REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	10 Sep 95	

4. TITLE AND SUBTITLE
Developmetn and Wvaluation of Over-Land Rain Rate Algorithms For the SSM/I

5. FUNDING NUMBERS

6. AUTHOR(S)
Mark David Conner

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
AFIT Students Attending:
Purdue University

8. PERFORMING ORGANIZATION REPORT NUMBER
95-091

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
DEPARTMENT OF THE AIR FORCE
AFIT/CI
2950 P STREET, BLDG 125
WRIGHT-PATTERSON AFB OH 45433-7765

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for Public Release IAW AFR 190-1
Distribution Unlimited
BRIAN D. GAUTHIER, MSgt, USAF
Chief of Administration

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

14. SUBJECT TERMS	15. NUMBER OF PAGES 67
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT

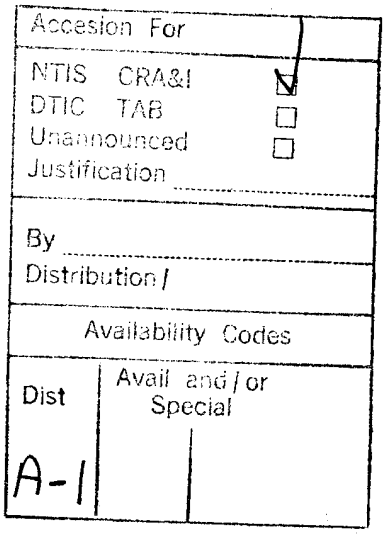
DEVELOPMENT AND EVALUATION OF OVER-LAND RAIN RATE ALGORITHMS FOR THE SSM/I

A Thesis

Submitted to the Faculty

of

Purdue University



by

Mark David Conner

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science

August 1995

Acknowledgements

First and foremost, I'd like to thank Dr. Grant Petty for the opportunity to study under his supervision these past two years. He has taken an area of meteorology, microwave remote sensing, which I had not previously considered pursuing and made it into an exciting field of study. From the beginning he has let me set the direction I wanted to go and let me proceed at my own pace. His availability to answer questions, even when up against deadlines of his own, was often used and greatly appreciated. He has taught me to think more critically when pursuing a scientific problem rather than accept the first convenient explanation that seems to fit the data at hand. I've learned a lot from him and look forward to applying this knowledge in my next assignments.

Few people are offered the chance to pursue a M.S. degree full-time while still on a regular salary, and I'm deeply grateful to the Air Force Institute of Technology's Civilian Institution program for the chance to do so here at Purdue. The AFIT meteorology program is one of the lesser-known and most valuable benefits of serving as an Air Force officer, and I hope many more are offered this possibility.

Next are thanks to the other students in our research group, Doug Miller, Nitin Gautam, and Dave Stettner. They answered my frequent substantive questions, and commiserated with me over the more rhetorical ones. Doug's morning newspaper helped me stay in touch with the world outside the fourth floor in Civil, though it meant he had to tolerate my unflattering comments about some of the paper's stories.

The rest of the group on the fourth floor were also invaluable resources, both in terms of my research and for outside activities. While not very much learning occurs on Friday and Saturday evenings, the stress reduction is just as important as the lab and classroom

time spent during the rest of the week. The camraderie felt here was just as strong (and just as important) as that I've experienced in my other Air Force assignments.

A special note of thanks to:

- Jon Schrage, who modified a NCAR graphics front end package so I could produce the HSS plots, thus saving me a considerable amount of time that would otherwise have been spent on a steep learning curve. He also got a load of questions on procedural matters related to thesis format and approval, and in general helped relieve a lot of the stress and uncertainty I was experiencing as I was writing all this.
- Dan Vietor, whose expertise in UNIX and his WXP programs kept me from having to reinvent the wheel and made the graphics for viewing the data a whole lot easier to produce.
- The other members of my committee, Drs. Phillip Smith and Dayton Vincent, whose comments helped clarify this writing for those less familiar with the microwave remote sensing field.
- Dr. John Snow, formerly at Purdue and now at the University of Oklahoma, for introducing me to research work as his assistant during my undergraduate days, and for shepherding my graduate school application while I was still in Germany.
- Drs. George Huffman and Ralph Ferraro, who provided the software used to implement their algorithms.

Finally, to my parents, Howard and Janice, and the rest of my family for their love and support. They have given me a most solid foundation upon which to build.

The Air Force Institute of Technology funded my stay here at Purdue. Funding for SSM/I datasets, administrative support, etc., came in part from NASA Grant NAGW-3944.

Table of Contents

List of Tables

List of Figures

Abstract

Conner, Mark David. M.S., Purdue University, August 1995. Development and Evaluation of Over-Land Rain Rate Algorithms for the SSM/I. Major Professor: Grant Petty.

Much work has been done in the past several years on rain-rate algorithms using the Special Sensor Microwave/Images (SSM/I) for over-water regions. However, many users of this sensor, including the Department of Defense (DOD), require rain/no-rain determinations and rain-rate estimates in over-land areas of the world not adequately covered by the present surface, upper-air, and radar observation network. Presented here is the development of three new over-land rain-rate algorithms and an evaluation of them against algorithms developed at the National Oceanic and Atmospheric Administration Satellite Research Laboratory (NOAA SRL) and at the National Aeronautics and Space Administration Goddard Space Flight Center (NASA GSFC). For ground truth, 10-cm radar data taken at six sites and hourly raingage reports from approximately 2700 locations were used. Prior to use, the radar data were compared to the gage reports to reject radar data likely to contain false echoes and to reduce site-to-site differences in how the radars observe rainfall.

The Heidke skill score (HSS) is introduced as an alternate method to determine a "best-fit" line to a set of data pairs. Least-squares linear regression, normally used for this application, requires some assumptions about the error distribution that cannot be made here. The HSS method produces reasonable results, while linear regression does not.

1. Introduction

One of the primary challenges to operational weather forecasting is a lack of reliable *in situ* meteorological data in some areas of the world. Military weather forecasters in particular must make predictions for data-sparse areas on short notice (e.g., for southwest Asia in 1990-91, northeast Africa in 1992-93, and central Africa in 1994). Remote sensing of atmospheric variables can fill in the gaps between conventional surface, upper-air, and radar observations and give the forecaster a more complete picture of the present weather conditions. Atmospheric modelers can also benefit from increased spatial resolution of data, allowing them to more reliably depict sub-synoptic-scale features.

Precipitation occurring at or near the surface has significant effects on military operations. Thunderstorms can cause aircraft to modify or abort their missions. Even light rain will degrade the accuracy and range of precision-guided weapons, both air-to-surface and surface-to-surface. Prolonged rain will affect how well vehicles can traverse an area in the absence of roads. For the military forecaster, the ability to remotely sense precipitation areas reliably has a large positive impact on the precision and accuracy of the forecast, and in turn increases the probability that the mission in question can be successfully completed.

Petty (1995) presents an overview of the current status of satellite-derived rainfall estimation over land. There are two major sensor types used: visible/infrared (VIS/IR) sensors and microwave sensors. A well-known example of the former is the Geostationary Operational Environmental Satellite (GOES) Precipitation Index (GPI), described by Arkin and Meisner (1987). The GPI algorithm uses IR cloud-top temperatures to estimate rainfall, a very indirect retrieval method. The GPI appears to reproduce climatological patterns over the tropics and subtropics when averaged over a sufficiently large temporal and spatial scale.

Microwave instruments, on the other hand, offer a method that is more physically based. Using properly-chosen frequencies, the instruments respond to precipitation-size water and ice particles, while remaining largely insensitive to non-precipitating clouds. Microwave imagers flown aboard spacecraft in the late 1970s and early 1980s, such as the Electrically Scanned Microwave Radiometer (ESMR) and the Scanning Multichannel Microwave Radiometer (SMMR), had channels between 6.6 and 37 GHz. These frequencies were sufficient to distinguish the thermal emission of rain from the radiometrically "cold" and highly polarized ocean background. Efforts to use the 37 GHz channels over land areas (Weinman and Guetter, 1977; Spencer et al., 1983; Spencer 1986) met with partial success, mainly in cases of heavier convective rainfall, for which brightness temperature depressions due to scattering by large ice particles are detectable against the strongly emitting land background.

The Special Sensor Microwave/Imager (SSM/I) was the first spaceborne microwave imager to include the 85.5 GHz frequency, making it possible to more reliably distinguish rainfall over land, owing to the increased sensitivity of higher frequencies to the presence of frozen precipitation aloft. Further, the increased spatial resolution and sampling interval (12.5 km for the 85.5 GHz channels, 25 km for the lower frequencies) improves the detection and delineation of mesoscale features, something not possible with the earlier imagers.

Microwave retrieval techniques are further subdivided into two major categories: physical inversion and empirical/statistical techniques. The former has the attraction of a rigorous theoretical treatment of the retrieval problem; however, such techniques have disadvantages such as (a) being far more complex and computationally expensive than the other techniques, and (b) many degrees of freedom are introduced into the possible solutions, requiring numerous [and sometimes arbitrary] constraints. The Kummerow

algorithm (Kummerow et al., 1989; Kummerow and Giglio, 1994a, b) and the Mugnai/Smith algorithm (Mugnai and Smith, 1988; Mugnai et al., 1993) are of this type.

Empirical/statistical techniques do not attempt to explicitly model all factors that affect microwave propagation. In general, these techniques take advantage of the fact that precipitation-sized ice particles (and, to a lesser extent, large raindrops) depress the 85 GHz brightness temperatures by reducing the emissivity of the cloud. Many algorithms such as the NOAA SRL (Grody, 1991, Wilheit et. al., 1994, Weng et al., 1994) employ thresholding techniques for the lower-frequency channels to screen out surface scatterers (snow cover, desert sand, etc.) from scattering associated with precipitation. The major drawback to all these techniques is the dependence on the brightness temperature depression at 85 GHz. Surface snow cover and ice particles aloft both cause this depression, and are difficult to discriminate when the earth's surface is relatively cold. Most algorithms flag the snow cover based on brightness temperatures of the lower frequencies and either assign a zero rain rate or do not attempt a retrieval. If relatively few significant ice crystals are produced, such as with orographic and shallow or warm convective precipitation, the algorithms will seriously underestimate the rain rate since there is little or no 85 GHz scattering.

In this work, three experimental empirical/statistical algorithms for rain rate retrieval using the SSM/I are evaluated against two algorithms developed at other laboratories. The principal feature distinguishing these new algorithms from the others is that the precipitation signal is detected by way of deviations of the observed brightness temperatures from the monthly average observed at a given location, with the objective of improved detection of light rain. Chapter 2 details the data sources used in this study and the processing done to prepare the data for intercomparisons. Chapter 3 describes the algorithms used for the comparisons, and Chapter 4 outlines the results of the comparisons. Chapter 5 gives a summary of the findings and Chapter 6 suggests areas for future work.

2. Data Sources and Processing

In order to compare the three data sources with differing resolutions, a common earth-referenced grid was used. The grid is a simple latitude-longitude grid over the conterminous United States (CONUS), extending from 60°W to 130°W and from 25°N to 50°N. The resolution is 0.25° in latitude and 0.33° in longitude, giving 21000 grid boxes. The size of each grid box varies somewhat with latitude and is 29 km by 28 km near the center of the grid. The box size was chosen to be approximately the same as the sample size of the low-resolution SSM/I channels. Two grids per day were produced: one containing data relating to the morning passes of the satellite, and one for the evening passes.

While radar and satellite estimates are instantaneous rain rate retrievals, the gages record an accumulation over a finite time period (in this study, one hour). Therefore, rain rates determined from gage data are necessarily time-averaged. While this is different than the other two methods, it will be shown later that gages can be used to validate satellite rain rate algorithms if radar data are not available.

All data contain latitude/longitude information to a resolution of 0.01 degrees, sufficient for accurate gridding. The navigation information in the SSM/I data may contain errors of up to approximately 10 km, and no attempt was made to correct for these errors prior to gridding. For the SSM/I and radar data, the center point of the pixel determined the grid box in which it was placed. A count was maintained of the number of values placed in the grid box and then an average taken once all data were entered for the time period in question.

a. Special Sensor Microwave/Imager (SSM/I) data

Microwave radiation ($\lambda \sim 1$ cm) interacts with the atmosphere in a different manner than the more familiar visible and infrared wavelengths. Microwaves are minimally scattered by the non-precipitating atmosphere and are only weakly attenuated by thin clouds such as cirrus. There are major absorption and emission bands in the microwave region, but the SSM/I channels are chosen so that they are in the spectral windows, wavelengths where the atmosphere is relatively transparent and the surface is not obscured from the satellite. One channel, however, was chosen to take advantage of the 22.235 GHz water vapor resonance line.

The SSM/I is flown aboard the Defense Meteorological Satellite Program's (DMSP) Block 5D spacecraft. The satellites are flown in sun-synchronous orbits with an inclination of 98.8°. A comparison of three of the DMSP satellites is shown in Table 1. This study utilized data collected by the F-11 satellite, whose sensor was fully functional throughout the study period.

	F-8	F-10	F-11
Launch date	June 1987	December 1990	November 1991
Altitude range (km)	830-882	740-853	841-876
Period (minutes)	101.8	100.7	101.9
Ascending equatorial crossing time (local)	0615	1942*	1704

*The F-10 did not achieve its desired orbit, and the crossing time increases by 45 minutes per year. Time shown is as of mid-January 1991.

Table 1. Key parameters for F-8, F-10, and F-11 satellites.

Seven channels are used by the SSM/I for sensing thermal emission at four different frequencies. Three frequencies are measured in horizontal and vertical polarization (19.35, 37.0, and 85.5 GHz), while the remaining frequency (22.235 GHz) is measured only in the vertical. For convenience, these channels will be referred to as 19V, 19H, 22V, 37V, 37H, 85V, and 85H.

The SSM/I sensor rotates about the satellite's vertical axis with the antenna maintaining a constant viewing angle at the earth's surface of 53.1°. Although the SSM/I rotates 360^0, only a 102° arc (fore or aft, depending on the satellite) centered on the satellite's subtrack is used. In this arc, the 85 GHz channels are sampled 128 times in the cross-track direction, while the lower-frequency channels are sampled 64 times. The 85 GHz channels are also sampled twice as frequently in the along-track direction. This gives the 85 GHz channels a sample interval of 12.5 km and the others 25 km in both the cross-track and along-track directions. The total swath width is about 1400 km.

The sun-synchronous orbit means a single satellite will be present over a given area only twice a day and at approximately the same local time each day. While this is not a problem for instantaneous rain rates, it could result in a biased result for climatological applications if there is a systematic diurnal component to rainfall occurrences. Furthermore, successive SSM/I swaths are not overlapping equatorward of 57° latitude. Therefore, a fixed point on the earth may not be sampled for a considerable period of time. Both problems can be reduced by using multiple satellites with different ascending times.

For this study, SSM/I data were obtained on Exabyte® tape format from Remote Sensing Systems, Inc. These tapes contain reformatted data from the Temperature Data Records produced by the Fleet Numerical Meteorology and Oceanography Center in Monterey, California. The data for each scan contain antenna temperatures, location, time, surface type, and calibration information. Wentz (1988) gives a complete descrip-

tion of the data contained on the tapes, and Hollinger et al. (1987) has a thorough discussion of the SSM/I instrument.

The data passed through a two-stage quality control check prior to gridding. First, if a scan contained any pixels with unphysical brightness temperatures, the entire scan was set to "missing". Second, if individual pixels were determined to be over water by

$$22V - 19V > 4.0\ K \qquad (1)$$

then the individual pixel data at that location were set to missing. Next, synthetic high-resolution scans were created for the low-resolution channels by interpolating from valid neighboring pixels. Finally, the pixels that passed all the checks were fitted to the common grid. Since the NASA GSFC algorithm had its own quality control checks built in, the original data were passed to that algorithm, and the rain rates returned by that routine were then gridded. Figure 1 shows the typical swath coverage over the conterminous United States (CONUS) for a 12-h period.

It was desired to have a background grid to compare the twice-daily grids against. This background grid should contain the "normal" brightness temperatures that the satellite would see under no-rain conditions at the same time of day, so that short-term (and probably meteorological) changes are emphasized. One grid was generated from the morning satellite passes and one from the evening passes. The background grids were calculated for each calendar month using the SSM/I data for that entire month. The same processing was done as for the twice-daily grids, except the pixel-by-pixel test rejected pixels failing (1) or

$$37V - 85V > 5.0\ K\ . \qquad (2)$$

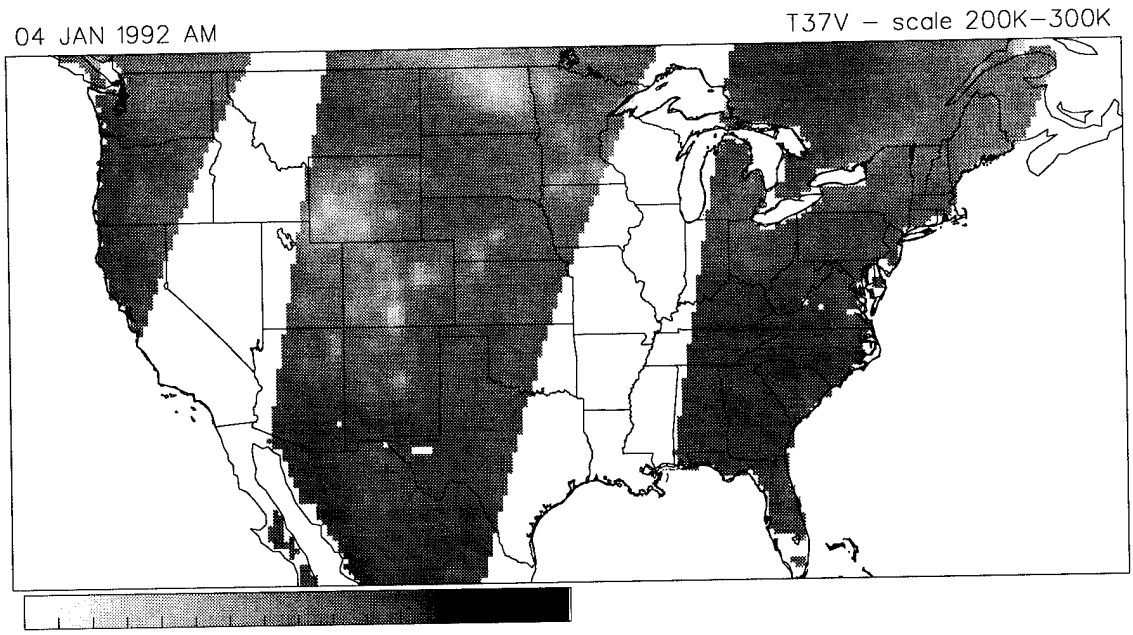


Figure 1. SSM/I data coverage for the morning overpasses on 4 January 1992. White areas indicate water was detected or no coverage.

