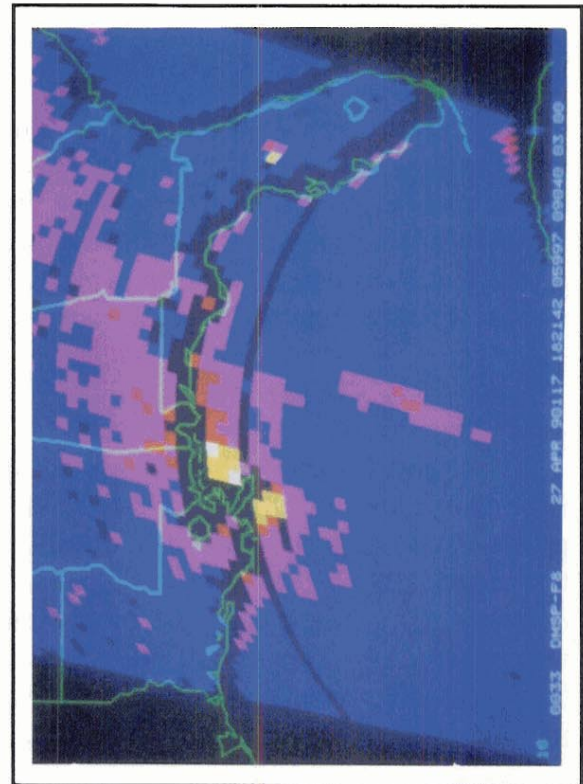
Figure 11.7 See text for explanation.

(b)

(d)

(a)

(c)

11-37
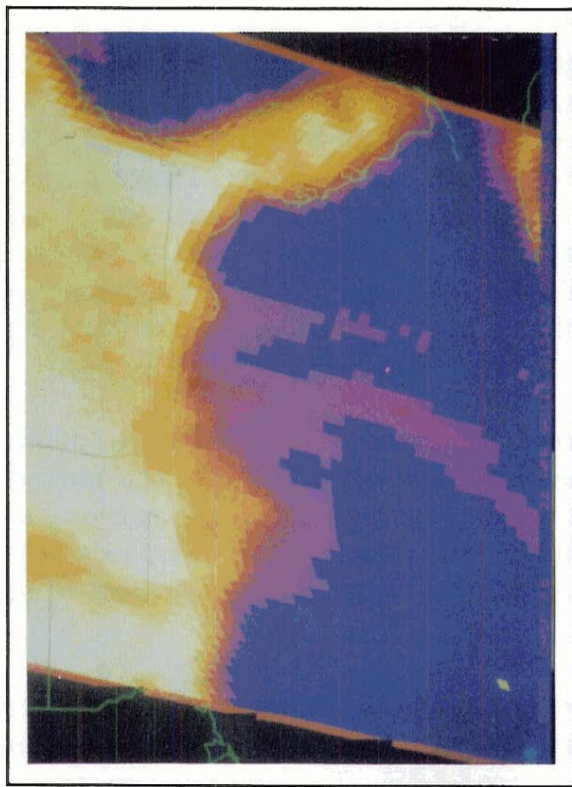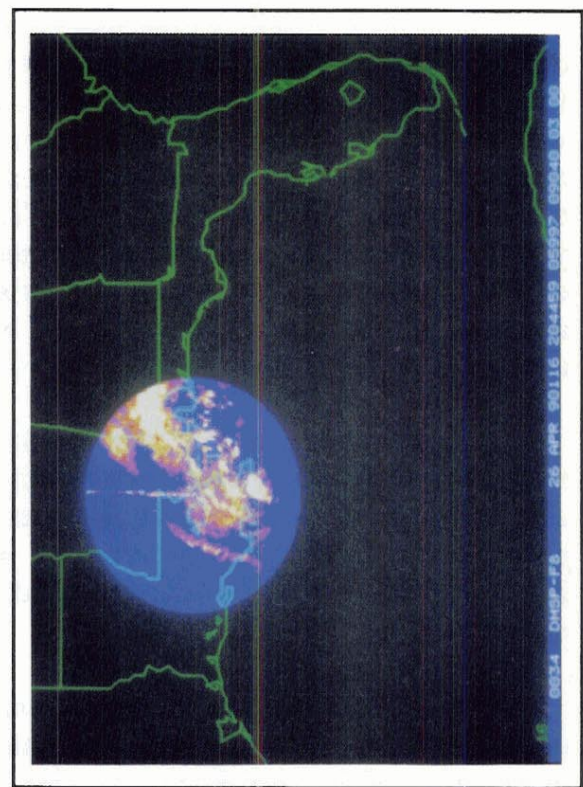
in the immediate vicinity of the coast contained significant contributions from both land and ocean backgrounds. The rain retrievals may be compared to the radar-derived rain rates obtained from the NWS WSR-57 station at Slidell, Louisiana in Figure 11.7d. The radar rainfall rates were corrected for range-dependent biases using a method suggested by Black [14]. Both the retrieved and radar-derived rain rate images utilize the same color enhancement. Purple indicates the 1 mm/hr rain rate threshold level. The color sequence from purple to red, orange, and yellow correspond to 1 mm/hr steps in the rain rate threshold. Regions where the rain rate exceeds 5 mm/hr are colored white. It should also be noted that the SSM/I estimates are averaged rain rates over 625 km^2 areas, whereas the radar-derived values are roughly 4 km^2 averages.

Figure 11-7c and d indicate a good spatial correlation between SSM/I retrieved rainfall rates and radar-derived rain rates within the observing range of the radar. The retrieval algorithm appears to overestimate rain rates just east of the Mississippi delta, while rain rates are underestimated in southwestern Alabama. Overall the SSM/I rain rate estimates are reasonable in comparison with the radar, if one takes into account the spatial averaging effect of the radiometer.

The SSM/I rain rate estimates from the current alternate algorithm show a much better correspondence to the radar-derived rain rates than the previous "midlatitude" algorithm described in Volume 1 of the Final Report (see Figure 1.20c on page 1-38). The improved performance of the current algorithm is attributed to the superior calibration of the Darwin and Kwajalein radars, upon which the algorithm is based.

11.8   CONCLUSION

Evaluations of the D-Matrix retrieval algorithm indicate that specified accuracies for derived rainfall rates are not being met over land at midlatitudes. Improvements in the algorithm based upon empirical relationships to the "ground truth" data set increase the accuracy of retrieved rainfall rates to within the requirement for both land and ocean situations. Application of the improved algorithm to tropical cyclone data yields rainfall rate estimates which are in reasonable agreement with coastal radar data.

11.9   ACKNOWLEDGMENTS

Division, who provided the radar data for the Hurricane Florence case study. Also Chris Velden of the University of Wisconsin lent his expertise in the analysis of the Florence data sets.

11.10 REFERENCES

[1] J. W. Wilson and E. A. Brandes, "Radar measurement of rainfall - a summary," <u>Bull. Amer. Meteor. Soc.</u>, vol. 60, pp. 1048-1058, 1979.

[2] C. G. Collier, P. R. Larke, and B. R. May, "A weather radar correction procedure for real-time estimation of surface rainfall," <u>Quart. J. R. Met. Soc.</u>, vol. 109, pp. 589-608, 1983.

[3] L. J. Battan, <u>Radar Observations of the Atmosphere</u>, Chicago: The University of Chicago Press, 324 pp., 1973.

[4] E. M. Sodano, "General non-iterative solution of the inverse and direct geodetic problems," <u>Bulletin Geodesique</u>, vol. 75, pp. 69-89, 1965.

[5] J. S. Marshall and W. K. M. Palmer, "The distribution of raindrops with size," <u>J. Meteor.</u>, vol. 5, pp. 165-166, 1948.

[6] T. W. Harrold, E. J. English, and C. A. Nicholass, "The accuracy of radar derived rainfall measurements in hilly terrain," <u>Quart. J. R. Met. Soc.</u>, vol. 100, pp. 331-350, 1974.

[7] R. W. Spencer, "Satellite passive microwave rain rate measurement over croplands during spring, summer, and fall," <u>J. Climate and Appl. Meteor.</u>, vol. 23, pp. 1553-1562, 1984.

[8] C. D. Kummerow, private communication, NASA/Goddard Space Flight Center, Greenbelt, MD, 1988.

[9] C. D. Kummerow, "Microwave Radiances from Horizontally Finite, Vertically Structured Clouds," Ph. D. Thesis, University of Minnesota, Minneapolis, MN, 146 pp., 1987.

[10] R. W. Spencer, "A satellite passive 37-GHz scattering-based method for measuring oceanic rain rates," <u>J. Climate and Appl. Meteor.</u>, vol. 22, pp. 2095-2099, 1986.

[11] W. S. Olson, "Estimation of Rainfall Rates in Tropical Cyclones by Passive Microwave Radiometry," Ph.D. Thesis, University of Wisconsin, Madison, WI, 282 pp., 1987.

[12]   W. S. Olson,"Physical retrieval of rainfall rates over the ocean by multispectral microwave radiometry-application to tropical cyclones," J. Geophys. Res., vol. 94, pp. 2267-2280, 1989.

[13]   J. Simpson, R. F. Adler, and G. R. North, "A proposed Tropical Rainfall Measuring Mission (TRMM) satellite," Bull. Amer. Met. Soc., vol. 69, pp. 278-295, 1988.

[14]   M. Black, private communication, NOAA/AOML/Hurricane Research Division, Miami, FL, 1988.

SECTION 12

CLOUD AMOUNT VALIDATION

by

Gerald W. Felde
Geophysics Laboratory (AFSC)
Hanscom AFB, MA

12.0   CLOUD AMOUNT VALIDATION

12.1   CLOUD AMOUNT ALGORITHMS

The Hughes Aircraft Company developed two algorithms for estimating cloud amounts (percent cloud coverage) from SSM/I brightness temperatures.   One is applicable over land backgrounds, the other over snow.   Hughes has not been tasked to develop a cloud amount estimation algorithm for ocean backgrounds.   In the initial formulation of the cloud amount algorithms, it was recognized that polarization characteristics at 85 GHz (i.e, vertical brightness temperature minus horizontal brightness temperature) should provide much of the information. Simulated values of 85 GHz polarization for a variety of land and snow background conditions for clear and cloudy cases were calculated.   For a given background condition, a smaller 85 GHz polarization value was associated with a cloudy atmosphere than with a clear atmosphere [1].   It was desired to retrieve information on the cloud coverage from the SSM/I as near as possible to the resolution of the approximately 45 km x 45 km area used by the Air Force Global Weather Central's (AFGWC) Real-Time Nephanalysis (RTNEPH) automated global cloud analysis.   So it was decided to base each individual estimate of cloud amount on a 3 x 3 array of adjacent 85 GHz samples with an all-channel scene at its center.   Figure 12.1 shows this array of 85 GHz footprints.   The array is framed by a 39 km (along scan) x 41 km (across scan) rectangle.   One 37 GHz footprint is also inside this rectangle.   Each 85 GHz footprint is 14 km (along scan) x 16 km (across scan) and the 37 GHz footprint is 29 km (along scan) x 36 km (across scan).   Further analysis also indicated a cloud signature in the 37 GHz brightness temperatures for land and snow-covered backgrounds.

In the final developmental phase of the cloud amount algorithms, for both land and snow backgrounds, simulated 37 and 85 GHz (vertical and horizontal polarizations) brightness temperature values for clear and overcast conditions were calculated by Hughes Aircraft Company [2] using the Air Force Geophysics Laboratory's RADTRAN atmospheric transmission model [3].   For snow backgrounds, its depth was varied between 4 and 20 cm in increments of 2 cm.   Surface emissivity for each snow depth value was calculated using the dry snow model of Ulaby and Stiles [4].   For land backgrounds, soil moisture was varied using values of 3, 5, 12, and 20 percent.   Surface emissivity for each value of soil moisture was calculated using Fresnel equations modified by the Choudhury et al. [5] correction factor of 0.6 to take into account surface roughness effects.

