A Statistical Technique for Determining Rainfall over Land Employing Nimbus 6 ESMR Measurements

EDWARD RODGERS

NASA/GSFC, Greenbelt, MD 20771

HONNAPPA SIDDALINGAIAH[1]

Computer Sciences Corporation, Silver Spring, MD 20910

A.T.C. CHANG AND THOMAS WILHEIT

NASA/GSFC, Greenbelt, MD 20771

(Manuscript received 6 December 1978, in final form 10 May 1979)

ABSTRACT

At 37 GHz, the frequency at which the Nimbus 6 Electrically Scanning Microwave Radiometer (ESMR 6) measures upwelling radiance, it has been shown theoretically that the atmospheric scattering and the relative independence on electromagnetic polarization of the radiances emerging from hydrometeors make it possible to monitor remotely active rainfall over land. In order to verify experimentally these theoretical findings and to develop an algorithm to monitor rainfall over land, the digitized ESMR 6 measurements were examined statistically.

Horizontally and vertically polarized brightness temperature pairs (T_H,T_V) from ESMR 6 were sampled for areas of rainfall over land as determined from the rain recording stations and the WSR 57 radar, and areas of wet and dry ground (whose thermodynamic temperatures were greater than 5°C) over the southeastern United States. These three categories of brightness temperatures were found to be significantly different in the sense that the chances that the mean vectors of any two populations coincided were less than 1 in 100. Since these categories were significantly different, classification algorithms were then developed. Three decision rules were examined: the Fisher linear classifier, the Bayesian quadratic classifier, and a non-parametric linear classifier. The Bayesian algorithm was found to perform best, particularly at a higher confidence level. An independent test case analysis showed that a rainfall area delineated by the Bayesian classifier coincided well with the synoptic-scale rainfall area mapped by ground recording rain data and radar echoes.

1. Introduction

Precipitation is a fundamental meteorological parameter and it functions as an indicator, determinant or component of the distribution and amount of latent heat release which is critical to the understanding of storm and global atmospheric energetics and of the total hydrological cycle. The ability to monitor the coverage and movement of rain over land areas is important because of the direct impact of rain on crop production and also its influence on insect breeding areas and migration (Idso *et al.*, 1975). Moreover, the destructive effects due to heavy rainfall could be reduced by advance warnings furnished by satellites that map regions of heavy rain.

Since the advent of the polar orbiting and geosynchronous satellites, quantitative techniques have been developed to estimate rainfall indirectly. Estimations of rainfall have been made by correlating rain rate and amount with either cloud cover and type, cloud brightness or cloud temperature, utilizing visible and infrared sensors on board these satellites (Barnett, 1970, 1973; Martin and Scherer, 1973; Martin *et al.*, 1975; Follansbee and Oliver, 1975; Scofield and Oliver, 1977; Griffith *et al.*, 1978). However, all these techniques suffer from being only indirectly related to rainfall.

The microwave technique developed by Wilheit *et al.* (1977) has a direct physical relationship with rain rate but only over ocean areas. This technique establishes a relationship between rain rate in the dynamic range of 1–20 mm h^{-1} and brightness temperatures (T_B) measured by the Electrically Scanning Microwave Radiometer on board Nimbus 5 (ESMR 5), which senses at 19.35 GHz upwelling radiation emitted by the earth and its atmosphere.

Meneely (1975) demonstrated that rainfall rate and coverage cannot be delineated using ESMR 5 mea-

[1] Present affiliation: OAO Corp., Beltsville, MD 20705.