Formula (1) rejected over-water pixels, while (2) rejected pixels where 85V was depressed enough to indicate possible precipitation. This also removed pixels where the ground was snow-covered as well, causing some locations in the winter season to have no pixels in a month-long period passing all the criteria. Figure 2 shows the morning and evening monthly averages from January and July 1992. Note the "missing" data over the northern portion of the grid and in the Rocky Mountains in the January averages, indicating persistent snow cover.

b. Raingage data

The raingage data were obtained from the National Climatic Data Center (NCDC) at Asheville, North Carolina. The data set is a compilation of hourly rainfall totals observed

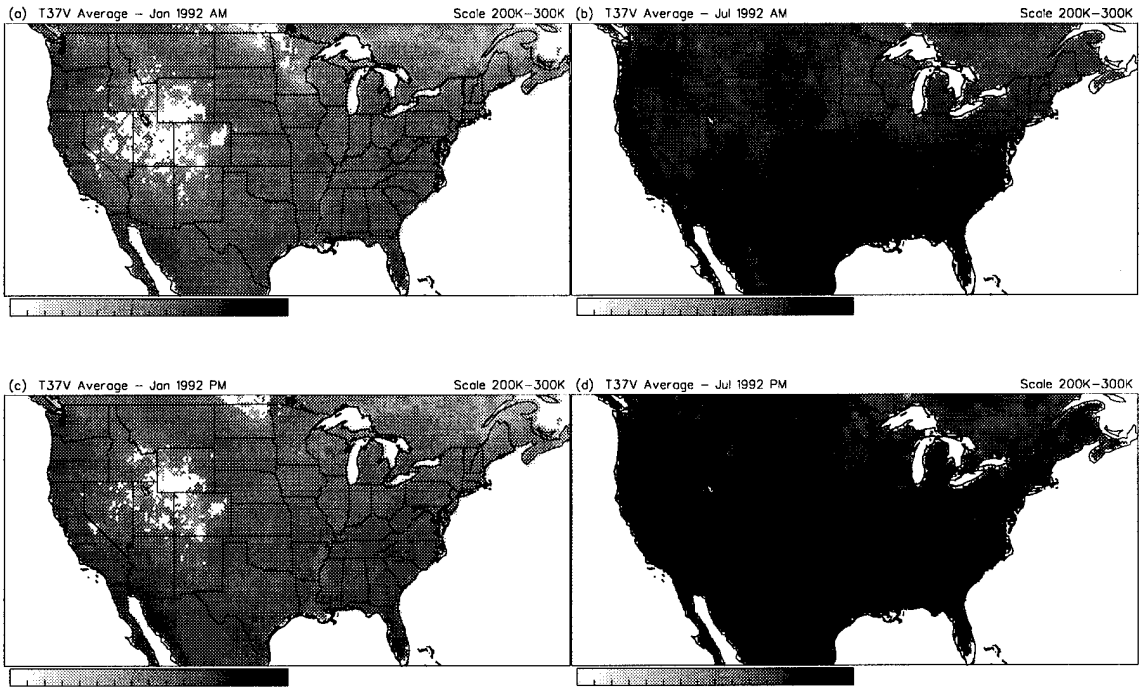


Figure 2. Monthly 37V averages (1992) for (a) January morning passes, (b) July morning passes, (c) January evening passes, and (d) July evening passes.

at approximately 2700 sites in CONUS. The dataset is available at no charge via the Internet.

There are two major types of raingages in the network. The first is the standard tipping bucket raingage used at most National Weather Service (NWS) first-order weather stations. This gage has a resolution of 0.254 mm (0.01 inch). For this gage type, the rainfall that occurred during a particular hour is recorded.

The second type is the Fisher-Porter gage. Instead of a tipping bucket, it has a weighing gage that is calibrated to punch a recording tape after 2.54 mm (0.1 inch) of rainfall has accumulated. When the dataset is compiled, the number of punches occurring during the hour is converted to a rainfall amount and entered. Due to the coarseness of the gage's resolution, it is possible for light rain to have occurred for a number of hours,

yet the gage's record will show only one 0.1-inch event after 0.1 inches of new rainfall is measured.

The raingage type is not reported in this dataset. However, it was possible to examine the entire period of record for each gage to see if any amounts besides zero or 0.1-inch increments were ever reported. If not, then the gage was determined to be of the Fisher-Porter type. Fisher-Porter gages comprised 83 percent of the total network, which is shown in Fig. 3.

The data were then fitted to the common grid. Of the grid boxes with gages in them, about 85 percent had only one gage, with the remaining boxes having up to a maximum of six. Multiple reports in a box were averaged. To examine the possible effects of the differing gage resolution of rainfall amount, two grids were created: one containing all gage reports, and the second containing only reports from non-Fisher-Porter gages.

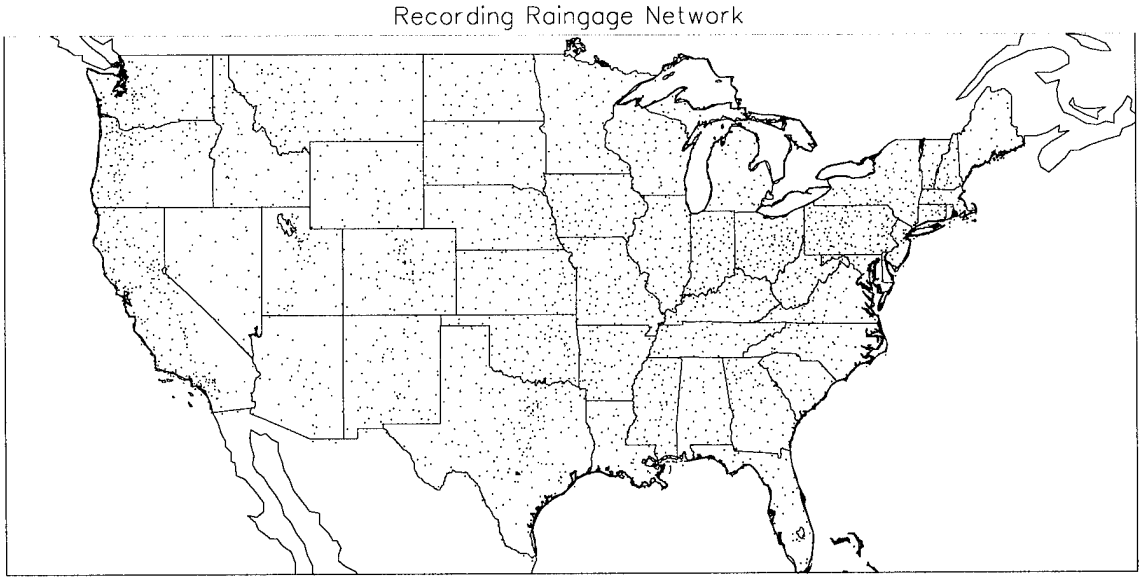


Figure 3. Recording raingages in the conterminous United States.

c. Radar data

The radar data used for this study came from the RADAP-II (Radar Data Processor version II) archive, which is also located at NCDC. The data were obtained through the US Air Force Environmental Technical and Application Center (USAFETAC) Operating Location A, which is the Air Weather Service liaison at NCDC. McDonald and Saffle (1994) cover the archiving and formatting process in detail. The RADAP-II archive project ended 30 September 1992.

The RADAP-II archive is digitized radar reflectivity data from 12 sites around the United States. One site ceased archiving prior to 1992, and five of the sites are located in or near mountainous areas, with significant ground clutter and mountain shadowing problems. The remaining six sites [Tampa Bay, FL (TBW); Nashville, TN (BNA); Monett, MO (UMN); Wichita, KS (ICT); Oklahoma City, OK (OKC); and Amarillo, TX (AMA)] were relatively free of persistent ground clutter that could affect the results.

The RADAP-II network (Fig. 4) contained two types of radars, the WSR-57 and the WSR-74C. Both have a 2.2° beam width and a 10 cm wavelength. The archive contains both base-level and tilt-sequence scans, but for this study only the base-level scans were used. Each scan is built from 180 radials of 2° width centered on even-numbered azimuths, covering the entire 360° field. Each radial was divided into 1.85 km (1 nautical mile) bins from 18.5 to 231 km (10 to 125 n.m.). Observations were taken every 10 or 12 minutes.

The reflectivity was coded as a value from 0 to 15, with each value corresponding to an entry in a lookup table included in each observation for conversion to dBZ (Table 2). Note that reflectivities less than the first threshold were coded as zero, or no precipitation. This could result in underrepresentation of very light precipitation. During the conversion from RADAP-II category to rain rate, the category value was converted to the

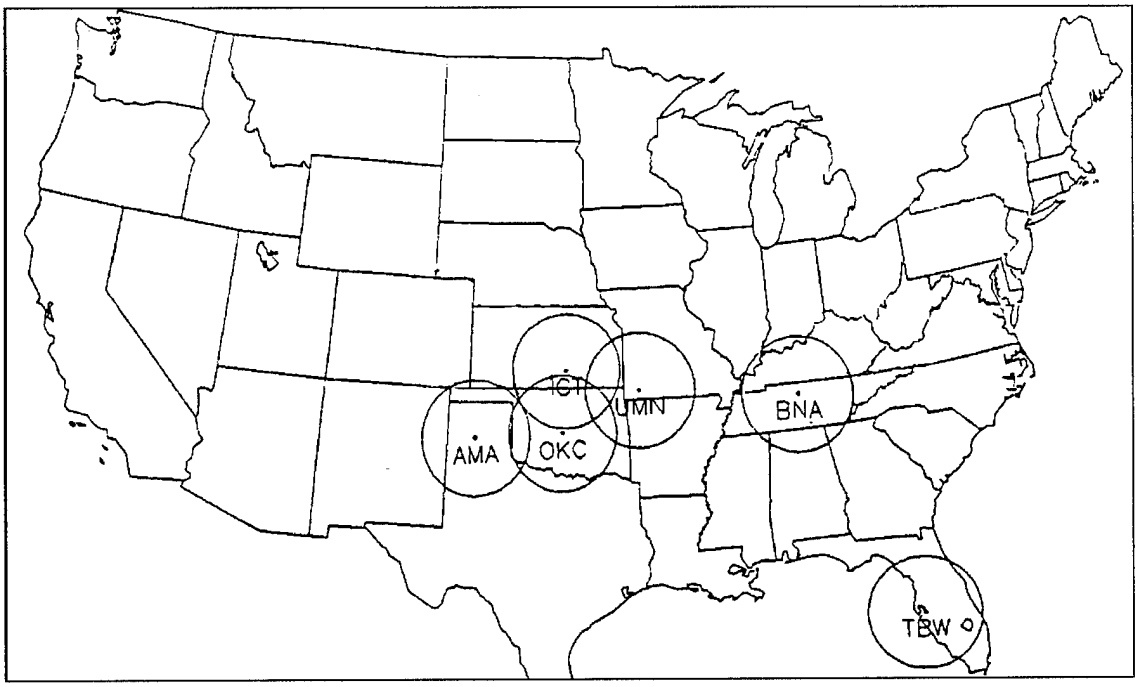


Figure 4. RADAP-II stations used in this study. Circles indicate 231 km (125 n.m.) surveillance region.

	RADAP-II category														
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Cool															
dBZ	18	20	22	24	26	28	30	32	34	36	38	40	42	44	46
Rain rate	0.49	0.65	0.86	1.15	1.54	2.05	2.73	3.65	4.86	6.48	8.65	11.5	15.4	20.5	27.3
Warm															
dBZ	18	24	30	35	38	41	43	44	46	47	49	51	53	55	57
Rain rate	0.49	1.15	2.73	5.61	8.65	13.3	17.8	20.5	27.3	31.6	42.1	56.2	74.9	99.9	133

Table 2: Typical RADAP-II reflectivity thresholds (dBZ) and associated rain rates (mm h^{-1}) for warm and cool seasons.

threshold dBZ value and then to a rain rate using the standard Z-R relationship for a Marshall-Palmer raindrop distribution (from Burgess and Ray, 1986):

$$R\ (\text{mm hr}^{-1}) = \left[\frac{10^{dBZ/10}}{200}\right]^{0.625} . \qquad (3)$$

Generally, it was desired to convert the azimuth-range format of the radar data into some kind of Cartesian coordinate system for display and analysis purposes. Since the range gates were 1 nautical mile in length, a natural choice was a latitude-longitude grid with squares 1/60 of a degree on a side. To grid the data, each azimuth was "tweaked" from plus to minus 1° in 0.25° increments from the nominal value; then, a conversion from azimuth/range to latitude/longitude was performed and the reported rain rate value was added to a sum for that grid box. After all azimuths had been processed, the grid boxes were averaged by the count for each box, and a 3x3 interpolation was done for any remaining missing values in the valid area.

To examine the possibility of persistent bias, either as a function strictly of range from the radar site or of position relative to the radar, average rain rates and counts of pixels at each threshold level were created for each radar site. If the occurrence and intensity of precipitation are assumed to be randomly distributed across the radar's surveillance region, and if the radar accurately senses the true rain rate, the average rain rate field should be relatively uniform and the pixel counts should not show a range dependence. Figure 5 shows these results for the Nashville, Tennessee (BNA) radar site, but similar results were observed for the other five sites in the study.

One problem is the annular structure visible in the average rain rate fields, which is also manifested by an oscillation with range in the pixel counts at each threshold, as shown in Fig. 6. It was hypothesized that the dBZ values produced by the radar were inadvertently modified by a range-dependent square wave function before being

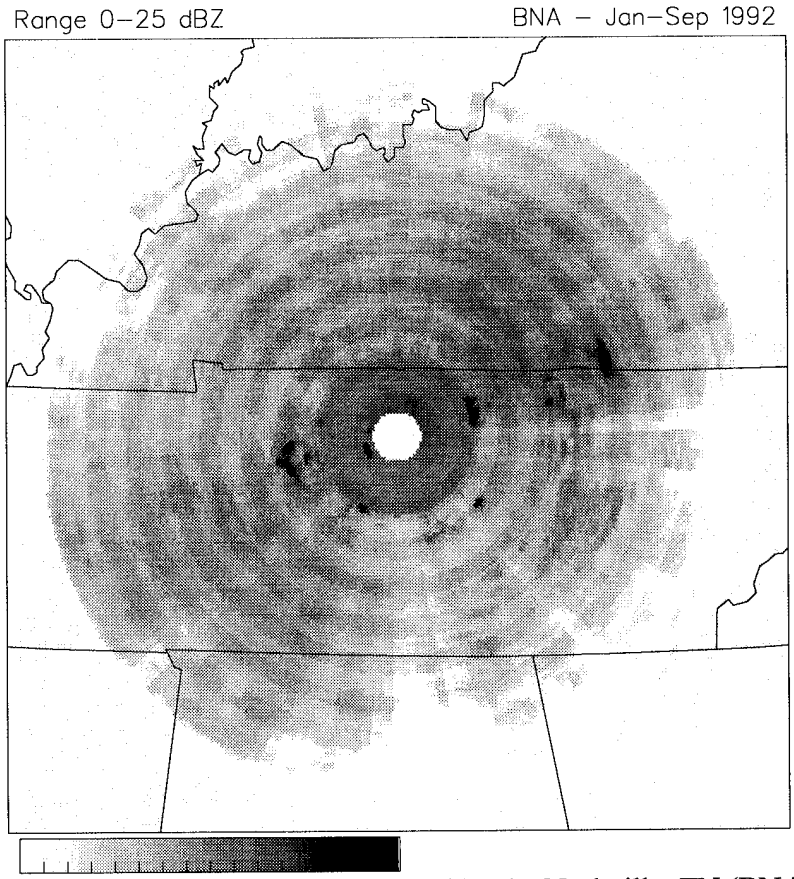


Figure 5. Average rain rate observed by the Nashville, TN (BNA) radar during January-September 1992. Light gray near the center and on the periphery indicates missing data.

thresholded by the RADAP-II processor and stored. To test this, smoothed pixel count curves were constructed by processing the counts for each level through a 13-n.m.-wide centered moving average (the approximate wavelength of the oscillation) and then compared to the raw counts. Figure 7 shows the method to estimate the dBZ change necessary in the RADAP-II input to make the raw counts conform to the smoothed counts, and Figure 8 shows the dBZ changes computed for each threshold level. Even though the threshold dBZ increments are uniform for the cool season, the warm season had a better distribution of rainfall intensities and a longer period of record: the problems with non-uniformity were not considered serious enough to offset these advantages.

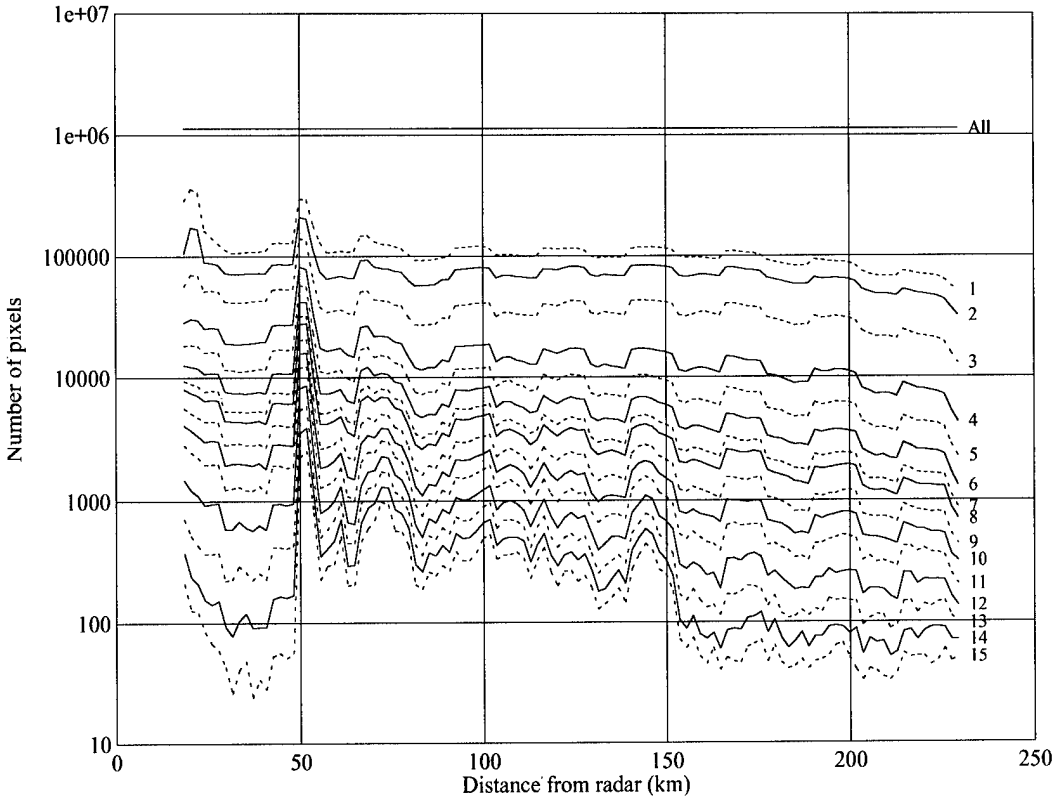


Figure 6. Cumulative count of number of pixels equal to or exceeding RADAP-II categories for the "warm season" scans at BNA during January-September 1992.

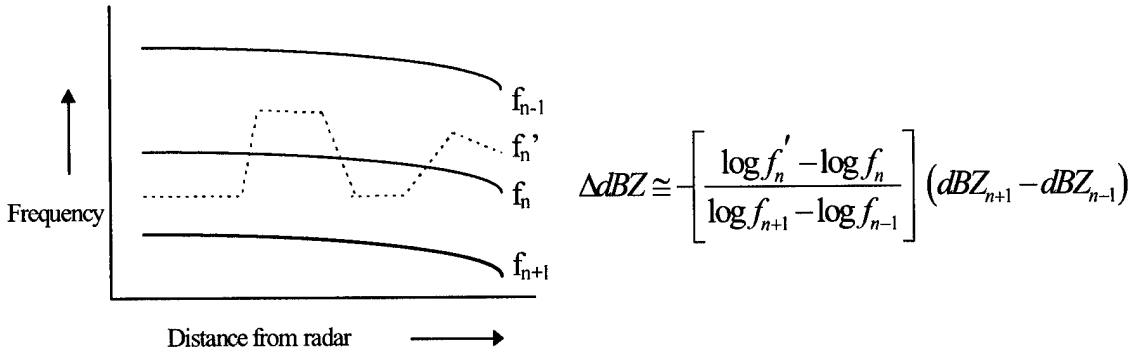


Figure 7: Estimating dBZ correction to remove square wave.