Interpolated values of 37 and 85 GHz simulated brightness temperatures were combined for each of the two surface backgrounds, using a random number generator to create clear fields of view (all nine 85 GHz footprints clear - 0 percent cloud cover), one 85 GHz footprint overcast (any one of the nine - 11.1 percent cloud cover), etc. through all nine 85 GHz footprints overcast (100 percent cloud cover).   Regression coefficients were then calculated from the simulation results [2].   For snow backgrounds, a four-step regression produced a percent cloud amount estimation equation that accounts for 95.9 percent of the modeled variance.   An error analysis of this estimation equation determined an rms error of the estimated percent cloud amount of 3.2 percent.   While for land backgrounds, a four-step regression produced a percent
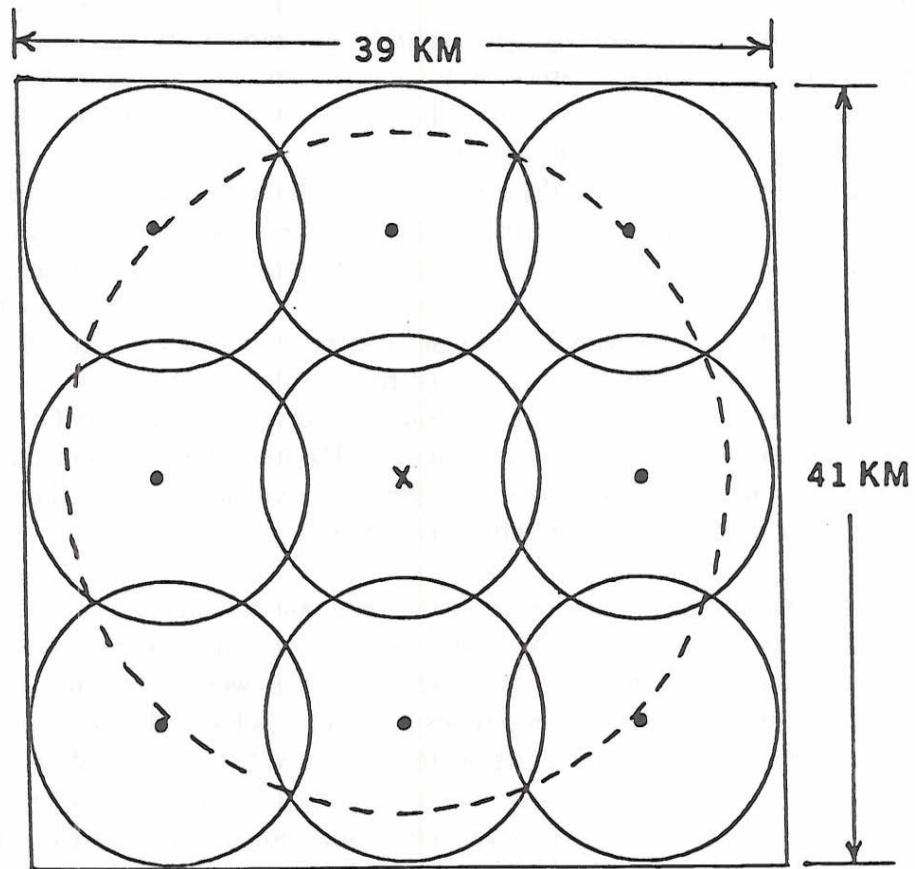
Figure 12.1   Area for Which a Single SSM/I Percent
Cloud Amount Value is Calculated

cloud amount estimation equation that accounts for 77.9 percent of the modeled variance. An error analysis of this estimation equation determined an rms error of the estimated cloud amount of 7.8 percent.

The final operational version of the two cloud amount algorithms developed by Hughes Aircraft Company [6, 7] are:

$$CAS = c_0 + (c_1 x T_{37V}) + (c_2 x T_{37H}) + (c_3 x \Sigma T_{85V}) + (c_4 x \Sigma T_{85H}) \qquad (12.1)$$

where CAS is the percent cloud amount over snow; T_{37V} and T_{37H} are the 37 GHz brightness temperatures - vertical and horizontal polarizations respectively; ΣT_{85V} and ΣT_{85H} are the sum of the nine 85 GHz brightness temperatures - vertical and horizontal polarizations respectively, at an all-channel scene and its eight surrounding 85 GHz scenes. The coefficients are c_0 = -189.5000, c_1 = -0.9710, c_2 = 0.7400, c_3 = -0.1987, c_4 = 0.3678.

$$CAL = c_0 + (c_1 x T_{37H}) + (c_2 x \Sigma T_{85V}) + (c_3 x \Sigma T_{85H}) \qquad (12.2)$$

where CAL is the percent cloud amount over land. The coefficients for the land equation are: c_0 = -638.9000, c_1 = -1.7050, c_2 = -0.2868, c_3 = 0.7457.

Note that the vertically polarized 37 GHz brightness temperature is not used in the cloud amount over land equation. This was the final brightness temperature in the four-step regression analysis, and it was found that its inclusion only provided an infinitesimal improvement to the estimation accuracy [2].

It is possible for either of the two cloud amount equations to produce results that are physically meaningless. To account for this, an "out-of-limits" flag is included in both algorithms. Out-of-limits is arbitrarily assigned to cloud amount estimates less than -20% and greater than 120%.

No cloud amount estimates are made for flooded or vegetative backgrounds. A dynamic determination of one of nine possible land types is made for each SSM/I data point tagged as having a land background [1]. Simulations indicated that the SSM/I would be unable to detect clouds over vegetated land because the high water content provides the same type of signature as a cloud. Flooded land is treated the same as an oceanic background. Since no SSM/I cloud amount algorithm for water backgrounds was required, these scenes are ignored.

12.2   VALIDATION METHODOLOGY

Manual cloud cover estimates were used to validate the automated SSM/I algorithm results. The manual analyses were performed on 3 nmi resolution visible (0.5 to 1.0 micrometers) and infrared (IR, 10 to 13 micrometers) imagery data obtained from the Operational Linescan System (OLS) sensor which is on board the same spacecraft as the SSM/I. The resolution of the OLS is considerably better than that of the SSM/I. Therefore, the ability

of the OLS to resolve clouds within the SSM/I footprints is excellent. Since the swath width of the OLS is twice that of the SSM/I and the sensors are on the same satellite, all SSM/I cloud amounts will have spatially coincident OLS cloud amounts. However, there will be a small temporal difference because the sensors have different scan geometries. The OLS scans in a straight line perpendicular to the satellite subtrack, while the SSM/I scans aft of the satellite with a constant angle of 45 degrees between satellite nadir and the antenna beam. For a given SSM/I scan, the OLS scan line that contains the center point of the SSM/I scan will be obtained 137 seconds prior to the SSM/I scan, while the OLS scan line that contains the endpoints of the SSM/I scan will be obtained 87 seconds prior to the SSM/I scan.

The Air Force Interactive Meteorological System (AIMS) at the Geophysics Laboratory (GL) was used as the test bed for this validation study. AIMS is a distributed system of mini- and micro-computers that was developed to support research in remote sensing at GL. Functional capabilities include the ability to receive, manage, store, display and interact with meteorological observations, radar and satellite data. Two identical image processing work stations are available on the system. (See reference [8] for a complete description of AIMS.) To obtain cloud truth data sets, a formalized procedure has been developed that involves interactive display and manipulation of the imagery on an AIMS image processing work station [9]. To assist in image interpretation, interactive image processing techniques are used to provide geometric and radiometric enhancements to the data and to provide for multispectral display. For example, an interactive piecewise linear stretch algorithm produces a different contrast enhancement over a number of selected brightness ranges in a monochrome (single channel) image by modifying the response of the display over each range. An input device on AIMS such as a mouse or graphics tablet is used to select interactively each brightness range and control the enhancement slope.

When performing a manual cloud analysis on OLS visible and infrared data, a number of display options are available. The most useful is a multiple image display generated by dividing the monitor into quadrants. Each quadrant can contain a separate monochrome or multispectral OLS image, each with a different enhancement. The OLS images have not been remapped; they are displayed in their original scan format to make use of the full resolution of the data. The analyst selects an area-of-interest on one target image to make a cloud boundary determination. This can often be a very small sub-region of the image. An iterative threshold blanking technique requires the analyst to select an intensity level that separates the clear and cloud regions in the area-of-interest. Regions below the threshold level are displayed as a color shade while the area above is displayed as a gray shade. This makes the boundary distinct while maintaining the detail below and above the threshold. The analyst then interactively raises or lowers the threshold until the proper level is obtained. The procedure is repeated until the entire target image has been classified. Two products are generated from this procedure, the first is a grayshade image that retains the original image characteristics above the cloud threshold and is black below, and the second is a binary image that simply delineates the cloud boundary from clear background. The first is used during the interactive threshold blanking process for visual comparison against reference images. The second is used for comparison with SSM/I algorithm results. Software was written to determine automatically the points in the OLS binary cloud

truth digital imagery data base corresponding to each area for which a single SSM/I cloud amount value is generated. This software also calculates the corresponding OLS cloud amounts by summing up the number of cloudy OLS pixels and dividing by the total number of OLS pixels within each SSM/I cloud amount area.

12.3 CASE STUDY DESCRIPTIONS

Four case study scenes were selected for the cloud amount algorithm validation study. The scenes contain several different cloud conditions and surface background types. For each case, the OLS data were first earth located and a binary synthetic image containing cloud truth information was generated using the techniques described in the previous section; then the SSM/I and OLS cloud amounts were compared.

Case 1 - Southern Africa: SSM/I and OLS data were collected for the late afternoon DMSP pass (satellite is descending) on 14 January 1988 over the southern part of Africa. Figure 12.2 depicts the area of coverage. This area includes desert, wet lands, cultivated regions, and forests. The OLS visible (Figure 12.3) and infrared (Figure 12.4) images show substantial areas of cumulus clouds. The OLS imagery data were manually analyzed using the interactive techniques described in the previous section to obtain a synthetic, binary image of the cloud cover (Figure 12.5). This was compared to the SSM/I algorithm results (Figure 12.6). These results will be discussed in detail in Section 12.4.

Case 2 - Central United States: This scene used the data from the morning DMSP pass on 14 January 1988 ascending over the central U.S. from coastal Gulf of Mexico up through Minnesota and the Dakotas into southern Canada. The northern quarter of the image was snow covered. The predominant cloud types were stratus and stratocumulus. The OLS visible imagery data were not usable in the manual cloud truth analysis because of the low light level in this scene during the early morning satellite crossing time.

Case 3 - Eastern United States: The data for this case is from the morning DMSP pass on 14 March 1988 ascending over the eastern third of the U.S. from Florida across the southeastern states up over the Great Lakes into southern Canada. The predominant clouds are stratus and stratocumulus which are associated with an upper level storm centered over southern Lake Huron. The northern part of the scene is snow covered and is approximately 10% of the total area of the scene. The OLS visible imagery data were of limited use in the manual cloud truth analysis because of light levels being too low for a sharp image.

Case 4 - North West South America: The data for this case were obtained from the same satellite pass as Case 3 but during an earlier time frame when the ascending satellite was still south of the equator. This scene contains northern Peru, Ecuador, Columbia, and Central America. The dominant cloud feature is a massive MCC (Mesoscale Convective Complex) over most of Ecuador. The main types of land surface backgrounds are rain forests and mountains. For this case, the light levels were high enough for the OLS visible data to be useful in the manual cloud truth analysis.
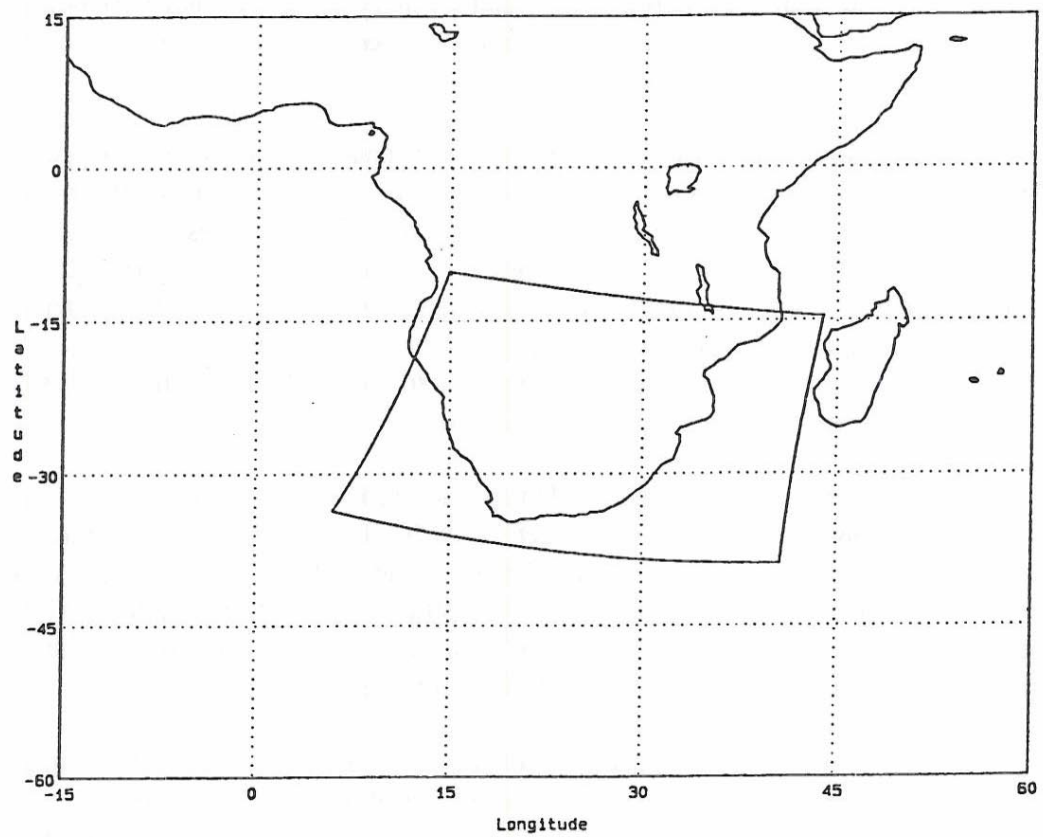
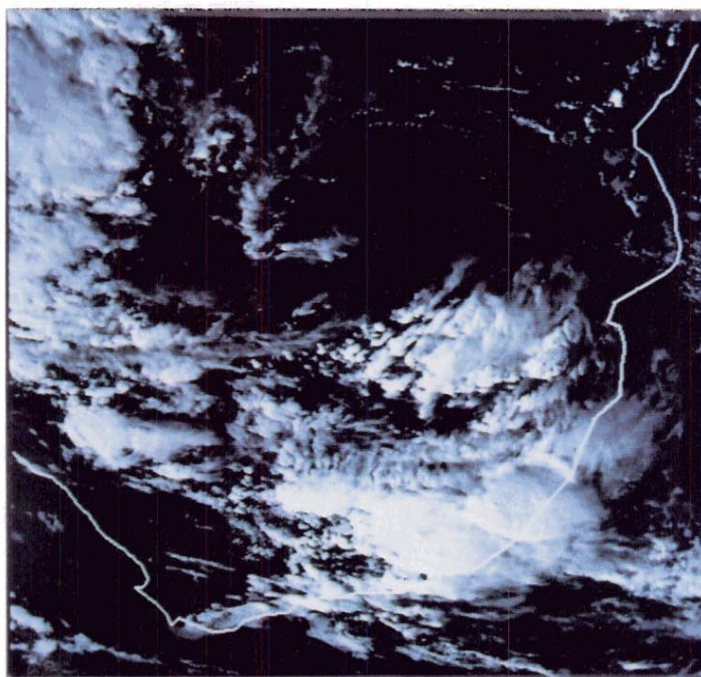Figure 12.2   OLS Coverage for Case 1
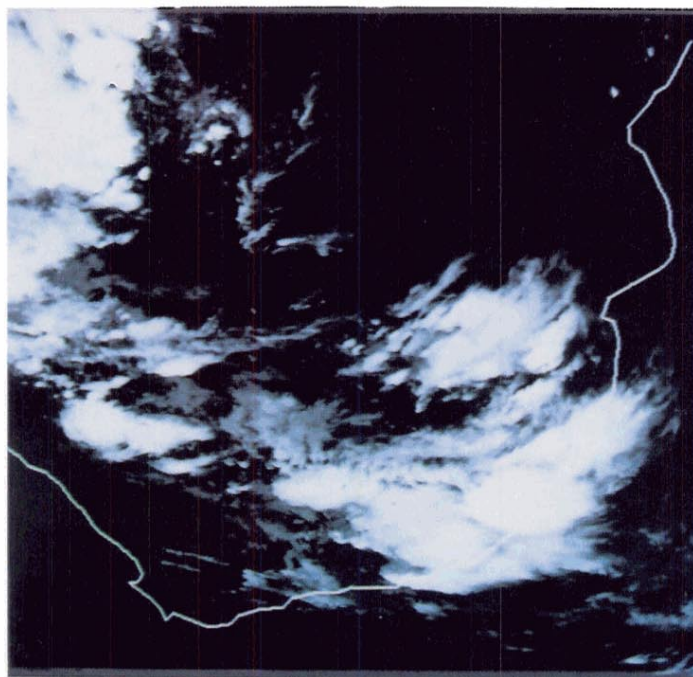
Figure 12.3   OLS Visible Image for Case 1