0021-8952/79/080978-14$07.50

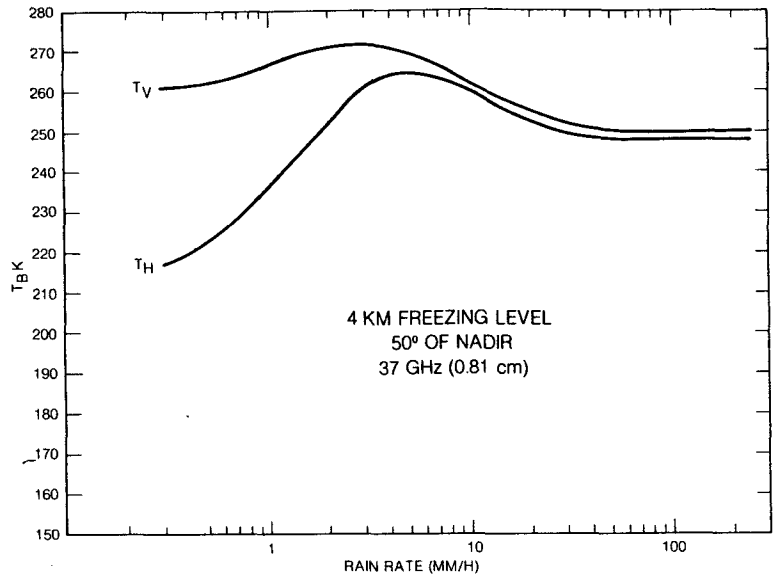

FIG. 1. Computed horizontally and vertically polarized brightness temperature at 37.0 GHz as a function of rain rate.

surements over land areas. This is because the rain has only a weak effect on the upwelling T_B from the land and the effect of soil moisture is comparable. Thus, although rain-like pattern can be discerned in the data, they correspond to both active rain areas and areas with moist soil. McFarland and Blanchard (1977), however, did demonstrate that rain amounts over land could be estimated indirectly by monitoring temporal changes in the ESMR 5 T_B.

Savage and Weinman (1975) and Savage *et al.* (1976) demonstrated theoretically that at 37.0 GHz (the frequency at which the Nimbus 6 ESMR sensor measures upwelling radiance) the scattering by hydrometeors is strong enough to provide a qualitative estimate of rain coverage over land. Furthermore, Weinman and Guetter (1977) demonstrated from a theoretical consideration that the upwelling radiation at 37.0 GHz emerging from rain clouds was essentially unpolarized and therefore was in contrast with the radiation emanating from wet surface background. According to the electromagnetic theory, if the emissivity of a surface is reduced by increasing its dielectric constant (as by adding moisture), then the emissivity will be highly polarized when the surface is viewed obliquely. These results are demonstrated in Fig. 1 which displays theoretically calculated bipolarized 37.0 GHz T_B at 40° incidence angle with the earth surface for a given rain rate. These T_B's were derived from a radiative transfer model with Lambertian reflection (Born and Wolf, 1975) from land surfaces at a thermodynamic temperature of 229.1 K and with a fixed dielectric constant and an atmospheric freezing level at 4 km (Wilheit *et al.*, 1977). It is seen from this figure that as rain rate increases (beyond 4 mm h^{-1}) T_B decreases due to strong backscattering by the large raindrops. Also, the polarization difference becomes smaller. Moreover, Hall *et al.* (1978) inferred theoretically that information analogous to that provided by the National Weather Service radar summary charts can be produced when both ESMR 6 and the Temperature Humidity Infrared Radiometer (THIR) 11.5 μm data on board Nimbus 6 are used.

Thus, the sum and substance of these theoretical investigations is that the obliquely viewed 37 GHz radiation emitted by wet soil surfaces is polarized ($T_V > T_H$), whereas radiation emanating from dry land or heavy rainfall areas is essentially unpolarized ($T_V \approx T_H$). Moreover, T_B's upwelling from dry land areas are distinguishably higher than those from heavy rainfall areas or wet land surfaces. Hence, according to these theoretical conclusions rainfall over land can be at least qualitatively delineated and therefore its coverage and movement can be monitored irrespective of the land background by employing 37 GHz measurements from the ESMR on board Nimbus 6. Quantitative measurement of rainfall over land using a 37 GHz radiometer, however, appears less promising.

It is the purpose of this paper to substantiate the above conclusions and to arrive at an algorithm for the detection of rain over land by statistically analyzing ESMR 6 data. This statistical analysis will be performed by first sampling three categories of ESMR 6 T_B's (representing rain over land, wet land surfaces without rain and dry land surfaces), then testing these populations for uniqueness and separability, and finally developing a classification algorithm to delineate rain over land.

TABLE 1. Dates of synoptic rain cases used to develop ESMR 6 classification algorithms.

Case	Date	Time (GMT)
1	31 Jul 1975	1620
2	4 Aug 1975	1635
3	1 Oct 1975	1700
4	7 Nov 1975	1700
5	12 Nov 1975	1700
6	29 Dec 1975	1717
7	3 Jan 1976	1715
8	6 Jan 1976	1655

2. The ESMR 6 system

The ESMR 6 system flown aboard Nimbus 6 (Wilheit, 1975) receives the thermal radiation upwelling from the earth's surface and atmosphere in a 250 MHz band centered at 37 GHz. The antenna beam scans electrically an arc of 70° in 71 steps ahead of the spacecraft along a conical surface with a constant earth incidence angle of 40° every 5.3 s. The nominal resolution is 20 km crosstrack and 45 km downtrack. The instrument measures both horizontal and vertical polarization components by using two separate radiometric channels. The data are calibrated using warm (instrument ambient) and cold (cosmic background) inputs to the radiometer.