Figure 8 shows a fairly uniform dBZ correction can be applied to all thresholds. The highest four thresholds show more variability, but that is likely due to small sample size (less than 1 percent of all pixels are in the highest four thresholds). While the graphs indicate about a 1 dBZ correction would be the most appropriate, empirical adjustments and similar analysis for other radar sites showed that a 0.75 dBZ (or about 18 percent) correction was appropriate for all sites. Figure 9 shows a composite similar to that in Fig. 5, except this correction has been applied.

Another problem, illustrated best in Fig. 9, is the discontinuity in the average rain rate about 50 km from the radar. To mitigate the effects of ground clutter, the RADAP-II

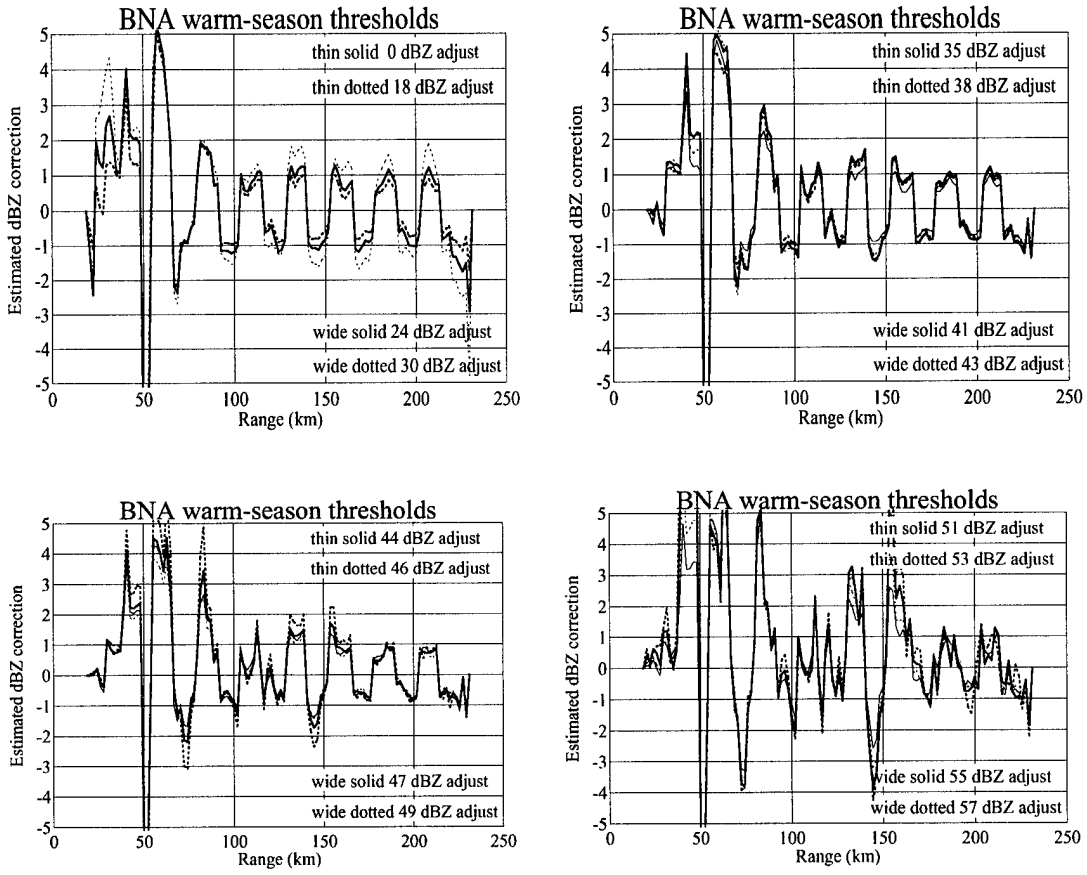


Figure 8. Estimated correction (dBZ) for all thresholds to remove anomalous square wave structure (BNA warm-season scans January-September 1992).

processor created a hybrid base-level scan. This scan consisted of data taken from higher antenna elevations within a certain site-dependent range (typically 40-60 km), and using the 0.5° scan for more distant range gates. The composite shows this causes an unphysical "jump" in the rain rates across that range threshold. To eliminate this, only data from range gates where the antenna elevation was 0.5° were used in subsequent comparisons.

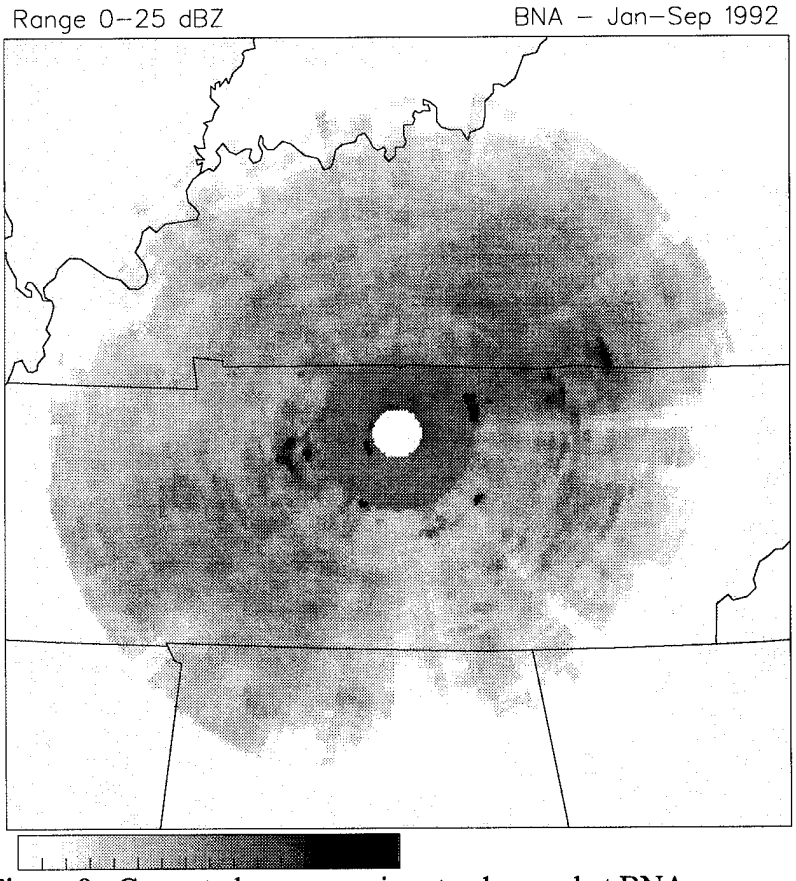


Figure 9. Corrected average rain rate observed at BNA.

3. Over-land rain rate algorithms using SSM/I

a. Purdue-EV (EV) algorithm

Eigenvectors, also known as empirical orthogonal functions (EOFs) or principal components (PCs), provide a means of explaining a covariance matrix through a few linear combinations of the original variables. In this study, the seven-channel output of the SSM/I at a particular location can be thought of as a vector $\vec{\mathbf{T}}$ containing seven elements, and for repeated observations a 7x7 covariance matrix $\mathbf{S}_T$ may be calculated. Principal components analysis can then be applied to this matrix, resulting in 7 eigenvectors $\hat{\mathbf{e}}_i$ and 7 eigenvalues $\hat{\lambda}_i$. The eigenvalues indicate the amount of variance explained by the associated eigenvector, and the sum of the eigenvalues is the total sample variance. The eigenvectors are all orthogonal to, and therefore linearly independent of, each other. Additionally, the loadings of the eigenvectors are forced to be uncorrelated with one another within the dataset from which the covariance matrix was generated. The first eigenvector (having the largest eigenvalue) lies along the axis of maximum variability in *n*-dimensional space (in this case, *n*=7). The second eigenvector lies along the axis of maximum variability that is *orthogonal* to the first. The third is orthogonal to the first two, and so on. Eigenvector directions are arbitrary, and the vector may be multiplied by $-\vec{\mathbf{1}}$ if desired. A more detailed discussion of PC analysis can be found in Johnson and Wichern (1992, pp. 356-395).

To isolate the precipitation signal from other influences, the morning and evening monthly averages for the SSM/I overpasses, as mentioned in Chapter 2, were used. For each valid grid box and all twelve months of 1992, the morning average was subtracted from the evening average, the covariance matrix computed, and then the eigenvectors. Since precipitation screening was performed in the averaging process, the first eigenvector

Channel	19V	19H	22V	37V	37H	85V	85H
$\hat{\mathbf{e}}_1$	0.3657	0.4205	0.3377	0.3910	0.4384	0.3248	0.3532

Variance explained by this eigenvector: 89%

Table 3. Eigenvector corresponding to temperature signal.

(which is aligned with the axis of maximum variability) should be representative of the temperature signal.

Table 3 shows the first eigenvector $\hat{\mathbf{e}}_1$ computed from this covariance matrix. The eigenvector's elements are all of similar magnitude, indicating all seven channels tend show the same difference between the morning and evening values. This is consistent with a temperature response vector, and will be denoted as $\hat{\mathbf{e}}_T$ from here on.

This temperature response vector can now be used to help isolate the precipitation signal from the background. The daily morning and evening grids are used, with the corresponding monthly averages subtracted to produce a set of 7-channel vectors $\delta\vec{\mathbf{T}}$ containing the departures from the average. These differences are then corrected for temperature effects by using

$$\delta\vec{\mathbf{T}}' = \delta\vec{\mathbf{T}} - \left(\delta\vec{\mathbf{T}} \cdot \hat{\mathbf{e}}_T\right)\hat{\mathbf{e}}_T . \tag{4}$$

Equation (4) has the effect of setting $\delta\vec{\mathbf{T}}'$ orthogonal to $\hat{\mathbf{e}}_T$. The filtered differences $\delta\vec{\mathbf{T}}'$ are presumed to contain short-term brightness temperature variations *not* associated with surface temperature. Since these variations were relative to a monthly average, they should contain variations on a time scale shorter than that period. Longer-term changes such as vegetation type or coverage are not expected to have significant effects within the month, though they would become obvious were a yearly average used. Shorter-term

variations due to changes in snow cover or consistency, soil moisture, precipitation, and other effects should be contained in these vectors.

To find the precipitation signal, PC analysis was used again on the covariance matrix computed from the $\delta\vec{T}'$ vectors. To guard against contamination by snow cover, which has a microwave signature similar to that of precipitation, only vectors from the months of May, June, July, and August 1992 were included that were east of longitude 91°W in the grid domain. Table 4 shows the resulting precipitation vector $\hat{\mathbf{e}}_P$. This vector is consistent with a precipitation signature over land, since it favors contrasts between the 85 GHz channels (negative eigenvector elements) and the lower-frequency channels (positive eigenvector elements).

It was also desired to isolate the precipitation signal from soil wetness effects. Since soil wetness, at its extreme, would be a water surface, a soil wetness vector was simulated by taking the 7-channel differences between typical averages over land and over water. The averages selected were the global land and water averages for 40° N for the month of April 1992 as observed by the F11 satellite. This vector was then converted to a unit vector $\hat{\mathbf{e}}_S$. This vector is sensitive to the differences between the horizontal and vertical channels, as shown by the horizontal channel vector elements being higher than the vertical channel elements, and is therefore sensitive to the polarization effects of a water

Channel	19V	19H	22V	37V	37H	85V	85H
$\hat{\mathbf{e}}_P$	0.2320	0.2010	0.2495	0.2112	0.1841	-0.6096	-0.6272
$\hat{\mathbf{e}}_S$	0.3851	0.6295	0.2907	0.2673	0.4903	0.0886	0.2264
$\hat{\mathbf{e}}'_S$	0.1221	0.6553	-0.0664	-0.2678	0.2253	-0.5904	-0.2834
$\hat{\mathbf{e}}'_P$	0.2020	-0.3156	0.3938	0.5222	0.0457	-0.2901	-0.5878

Table 4. Precipitation and soil moisture vectors.

surface. However, because all the elements are positive, it is also somewhat sensitive to temperature. To reduce this, the $\hat{\mathbf{e}}_S$ is then set orthogonal to the temperature effects vector $\hat{\mathbf{e}}_T$ by using (4), becoming $\hat{\mathbf{e}}'_S$. This vector is now far less sensitive to temperature changes, since a joint change in all seven channels (surface temperature change) produces far less of a change in $\hat{\mathbf{e}}'_S$ than it does for $\hat{\mathbf{e}}_S$. Finally, $\hat{\mathbf{e}}_P$ was set orthogonal to $\hat{\mathbf{e}}'_S$, becoming $\hat{\mathbf{e}}'_P$. This last vector should be as independent as possible from temperature and soil moisture effects. It is more difficult to evaluate empirically, however, because it is now orthogonal to two other vectors. The $\hat{\mathbf{e}}'_P$ vector is used with the brightness temperature differences $\delta\vec{\mathbf{T}}$ to compute a scalar "precipitation" field as

$$p_{EV} = \hat{\mathbf{e}}'_P \cdot \delta\vec{\mathbf{T}}. \tag{5}$$

Equation (5), the dot product of the corrected precipitation vector and the vector differences between the observed brightness temperatures and the monthly averages, is the uncalibrated "eigenvector algorithm" (EV) used for the comparisons in Chapter 4. These scalar values are computed for each high-resolution pixel value and then averaged and gridded in the same fashion as the brightness temperature data. The resulting values are in units of degrees Kelvin, with positive values indicating detection of a precipitation signature. These values can be multiplied by a calibration constant (to be determined later) to yield a rain rate in mm h^{-1}.

b. Purdue-2Channel (2C) algorithm

The Purdue-2Channel algorithm can be thought of as a simplified version of the precipitation eigenvector algorithm. Since the 37 GHz channels are less sensitive to precipitation effects than the 85 GHz channels, but still respond to variations in the surface

temperature, it was hypothesized that if the 85 GHz departure from average exceeded the 37 GHz departure from average, precipitation was likely. Also, vertical polarization channels are less sensitive to specular reflection of the cold sky from surface water for the viewing angle of the SSM/I, so the vertical polarization of the 85 GHz and 37 GHz channels were chosen.

In vector form, the uncalibrated algorithm is

$$p_{2C} = -\left(\vec{\mathbf{e}}_{2C} \cdot \delta\vec{\mathbf{T}}\right), \tag{6}$$

where

$\vec{\mathbf{e}}_{2C} = \{0,0,0,-1,0,1,0\}$ and

$\delta\vec{\mathbf{T}}$ is the 7-channel vector difference between the observed brightness temperatures and the monthly averages.

The negative sign in (6) is to associate positive values of p_{2C} with areas of precipitation. Units are in degrees Kelvin, as for the EV algorithm.

c. Purdue-4Channel (4C) algorithm

The Purdue-4Channel algorithm (4C) is more complex. It uses both polarizations of the 37 GHz and 85 GHz channels to minimize temperature and surface water effects on the rain rate retrieval.

Open water areas (and areas such as very moist soil that mimic open water), exhibit strong polarization differences in their emission at all SSM/I frequencies, with the emissivity higher for vertical polarization than for horizontal (Petty, 1990). Drier land surfaces have largely unpolarized emissions. Over land areas and in the absence of atmospheric effects, it is hypothesized that as the fraction of the sample area covered by wet surfaces

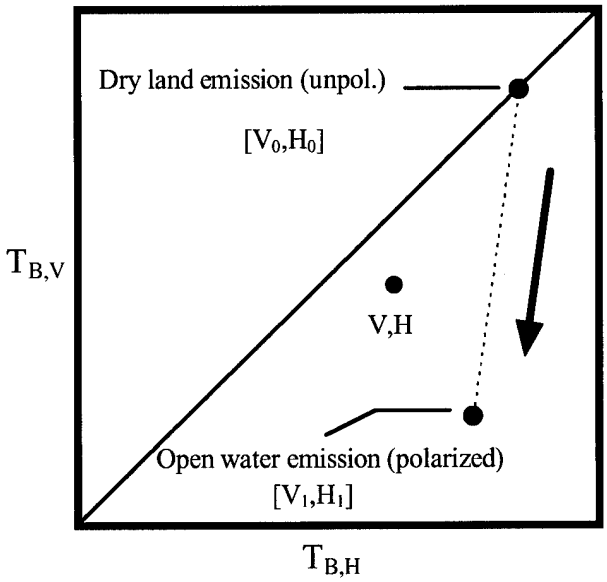


Figure 10. Microwave brightness temperature changes (arrow) as the fraction of a sampled land surface that is "wet" increases.

increases, the brightness temperatures will change in a linear fashion as indicated by the arrow in Fig. 10. This principle was first exploited by Weinman and Guetter (1977).

In the presence of wet soil and/or precipitation, the nonprecipitating-sky values indicated by (V_0, H_0) and (V_1, H_1) in Fig. 10 are not known. However, climatological values can be substituted with some success. Since precipitation-sized ice particles cause unpolarized scattering of the surface emission at 85 GHz, as precipitation increases, the 85H and 85V pair will depart from the dashed line (non-precipitating-sky value) down and left at a 45° angle. Thus, the linear distance of an arbitrary point (V,H) from the dashed line can be related to a precipitation rate and is calculated by

$$d = \left[V_1 - \frac{H_1 m}{1-m} \right] - \left[V - \frac{Hm}{1-m} \right], \tag{7}$$

where m is the slope of the dashed line in Fig. 10, or $\frac{V_0 - V_1}{H_0 - H_1}$.

A colder surface temperature, however, will also cause such a departure, though usually of a smaller scale. Dry land surfaces are nearly blackbody emitters in the microwave, so surface temperature changes will cause brightness temperature changes of approximately the same scale. Moderate to heavy precipitation, on the other hand, will depress the brightness temperatures around 10-30 K. The precipitation signal is, therefore, often stronger than the departure from climatology, but obviously that departure affects the retrieval.

The 37 GHz channels are not nearly as sensitive to precipitation-sized ice particles as the 85 GHz channels, but are affected by temperature in much the same fashion. Therefore, applying (7) above to both frequencies and taking the difference of the results should yield a measure of the precipitation rate that is fairly independent of temperature or soil wetness.

The Purdue-4Channel algorithm is then

$$p_{4C} = d_{85} - d_{37} \tag{8}$$

where d_{85} and d_{37} are (7) applied to the 85 GHz and 37 GHz frequencies, respectively. Units are degrees Kelvin.

d. NOAA SRL (SRL) algorithm

The NOAA SRL algorithm is based on a surface classification procedure developed by Grody (1991) and has been updated since then (Wilheit et. al., 1994, Weng et. al., 1994, Ferraro, 1995, personal communication). The algorithm uses a scattering index (SI) to identify the scattering signal at 85 GHz from precipitation. The index is defined as

$$\mathrm{SI} = F - 85\mathrm{V}, \tag{9}$$

where F is the non-scattering (emission) component of 85V and is estimated from 19V and 22V by

$$F = 451.9 - 0.44 \times 19\text{V} - 1.775 \times 22\text{V} - 0.00574 \times (22\text{V})^2 . \quad (10)$$

If SI is less than 10 K, a zero rain rate is assumed. To guard against snow surfaces and desert sand areas returning false positive rain rates, two further checks are performed. If

$$22\text{V} < 264 \text{ and } 22\text{V} < 175.0 + 0.49 \times 85\text{V} \quad (11)$$

or

$$85\text{V} > 253.0 \text{ and } (19\text{V}-19\text{H}) > 7.0, \quad (12)$$

then the surface is assumed to be snow or desert sand, respectively, and the rain rate is set to zero. If both tests are passed, then the rain rate p_{SRL} (mm h^{-1}) is computed from SI as

$$p_{\text{SRL}} = 0.00513 \times (\text{SI})^{1.9468} . \quad (13)$$

The values of p_{SRL} are limited to 35 mm h^{-1} to prevent spurious data from affecting the retrievals.

e. NASA GSFC (GSFC) algorithm

The NASA Goddard Space Flight Center (GSFC) algorithm is known as the Goddard Scattering Algorithm, Version 3 (GSCAT3). It is an update to the GSCAT2 algorithm discussed in Adler, et al. (1994). Huffman (1995, personal communication) provided the GSCAT3 software.