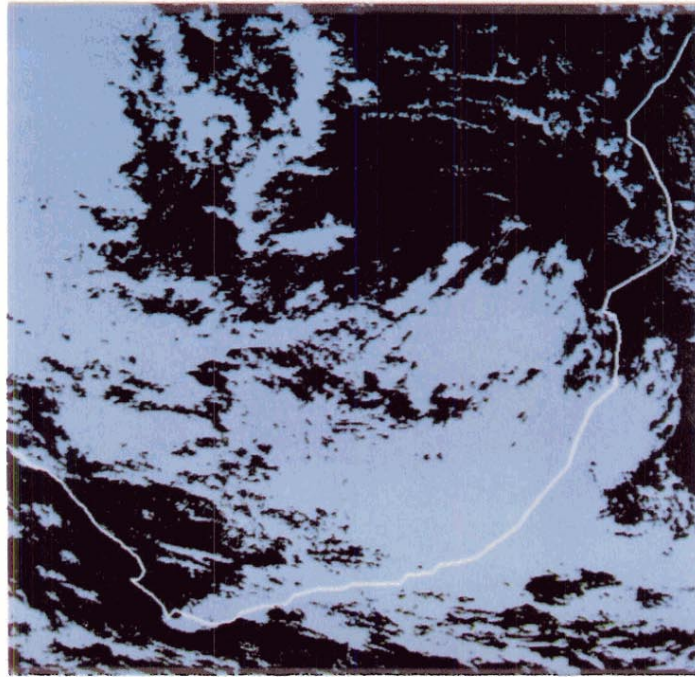Figure 12.4   OLS Infrared Image for Case 1

Figure 12.5   Binary Synthetic Cloud Cover Image Derived
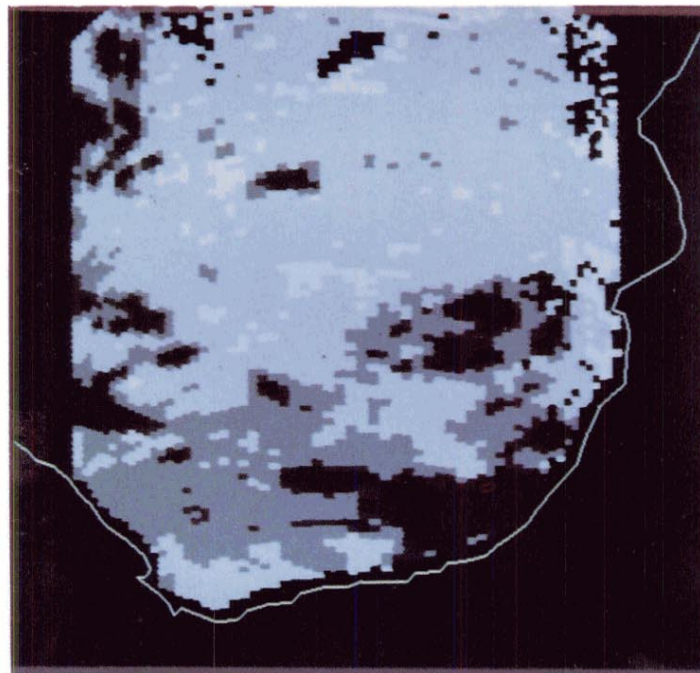from OLS Visible and IR Imagery for Case 1

Figure 12.6   SSM/I Cloud Amount Image for Case 1
Black: Out-of-Limits
Gray Shades: Increasing Cloud Amount from
Dark to Light Gray Shades

12.4 CASE STUDY RESULTS

The cloud amount values calculated by the SSM/I algorithms for the four case study scenes described in Section 12.3 were statistically compared to the cloud truth values obtained from the manual interactive computer analysis of OLS data. The results are presented in Table 12.1 and are stratified for land and snow backgrounds for each case. This is done to assess the performance of each of the two separate SSM/I cloud amount algorithms; one for land backgrounds and one for snow backgrounds. Recall SSM/I cloud amounts are not calculated when the land background is vegetated or flooded. Also, there is no Hughes SSM/I cloud amount algorithm for oceanic backgrounds.

Table 12.1 contains the mean, standard deviation about the mean, and the range (minimum and maximum) for both the SSM/I and the corresponding OLS derived cloud amounts. N in this table is the number of SSM/I cloud amounts in the allowed range of -20 to 120 percent. Values outside this range are tagged as "out-of-limits" and are listed in the last column of the table. The root-mean-square differences (rms) and the linear correlation coefficient (r) between the individual OLS and SSM/I cloud amount values are also given. The

			Mean CA	σ	Min CA	Max CA	N	RMS	r	Out-of-Limits
Case	Background									
1	Land	OLS	49.3	39.4	0	100	4343	49.3	-.44	414
		SSM/I	43.6	16.5	-20	74				
2	Land	OLS	52.0	40.1	0	100	226	72.7	.27	2947
		SSM/I	-9.7	10.9	-20	55				
2	Snow	OLS	58.1	44.2	0	100	1145	58.5	-.18	1
		SSM/I	53.0	30.6	-16	120				
3	Land	OLS	38.9	43.0	0	100	1053	65.5	-.15	2038
		SSM/I	-7.4	12.0	-20	64				
3	Snow	OLS	92.3	17.7	0	100	243	18.3	.26	0
		SSM/I	94.1	11.0	40	114				
4	Land	OLS	71.3	39.0	0	100	1084	61.4	-.17	138
		SSM/I	34.2	23.7	-20	67				

TABLE 12.1

STATISTICAL COMPARISON OF SSM/I AND OLS TRUTH CLOUD AMOUNTS

rms value is a measure of the amount of error between the individual SSM/I and corresponding "truth" OLS cloud amount values. An rms value of zero would mean that there is no error (i.e., all corresponding SSM/I and OLS cloud amounts are equal). The correlation coefficient is a measure of the linear relationship between a set of SSM/I and OLS cloud amount distributions.

There are no snow covered land backgrounds for Case 1 (Southern Africa). The mean OLS and SSM/I cloud amounts are close, but the OLS standard deviation is twice as large as that of the SSM/I and the maximum SSM/I value is considerably smaller than that of the OLS. The large rms and negative correlation coefficient indicate that the SSM/I algorithm has serious deficiencies.

There are both snow-free and snow-covered land backgrounds for Case 2 (Central United States). For land backgrounds, the SSM/I mean and standard deviation are much smaller than the OLS values. The maximum SSM/I value is about half that of the OLS. The rms for land backgrounds is quite large and the correlation coefficient is close to zero, indicating the two results are uncorrelated. For snow backgrounds, the mean and standard deviations are comparable. However, the rms value is large and r is close to zero which shows again there is very little relationship between the OLS and SSM/I cloud amounts.

Case 3 (Eastern United States) like Case 2 has both snow-free and snow-covered land backgrounds. The SSM/I mean and standard deviation are much smaller than those for the OLS over land backgrounds and the maximum SSM/I value is about two-thirds that of the OLS. Also for land backgrounds, the rms is large and the correlation coefficient is close to zero. The SSM/I and OLS mean and standard deviation are comparable for snow backgrounds, but the minimum SSM/I cloud amount is considerably larger than that of the OLS. Also, the rms value is large and the r value is close to zero.

There are no snow covered land backgrounds for Case 4 (North West South America). The SSM/I mean and standard deviation values are much smaller than those for the OLS and the maximum SSM/I value is two-thirds that of the OLS. The rms value is large and the r value is close to zero.

The values of the root-mean-square difference between the OLS and SSM/I cloud amounts for both land and snow backgrounds for all four cases are large. This indicates that the cloud amount estimates calculated by both SSM/I algorithms are poor. All the values of the linear correlation coefficients indicate that no significant linear relationship exists between the SSM/I and OLS cloud amounts. For land backgrounds, the mean and maximum OLS and SSM/I cloud amounts show the SSM/I values are consistently lower than the OLS values. The number of cases flagged as out-of-limits over land backgrounds for Cases 2 and 3 are very large. The significance of this is addressed in Section 12.6.

12.5   OTHER RESULTS

The frequency distribution of SSM/I-derived cloud amount values were examined for several orbits.   All the distributions were found to have similar characteristics.   Table 12.2 shows the distribution for revolution 655 which occurred on 5 August 1987.   The results in the table are stratified into land and snow backgrounds and shown are the total number of SSM/I cloud amount values calculated and the percentage of the total number that are within various categories.   For both backgrounds, there are few cloud amount values greater than 40% and many cloud values tagged as "out-of-limits".

The 37 and 85 GHz brightness temperatures for all "out-of-limits" cases for revolution 655 were put into the SSM/I cloud amount equations to determine the specific numerical values generated by the algorithms.   Table 12.3 shows the total number of land background "out-of-limits" cases with SSM/I cloud amount values within various categories.   It also shows this information for snow backgrounds.   For both backgrounds, all the values are negative.

<table>
<tr><td colspan="6">TABLE 12.2

SSM/I CLOUD AMOUNT VALUES (REV. 655 - 5 AUG 1987)</td></tr>
<tr><td>Surface Type:</td><td>Number of cases</td><td>0 - 30%</td><td>40 - 70%</td><td>80 - 100%</td><td>Out-of-Limits</td></tr>
<tr><td>Land</td><td>29403</td><td>53</td><td>14</td><td>0</td><td>33</td></tr>
<tr><td>Snow</td><td>3151</td><td>24</td><td>3</td><td>2</td><td>72</td></tr>
</table>

<table>
<tr><td colspan="6">TABLE 12.3

OUT OF LIMITS VALUES (REV. 655 - 5 AUG 1987)</td></tr>
<tr><td>Surface Type:</td><td>Number of cases</td><td>< -100%</td><td>-100 to -50%</td><td>-50 to -20%</td><td>> 120%</td></tr>
<tr><td>Land</td><td>9651</td><td>61</td><td>11</td><td>28</td><td>0</td></tr>
<tr><td>Snow</td><td>2269</td><td>0</td><td>35</td><td>65</td><td>0</td></tr>
</table>

12.6   DISCUSSION

In Section 12.4, a great number of "out-of-limits" cloud amount values over land backgrounds for Cases 2 and 3 were noted. The 37 and 85 GHz brightness temperatures for each of these occurrences, when put into the cloud amount equation, produced a negative cloud amount value. In Section 12.5, there were many "out-of-limits" cloud amount values over both land and snow backgrounds during revolution 655. Again the actual numerical values produced by the SSM/I equations for all these "out-of-limits" occurrences were negative. From Equations 12.1 and 12.2 (Section 12.1), negative values occur when the 85 GHz polarization values actually observed are significantly larger than those predicted by the simulations. McFarland [10], in a similar study of the SSM/I algorithm used to determine specific land surface types, noted that the actual SSM/I polarization values at 19 and 37 GHz are often larger than the simulated values.

Based on the preflight simulations discussed in Section 12.1, the accuracy of the SSM/I algorithms were expected to be good. However, the statistical comparisons of OLS derived "truth" cloud amounts to SSM/I cloud amounts for four cases containing a variety of cloud types and land backgrounds (see Section 12.4) indicate that both algorithms have no skill at estimating the correct cloud amounts. Even if the SSM/I algorithms had shown some skill, their use would have been limited because of the large percentage of "out-of-limits" values they generate.