Examination of the data revealed two problems affecting calibration, apparent modulation of the loss of the antenna during the orbit and excess noise when measuring the radiation from the warm calibration load. The modulation of the antenna loss was found to be consistent with respect to sun angle. It showed its maximum rate of change as the spacecraft entered the sunlight with the instrument facing the sun and fairly rapid changes whenever the spacecraft was in the sunlight. The changes were rather gradual whenever the spacecraft was in darkness. The most reasonable explanation seems to be thermomechanical warping of the radome modulating the coupling among the radiating elements in the antenna. An empirically derived correction as a function of beam position has been applied to the data to correct the biases in the data due to antenna losses. No cause is easily discerned for excess noise in the warm calibration load However, the data were rejected whenever the warm calibration load was too high (>310 K) or too low (<290 K). Otherwise, an empirical correction was applied to the data to mitigate excessive noise due to warm calibration load. It was found in a more subtle study by Wilheit (1978), where the effects of wind on ocean surface emission at 37 GHz were estimated, that the residual errors of these problems were excessive whenever the instrument was in the sun. Therefore, only nighttime data were used for that study.

The T_B as observed from the satellite is dependent upon the emission from the earth's surface modified by the intervening atmosphere. The emissivity, a function of the dielectric constant, is variable over land surfaces (depending on vegetation, soil type, soil moisture, etc.) and generally is large (~ 0.9). In rain situations three constituents contribute significantly to the absorption: molecular oxygen (Meeks and Lilley, 1963), water vapor (Staelin, 1966) and liquid water droplets (Mie, 1908; Gunn and East, 1954). Water droplets contribute more significantly to absorption and reemittance than the other constituents and are the only source of scattering at this frequency. Ice crystals are essentially transparent at this frequency.

3. Data sampling

Simultaneous ground station and radar measurements of rain and ESMR 6 T_B were needed in order to develop an algorithm which classified a given ESMR 6 instantaneous field of view (IFOV) as rain over land, dry land surface or wet land surface. Eight daytime synoptic-scale rainfall cases over the southeastern United States were used where surface rainrate data taken from stations reporting hourly rainfall amounts and from the WSR 57 radar coincided with Nimbus 6 overpass to within 5 min. The surface temperature in each of these cases was not less than 5°C. Rain areas were sampled within areas delineated as rain by either the WSR 57 radar (rain rates $\geqslant 2.5$ mm h^{-1}) and/or the stations reporting hourly rainfall amounts. The dates and time of the occurrence of these cases are given in Table 1. Wet land surfaces were sampled upwind and adjacent to the raincells observed on the WSR 57 radar and dry land surfaces were sampled over areas where rain had not fallen within a 24 h period previous to the Nimbus 6 pass.

Fig. 2 illustrates the sampling technique. The figure shows the ESMR 6 horizontally polarized T_B's (K) measured at approximately 1655 GMT 6 January 1976 together with rainfall as delineated by the WSR 57 radar (located at Waycross, Georgia at 1700 GMT) and by stations reporting hourly rainfall

TABLE 2. Elementary statistics of sampled data (surface temperature $\geqslant 5$°C).

Sample size (N)	Rain area 216		Dry ground 189		Wet soil 66	
	T_{HR}	T_{VR}	T_{HD}	T_{VD}	T_{HW}	T_{VW}
Mean (μ)	254.53	260.98	271.46	278.18	252.05	268.86
Mean brightness temperature difference	6.45		6.72		16.81	
Standard deviation (d)	7.21	5.81	6.18	7.20	9.41	7.64
Sample correlation coefficient between T_H and T_V (ρ)	0.55		0.37		0.82	