The GSFC algorithm is similar to the SRL algorithm in that it uses simple threshold checks to screen out pixels that are not likely to be experiencing precipitation, though the screening is somewhat more complex. For example, the local standard deviation of 85H in a 5x5 box around the pixel is used to pare down "ambiguous" areas (pixels near the 10 K SI computed from (9) above). When this standard deviation is high, convective cores, and hence rain, are implied. Low standard deviations coupled with a cold 22V temperature indicate a "cold" and possibly snow-covered surface.

If all the checks are passed, the rain rate is calculated pixel-by-pixel using

$$p_{\mathrm{GSFC}} = \frac{262.0 - (85\mathrm{H})}{4.188} \times r,$$

where p_{GSFC} = rain rate in mm h^{-1} and
r = correction factor. (14)

The correction factor r is based on the surface type, which is inferred from the navigation data and a surface-type database. This ratio is set to 0.8 over land areas, 1.2 over coastal areas, and 1.6 over ocean areas (Huffman, personal communication). Values of p_{GSFC} less than 1 mm h^{-1} are set to zero for land and coastal areas to reduce spurious values.

4. Comparisons of satellite, radar, and raingage derived rain rates

a. Selecting and calibrating radar data based on gages

Radar reflectivity data often contain echoes that are not associated with precipitation. The most common contamination is "ground clutter", which are echoes near the radar site that are reflections of objects on the ground surface. As described earlier, the RADAP-II dataset used a hybrid scan to attempt to eliminate ground clutter. However, this introduced a discontinuity in the mean radar reflectivity at the range where the low- and high-antenna-elevation scans were merged. For this study, all data within this range were eliminated, effectively removing nearly all ground clutter.

Anomalous propagation (AP) is a second cause of contamination. Normally, a radar beam aimed at a slight upward angle relative to the horizon will propagate through the atmosphere at higher and higher levels relative to the earth's surface. Battan (1973, pp 17-28) describes the effects of nonstandard temperature and moisture gradients on radar propagation. These gradients produce anomalous gradients in the refractive index of the atmosphere, leading to AP. This can be in the form of "subrefraction", where the radar beam is bent upwards from its normal path, or "superrefraction", where it's bent downwards back towards the earth's surface. The form of AP of most concern here is "superrefraction", since it causes false echoes.

Meteorological conditions that favor superrefraction are not normally associated with precipitation. The most common condition for superrefraction is a surface-based inversion due to subsidence or radiational cooling. Subsidence generally indicates high pressure and anticyclonic flow, and strong radiational cooling indicates a lack of cloudiness. Neither are common in areas of precipitation, yet the resulting superrefraction can cause strong echoes on a radar display.

The RADAP-II data record format has a flag the radar operator can set to indicate AP, and the data processing rejected all scans where this flag was set. Subjective review of several time series of images taken from the six radar sites used in this study indicated this flag was not always set properly, and early attempts at satellite-radar and gage-radar correlations indicated this happened often enough to significantly affect them. Figure 11 shows the dBZ value associated with the mean rain rate for one hour observed at Tampa Bay, FL (TBW) with raingage and surface weather reports for the same hour superimposed. None of the radar observations during this hour was flagged by the operator as having anomalous propagation, even though none of the surface weather

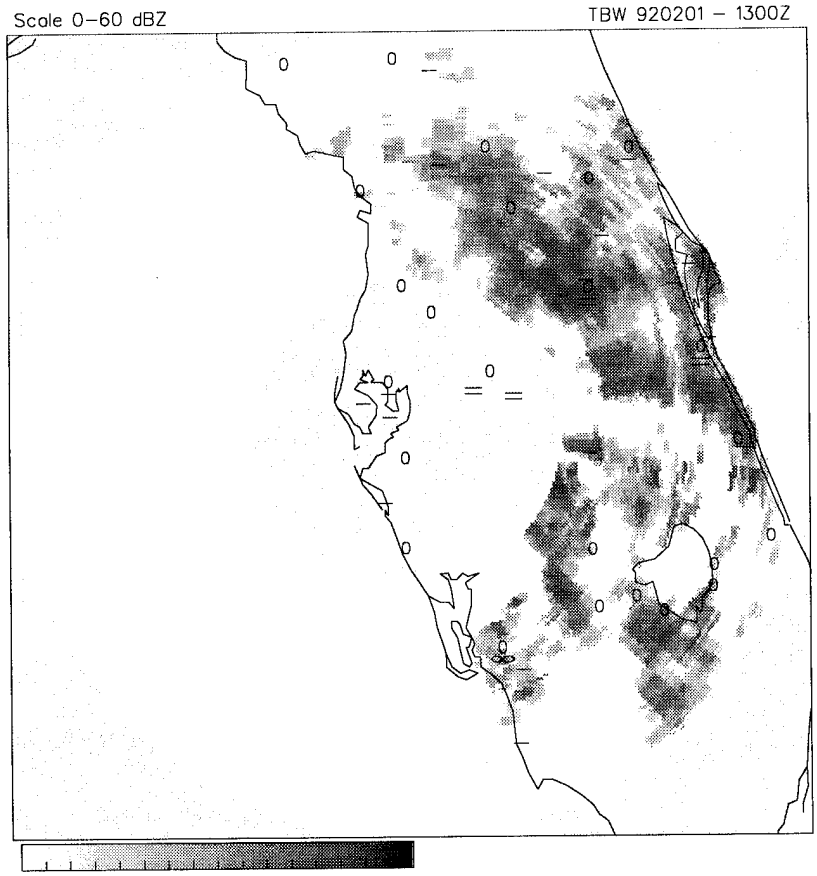


Figure 11. Reflectivity (dBZ) corresponding to mean rain rate observed at TBW for the hour ending 1300 UTC 1 February 1992. Weather symbols are surface observations (single dash indicates no weather reported) and numbers indicate gage precipitation totals (mm) associated with the same time period.

reports indicated any precipitation during that hour.

Composite radar images were made for each hour by computing the mean rain rate for all the radar scans during the previous hour. For example, the 0200 UTC composite would be the mean reflectivity observed of all radar scans from 0101 to 0200 UTC. This scheme was chosen to coincide with the observation method used for the raingages. No composite was made for hours where more than one observation was missing.

Since the gage reports were already one-hour totals, no temporal averaging was needed. To minimize possible spatial mismatches, the rain rate derived from the radar composite was the mean of a 5x5 grid (46 km^2) surrounding the gage's location in the radar grid. Because of the sparsity of the non-Fisher-Porter gages and the small number of gages within each radar's domain, there was no attempt to segregate by gage type.

To measure the "goodness" of the radar-gage matchups, the standard linear correlation coefficient and the Heidke skill score (HSS) were used. The correlation coefficient shows how well a linear relationship fits the data, while the HSS shows how well one variable (radar) predicts the other (gages) using a 2x2 contingency table (Table 5). The HSS can range from -1 to 1, with -1 indicating perfect negative skill, zero indicating no skill compared to chance, and 1 indicating perfect positive skill. Lee and Passner (1993) discuss HSS and other measures of skill utilizing contingency tables for a forecast verification problem. In this case, it was desired to see how well the radar was "predicting" the gage rainfall. Non-zero rain rates for either system were counted as a

	Forecasted YES	Forecasted NO
Observed YES	*A*	*B*
Observed NO	*C*	*D*

Table 5. Standard 2x2 contingency table for forecast verification.

"yes", while a zero rain rate was a "no". For AP screening purposes, the HSS serves well to measure how well precipitation areas on the radar correlate to gage precipitation reports.

Based on Table 5, HSS is calculated by

$$\text{HSS} = \frac{2(AD - BC)}{B^2 + C^2 + 2AD + (B + C)(A + D)} \quad . \tag{15}$$

Table 6 shows the correlation coefficients and HSS for the six radar sites. Even with 1-h temporal averaging and 46 km^2 spatial averaging for the radar, the statistics were surprisingly poor. To further filter out temporal and spatial variations, these hour composites were compiled to compute a mean rain rate over 6 h. At least 4 of the 6 h had to contain valid hourly composites for a 6-h composite to be created. Requiring the 6-h composites to have five or six valid composites greatly reduced the dataset without having a significant effect on the results.

Within the valid area of the radar coverage, two counts of raingages were made: (1) G_0, a gage composite of zero but a radar composite greater than zero, and (2) G_P, a gage composite greater than zero and a radar composite greater than zero. The composite was rejected if

$$G_0 > G_P \quad \text{and} \quad G_0 + G_P > 2 \, . \tag{16}$$

This formula rejects composites (a) for radar-indicated rain areas, more than half of the gages in that area show no rain, and (b) where the radar indicated rain was occuring at more than two gage locations.

The threshold of 2 in (16) was found through experimentation using the HSS as a measure of goodness. Increasing the G_0+G_P threshold above 2 decreased the HSS so the

	Radar site					
	TBW	BNA	UMN	ICT	OKC	AMA
1-h composites (N=)	69129	176815	142044	153086	196508	114274
HSS	0.219	0.353	0.373	0.309	0.326	0.174
Correlation	0.074	0.152	0.249	0.130	0.211	0.059
6-h composites (N=)	11091	25190	18827	20393	28318	17904
HSS	0.267	0.476	0.485	0.413	0.422	0.239
Correlation	0.072	0.176	0.402	0.207	0.364	0.102
6-h selected composites (N=)	6694	19287	12056	14040	20351	11486
HSS	0.573	0.599	0.685	0.621	0.594	0.463
Correlation	0.544	0.458	0.636	0.571	0.594	0.408

Table 6. Heidke skill scores (HSS) and correlation coefficients (r) for one- and six-hourly composites. *Selected composites* are those that pass the criteria outlined in the text.

threshold was left at 2. Table 6 shows that marked improvement in HSS and correlation was obtained by selecting only those cases meeting the criteria, at the cost of removing about 30 percent of the radar-gage matchups. The improvement at TBW was especially noteworthy: subjective review of the radar reflectivity images such as those in Fig. 11 indicated this site had many more instances of widespread false echoes than any of the other sites. Eliminating time periods where this appeared to be prevalent improved the quality of the dataset.

The relationship between G_0 and G_P showed that (16) was sufficient to reject composites where the gages recorded no rainfall and a significant area of the radar had echoes, yet not unnecessarily reject cases where only a small area indicated a misidentifi-

cation by the radar (the reason the G_0+G_P threshold was not set less than 2). The 6-h time periods corresponding to the rejected composites were recorded and radar data from those time periods were excluded from future satellite-radar and gage-radar comparisons.

Mean rain rates were also computed for the 6-h gage composites and radar composites (using the 5x5 spatial average for the radar as mentioned above) where both methods indicated rainfall during the composite time period. This was done to remove the effects of temporal and spatial mismatches, and to reject occurrences where either the gage or the radar indicated precipitation that was unconfirmed by the other. Since rain rates are not normally distributed about some mean but are instead highly skewed toward low values, a ratio of mean rain rates should be more statistically valid than using the bias obtained from a linear regression analysis (a few points near the high end of the distribution have a disproportionately large effect on the bias obtained by linear regression). Table 7 shows the ratio of these mean rain rates at the six radar sites used in this study for the nine-month period of the study.

The temporal variation of the radar/gage ratio is considerable. A more consistent change in the ratio would lead to the hypothesis that the mean *Z-R* relationship changes with the season as the weather transitions from predominantly stratiform precipitation to more convective precipitation. In this case, the temporal variation is so great that there is no sound meteorological or physical basis for applying these corrections month-by-month to the radar data. However, applying a correction by site only would reduce the site-to-site bias relative to the gage totals, and allow valid comparisons of all radar sites simultaneously. To correct the radar-derived rain rates so that they correspond to the gage-derived rates, the radar rain rates were divided by the "Mean" in Table 7.

The standard deviation of the logarithm of the monthly ratios is an indication of the temporal variability in the radar/gage ratio, and perhaps an indication of the overall reliability of the data from that radar site. Large variations in the monthly ratios are

	Radar site					
	TBW	BNA	UMN	ICT	OKC	AMA
Gages in radar domain	17	46	58	44	51	33
January	0.1907	0.2899	*	0.2071	0.1387	0.1089
February	0.3996	0.3104	0.3246	0.1077	0.1749	0.0843
March	0.2956	0.2047	0.3636	0.5963	0.1076	0.4462
April	0.5394	0.2760	0.3381	0.4479	0.1366	0.8253
May	0.3980	0.3546	0.4840	0.3628	0.1143	0.2114
June	0.4775	0.2924	0.4306	0.4556	0.2443	0.8028
July	0.8167	0.3748	*	0.4300	0.0810	1.2560
August	0.8201	0.3444	*	0.5512	0.1138	0.5173
September	0.6174	0.2772	*	0.5053	0.2413	0.4326
Mean	0.5015	0.3125	0.4240	0.4455	0.1840	0.5568
Std. Dev. of log of monthly means	0.4728	0.1798	0.1689	0.5528	0.3722	0.9288

Table 7. Ratio of radar/gage rain rates for RADAP-II sites during January-September 1992. An asterisk (*) indicates no data for that month.

probably unphysical and could be due to problems with the radar or poor sampling by either the radar or the gages. Both are possible, as shown later in the satellite/radar and satellite/gage comparisons.

b. Use of linear regression and Heidke skill score to determine best-fit lines

Assuming the relationship between an algorithm's output and ground truth is linear, the usual method of calibration is to perform a linear regression analysis using the output as the dependent variable X and the ground truth as the independent variable Y. Linear regression makes some underlying assumptions about the data, namely that in the model

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \tag{17}$$

the error terms ε_i are independent of the X_i and normally distributed with a mean of zero and a constant variance. Also, it is assumed that measurement error is associated with the measurement of Y_i, i. e., the X_i are known without measurement error.

The latter constraint is the most serious. Neter et al. (1990) discuss how the classic linear regression model (17) is not valid if there is a measurement error associated with each X_i. In this case, the X_i have errors associated with the measurements of the brightness temperatures for each of the seven channels, and these errors cannot be assumed to be small enough to be neglected.

The error terms ε_i cannot be assumed to be normally distributed with a mean of zero in all cases. A sizable fraction of the data set has Y_i =0, and since rain rate must be non-negative, the error term cannot be greater than zero, leading to a non-zero mean. This again violates one of the major assumptions of the linear regression model.

The distribution of the rain rates is also problematic. Figure 12 shows a scatterplot of rain rates obtained by the SRL algorithm and the radar-derived "ground truth" for the BNA radar site. The vast majority of the points are near the origin, and in fact over 90 percent of the points in this sample lie directly on the origin. The few points in the upper

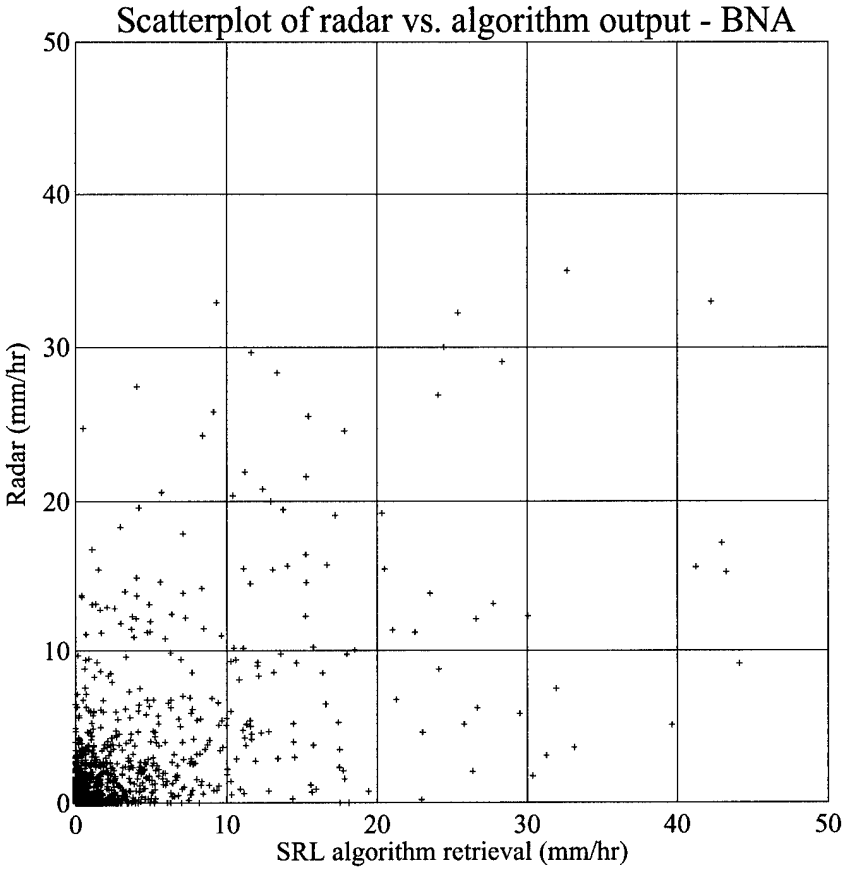


Figure 12. Scatterplot of radar vs. SRL algorithm rain rates for BNA for January-September 1992.

right portion of the graph have a substantial influence on any least-squares linear regression line.

Box-Cox transformations (Neter et al., 1990) can correct for non-linear relationships, skewness of error terms, and unequal error variances. Since there is no evidence of nonlinearity for the new Purdue algorithms, and the GSFC and SRL algorithms are already converted to rain rates, the Box-Cox approach can be used to attempt to correct for problems with the error terms by using the transformation on both the X and Y variables. Box-Cox transformations are simply power transformations of the form

$$\begin{aligned} Y' &= Y^{\lambda} \qquad (\lambda \neq 0), \\ Y' &= \log_e Y \quad (\lambda = 0). \end{aligned} \tag{18}$$

For $\lambda \leq 0$, Y' is undefined when Y is not greater than zero, which is undesired. For continuous distributions above and below zero, this can be solved by adding constants so that all X and Y are positive. However, rain rate distributions are not continuous below zero. An acceptable solution is to use a value of λ that has a defined value for zero and deemphasizes the error variance for high numerical values of X and Y. A simple and somewhat arbitrary choice is to use λ=0.5, or $\sqrt{Y}$.

While a square-root transformation removes some of the skewness of the error terms, it does not address the problems of measurement error in the dependent variable nor that of non-normal distribution of error variance in the independent variable. This transformation is still useful in that the correlation coefficient r for such transformed variables may give a better indication of the linearity of the relationship between ground truth and the algorithm in question, without allowing a handful of points at the far end of the distribution to have an undue influence on the regression and r.

In this work, we explore the possibility of using the Heidke skill score (HSS) as a calibration tool in place of linear regression. While this use of HSS is somewhat nonstandard and subjective, it does offer advantages over linear regression in cases where the data are highly skewed because there are no underlying assumptions on how the data and errors are distributed. An additional advantage is that HSS is a valid measure of the skill of an algorithm at distinguishing rain exceeding a specified intensity, a useful performance measure in an operational setting. The HSS varies from -1 to +1, with -1 indicating perfect negative skill, zero indicating no skill relative to chance, and +1 perfect positive skill.

In discussions of linear regression and algorithm calibration, there is a difference in the use of certain terms. In statistics, the term *bias* is the slope of a linear regression line, while most works discussing algorithms use the term to mean a constant over- or underestimate, which is the *intercept* in linear regression. To avoid confusion, the terms *slope* and *intercept* will be used.

Referring back to Table 5, it was shown that HSS is based on a yes or no determination of a forecast and an observed variable. This yes/no decision can be applied to a variety of parameters. Lee and Passner (1993) used HSS to score how thunderstorm forecasts verified, with the occurrence/nonoccurrence of thunderstorms as the yes/no criteria. For rain rates, the yes/no criteria can simply be whether the algorithm or verification indicates rain or no rain.