The small correlation coefficients for all four cases indicate no relationship between SSM/I and OLS cloud amount values. In other words, most of the individual SSM/I cloud amount values were either considerably larger or smaller than the corresponding OLS cloud amount values. For example, compare the OLS cloud truth image (Figure 12.5) to the SSM/I cloud amount image (Figure 12.6) over the land areas for Case 1 (Southern Africa). In the SSM/I image, the black represents "out-of-limits" values, the dark gray represents values of 0 to 40%, and the light gray represents values of 40 to 74%. Recall 74% was the maximum SSM/I value for this case (see Table 12.1). The cloud coverage in the OLS cloud truth image ranges from clear to overcast; most of the clear to partly cloudy areas do not match the dark gray areas (0 to 40% cloud amounts) of the SSM/I image; most of the partly to mostly cloudy areas do not match the light gray areas (40 to 74% cloud amounts) of the SSM/I image; and there are no overcast areas in the SSM/I image.

Recall that loss of polarization at 85 GHz over land and snow backgrounds in the presence of cloud was the basis of the SSM/I cloud amount algorithm. It is concluded from the preceding discussion that there is no discernable cloud signature from 85 GHz polarization values over land and snow backgrounds when no distinction is made between the many different types of land and snow surfaces which occur in nature. Several factors probably contributed to the failure of the technique. The SSM/I cloud amount algorithm was based entirely on simulated data. This was necessary since no previous microwave satellite sensor had measured radiation at frequencies as high as the 85 GHz channel on the SSM/I. All simulations contain inherent errors due to an incomplete modeling of the atmosphere and the earth surface. During the algorithm development, several simplifications were made. For snow backgrounds, only one

type of cloud (stratus/stratocumulus), one type of temperature profile (mid-latitude winter), one type of humidity profile (mid-latitude winter), and one type of precipitation state (rain-free) were used in the simulation calculations. This was also the case for land backgrounds, where the cloud type was stratus/stratocumulus, the temperature and humidity profiles were both mid-latitude summer, and the atmosphere was rain-free. It should be noted that the stratus/stratocumulus cloud used for the land and snow background simulations were identical. The cloud layer was between 0.5 and 2 km in altitude with a liquid water content 0.15 g/m^3. Clouds exhibit a wide range of liquid water contents, altitudes, and thicknesses which can be quite different from the one set of values used in the simulations. A more complete set of simulations containing a better representation of atmospheric temperature and humidity profiles, cloud conditions, precipitation states, and land and snow surface types could have produced more realistic expectations.

In order to obtain a more complete quantitative understanding of the effects of different types of clouds over various land backgrounds on 85 GHz microwave radiation, additional simulated brightness temperatures were calculated from the Geophysics Laboratory's RADTRAN atmospheric transmission model [3]. Table 12.4 contains the simulated 85 GHz values (in degrees K) for the horizontal polarization (85H) for several cloud conditions and land types. In this set of simulations, the following conditions were selected and kept constant: land surface skin temperature of 290 K, rain-free, and mid-latitude summer temperature and humidity profiles. Table 12.5 is similar to Table 12.4, but contains the values of the difference between the 85 GHz brightness temperatures for the vertical and horizontal polarizations (85D = 85V - 85H).

TABLE 12.4
SIMULATED 85H BRIGHTNESS TEMPERATURES (K) FOR VARIOUS CLOUD AND LAND TYPES UNDER MID-LATITUDE SUMMER ATMOSPHERIC CONDITIONS

Land/Cloud Types	No cloud	Stratus/stratocu	Altostratus	Cumulus
wet soil	251.0	267.9	271.3	261.6
dry soil	272.8	278.3	277.6	261.6
light veg.	274.4	279.1	278.1	261.6
moderate veg	284.1	283.7	280.9	261.6

First focus on the first three land types listed in Tables 12.4 and 12.5. For a given land type, the 85H value is larger for any of these three cloud types compared to the no cloud condition. An exception is for the cumulus condition, where for dry soil and light vegetation surfaces, the 85H value is smaller compared to that for the no cloud condition. The total columnar liquid water content increases while the 85D values decrease from left to right. Three of the five possible RADTRAN cloud models are included in these tables. The smallest and largest columnar cloud water amounts available in the RADTRAN cloud models are the stratus/

TABLE 12.5

SIMULATED 85D BRIGHTNESS TEMPERATURES (K) FOR VARIOUS CLOUD AND LAND TYPES UNDER MID-LATITUDE SUMMER ATMOSPHERIC CONDITIONS

Land/Cloud Types	No cloud	Stratus/stratocu	Altostratus	Cumulus
wet soil	21.8	10.4	6.3	0.0
dry soil	11.3	5.4	3.3	0.0
light veg.	8.1	3.8	2.4	0.0
moderate veg	0.0	0.0	0.0	0.0

stratocumulus and cumulus clouds, respectively and are included in these tables. For moderate (and greater) density vegetation, 85D values are zero for all cloud conditions because the horizontal and vertical emissivity of vegetation are equal. This indicates that at 85 GHz clouds are not detectable over land covered with moderate or greater density vegetation. It is noted in these two tables that the 85H surface emissivity values increase and that the difference between the 85V and 85H surface emissivities decrease from top to bottom. Thus for a given cloud condition, the 85H values increase while the 85D values decrease from top to bottom. An exception is for the cumulus cloud condition where the 85H values are constant and the 85D values are all zero no matter what the land type which indicates that the cumulus cloud is completely masking the surface.

Another set of RADTRAN simulations were generated for the same set of cloud conditions and land types as those presented in Tables 12.4 and 12.5. However, in this set of simulations, colder and drier conditions were used - a land surface skin temperature of 280 K and mid-latitude winter temperature and humidity profiles. The 85H results are presented in Table 12.6 and the 85D results are presented in Table 12.7. These results are similar to those for the mid-latitude summer profiles (Tables 12.4 and 12.5). The main difference for a given cloud condition and land type is that the 85H values are smaller and the 85D values are larger for the winter simulation set compared to the summer. It was also noted when the rain-free condition used for the two sets (summer and winter) of simulations was changed to light or heavier intensity rain that all 85D values were zero for any of these cloud types and surface conditions indicating that the rain completely masks the surface.

Tables 12.4 - 12.7 illustrate that clouds over land backgrounds are expected to have a distinct effect on the upwelling 85 GHz microwave radiation. However, the quantitative effect depends on the land surface type, type of cloud (columnar liquid water), the presence or absence of rain, and the atmospheric temperature and humidity profiles. For a cloud amount algorithm to be feasible, all these factors would have to be accounted for which was not the case in the Hughes cloud amount algorithm. Climatological temperature and humidity profiles might provide sufficient temperature and water vapor information. If not then perhaps radiosonde

measured profiles could be used. The presence or absence of rain can be determined by using the SSM/I rain screening algorithm developed by the DOD SSM/I land parameters validation team. Also, this team developed land surface classification and soil moisture algorithms which produce reliable land surface type and soil moisture information. The 85 GHz RADTRAN simulated values show that the ability to identify cloud types with 85 GHz SSM/I data, even when the land surface type and atmospheric profiles are known, does not appear to be likely (except for cumulus covering the entire footprint), especially when cloud types are mixed and/or only partially cover the footprints. However, the maximum and range of the 85D values over the various cloud types for a given land type and temperature and moisture profile (see Tables 12.5 and 12.7) are both small compared to the 85D value for the no cloud condition so that the maximum or average 85D value for all five possible RADTRAN cloud types would probably be adequate for use in the development of a reasonably accurate SSM/I cloud amount algorithm. An SSM/I cloud amount algorithm possibly is feasible over land surfaces that are homogeneous, except for surfaces covered with moderate or greater density vegetation. The development of a new SSM/I cloud amount algorithm using the recently developed SSM/I algorithms for land surface classification, soil moisture, and rain screening should be explored.

TABLE 12.6

SIMULATED 85H BRIGHTNESS TEMPERATURES (K) FOR
VARIOUS CLOUD AND LAND TYPES UNDER MID-LATITUDE
WINTER ATMOSPHERIC CONDITIONS

Land/Cloud Types	No cloud	Stratus/stratocu	Altostratus	Cumulus
wet soil	210.0	243.4	249.7	245.8
dry soil	251.0	260.5	260.6	245.8
light veg.	254.0	261.8	261.5	245.8
moderate veg	272.2	269.4	266.3	245.8

TABLE 12.7

SIMULATED 85D BRIGHTNESS TEMPERATURES (K) FOR
VARIOUS CLOUD AND LAND TYPES UNDER MID-LATITUDE
WINTER ATMOSPHERIC CONDITIONS

Land/Cloud Types	No cloud	Stratus/stratocu	Altostratus	Cumulus
wet soil	41.0	17.1	10.9	0.0
dry soil	21.2	8.9	5.7	0.0
light veg.	15.1	6.4	4.0	0.0
moderate veg	0.0	0.0	0.0	0.0

It is expected for snow backgrounds that its water equivalent and type of snow surface (dry, wet, stage of ripening, etc.), type of cloud, presence or absence of precipitation, and the atmospheric temperature and humidity profiles will have to be considered for the possible development of a cloud amount algorithm. A new cloud amount algorithm for snow backgrounds should be explored when the DOD SSM/I land parameters validation team's snow type and water equivalent algorithm is perfected.

12.7 OTHER CONSIDERATIONS

12.7.1 Composite SSM/I and OLS Imagery

The powerful capability of AIMS to generate a false color composite multispectral image proved to be fruitful in regards to OLS and SSM/I imagery data. There are three color guns on AIMS; red, green, and blue. The intensity of each color gun is controlled by 8 bits. Individual channels of a composite image are simultaneously directed to one of the three color guns. In regions of the image where the response of each channel is approximately equal; the red, green, and blue color intensities will be about the same and produce a shade of gray. In other regions where the spectral response of one channel is different than another, the image will be a distinctive color depending on the relative strength of the signal at the individual wavelengths. A useful display over land backgrounds uses the OLS visible channel, IR channel, and SSM/I horizontally polarized 85 GHz channel to drive the red, green, and blue guns, respectively. The resulting false color composite image (an example is shown in Figure 1.22 in Volume I of this report) shows low altitude water clouds in red because of their high visible reflectivity (large red contribution) and warm IR and microwave brightness temperatures (small green and blue contributions); thick cirrus clouds as yellow because of their high visible reflectivity and cold IR brightness temperatures (large red and green contributions) but warm microwave brightness temperatures (small blue contribution); thin cirrus clouds as green because of their cold IR brightness temperatures (large green contribution) but weak visible reflectivity and warm microwave brightness temperatures (small red and blue contributions); and strong convective cells as white because of their high visible reflectivity and cold IR and microwave brightness temperatures (large red, green and blue contributions). Thus, combining OLS visible and IR data with SSM/I brightness temperature data yields useful cloud type information. This false color composite technique also works over ocean backgrounds but the color/cloud type interpretation is not the same as for land because the ocean surface microwave, visible, and IR signatures are different.

12.7.2 Convective Clouds

It is noted that well developed convective clouds have a distinct signature at 85 GHz over land backgrounds. The 85 GHz brightness temperatures are very low under these conditions. The horizontally polarized 85 GHz brightness temperature image for Case 1 (Southern Africa - see Section 12.3) is shown in Figure 12.7. In this image, the brightness temperatures decrease as gray shades go from dark to light. The white areas indicate where the coldest brightness temperatures are located (minimum brightness temperature in this image is 136 K), and
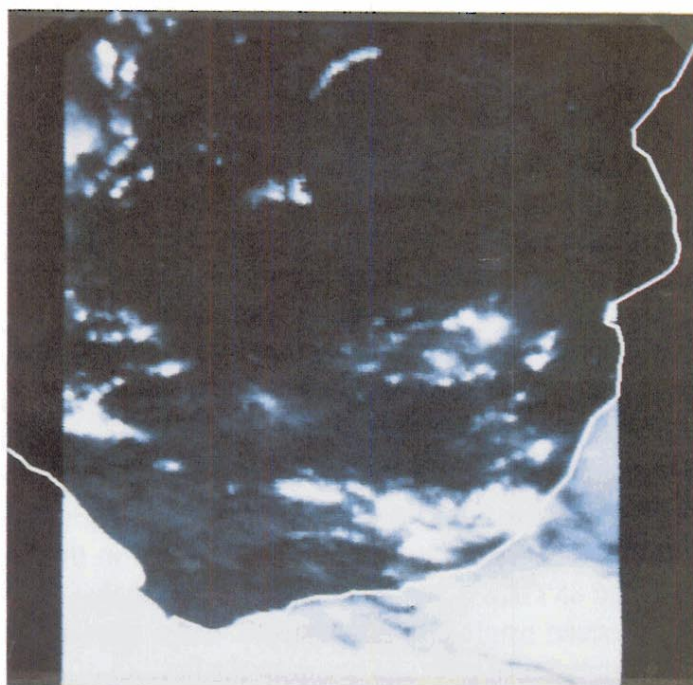
Figure 12.7   Horizontally Polarized 85 GHz Brightness
Temperature Image for Case 1

comparison of this image to the OLS visible (Figure 12.3) and IR (Figure 12.4) show that these cold 85 GHz horizontally polarized brightness temperatures are within the overcast convective regions. This cold signature of well developed convective clouds at 85 GHz has also been observed by the author in several other SSM/I brightness temperature images in various locations and seasons over several different surface backgrounds including oceans. The 85 GHz horizontally polarized brightness temperature for a well developed convective cloud can be very low. For example, a value of 95K was observed for an evening DMSP pass over India on 28 June 1987. The cold signature is due to large raindrops and ice particles in the upper portions of well developed convective clouds which scatter the upwelling radiation emitted from the lower portions of the clouds out of the SSM/I's field of view. This was first observed by Wilheit [11] with a 92 GHz radiometer flown on an aircraft.

12.7.3 Clouds over Ocean

The ocean surface in general is much more homogeneous, and has much lower microwave emissivity and much greater microwave polarization than land which indicates that clouds (as well as other atmospheric parameters) should be more easily discernable over ocean compared to land in microwave imagery data. Examination of several SSM/I 37 and 85 GHz brightness temperature images over oceanic backgrounds containing various cloud types (which were verified with coincident or near-coincident visible and IR satellite data and synoptic data) indicated that all cloudy regions, no matter what the cloud type (except for cirrus), were evident. As discussed above, the cloud signature is very cold at 85 GHz for convective clouds containing large raindrops and ice particles. Other clouds have a warm brightness temperature signature at both 37 and 85 GHz compared to the cold brightness temperature signature due to the low emissivity ocean surface and relatively small atmospheric attenuation in the absence of clouds. The emissivity of clouds at 37 and 85 GHz is significantly greater than that of the ocean surface. In the 85 GHz horizontally polarized brightness temperature image shown in Figure 12.7, the dark bands (warm brightness temperature) over the ocean in the bottom of the image are cloudy areas as can be verified by comparison with the corresponding OLS visible (Figure 12.3) and IR (Figure 12.4) images. Additionally, it is seen that the 85 GHz brightness temperature polarization values (85V - 85H) over the ocean in cloudy regions are much smaller than those for clear regions.

In order to obtain a better quantitative understanding of the effects of clouds over ocean backgrounds on microwave radiation, simulated 85 GHz brightness temperatures were calculated from the Geophysics Laboratory's RADTRAN atmospheric transmission model [3]. Table 12.8 contains the simulated 85 GHz brightness temperature values (in degrees K) for the horizontal polarization (85H) and the difference between the vertical and horizontal polarizations (85D = 85V - 85H) for several cloud types (all available cloud models in RADTRAN) with conditions of no rain and light rain (5 mm/hr at the surface). In this set of simulations, the following conditions were selected and kept constant: 85 GHz vertical and horizontal emissivity values typical for a calm ocean surface; ocean surface temperature of 290 K; and mid-latitude summer temperature and humidity profiles. Under rain-free conditions, the 85H values when any of the cloud types is present is considerably warmer than the cloud-free condition. Also, the 85D

values when clouds are present range from 0 to 14 K which is much smaller than the value for the cloud-free condition. The total columnar cloud liquid water value increases from top to bottom in the table. As the cloud water increases, the 85D value decreases indicating that the ocean surface emission and reflection of 85 GHz radiation is more heavily attenuated by the atmosphere. Under light rain conditions, the 85D values are all zero, indicating that the ocean surface is completely masked by the atmosphere at 85 GHz. This is also true under moderate and heavy rain conditions (not shown in the table).

TABLE 12.8

SIMULATED 85H AND 85D BRIGHTNESS TEMPERATURES (K) FOR SEVERAL CLOUD TYPES OVER A CALM OCEAN SURFACE UNDER MID-LATITUDE SUMMER ATMOSPHERIC CONDITIONS

Cloud Types	No Rain 85H	No Rain 85D	Light Rain 85H	Light Rain 85D
no cloud	238.0	29.1	--	--
stratus/stratocu	261.8	13.8	270.3	0
alto stratus	267.5	8.5	270.1	0
nimbostratus	275.7	5.7	270.4	0
stratocumulus	275.6	5.2	270.4	0
cumulus	261.6	0	260.8	0

Table 12.9 shows the 85H values in degrees K calculated with RADTRAN using various atmospheric temperature and humidity profiles under clear and cloudy (stratus/stratocumulus cloud with no rain) conditions for calm and rough ocean surfaces. The atmospheric profiles become colder and drier from top to bottom in the table. For a given atmospheric profile over a calm ocean surface, the 85H brightness temperature is larger for cloudy than for clear conditions. This is also true over a rough ocean surface, but the amount of brightness temperature increase with cloud is approximately half that as for the calm surface. Also for a given atmospheric profile and cloud condition, the 85H brightness temperature is larger over a rough surface than a calm one. For each surface and cloud condition, the 85H brightness temperatures decrease and the amount of increase of brightness temperature with cloud compared to no cloud becomes greater as the atmosphere becomes colder and drier.

TABLE 12.9

SIMULATED 85H BRIGHTNESS TEMPERATURES (K) FOR CLEAR VS. CLOUDY UNDER VARIOUS ATMOSPHERIC CONDITIONS OVER CALM AND ROUGH OCEAN SURFACES

Atmos. Profiles	Ocean Temps.	(Calm Sfc.) no cloud	(Rough Sfc.) no cloud	(Calm Sfc.) stratus/ stratocu	(Rough Sfc.) stratus/ stratocu
tropical	300	258.2	271.7	272.8	279.6
mid-lat. summer	290	238.0	257.4	261.8	271.0
sub-arctic summer	285	217.7	243.2	250.5	261.7
mid-lat. winter	280	185.7	222.2	233.2	248.4
sub-arctic winter	275	170.4	210.9	220.9	238.5

Table 12.10 shows the 85D values in degrees K calculated with RADTRAN for the same conditions as those for the 85H values shown in Table 12.9. For a given atmospheric profile over a calm ocean surface, the 85D value is smaller for cloudy than for clear conditions. This is also the case over a rough ocean surface, but the amount of decrease of the 85D value with cloud is approximately half that as for the calm ocean. Also for a given atmospheric profile and cloud condition, the 85D value is larger over a calm surface than over a rough surface. The 85D values increase for a given surface and cloud condition and the amount of decrease of 85D with cloud compared to no cloud becomes greater as the atmosphere becomes colder and drier.

The simulated 85 GHz RADTRAN values given in Tables 12.8 - 12.10 indicate that clouds have a distinct effect on the upwelling 85 GHz microwave radiation at the top of the atmosphere over oceanic backgrounds. The amount of columnar liquid water (type of cloud) has an important influence on the 85 GHz brightness temperatures. Other important factors are the degree of roughness of the ocean surface, columnar water vapor (moisture and temperature profiles), and the presence or absence of rain. Other members of the DOD SSM/I geophysical

TABLE 12.10

SIMULATED 85D BRIGHTNESS TEMPERATURES (K) FOR CLEAR VS. CLOUDY UNDER VARIOUS ATMOSPHERIC CONDITIONS OVER CALM AND ROUGH OCEAN SURFACES

Atmos. Profiles	Ocean Temps.	(Calm Sfc.) no cloud	(Rough Sfc) no cloud	(Calm Sfc.) stratus/ stratocu	(Rough Sfc) stratus/ stratocu
tropical	300	20.3	12.4	10.2	6.2
mid-lat. summer	290	29.1	17.8	13.8	8.5
sub-arctic summer	285	38.2	23.4	16.9	10.4
mid-lat. winter	280	54.6	33.3	22.9	14.0
sub-arctic winter	275	60.8	37.2	26.4	16.2

parameter algorithm validation team have shown that these atmospheric and surface conditions can be determined from the SSM/I data since they have developed SSM/I algorithms for ocean backgrounds which calculate columnar cloud liquid water and water vapor, ocean surface wind speed (which is related to the surface roughness), and surface rain rates. Therefore, an accurate SSM/I cloud amount algorithm for ocean backgrounds is plausible. Recall that Hughes Aircraft Company has not been tasked to develop one. However, Rubinstein [12], a member of the sea ice validation team, recently developed an SSM/I cloud amount algorithm as a spin-off of her work. The accuracy of this algorithm requires validation. It is important to note that cirrus clouds are transparent at SSM/I frequencies over all backgrounds and so any SSM/I cloud amount algorithm will lack cloud coverage information in areas containing only cirrus type clouds.

12.7.4 Potential SSM/I Contributions to the RTNEPH Cloud Analysis

The RTNEPH cloud analysis done at AFGWC uses conventional ground-based cloud observations, and OLS IR and visible satellite data. The RTNEPH produces operational global estimates of cloud cover, altitude, and type. (See Keiss and Cox [14] for a complete description of RTNEPH.) There is good potential to improve the RTNEPH analysis by incorporation of new algorithms which use SSM/I data by itself and also in conjunction with other types of data.

If the development and validation of new SSM/I cloud amount algorithms over some types of land, snow, and ocean surfaces is successful, then these cloud amounts should be examined to determine if they are more accurate under certain or all situations than the RTNEPH cloud values determined from conventional observations, and IR and visible satellite data. For instance, there are few conventional cloud observations over the ocean and under certain common oceanic conditions it is difficult to detect clouds with IR satellite data and if it is night time then no visible satellite data are available. An example is that often in the IR there is little contrast between stratocumulus clouds in the marine boundary layer and the ocean background in the presence of the commonly occurring temperature inversion in this layer.

The extraction of cloud type information available from false color composite OLS and SSM/I images described in Section 12.7.1 probably could be automated and incorporated into the RTNEPH analysis. If this can be done, then improvement to the cloud type portion of the RTNEPH analysis would probably result.

Combined use of SSM/I microwave data with OLS IR data for determination of cloud amounts over land backgrounds is promising. Savage et al. [13] have found that the expected surface IR brightness temperature for clear conditions over vegetated land backgrounds can be predicted from SSM/I brightness temperatures with sufficient accuracy (rms of 2.5 K) to be used as input for a cloud analysis. The observed IR values are compared to the expected IR to estimate cloud. Observed IR values less than the expected IR indicate cloud. The regression equation used to estimate an expected surface IR value for clear conditions was developed from an analysis of observed IR values in clear areas, using the SSM/I 19 and 22 GHz channels as predictors. The two lowest frequency SSM/I channels were used since most clouds are transparent at these frequencies. This method of comparing observed IR values to the expected surface IR temperature to estimate cloud is comparable to the technique presently used by the RTNEPH cloud analysis model at AFGWC. However, the RTNEPH estimates the expected IR temperature from surface air temperature reports for comparison to the observed OLS IR temperatures. The technique based entirely on satellite data is potentially more accurate because there is error resulting from estimating IR background temperatures from the surface temperature report which is a shelter air temperature (several feet above ground level). Another advantage of the all-satellite technique is that it requires less data processing and produces more timely results. It is also noted that the SSM/I land parameter validation team has developed algorithms for the determination of surface skin temperatures for several land types in addition to vegetated land which should be useful for estimation of IR background temperatures.

The all-satellite technique is expected to be successful over all surfaces whose microwave emissivity is high and relatively constant. Vegetated land, as well as desert, have these emissivity characteristics. However, snow, glacial, and ocean surfaces have low emissivities (high reflectivities). Savage et al. [13] found that IR brightness temperatures for clear conditions could not be accurately estimated from SSM/I observations over snow-covered land backgrounds because of the physical properties of snow. However, they found that an approach (differing from the algorithm GL validated) based entirely on SSM/I observations for recognition of clouds over snow showed good promise. They resolved a set of SSM/I data which was stratified into

cloudy and clear groups, into eigenfunctions and then formed a discriminant function. Only a few of the largest discriminant scores of the cloudy group overlapped with a few of the smallest discriminant scores of the clear group. Thus, the two groups were quite well distinguished.

The SSM/I is good at snow and ice cover detection because of the strong microwave signatures of these surfaces. Use of this timely and accurate information in the RTNEPH cloud analysis would improve it. The RTNEPH satellite data processor consists of two parts - one for OLS visible data and one for OLS IR data. The visible data processor is not allowed to make a cloud amount calculation over grid points where snow or ice cover is believed to be present. This is because cloud-free snow and ice covered areas have approximately the same brightness as clouds. The snow-cover data base is of particular concern. It is based on surface reports and climatology and may not represent the true snow cover condition over many gridpoints, especially in sparsely populated regions where the surface weather observing stations are far apart. If snow or ice is actually present when RTNEPH believes it not to be, then the visible data processor will be used and RTNEPH's estimates of cloud amounts will probably be too large. On the other hand, if snow or ice is really absent when RTNEPH believes it to be present, then the IR satellite processor will be used and low clouds, that are easily found by the visible processor, may be poorly analyzed.

12.8   CONCLUSIONS

The present Hughes SSM/I cloud amount algorithms over land and snow backgrounds do not work because the variability of the land and snow surface types, cloud types, and atmospheric temperature and humidity profiles, and the presence or absence of rain were not all taken into account. An SSM/I cloud amount algorithm possibly is feasible over land surfaces that are homogeneous, except for surfaces covered with moderate or greater density vegetation. The development of a new SSM/I cloud amount algorithm using climatological temperature and humidity profiles and the recently developed SSM/I algorithms for land surface classification, soil moisture, and rain screening should be explored. Also, a new cloud amount algorithm for snow backgrounds should be explored when the DOD SSM/I land parameters validation team's snow type and water equivalent algorithm is perfected.

Hughes Aircraft Company has not been tasked to develop an SSM/I cloud amount algorithm over ocean backgrounds. Investigation to date indicates an accurate algorithm over ocean is plausible. In fact, an algorithm has recently been developed but requires validation. It is important to note that cirrus clouds are transparent at SSM/I frequencies over ocean and all other backgrounds and so any SSM/I cloud amount algorithm will lack cloud coverage information in areas containing only cirrus type clouds.

SSM/I data combined with other types of data and several SSM/I geophysical parameter algorithms offer the opportunity for improvement to the Air Force's RTNEPH operational global cloud analysis. Cloud amount estimates from potential SSM/I algorithms might prove to be more accurate in certain situations than those obtained from OLS data by the RTNEPH. The extraction of cloud type information available in color composite SSM/I and OLS (visible and

IR) imagery probably could be automated and incorporated into the RTNEPH analysis. Other promising SSM/I contributions to the RTNEPH include improvements to its snow and ice cover data base and more accurate and timely estimation of expected IR temperatures for clear conditions over vegetated land, moist soils, desert, and arable land.

12.9 REFERENCES

[1]   Hughes Aircraft Company, "Special Sensor Microwave/Imager SSM/I Technical Operating Report," Contract F0470-1084-C-0036, Los Angeles, CA, 1985.