However, HSS can be considered a measure of the skill of an algorithm to determine whether the verification exceeded a specific threshold. In the above case, the threshold was zero. If the threshold were set at 5.0 mm h^{-1}, for example, the HSS would reflect the algorithm's skill at determining whether the rain rate at a particular location exceeded 5.0 mm h^{-1}. Now suppose that a number of skill scores were computed with the observed threshold set at 5.0 mm h^{-1}, but the forecast threshold set at intervals from, say, 0 to 10.0 mm h^{-1}. If the algorithm was calibrated properly, the skill score at 5.0 mm h^{-1} would be the highest. However, if it had a slope of 0.5 (algorithm output indicated half the actual rainfall rate), then the peak skill score would occur at a forecast threshold of 2.5 mm h^{-1}.

If both the forecast and observed skill scores were computed at intervals, a properly calibrated algorithm would have the highest skill scores where the forecast threshold equaled the observed threshold. If the algorithm perfectly replicated the observations, the skill score would be 1.0 at those points, and a 2-D contour plot of the results would indicate a maximum along the axis where the forecast threshold equals the observed threshold.

Figure 13 shows an idealized case where it is possible to determine slope and intercept from the skill score plots. Note that the slope of the axis of maximum skill score is equal to the slope obtained by linear regression and that where a line through this axis intercepts the *y*-axis of the plot is equal to the intercept. Therefore, this technique holds promise as a method of obtaining a valid slope and intercept for an arbitrary dataset.

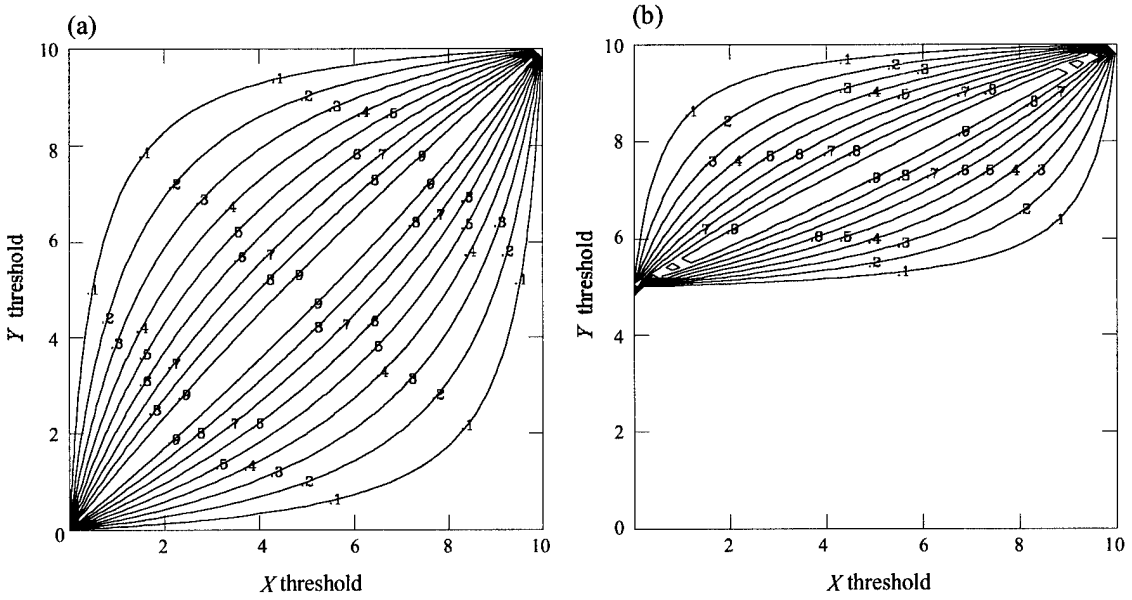


Figure 13. Two-dimensional HSS plots for a uniformly distributed and perfectly correlated dataset. (a) $y_i=x_i$. (b) $y_i=5 + 0.5x_i$.

c. Comparison of satellite-derived rain rates to radar-derived rain rates

Figure 14 shows 2-D HSS plots for rain rates for (5), (6), (8), (13), and (14), which are the three Purdue algorithms, the SRL algorithm, and the GSFC algorithm, respectively, versus the radar-observed rain rates at BNA. The linear relationship becomes indistinct above rain rates of 8 mm h^{-1}. In fact, the maxima trend to a vertical line at or above this rain rate, indicating (a) the algorithms are insensitive to rain rates above this

value, or (b) the radar data does not faithfully represent such rain rates. Since all the algorithms show this insensitivity at nearly the same rain rate, it was deemed more likely that the fault was with the radar data.

To evaluate the other radar sites for possible inclusion into a composite dataset, HSS plots were made for the EV algorithm versus radar rain rates at the other five sites (Fig. 15). The AMA and ICT radar sites showed substantially lower skill than the other four sites even at lower rain rates. This, combined with the month-to-month variability noted in Table 7 earlier led to removal of these two sites from the composite data set. These distributions were not peculiar to the EV algorithm. Figure 15 also shows that OKC has essentially no data points above 7 mm h^{-1}, and that UMN have very few data points above 12 mm h^{-1}, again something not unique to the EV algorithm.

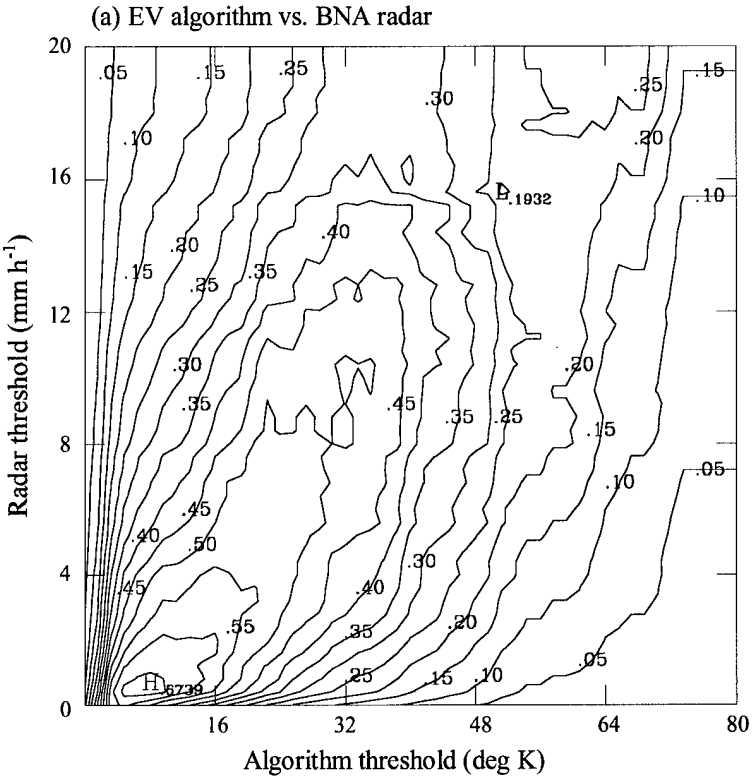


Figure 14. HSS plots for algorithm rain rate thresholds vs. BNA radar rain rate thresholds. (a) Purdue-EV (b) Purdue-2C (c) Purdue-4C (d) GSFC (e) SRL

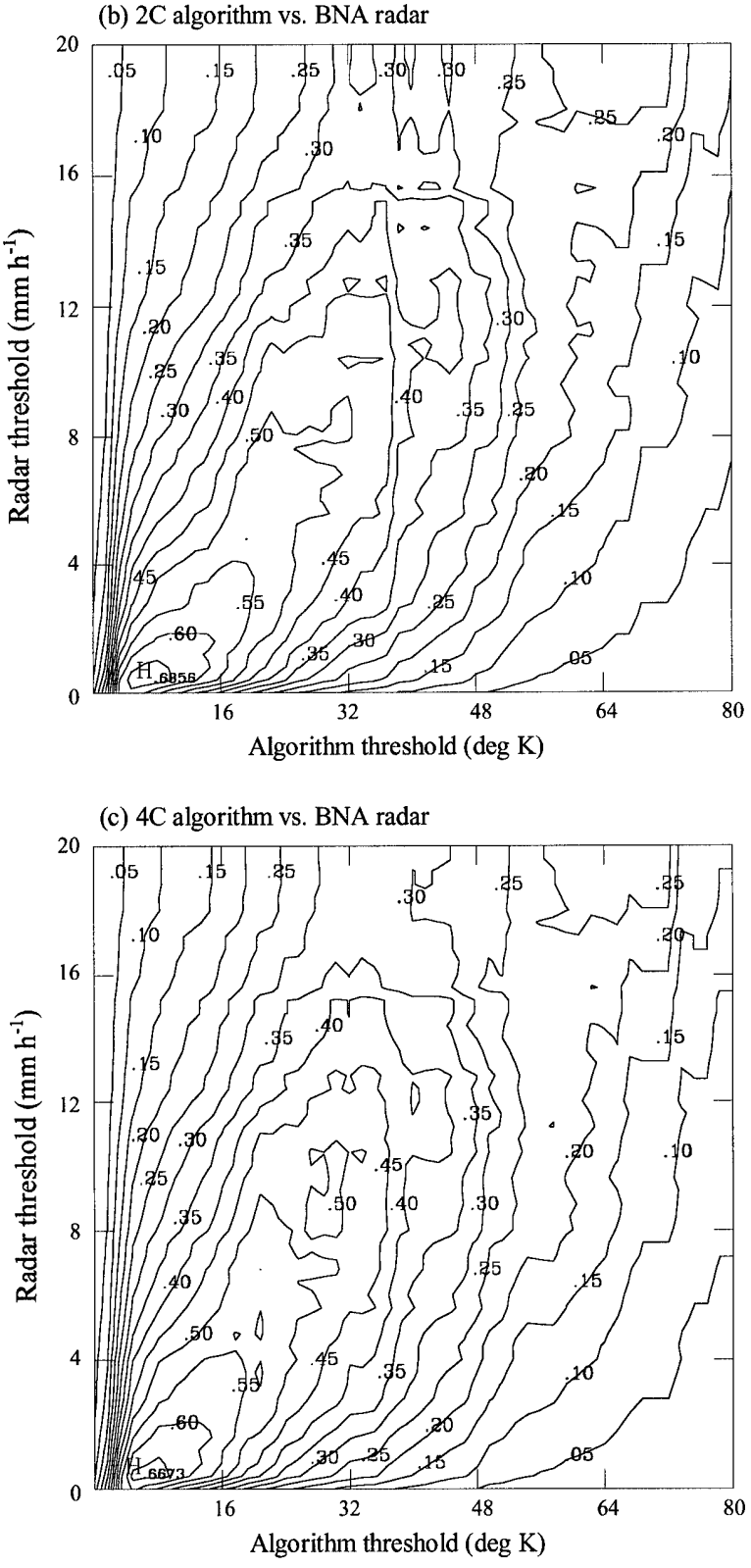
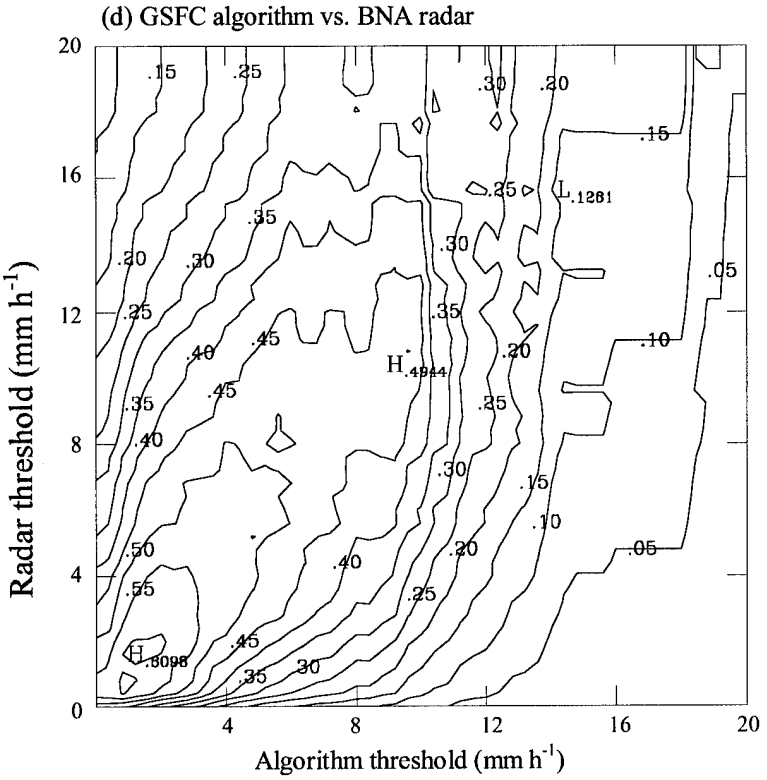


Figure 14 (continued).



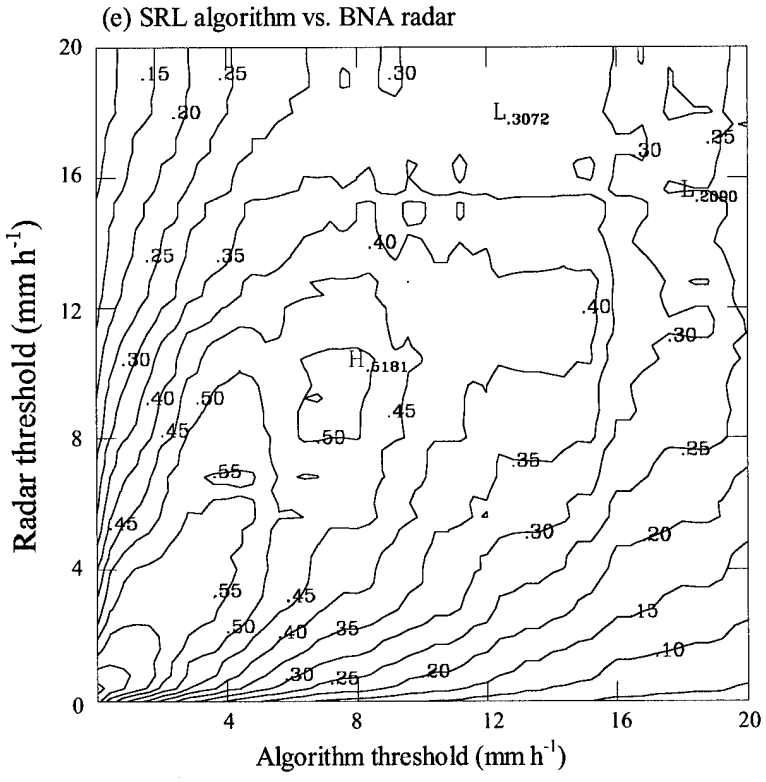


Figure 14 (continued).

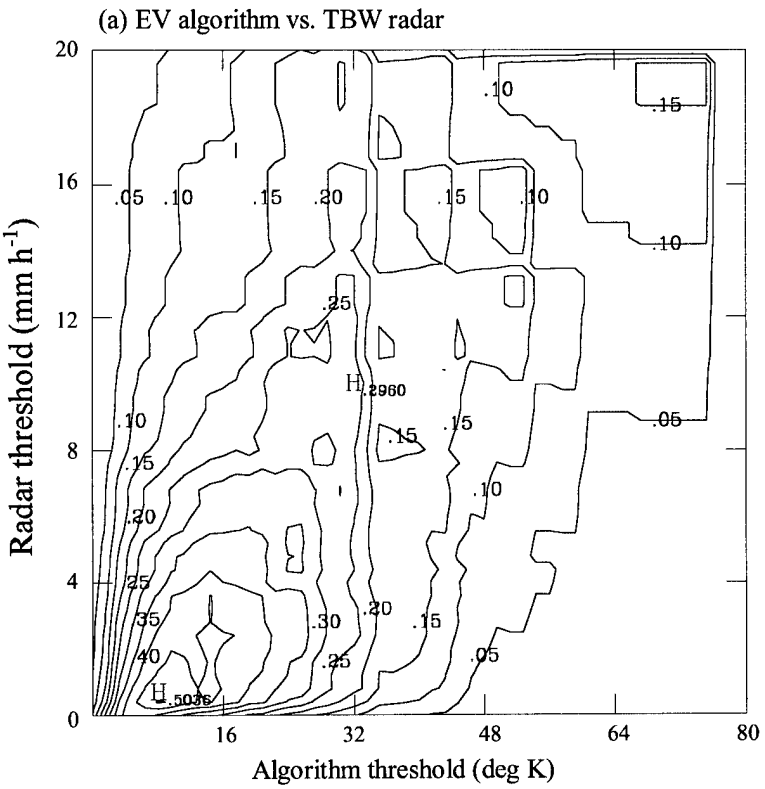


Figure 15. Purdue-EV algorithm rain rate threshold vs. rain rate at other radar sites. (a) TBW (b) UMN (c) ICT (d) OKC (e) AMA

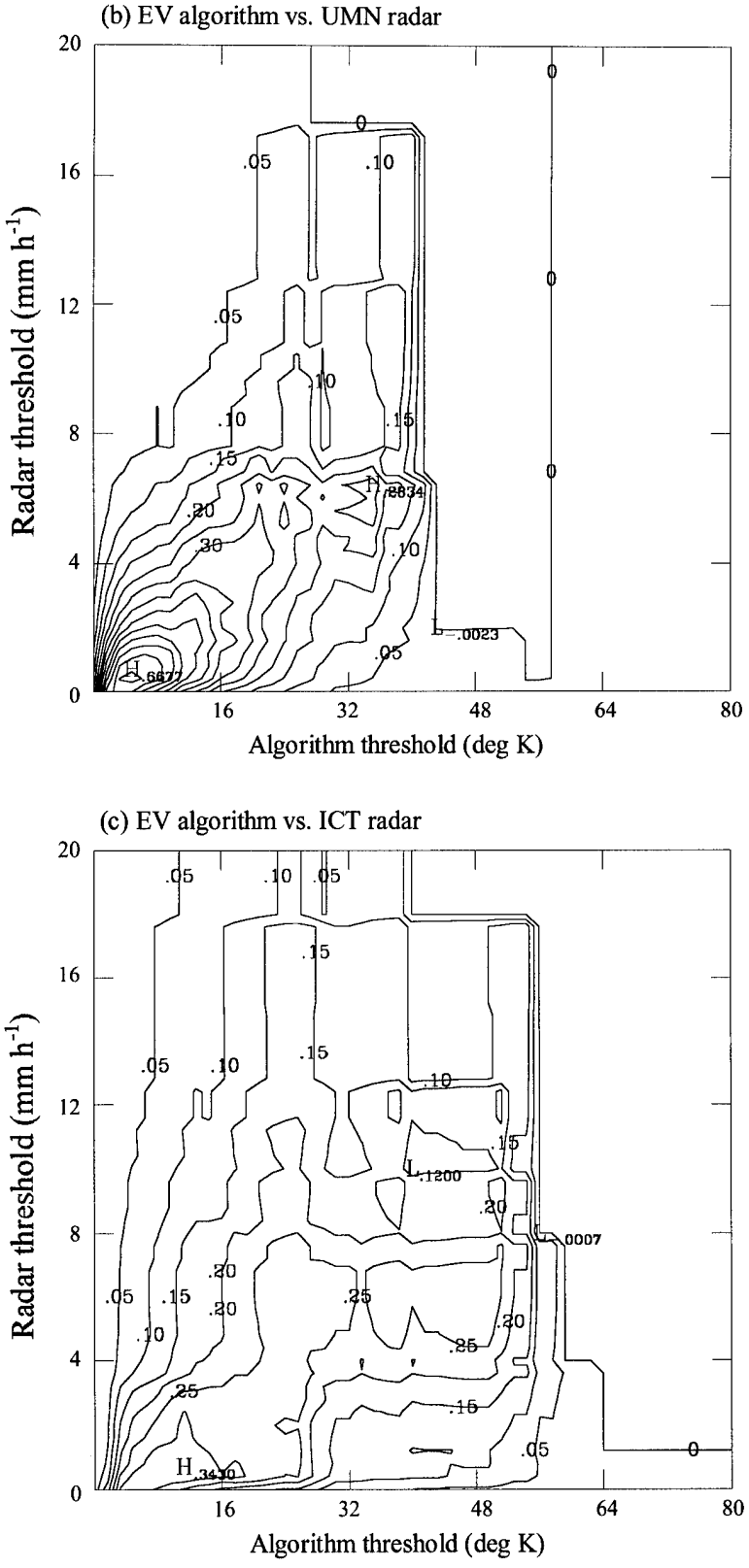


Figure 15 (continued).

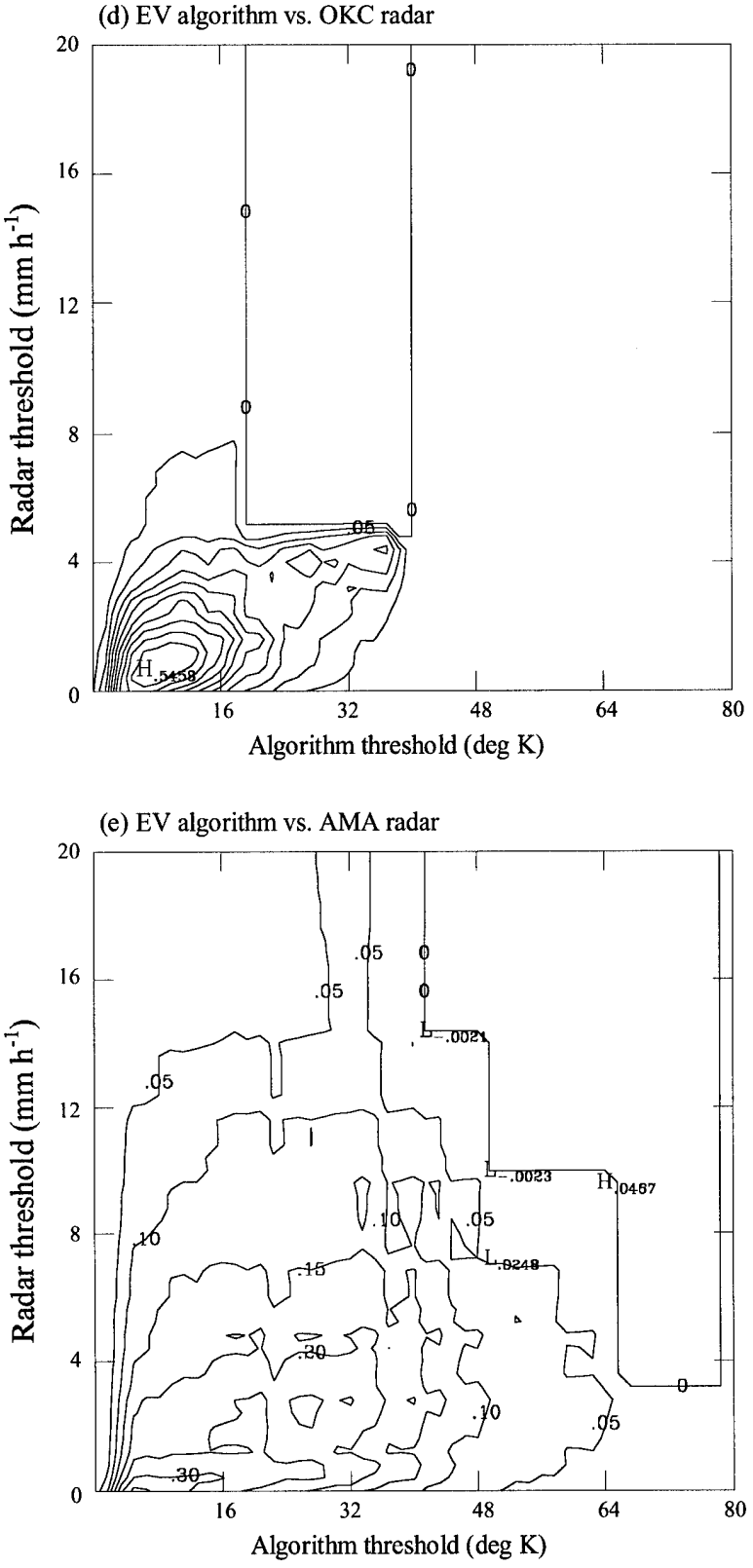


Figure 15 (continued).

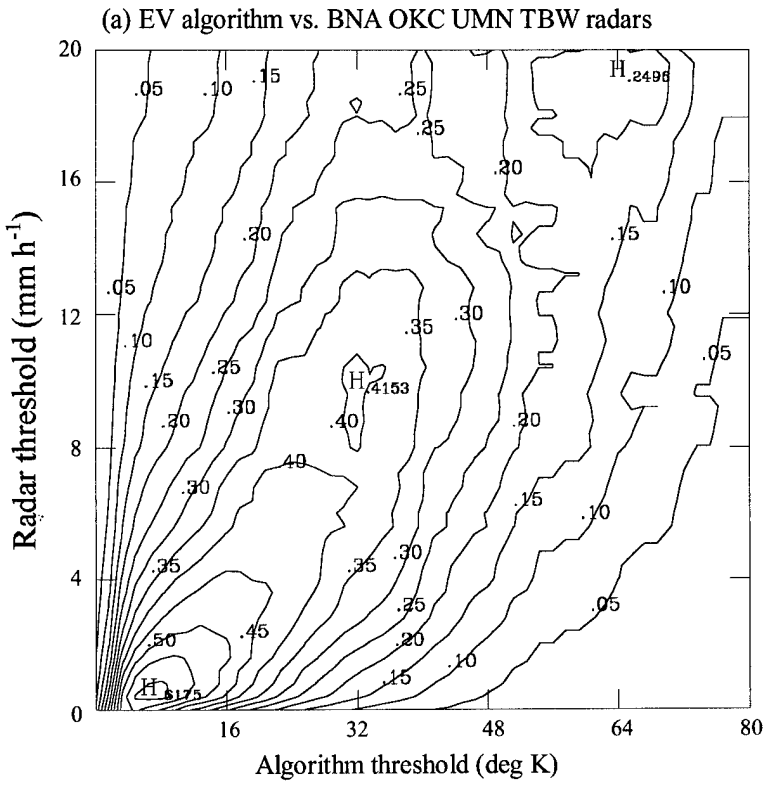


Figure 16. Algorithm rain rate thresholds vs. radar rain rate thresholds from a combined dataset from the BNA, OKC, UMN, and TBW radars. (a) Purdue-EV (b) Purdue-2C (c) Purdue-4C (d) GSFC (e) SRL

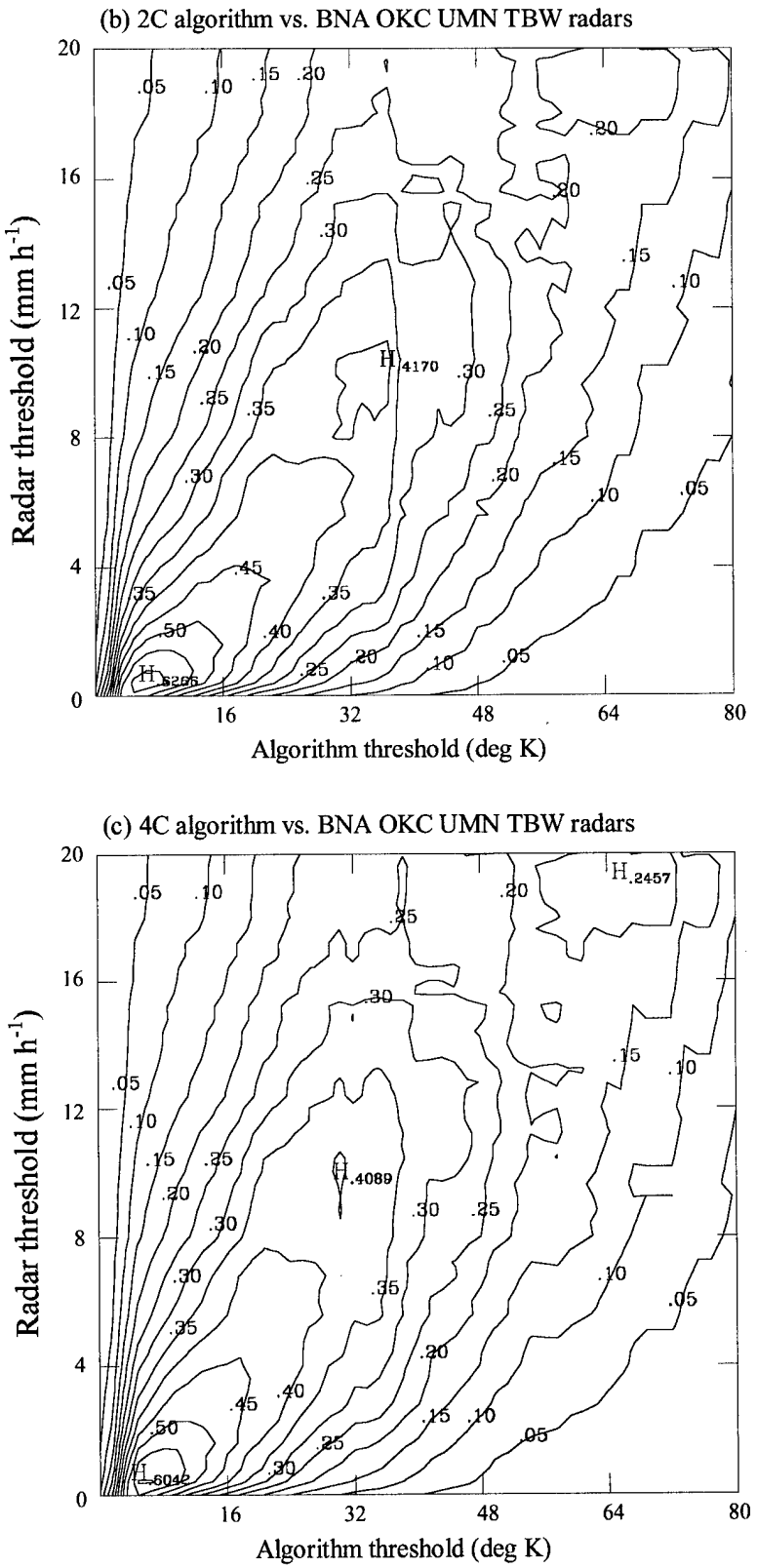


Figure 16 (continued).

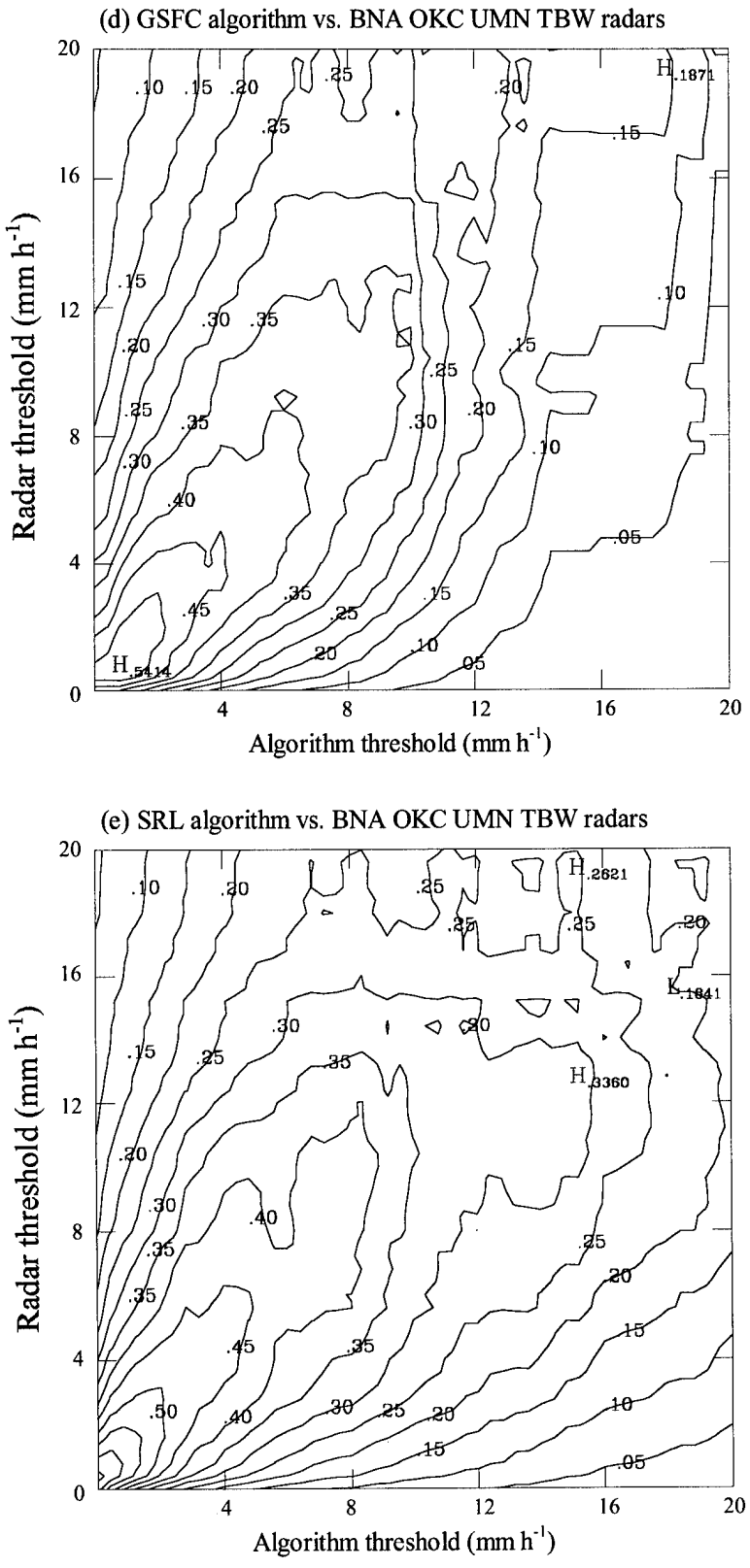


Figure 16 (continued).

Figure 16 shows the HSS plots for a composite dataset consisting of the algorithm output and radar observations from BNA, TBW, UMN, and OKC. Table 8 shows the best-fit lines obtained by linear regression and from a subjective line drawn along the maxima of the skill score plots, and correlations for the untransformed data and for a $x^{0.5}$-$y^{0.5}$ transformation. Since the additional radar sites had little data for rain rates above 10 mm h^{-1}, there was correspondingly little change in that portion of the graph. Because of the uncertainty of the data at the higher rain rates, the subjective best-fit line on the HSS plots was chosen to fit the more definite trend shown below 10 mm h^{-1}.

For the linear relationship $y = a + bx$, normally the y-intercept a is reported. However, for the calibration, the x-intercept is needed instead, which is a' in the relationship $y = b(x - a')$. The calibrated results $\hat{y}_{ij}$ for each Purdue algorithm j were calculated by:

$$\begin{aligned} \hat{y}_{ij} &= b_j\left(x_{ij} - a_j'\right) && \text{if } x_{ij} \geq a_j', \text{ or} \\ \hat{y}_{ij} &= 0 && \text{if } x_{ij} < a_j' \end{aligned} \tag{19}$$

The linear regression and correlation coefficients were recalculated for the calibrated Purdue algorithms. The slopes should become near unity and the x-intercepts near zero for an ideal case. Because the linear regression model is not usable in this case, and because the best-fit line applied was not from linear regression, the parameters do not become the ideal. However, both have trends in the right direction.

	Linear regression		HSS-determined		Correlation coefficient	
	Slope	*x*-intercept	Slope	*x*-intercept	Linear	Sq. root
EV (uncal.)	0.099	-0.567	0.34	4.48	0.521	0.581
EV (cal.)	0.715	-0.063	*	*	0.563	0.696
2C (uncal.)	0.102	-0.355	0.34	4.48	0.503	0.562
2C (cal.)	0.681	-0.063	*	*	0.558	0.696
4C (uncal.)	0.096	-0.726	0.39	4.80	0.508	0.554
4C (cal.)	0.641	-0.086	*	*	0.557	0.685
GSFC	0.306	-0.222	1.43	0.00	0.542	0.644
SRL	0.415	-0.165	1.20	-0.40	0.555	0.712

Table 8. Slopes and *x*-intercepts from linear regression and HSS method and correlation coefficients for untransformed and square-root-transformed data for algorithms vs. a composite dataset from the BNA, OKC, UMN, and TBW radars. Linear regression and correlation coefficient results also shown for calibrated Purdue algorithms.

d. Comparison of satellite-derived rain rates to raingage-derived rain rates.

As noted earlier, the satellite-derived rain rates are instantaneous snapshots of the rain rate field with some spatial averaging due to the pixel sample size and the subsequent gridding process. This is also true for the radar-derived rain rates. Raingages, on the other hand, are point estimates of the rain rate averaged over some temporal domain (in this case, over one hour). While this is a disadvantage, the gage data are superior to the radar data in other respects: (1) the gages are a more direct measure of the true rain rate

and thus less susceptible to calibration errors, and (2) the gage data cover the entire CONUS instead of small portions of it.

Figure 17 shows the HSS plots for all five algorithms versus all gage reports (around 640 000 pairs). The stairstep pattern at intervals along the vertical axis is an artifact of the Fisher-Porter gage reports. These gages report in 2.54 mm (0.10 inch) increments, and this gage type comprises about 85 percent of the network. While the skill scores are lower than for comparable points in Fig. 15 or Fig. 16, the linearity of the relationships is encouraging.

To eliminate the stairstep effect, and to test if the Fisher-Porter gages detracted from the skill observed, the Fisher-Porter gage reports were removed from the dataset. Figure 18 shows the HSS plot for the EV algorithm vs. the non-Fisher-Porter gages. Removing the Fisher-Porter gages eliminated the stairstep effect, but improved the skill scores only a small amount.

Unlike the SRL and GSFC algorithms, the Purdue algorithms have no explicit screening for the detection of likely snow-covered surfaces. These snow surfaces exhibit a microwave signature similar to that of precipitation areas, and must be detected in order to prevent the return of a false non-zero rain rate. To detect these snow areas, a snow-cover flag was created based on the gridded brightness temperatures and (9), (10), and (11) from the SRL algorithm. The grid boxes where this flag indicated a snow surface were removed from the dataset. Snow detection was found to be unnecessary for the radar comparisons because the radar sites had few instances of snow cover within the radar domain.

Figure 19 shows the HSS plots for the "non-Fisher-Porter, no-snow" data. The linearity of the algorithm-verification relationship is maintained, but the pattern has become more complex. The skill scores have increased, especially at lower rain rates. This was expected because (1) spurious Fisher-Porter gage results give errors of 2.54 mm

rather than 0.254 mm and (2) the Purdue algorithms had no snow screening. Table 9 shows the best-fit line parameters and correlations similar to that for the radar comparisons.