[2]   R. C. Savage, private communication, Hughes Aircraft Company, Aurora, CO, 1986.

[3]   V. J. Falcone, L. W. Abreu, and E. P. Shettle, "Atmospheric Attenuation of Millimeter and Submillimeter Waves: Models and Computer Code," AFGL-TR-79-0253, ADA084485, Hanscom AFB, MA, 1979.

[4]   F. T. Ulaby and W. H. Stiles, "The active and passive microwave response to snow parameters," J. Geophys. Res., vol. 85, no. C2, pp. 1037-1049, 1980.

[5]   B. J. Choudhury, T. J. Schmugge, A. Cheng, and R. W. Newton, "Effect of surface roughness on the microwave emission from soils," J. Geophys. Res., vol. 84, no. C9, pp. 5699, 1979.

[6]   Hughes Aircraft Company, "Special Sensor Microwave/Imager SSM/I Computer Program Product Specification," Vol. III Environmental Parameter Extraction Computer Program Specification Component (SMIEPE), Contract F04701-84-C-0036, Los Angeles, CA, 1986.

[7]   Hughes Aircraft Company, "Special Sensor Microwave/Imager SSM/I Data Requirements Document," Contract F04701-84-C-0036, Los Angeles, CA, 1986.

[8]   G. B. Gustafson, D. Roberts, C. Ivaldi, R. Schecther, T. Kleespies, K. Hardy, and R. Lynch, "The AFGL Interactive Meteorological System," Third International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology, AMS, Boston, MA, pp. 151-154, 1987.

[9]   G. B. Gustafson and G. W. Felde, "Interactive Satellite Image Processing Applied to Cloud Detection," Fourth International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology, AMS, Boston, MA, pp. 100-103, 1988.

[10]   M. J. McFarland, private communication, Texas A&M University, College Station, TX, 1987.

[11] T. T. Wilheit, A. T. C. Chang, J. L. King, E. B. Rodgers, R. A. Nieman, B. M. Krupp, A. S. Milman, J. S. Stratigos, and H. Siddalingaiah, "Microwave radiometric observations near 19.35, 92, and 183 GHz of precipitation in tropical storm Cora," J. Appl. Meteor., vol. 21, pp. 1137-1145, 1982.

[12] I. G. Rubinstein, private communication, York University, North York, Ontario, Canada, 1988.

[13] R. C. Savage, J. F. Heinrichs, and A. J. Koscielny, "Cloud Identification Using IR and Microwave Imagery," Fifth International Conference on Interactive Information and Processing Systems for Meteorology, Oceanography, and Hydrology, AMS, Boston, MA, 1989.

[14] R. B. Keiss and W. M. Cox, "The AFGWC Automated Real-Time Cloud Analysis Model," AF Global Weather Central Technical Note 88-001, Offutt AFB, NE, 1988.

APPENDIX A

GEOLOCATION

by

James P. Hollinger
Radio, Infrared and Optical Remote Sensing Branch
Center for Advanced Space Sensing
Naval Research Laboratory
Washington, DC 20375-5000

and

Dawn K. Conway, Robert W. Conway and David J. Spangler
Bendix Field Engineering Corporation
Washington, DC 20375-5000

A.0 GEOLOCATION

A.1 INTRODUCTION

The process of geolocating SSM/I pixels was investigated and described in detail by Gene A. Poe and Robert W. Conway in Section 6.0, Volume 1 of this report and also in [1]. In brief they found geolocation errors of the order of 20 to 30 km and discovered that approximately one half of this error was due to the use of a seven-day predictive ephemeris in data processing at Fleet Numerical Oceanographic Center (FNOC). This variable error was removed when the satellite ephemeris, contained in the down-link data stream, was used in place of the predictive ephemeris in the data processing beginning with revolution 10048 on May 31, 1989. Due to system testing the predictive ephemeris was used for short periods of time until revolution 10647 July 12, 1989. Poe and Conway found indications that the remaining error of about +/- 13 km could be reduced to within specification of +/- 7 km by the use of a constant or slowly varying and predictable adjustment to the apparent spin axis of the SSM/I in the geolocation software. This does not necessarily mean that the SSM/I axis is misaligned with respect to the spacecraft, but only that it is possible to compensate by this means for some other error or errors in the overall system. They concluded that indications were that spacecraft attitude biases were not the main contributor to the remaining error.

Their initial work to determine the software adjustment to the spin axis was hampered by the difficulty in obtaining the spacecraft ephemeris, matching it to the corresponding SSM/I data and then recalculating the geolocation. It was necessary to wait until the satellite ephemeris was used in operational data processing at FNOC in order that a sufficiently large number of cases could be examined to ensure that the residual error was indeed constant and could be removed. This has now been done and a constant software correction to the apparent spin axis has been determined which reduces the geolocation error to less than the DMSP SSM/I geolocation accuracy specification of +/- 7 km. The procedure, data selection and results are described in the following sections.

A.2 PROCEDURE

The accuracy of the SSM/I geolocation was determined by a visual comparison of SSM/I 85 Ghz horizontally polarized brightness temperature (85H) images with superimposed World Data Banks II (WDB2) coastlines on a color monitor. Incremental pitch, roll and yaw corrections were estimated by trial and error and the SSM/I image geolocation repeated until the SSM/I and WDB2 coastlines coincided. The WDB2 coastlines are believed to be accurate to better than 1 km over 90% of all identifiable shoreline features and introduced no significant error in the comparison. The 85H has a resolution of 13 km and is sampled each 12.5 km along scan. Successive scans are separated by 12.5 km. Interpolation of these data using the procedure developed by Poe [2] produced an additional three equally spaced samples between each original pair in the scan direction and an additional three scans between successive scans; a sixteen fold increase in data density. The regions selected were 20 degree by 20 degree boxes. This resulted in a pixel separation on the monitor of the geolocated 85H image varying from

about 4.2 km at the equator to about 2.8 km at high latitudes. The WDB2 coastline was one pixel wide on the monitor. It was generally possible to obtain agreement between the 85H and WDB2 coastlines to one pixel or about 3 to 4 km. The use of higher resolution on the monitor would not have improved this precision significantly.

The 85H was chosen because it has the highest spatial resolution of any of the channels and a high brightness temperature contrast between land/water boundaries. This contrast varied between about 50 and 100 K depending upon atmospheric conditions. Only images containing clear sharp land/water boundaries not obscured by heavy clouds or rain were selected. The RMS noise output of the 85H was generally less than 1 K as shown in Figure A.1 until about February 1990 when it began increasing. It increased to as much as 10 K before failing entirely in February 1991. Images with RMS noise up to 5 K were used resulting in a land/boundary signal-to-noise of from 10 to 100. Thus it was readily possible to locate the coastline in the 85H image to one fourth of a half power beam-width or about one pixel as stated above.
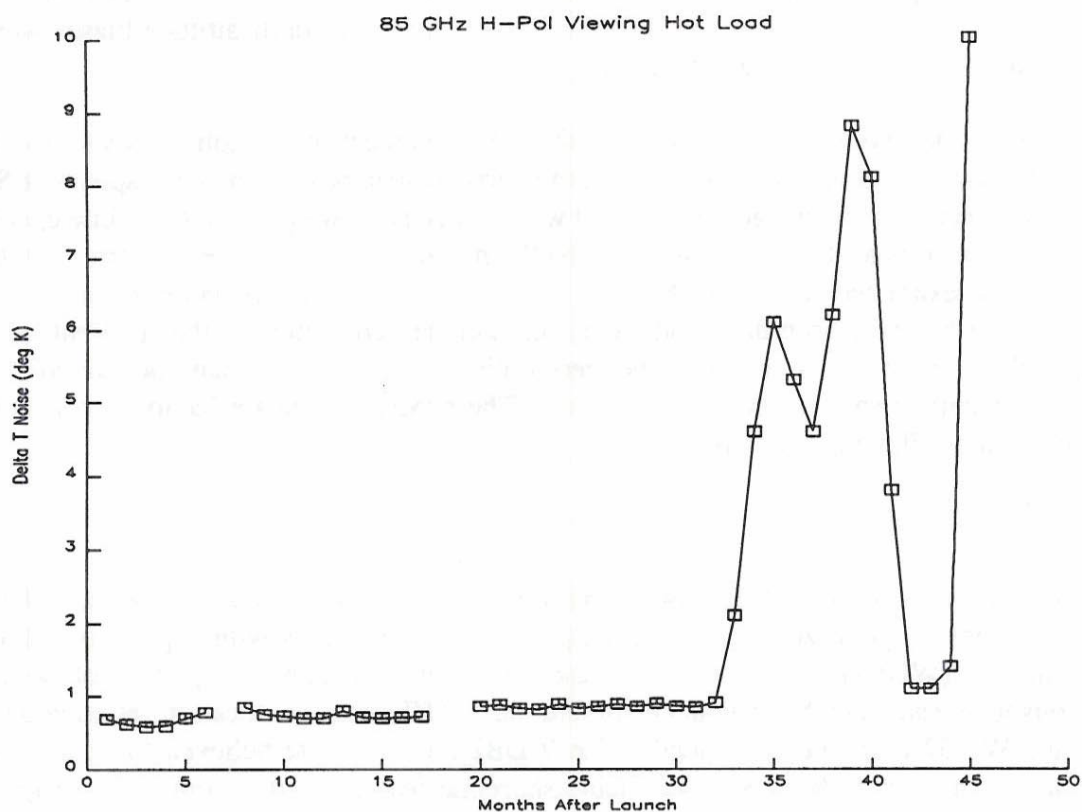
Figure A.1 SSM/I Delta T Noise

Hughes Aircraft Company measured the antenna beam positions relative to bore-sight for all seven channels prior to launch on DMSP F-8. All channels were within 0.03 degrees of the 85H except the 37H which was displaced by 0.07 degrees. This would result in relative shift of the two beams of about 2 km at the earth's surface. The co-registration was checked by overlaying the 85H and each of the other channel images. They were aligned to within better than 3 km; the accuracy of the measurement.

The algorithm used to geolocate the 85H data with different pitch, roll and yaw offsets of the spin axis is fully described in Volume I of this report and in [1]. The computations used the satellite ephemeris position vectors stored approximately each minute in the TDR archival tapes produced at FNOC. These computations would normally not introduce significant error. However when the satellite ephemeris was incorporated into the SSM/I processing at FNOC the ephemeris time was truncated to integer seconds. This error was not corrected until revolution 17057 on October 9, 1990. The error was minimized by adding one half second to the truncated time resulting in an error of up to one half second in the geolocation computations. Since this is a timing error it primarily affects the intrack position and closely resembles a pitch error. A half second timing error is equivalent to a 3.3 km position error or about a tenth of a degree pitch error.

Consideration of the above sources of error indicates that the geolocation procedure used here to obtain pitch, roll and yaw corrections for a software realignment of the apparent SSM/I spin axis to correct geolocation errors is accurate to better than 6 km. This is consistent with the geolocation accuracy specification of +/- 7 km.

A.3   DATA SELECTION

The images for coast line comparison were selected to allow possible systematic variations due to the time of year, sun angle, latitude, longitude and ascending/descending orbits to be examined. This requires data covering a range of more than 5000 orbits. In order to use only data processed with the satellite ephemeris only data from revolutions following number 10048 were selected. The noise of the 85H channel was below 1 K until about revolution 13500 on January 30, 1990 after which it began increasing; see Figure A.1. Therefore it was necessary to use some images with noise as large as 5 K. Data from 203 orbits during the period June 2, 1989, revolution 10070, to July 29, 1990, revolution 16036, were used for coastline comparisons. The center latitude and longitude of the 20 degree by 20 degree regions, the number of ascending and descending revolutions, the pitch, roll and yaw corrections which eliminate the geolocation error and the Julian dates of the images are given in Table A.1. The distribution of the data as a function of the day of year, center latitude and center longitude is given in Figures A.2, A.3 and A.4 respectively. The sun angle for revolutions 10000 through 16000 is shown in Figure A.5.