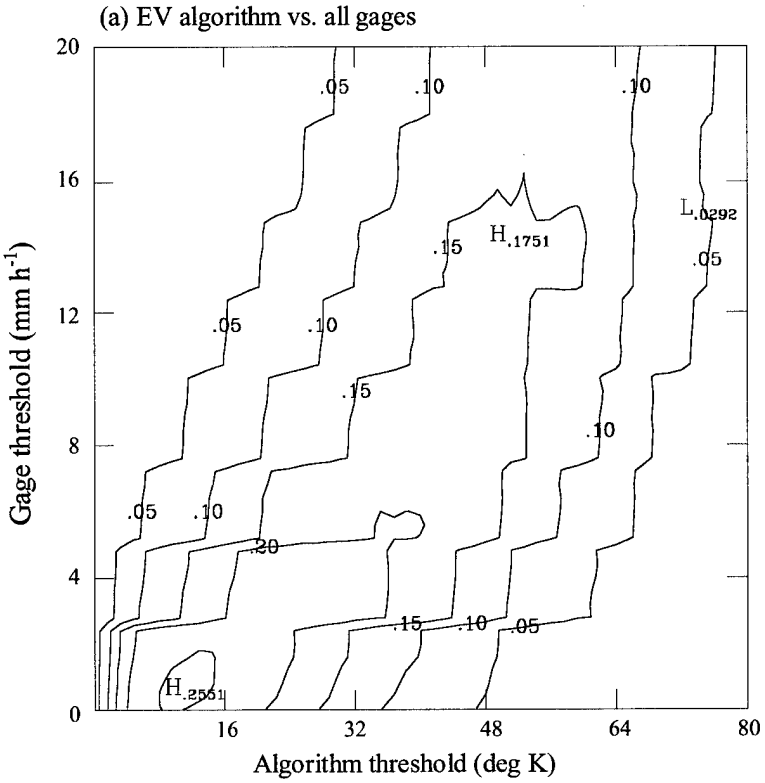
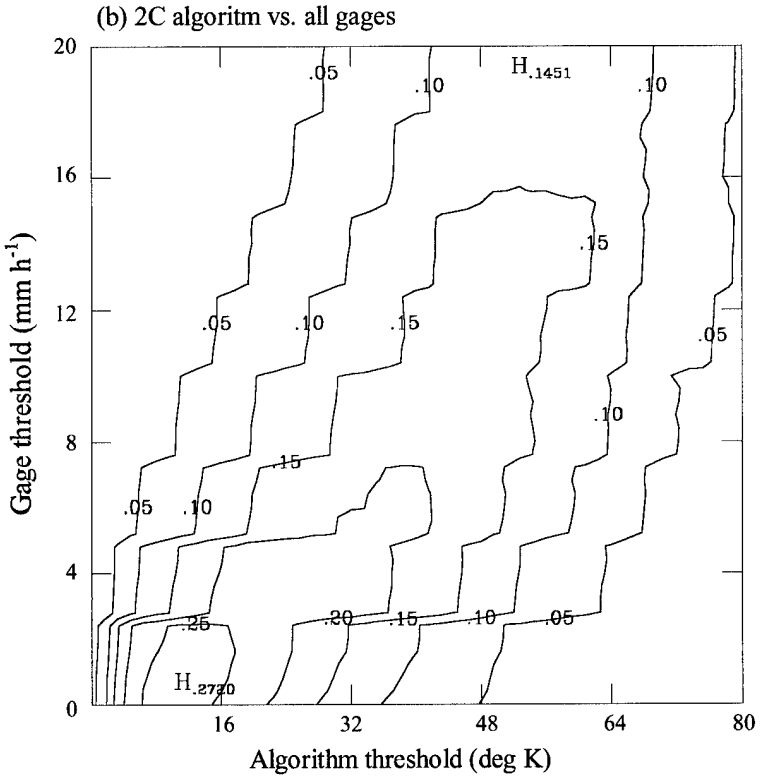


Figure 17. Algorithm rain rate thresholds vs. raingage rain rate thresholds for all gages. (a) Purdue-EV (b) Purdue-2C (c) Purdue-4C (d) GSFC (e) SRL



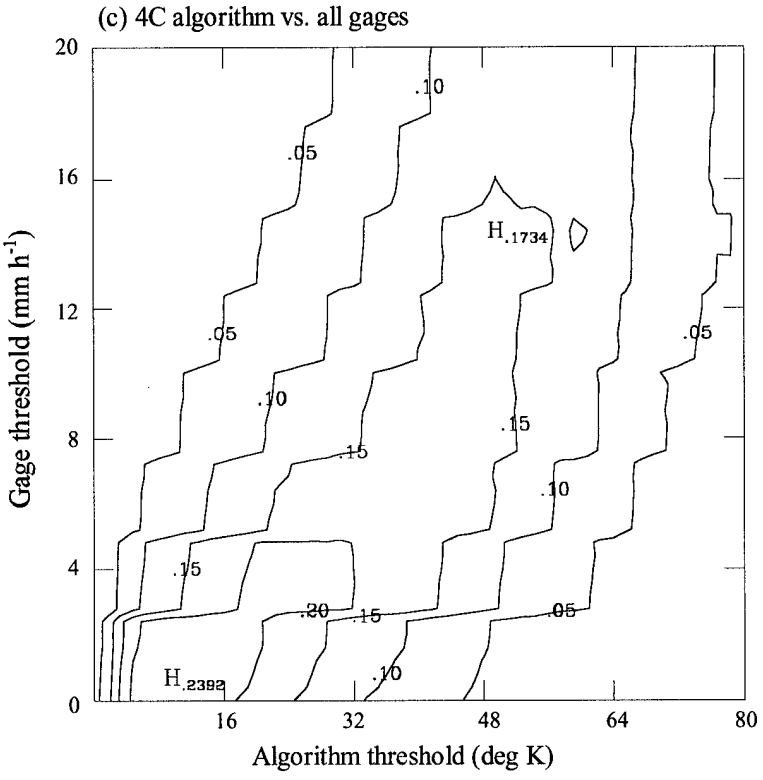


Figure 17 (continued).

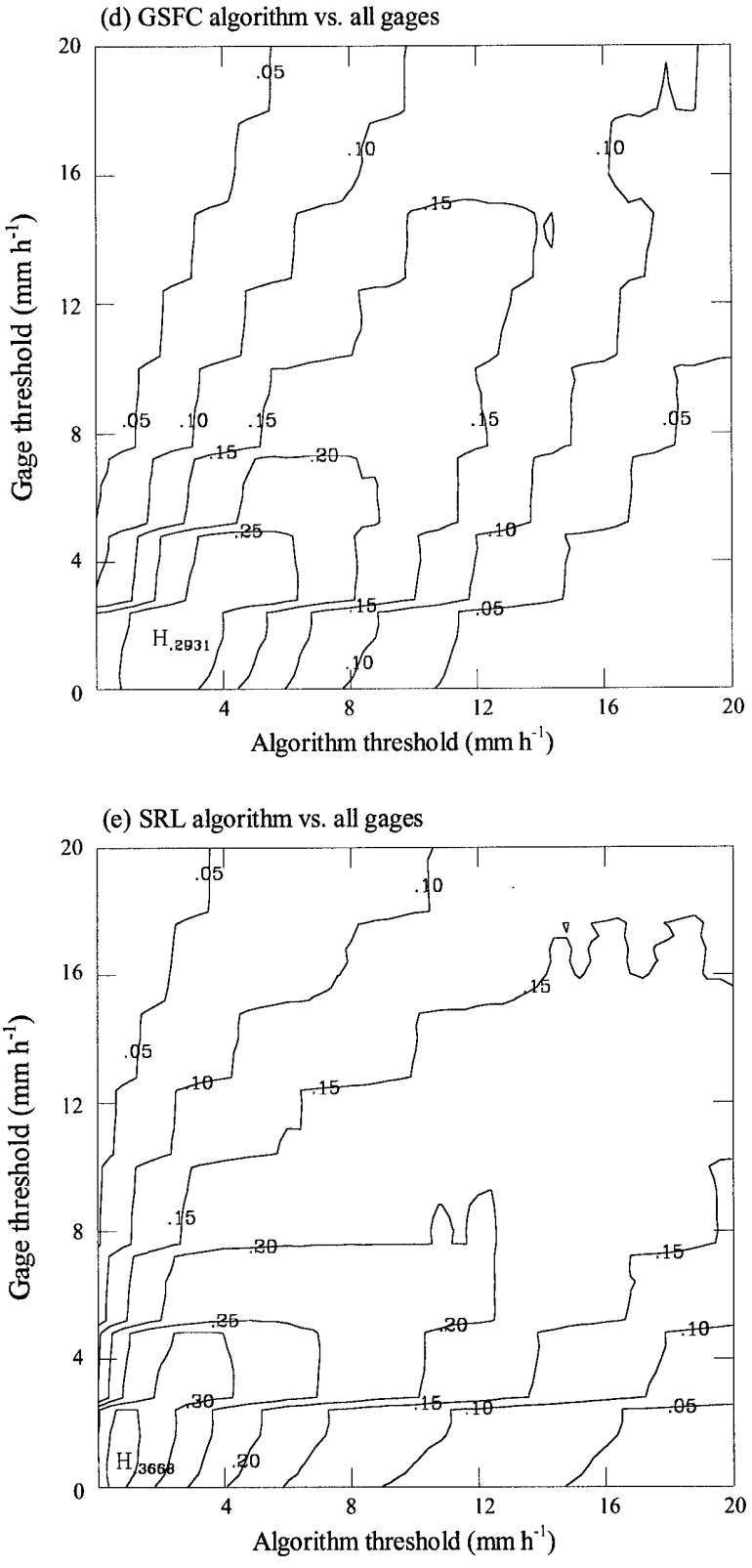


Figure 17 (continued).

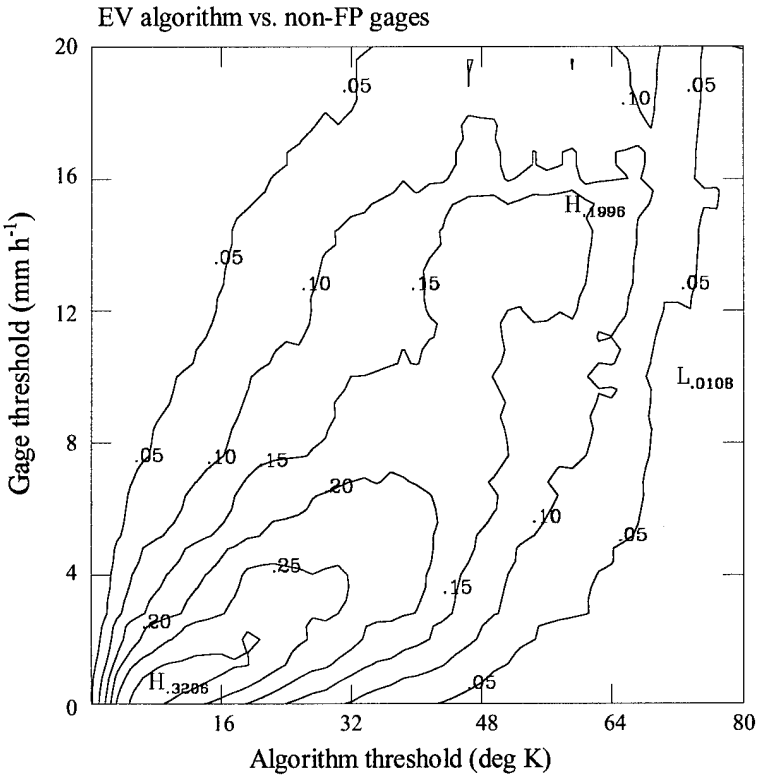


Figure 18. Purdue-EV algorithm rain rate thresholds vs. raingage rain rate thresholds for non-Fisher-Porter gages.

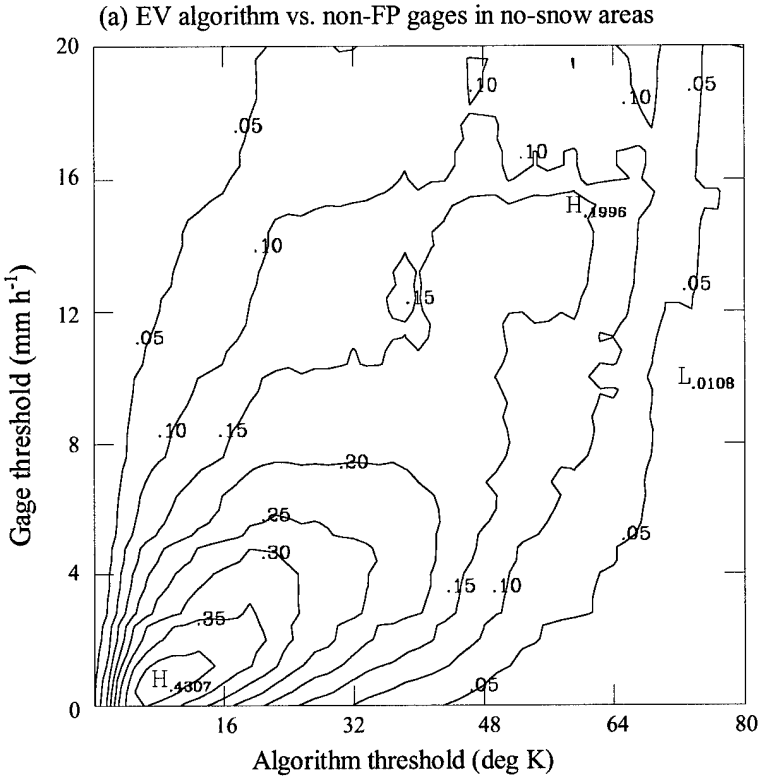
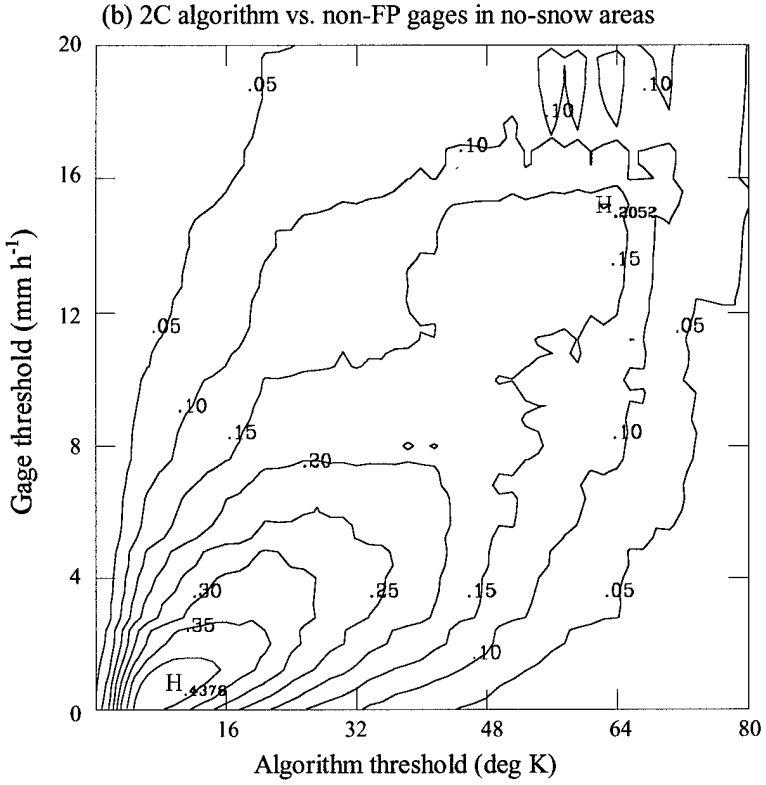


Figure 19. Algorithm rain rate thresholds vs. raingage rain rate thresholds for non-Fisher-Porter raingages in grid boxes not flagged as containing snow (see text). (a) Purdue-EV (b) Purdue-2C (c) Purdue-4C (d) GSFC (e) SRL



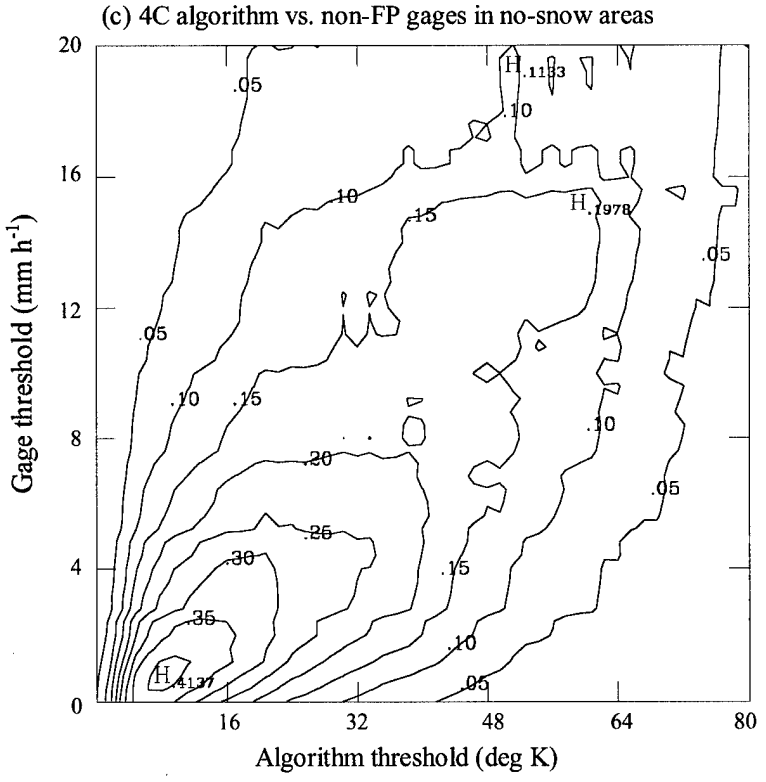
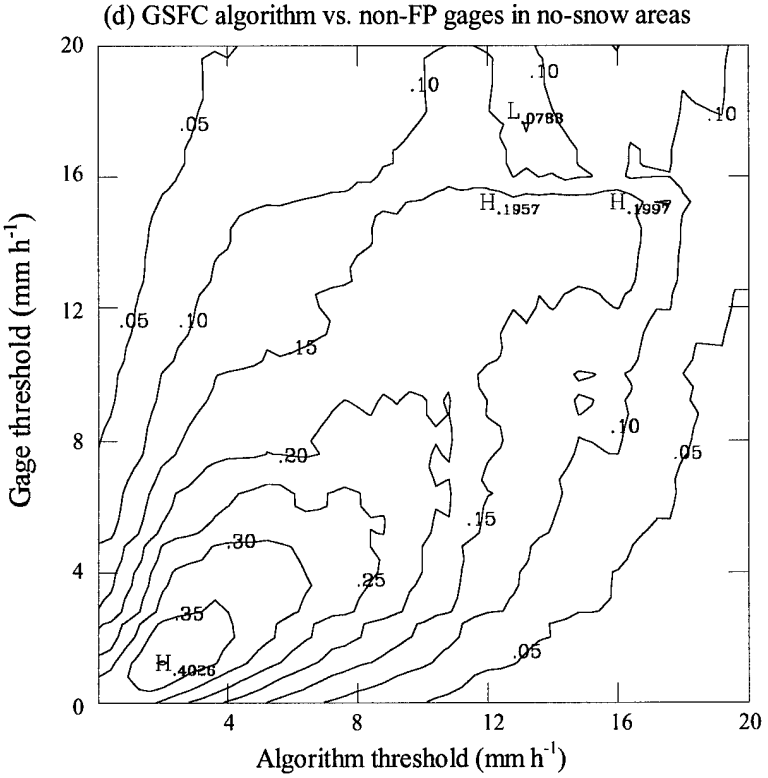


Figure 18 (continued).



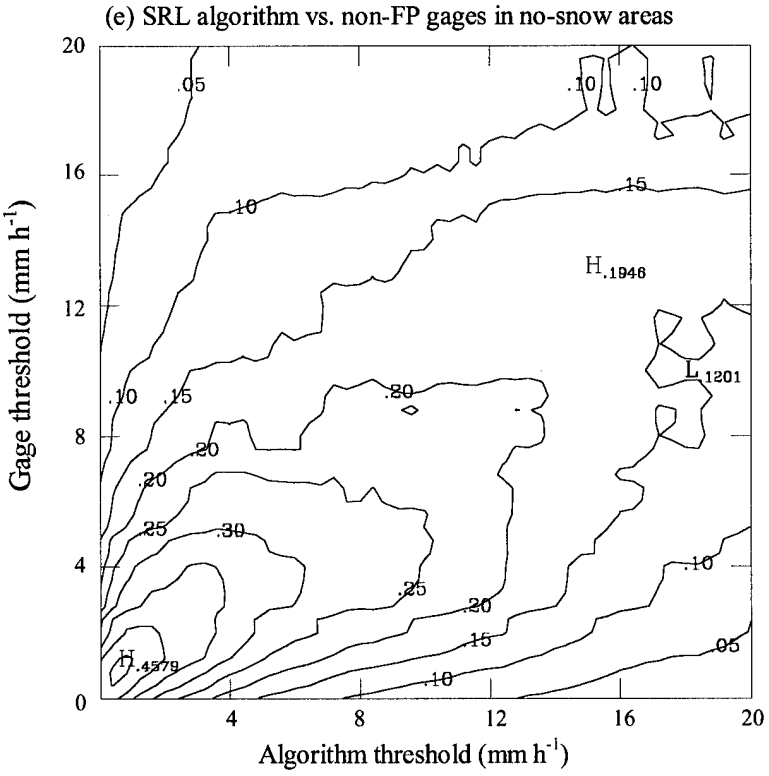


Figure 18 (continued).

	Linear regression		HSS-determined		Correlation coefficient	
	Slope	*x*-intercept	Slope	*x*-intercept	Linear	Sq. root
EV (uncal.)	0.071	-0.724	0.28	6.40	0.369	0.399
EV (cal.)	0.455	-0.121	*	*	0.404	0.506
2C (uncal.)	0.078	-0.834	0.26	5.76	0.375	0.413
2C (cal.)	0.458	-0.111	*	*	0.410	0.520
4C (uncal.)	0.074	-1.034	0.27	5.44	0.357	0.375
4C (cal.)	0.451	-0.121	*	*	0.396	0.493
GSFC	0.458	-0.102	0.77	-0.80	0.399	0.534
SRL	0.439	-0.151	0.92	0.04	0.349	0.457

Table 9. Slopes and *x*-intercepts from linear regression and HSS method and correlation coefficients for untransformed and square-root-transformed data for algorithms vs. non-Fisher-Porter gages in grid boxes not flagged as containing snow (see text). Linear regression and correlation coefficient results also shown for calibrated Purdue algorithms.

5. Summary and Conclusions

Three new algorithms for determining rain rate over land from SSM/I brightness temperatures were developed and evaluated against two algorithms from other research groups. The algorithms were calibrated against both radar and raingage reports, and the raingage data were used to filter out bad radar data and to apply a site-specific calibration factor to the radar-derived rain rates. Figure 20 shows an example of the Purdue-EV algorithm output along with the corresponding gage reports and radar scan from BNA. The other algorithms showed patterns similar to that of the Purdue-EV algorithm.

The use of the Heidke skill score (HSS) as a proxy for linear regression for obtaining a best-fit line was introduced. The results obtained by this method appear to be superior to those obtained by linear regression for the datasets used in this study. Linear regression assumes that (1) the dependent variable is known to a high degree of accuracy (no significant error) and (2) the error terms for the independent variable are normally distributed with a mean of zero. These assumptions are known to be false for the datasets used and this, combined with the highly skewed distribution of the data, leads to the conclusion that linear regression is an inappropriate method to determine a best-fit line.

The utility of the RADAP-II dataset as a calibration tool was disappointing, while the raingage dataset was surprisingly valuable. All five algorithms exhibited non-linear tendencies above ~10 mm h^{-1} using the radar data, while the relationships were relatively linear up to ~ 20 mm h^{-1} using the gage data. While this does not prove the radar data are faulty, the gage data are preferred as a calibration tool not only because of the linearity, but also because gage measurements are more direct, the data cover a wider area, and the sample size is larger (about 100 000 gage/algorithm pairs for the final dataset). The higher skill scores at low rain rates for the radar dataset indicate that radar estimates are probably more useful for rain/no-rain indications than the gages.

(a)

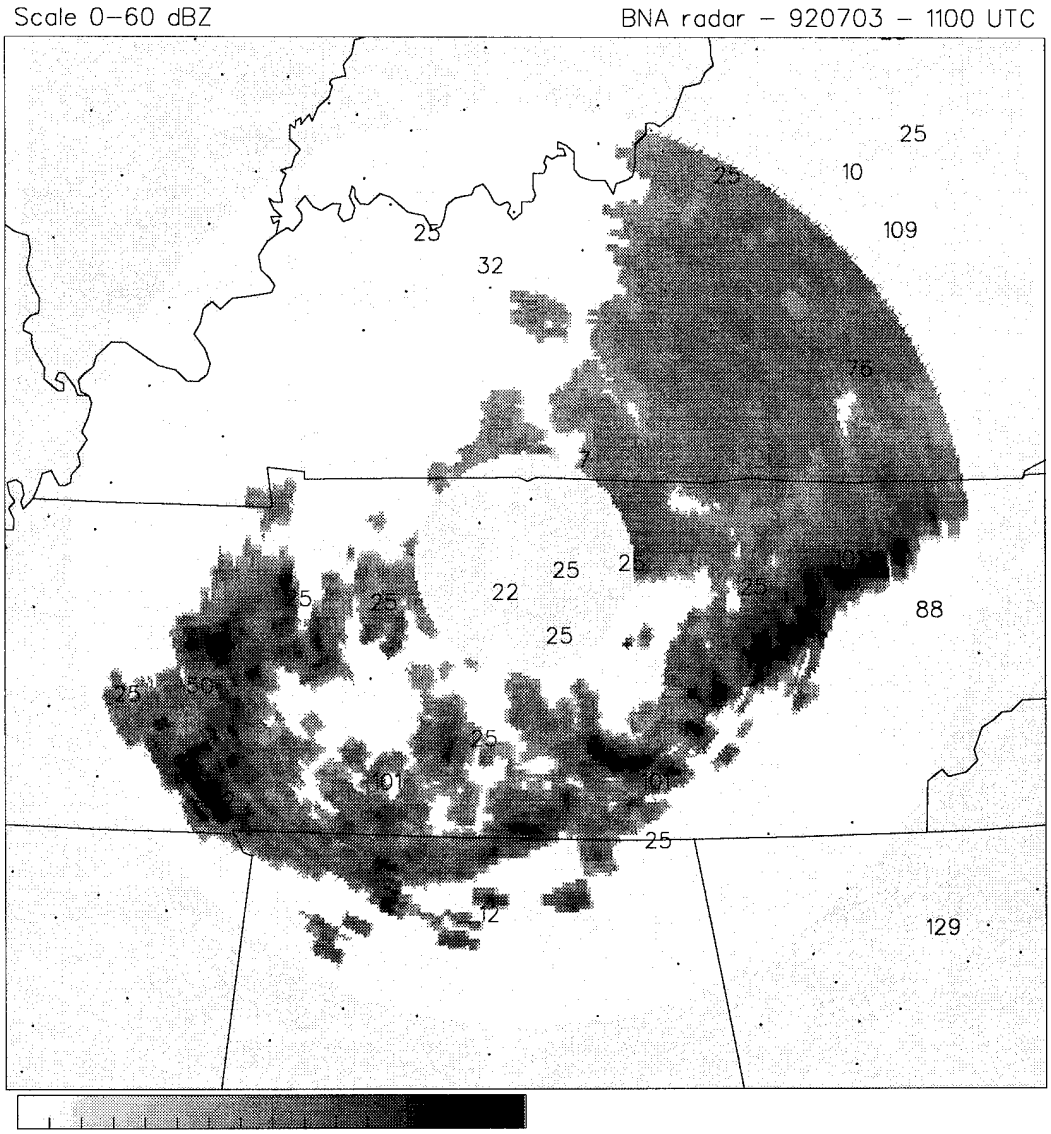


Figure 20. Raingage totals (mm x 10) for the hour ending 1300 UTC 3 July 1992. Grayscale overlay is (a) corrected reflectivity (dBZ) observed by the BNA radar at 1300 UTC (b) reflectivity associated with the the Purdue-EV algorithm rain rate from the satellite overpass at 1309 UTC.

(b)

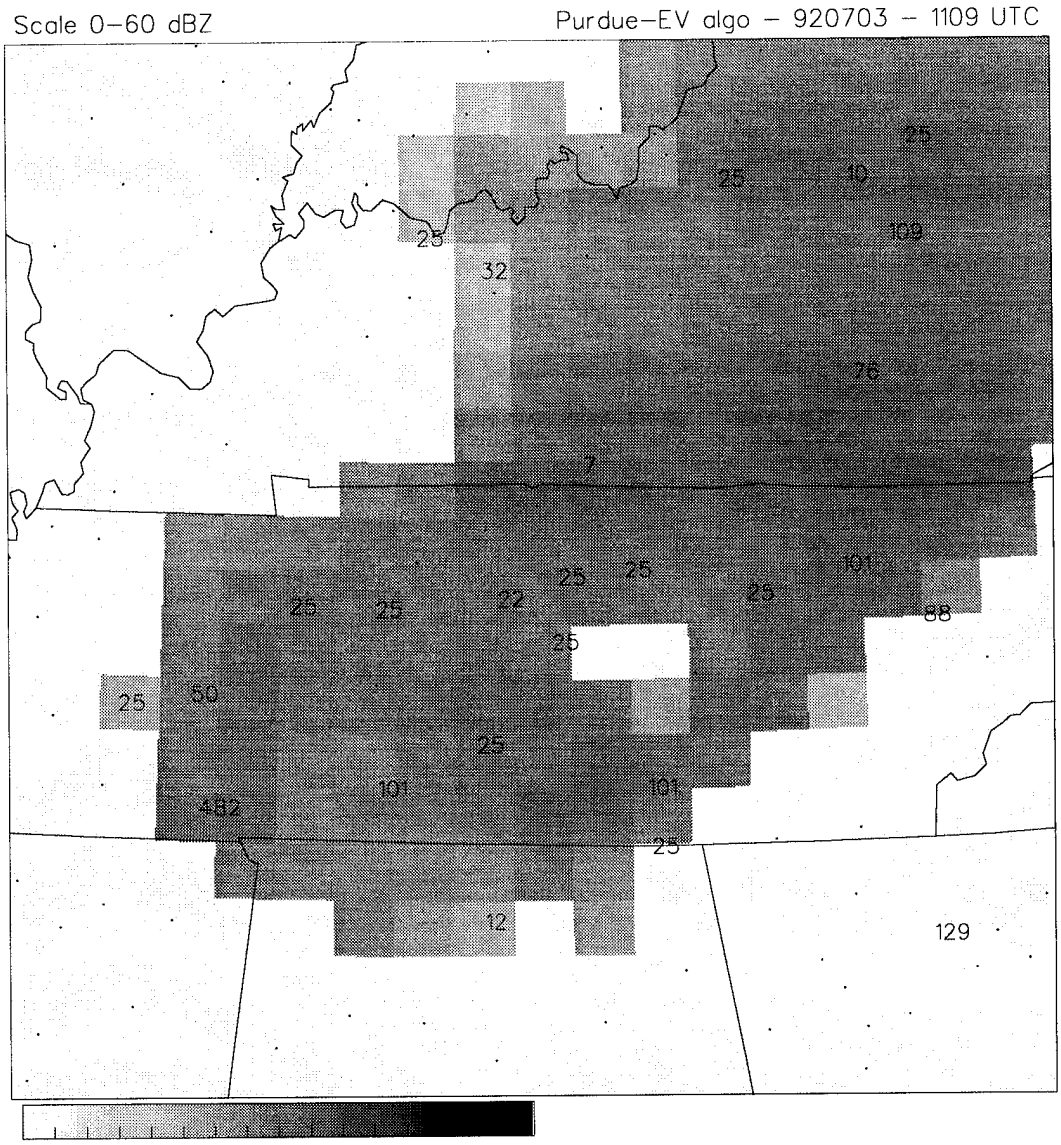


Figure 20 (continued).

The similarity of the Purdue algorithms is striking. Since a common reference for all three is the averages for a non-precipitating sky, it is hypothesized that the "brains" to these algorithms lie in the negative departure from normal of the brightness temperatures. It appears that as long as the 85 and 37 GHz channels are used, the remaining channels add little to the overall performance.

The performance of the algorithms relative to one another is somewhat difficult to assess. Using correlation coefficients from the square-root-transformed radar data, it would appear the SRL algorithm is the best, then the Purdue algorithms, then GSFC. When compared with the gage data, the order is reversed. Using the maxima on the HSS plots, an indication of the optimum skill of the prediction of some rain rate, the ranking is SRL, Purdue algorithms, then GSFC for both gage and radar as verification. In any case, the Purdue algorithms are competitive with the others, though not provably superior to either.

6. Future Work

The testing and verification of these algorithms are by no means complete. The SRL snow flag was used to remove suspected snow cover from the calibration dataset, but at the cost of removing significant portions from the study area during the winter months. A more rigorous test for snow cover may better delineate areas of snow without unnecessarily removing areas from evaluation. Detection of precipitation over snow poses a major challenge, since the brightness temperatures of airborne snow and surface snow may be nearly identical. It may be possible to accumulate a time series of brightness temperatures for a point on the surface and then check for anomalous depressions from a shorter-term trend to detect precipitation.

The monthly non-precipitating-sky averages were computed using very simple thresholds to reject times when a water or snow surface was detected or when precipitation was apparent. Using any one of the precipitation algorithms as a basis for rejection may improve the monthly averages, and if one of the Purdue algorithms is used, an iterative approach might be taken to see if there is convergence towards some optimum set of values.

The use of 2-D HSS plots should be subjected to a more rigorous statistical proof of its validity. While the results for this study were quite useful (definitely more so than linear regression), there may be unanticipated problems for some data distributions. Since it gave reasonable results for some rather unreasonable distributions, the method's utility cannot be discarded simply because it has not been used elsewhere. The use of linear regression found elsewhere in the literature for similar distributions, similarly, does not validate that method's applicability here.

While the domain used in this study contained a variety of terrain types, the algorithms should be tested in other areas of the world. This study focused on CONUS because a

large quanitity of data was easily obtained for this area. The Algorithm Intercomparison Project (AIP) has several high-quality radar and raingage datasets covering western and central Europe and could be used for additional comparisons.

The NOWRAD data received at Purdue from WSI Corporation are being archived for future use. One subset of these data is a CONUS-wide composite of WSR-88D and older radars on an 8 km by 8 km grid that is available every 15 minutes. The most exciting aspect of these data is that the reflectivity values are far more rigorously quality controlled than for any other dataset based upon operational (non-research) radars. Rain rates derived from this dataset should have a much better correlation to true rainfall and should serve quite well as a verification tool. Unfortunately, there are no SSM/I data available at Purdue yet to coincide with these radar data since the archive was started quite recently (April 1995). Future work on calibration of SSM/I algorithms should be able to make excellent use of these high-quality radar data.

List of References

Adler, R. F., G. J. Huffman, and P. R. Keehn, 1994: Global tropical rain rate estimates from microwave-adjusted geosynchronous IR data. *Remote Sens. Rev.*, **11**, 125-152.

Arkin, P. A., and B. N. Meisner, 1987: Spatial and annual variation in the diurnal cycle of large scale tropical convective cloudiness and precipitation. *Mon. Wea. Rev.*, **115**, 1009-1032.

Battan, L. J.: 1973: *Radar observation of the atmosphere.* Univ. of Chicago Press, Chicago, IL, 324 pp.

Burgess, D., and P. S. Ray, 1986: Principles of radar. *Mesoscale Meteorology and Forecasting,* P. S. Ray, Ed., Amer. Meteor. Soc., Boston, MA, 85-117.

Grody, N. C., 1991: Classification of snow cover and precipitation using the Special Sensor Microwave Imager. *J. Geophys. Res.*, **96**, 7423-7435.

Hollinger, J., R. Lo, G. Poe, R. Savage, and J. Peirce, 1987*: Special Sensor Microwave/Imager User's Guide.* Naval Research Laboratory, Washington, DC, 177pp.

Johnson, R.A., and D. W. Wichern, 1992: *Applied Multivariate Statistical Analysis.* Prentice-Hall, Inc., Englewood Cliffs, NJ, 642pp.

Kummerow, C. D., and L. Giglio, 1994a: A passive microwave technique for estimating rainfall and vertical structure information from space. Part I. Algorithm description. *J. Appl. Meteor.*, **33**, 3-18.

------- , and --------, 1994b: A passive microwave technique for estimating rainfall and vertical structure information from space. Part II. Applications to SSM/I data. *J. Appl. Meteor.*, **33**, 19-34.

------- , R. A. Mack, and I. M. Hakkarinen, 1989: A self-consistency approach to improve microwave rainfall estimates from space. *J. Appl. Meteor.*, **28**, 869-884.

Lee, R.R., and J. E. Passner, 1993: The development and verification of TIPS: an expert system to forecast thunderstorm occurrence. *Wea. Forecasting*, **8**, 271-280.

McDonald, M., and R. E. Saffle, 1994: *Revised RADAP-II Archive Data User's Guide.* TDL Office Note 94-2, National Oceanic and Atmospheric Administration, Washington, DC, 87pp.

Mugnai, A., and E. A. Smith, 1988: Radiative transfer to space through a precipitating cloud at multiple microwave frequencies. Part I: Model description. *J. Appl. Meteor.*, **27**, 1055-1073

--------, --------, and G. J. Tripoli, 1993: Foundations for physical-statistical precipitation retrieval from passive microwave satellite measurements. Part II: Emission source and generalized weighting function properties of a time dependent cloud-radiation model. *J. Appl. Meteor.*, **32**, 17-39.

Neter, J., W. Wasserman, and M. H. Kutner, 1990: *Applied linear statistical models: regression, analysis of variance, and experimental designs.* Richard D. Irwin, Inc., Homewood, IL, 1181 pp.

Petty, G. W., 1990: *On the response of the Special Sensor Microwave/Imager to the marine environment: implications for atmospheric parameter retrievals.* Ph. D. dissertation, Department of Atmospheric Sciences, Univ. of Washington, Seattle, WA, 291 pp.

--------, 1995: The status of satellite-based rainfall estimation over land. *Remote Sens. Environ.*, **51**, 125-137.

Spencer, R. W., 1986: A satellite passive 37 GHz scattering-based method for measuring oceanic rain rates. *J. Climate Appl. Meteor.*, **25**, 754-766.

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

Weinman, J. A., and P. J. Guetter, 1977: Determinations of rainfall distributions from microwave radiation: measured by the Nimbus 6 ESMR. *J. Appl. Meteor.*, **16**, 437-442.

Weng, F., R. R. Ferraro, and N. C. Grody, 1994: Global precipitation estimates using Defense Meteorological Satellite Program F10 and F11 special sensor microwave imager data. *J. Geophys. Res.*, **99**, 14,493-14,502.

Wentz, F. J., 1988: *User's Manual SSM/I Antenna Temperature Tapes.* RSS Technical Report 032588, Remote Sensing Systems, Santa Rosa, CA, 23pp.

Wilheit, T., R. Adler, S. Avery, E. Barrett, P. Bauer, W. Berg, A. Chang, J. Ferriday, N. Grody, S. Goodman, C. Kidd and D. Kniveton, C. Kummerow, A. Mugnai, W. Olson, G. Petty, A. Shibata, and E. Smith, 1994: Algorithms for the retrieval of rainfall from passive microwave measurements. *Remote. Sens. Rev.*, **11**, 163-194.