							TABLE A.1 GEOLOCATION REGIONS	
REGION	CENTER LAT/LON	TOTAL ASC	TOTAL DES	AVG PITCH	AVG ROLL	AVG YAW	JULIAN DAY 1989	JULIAN DAY 1990
ENGLAND	50N/0E	9	4	−0.2	−0.1	0.7	217,229,270, 277,307,342	20,53,81,121, 128,145,166
GREAT LAKES	45N/275E	3	8	−0.2	−0.1	0.7	232,248,280, 285,325,355	97,132,162, 167,185
ITALY	40N/15E	3	7	−0.2	−0.1	0.7	232,243,255, 324,327	30,56,142, 143,166
BLACK SEA	40N/30E	2	3	−0.3	−0.1	0.7	335	79,120,147, 182
KOREA	40N/125E	5	3	−0.3	−0.1	0.7	222,258,293	03,79,80,108, 145
SPAIN	40N/0E	3	4	−0.1	−0.1	0.7	232,262	02,24,30,93, 145
PERSIAN GULF	23N/60E	4	9	−0.2	−0.1	0.7	220,233,254, 290	10,46,84,94, 102,122,171, 191
BAJA PENINSULA	30N/245E	4	8	−0.3	−0.1	0.7	232,270,301, 325	07,16,39,98, 153,160,167, 197
FLORIDA, CUBA	30N/280E	4	5	−0.2	−0.1	0.7	298,327,346	26,43,80,87, 115,145
GULF OF MEXICO	25N/265E	6	4	−0.2	−0.1	0.7	232,277,279, 326	05,30,68,80, 105,118
PAKISTAN	20N/75E	3	0	−0.2	−0.1	0.7		175,194,205
MEXICO, FLORIDA	20N/275E	1	0	−0.2	−0.1	0.7		31

				TABLE A.1 continued				
REGION	CENTER LAT/LON	TOTAL ASC	TOTAL DES	AVG PITCH	AVG ROLL	AVG YAW	JULIAN DAY 1989	JULIAN DAY 1990
INDIA	15N/85E	2	5	-0.2	-0.1	0.7	270,338	05,27,92,143, 144
ISTHMUS OF PANAMA	5N/275E	3	2	-0.2	-0.1	0.7	219,263	26,31,173
COLOMBIA	5N/285E	5	2	-0.2	-0.1	0.7	265,305,345, 345	96,112,153
WEST COAST AFRICA	0 /5E	1	0	-0.2	-0.1	0.7	257	
BRAZIL	5S/320E	1	1	-0.2	-0.1	0.7		200,210
EAST COAST AFRICA	10S/40E	1	0	-0.2	-0.1	0.7	153	
NEW GUINEA	10S/140E	3	7	-0.2	-0.1	0.7	259,276,328	15,81,85,88, 123,143,167
PERU	10S/285E	2	4	-0.2	-0.1	0.7	225,253,315, 344	61,137
MADAGASCAR	20S/50E	3	5	-0.2	-0.1	0.7	214,230,232, 324	07,28,85,166
SOUTHERN AFRICA	25S/30E	3	0	-0.2	-0.1	0.7	274,275	145
AUSTRALIA	35S/140E	8	4	-0.2	-0.1	0.7	234,270,288, 314,323	10,50,72,92, 135,153,200
NEW ZEALAND	40S/170E	6	4	-0.3	-0.1	0.7	239,268,324, 346	03,41,59,95, 143,201
SOUTH AMERICA, FALKLAND ISLANDS	45S/295E	7	10	-0.2	-0.1	0.7	227,229,260, 310,318,358, 360,363	05,36,74,87, 113,126,135, 161,201
SOUTH AMERICA, FALKLAND ISLANDS	55S/300E	1	0	-0.2	-0.1	0.7	258	
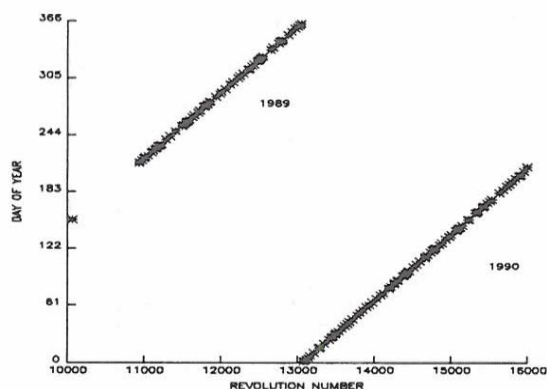
Figure A.2 Julian Day Distribution of SSM/I Geolocation Data
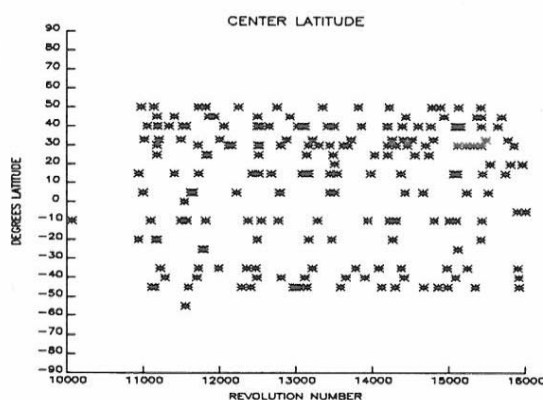
Figure A.3 Latitude Distribution of SSM/I Geolocation Data
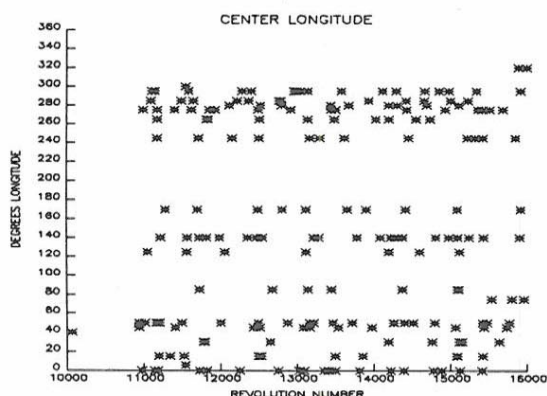
Figure A.4 Longitude Distribution of SSM/I Geolocation Data
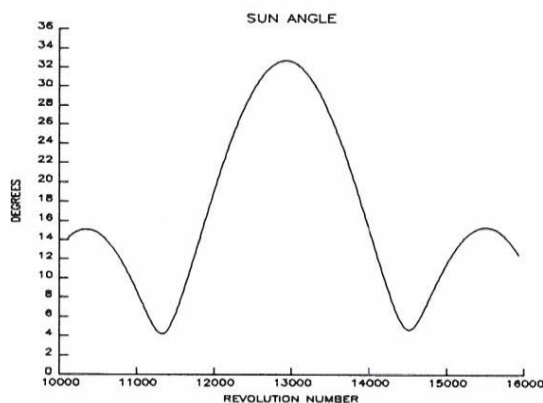
Figure A.5 Sun Angle Variation For SSM/I Geolocation Data

A.4   RESULTS

The pitch, roll and yaw corrections which eliminate the geolocation error are shown as a function of revolution number for all 203 cases in Figures A.6, A.7 and A.8 respectively. There is almost no variation in the value of the roll or yaw correction. The pitch correction shows a greater variation than either the roll or yaw with a standard deviation of 0.11 degrees. This is not surprising since timing errors result in geolocation errors very similar to those caused by pitch errors. A timing error of 1 second produces a 6.6 km intrack geolocation error which is roughly the same as a 0.18 degree pitch error. Thus the 0.11 degree pitch error resembles a 0.6 second timing error. As mentioned earlier the truncation of the ephemeris time for all of the data used here results in timing errors of up to 0.5 seconds. There is just a hint that the pitch error may be slightly larger (more positive) at larger sun angles when the sensor is warmer but no sun angle dependent correction is necessary or justifiable. The pitch correction is also

independent of latitude and longitude and is the same for ascending and descending passes as is shown in Figures A.9 through A.12. Thus a constant correction is possible.
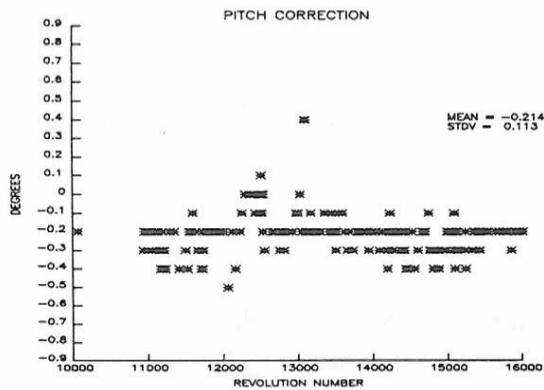
Figure A.6 SSM/I Geolocation Pitch Correction
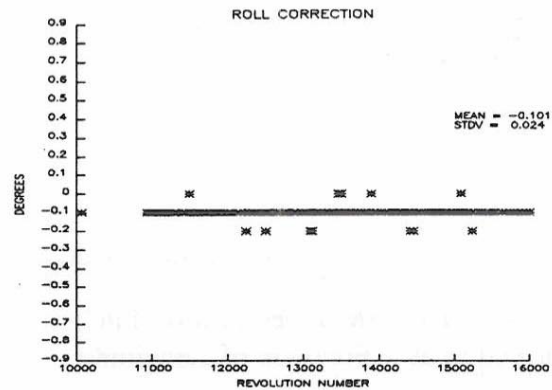
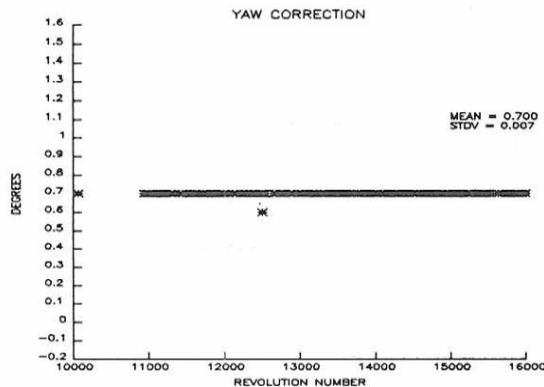Figure A.7 SSM/I Geolocation Roll Correction
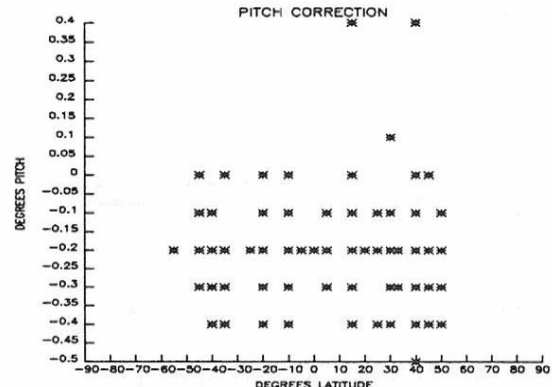

Figure A.8 SSM/I Geolocation yaw Correction

Figure A.9 SSM/I Geolocation Pitch Correction as a Function of Latitude

The pitch, roll and yaw correction to the apparent SSM/I spin axis which brings the SSM/I geolocation within the specification of +/- 7 km is:

Pitch = -0.21 degrees
Roll = -0.10 degrees
Yaw   = +0.70 degrees.

The sign of these coefficients is according to the DMSP convention which is shown in Figure A.13. The geolocation shift imposed by this realignment as a function of scan angle is shown in Figure A.14. In the figure the positive cross track direction is to the port side and the positive in track direction is aft of the spacecraft. The bottom curve is the scan track with no correction and the curve displaced towards the upper left is the shifted scan track resulting from

A-7
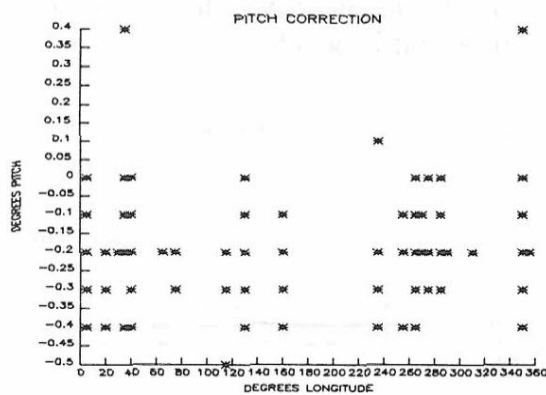
Figure A.10 SSM/I Geolocation Pitch
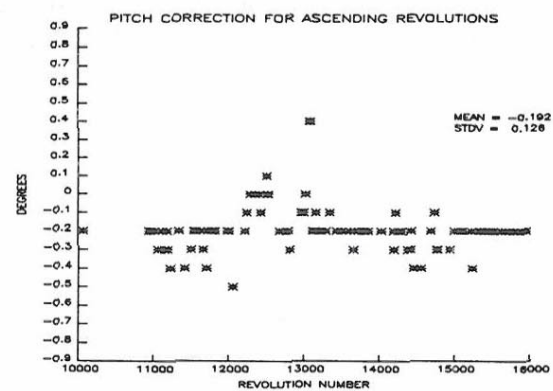Correction as a Function of Longitude

Figure A.11 SSM/I Geolocation Pitch
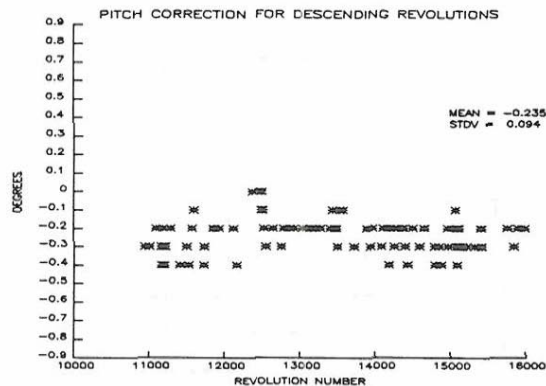Correction for Ascending Revolutions

Figure A.12 SSM/I Geolocation Pitch
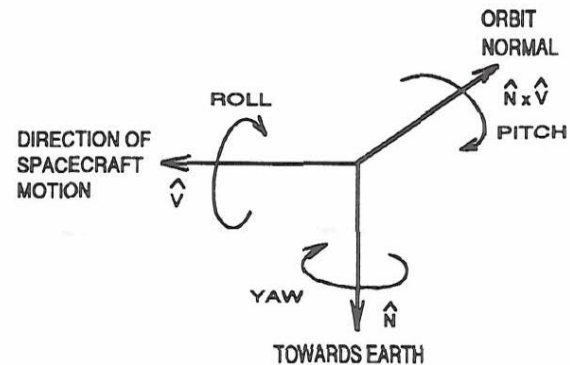Correction For Descending Revolutions

Figure A.13 DMSP Reference Axes

the above correction. The magnitude of the geolocation shift in both the cross track and in track direction is given in Figure A.15 as a function of scan angle. Note that the total geolocation shift ranges from about 8 to 15 km. The change in incidence angle resulting from the realignment of the spin axis as a function of scan angle is given in Figure A.16. The calculations in Figure A.14 through A.16 are with respect to a spherical earth. They will change slightly in detail with latitude for an oblate spheroidal earth model and with the rotation of the argument of perigee of the slightly elliptical F-8 orbit. In order to determine the magnitude of the incidence angle variation for an oblate spheroid model of the earth and the extremes of the elliptical orbit, calculations were made for revolutions 15106 and 15563 for which the argument of perigee is 90 and 0 degrees respectively. The maximum change of incidence angle during a single scan was 0.30 degrees. The maximum incidence angle variation over the two orbits was 0.88 degrees. It should be noted that a change of 0.72 degrees results for these same orbits for a zero pitch, roll and yaw correction. The primary cause of the incidence angle variation, for

A-8

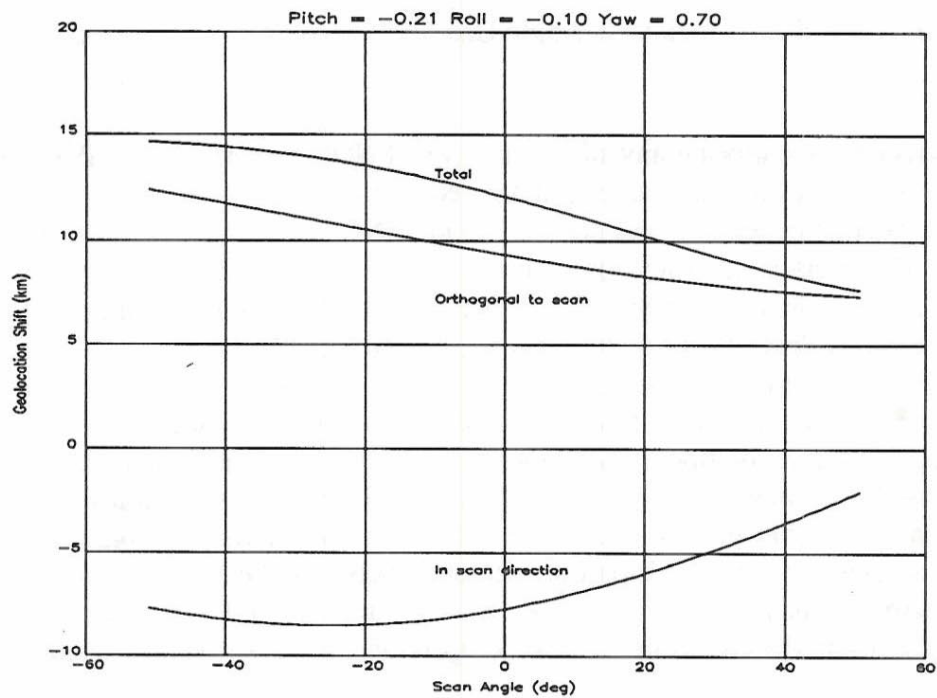Figure A.14 SSM/I Geolocation Shift Due to Spin Axis Realignment

Figure A.15 SSM/I Geolocation Shift In and Orthogonal to Scan Direction

A-9

this alignment correction, is the variation of spacecraft altitude which ranged from 837 to 885 km over these orbits. Incidence angle variations of +/- 0.5 degrees can result in brightness temperature changes of +/- 1 K or more and refined environmental retrieval algorithms must take the actual incidence angle at each scan position into account.
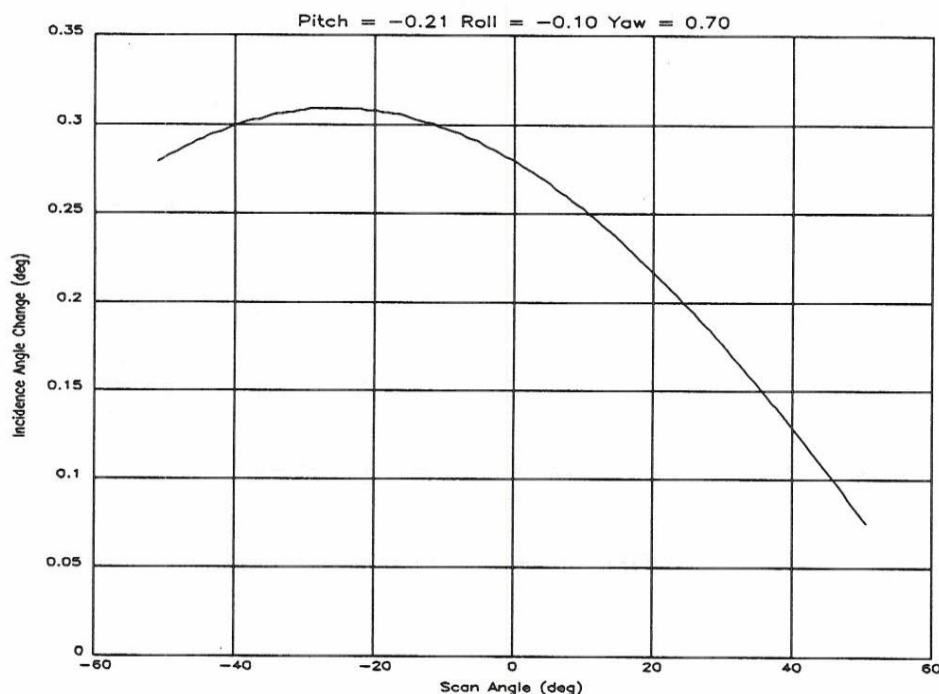

Figure A.16 SSM/I Incidence Angle Change Due to Spin Axis Realignment

It is possible to geolocate any of the archived FNOC processed F-8 SSM/I data to an accuracy of +/- 7 km using the above spin axis correction. For those data prior to July 12, 1989 when the satellite ephemeris was not used in the FNOC processing a new ephemeris must be generated. This is possible using, for example, orbital elements from the Space Surveillance Center (SSC) of the United States Space Command (formerly NORAD), Cheyenne Mountain, Colorado or the Naval Space Surveillance System (NAVSPASUR), Dahlgren, Virginia and their respective orbital prediction programs. The difference in geolocation obtained by using the satellite ephemeris and that using NAVSPASUR orbital elements and the PPT7 ephemeris prediction program for revolution 10121 is given in Figure A.17 as a function of time. The error is at most 6 km with a standard deviation of 2.5 km. This accuracy is not strongly dependent upon the number of revolutions over which the ephemeris is propagated. The mean and standard deviation of the geolocation difference between the two ephemerides for nine different comparisons is given in Figure A.18 as a function of the number of revolutions propagated. The mean difference of all nine comparisons is 2.3 km with a standard deviation of 1.4 km.
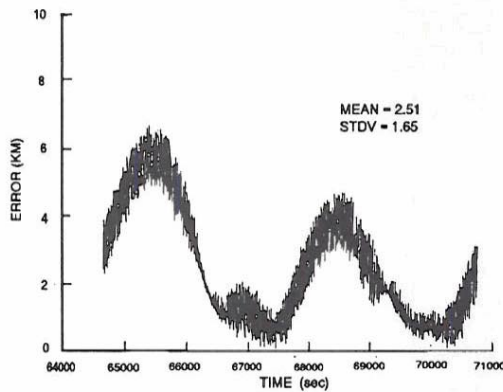
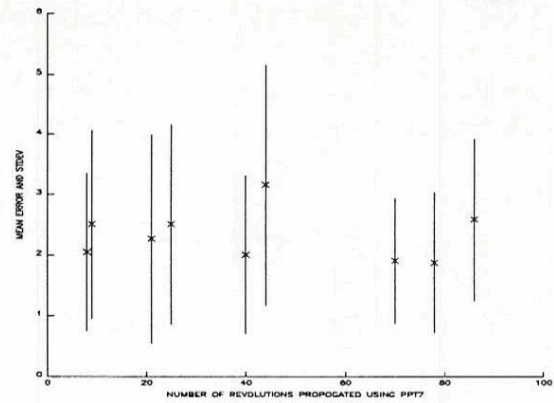Figure A.17 DMSP - PPT7 Ephemerides for Revolution 10121

Figure A.18 Mean Error of PPT7 - OLS Ephemerides for Various Propagations

It should be noted that a yaw correction may result from a timing or other error in the start of scan signal and need not be an alignment error of the spin axis. The pitch and roll corrections of -0.21 and -0.10 degrees respectively are entirely consistent with the Hughes antenna beam alignment error specification with respect to the spacecraft of +/- 0.2 degrees in all three axes. Therefore the above correction is consistent with the SSM/I design alignment tolerances but does not necessarily mean that the SSM/I axis is misaligned with respect to the spacecraft. However it is possible, by using this correction to the apparent spin axis, to compensate for alignment error or other errors in the overall system and geolocate the SSM/I to an absolute accuracy of +/- 7 km.

An example of the improvement in geolocation resulting from the use of this software correction of the spin axis alignment is given in Figure A.19. The 85H image of the southern part of South America obtained from revolution 11155 on 17 August 1989 without the correction is shown on the left and with the correction on the right. The red areas are primarily lower elevation land. The light and dark blue areas in the Andes are due to snow. The light blue areas in the vicinity of the Falkland Islands are heavy clouds. Note the excellent agreement between the 85H image and the WDB2 coastlines and lakes throughout the image. It should be noted that this uniform spatial fidelity can only be obtained with the three angle spin axis correction and cannot be duplicated by a simple two dimensional translation of the image.

Accurate geolocation is very important for the delineation and recognition of small atmospheric and terrain features. This is especially true if successive passes over a specific region are to be averaged for the study of slowly varying phenomena. It is also essential for algorithm development and validation; particularly in the case of precipitation, sea ice edge and land surface type. Now that the geolocation problem has been solved and the methodology for determining the correction established the accurate geolocation of SSM/I's on future DMSP satellites can be readily accomplished.
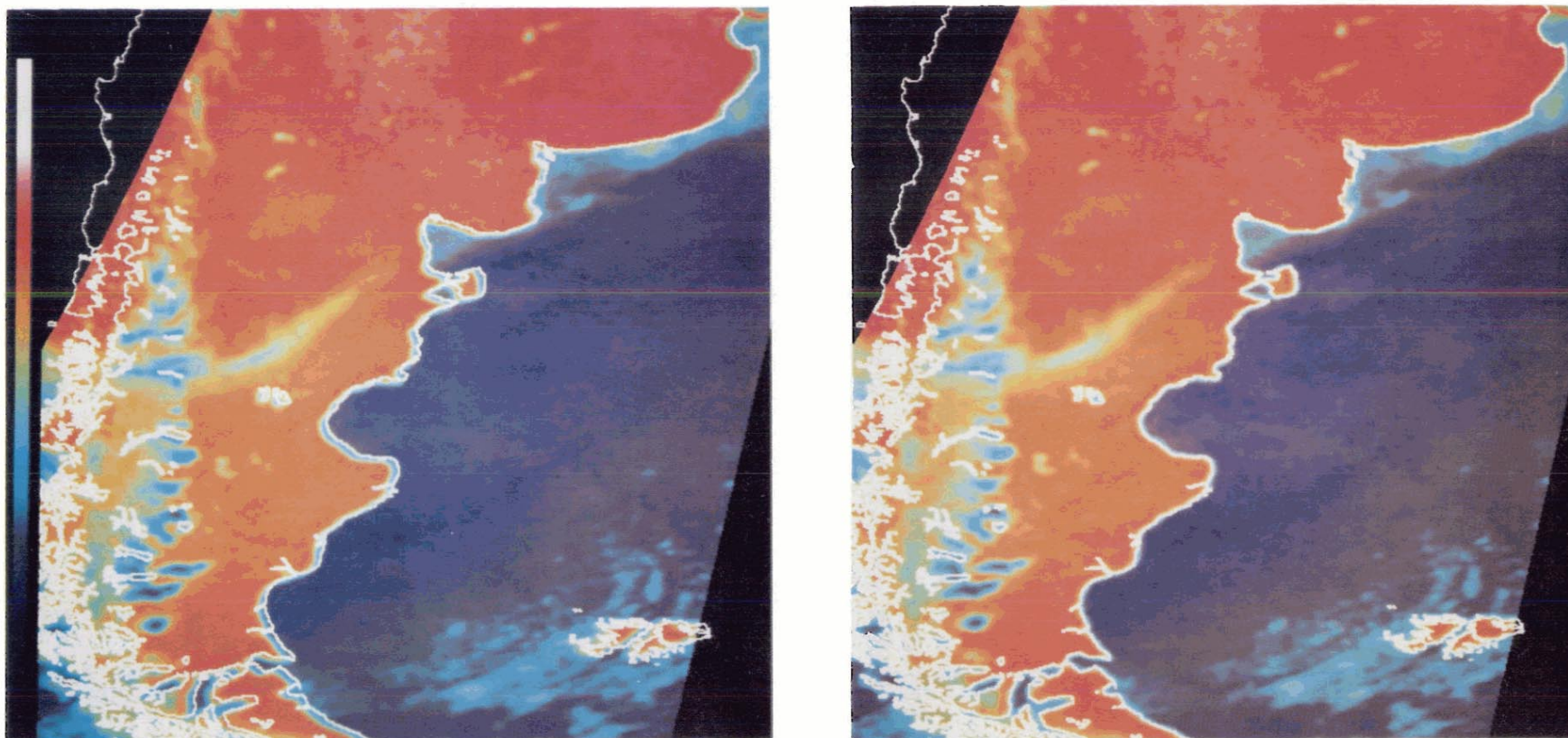
Figure A.19 85H Image of South America Obtained From Revolution 11155 on 17 August 1989. With No Correction on the Left and With the Spin Axis Correction of Pitch = -0.21 deg, Roll = -0.10 deg, and Yaw = 0.70 deg on the Right.

A.5   REFERENCES

[1]   G. A. Poe and R. W. Conway, "A Study of the Geolocation Errors of the Special Sensor Microwave/Imager (SSM/I)," IEEE Trans. Geosci. Remote Sensing, vol. 28, no. 5, pp. 791-799, 1990.

[2]   G. A. Poe, "Optimum Interpolation of Imaging Microwave Radiometer Data," IEEE Trans. Geosci. Remote Sensing, vol. 28, no. 5, pp. 800-810, 1990.