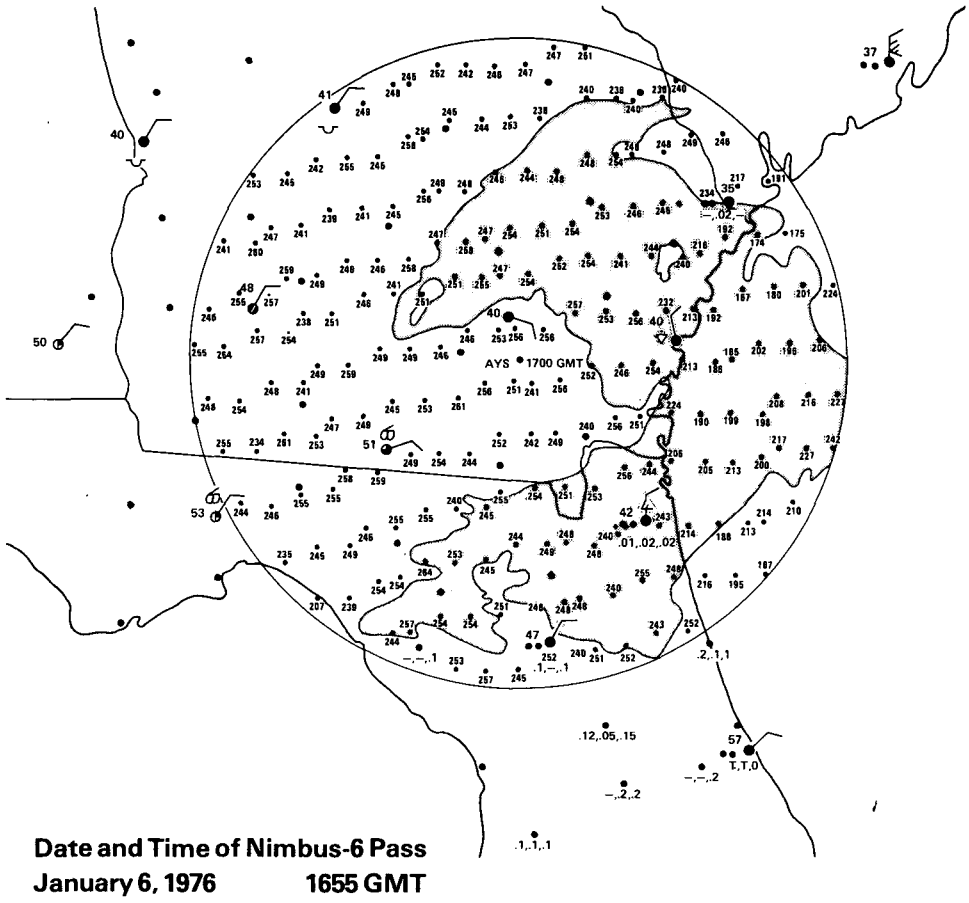

FIG. 2. ESMR 6 horizontally polarized T_B (1655 GMT 6 January 1976) superimposed on the PPI WSR 57 radar image at Waycross, Georgia (1700 GMT 6 January 1976)

amounts. The ESMR 6 T_B's are within the field of view of the radar where the circle shows the outer bounds of the PPI image at a 232 km radius. The shaded area represents rain (rain rate ≥ 2.5 mm h^{-1}). The large dots are hourly rain recording stations where rain amounts (in inches) for hours ending at 1700, 1600 and 1500 GMT are displayed according to model in the figure. If no rain has fallen during that period, no measurements are shown. Station models reporting temperature, present weather, cloud type and amount, and wind direction and speed for 1800 GMT are also given. The small dots locate the center of the ESMR 6 footprints. For this case, ESMR 6 T_B's representing rain over land was sampled within the shaded area. The T_B's representing wet land surfaces were sampled southwest of the shaded area since the rain area was moving northeast, and T_B's representing dry land surfaces were sampled over western Georgia where rain had not fallen within 24 h of the Nimbus 6 pass. It should be noted that the horizontally polarized T_B's over the radar echoes and the wet or dry land areas outside the radar echoes are all about the same. The reason for this non-variability of T_B's is that the rain in this case was light and did not significantly influence the horizontally polarized T_B's at 37 GHz.

4. Statistical analysis

Elementary statistics of the total sampled data (ESMR 6 measurements where surface thermodynamic temperatures were greater than 5°C) are presented in Table 2. The table gives for each category the sample size, the mean and standard deviation of the horizontally and vertically polarized T_B, and the correlation and the mean difference between horizontally and vertically polarized T_B's. These data are also shown as a scatter plot in Fig. 3. In this figure the C represents the mean of the population and each frequency concentration ellipse encompasses 68% (one

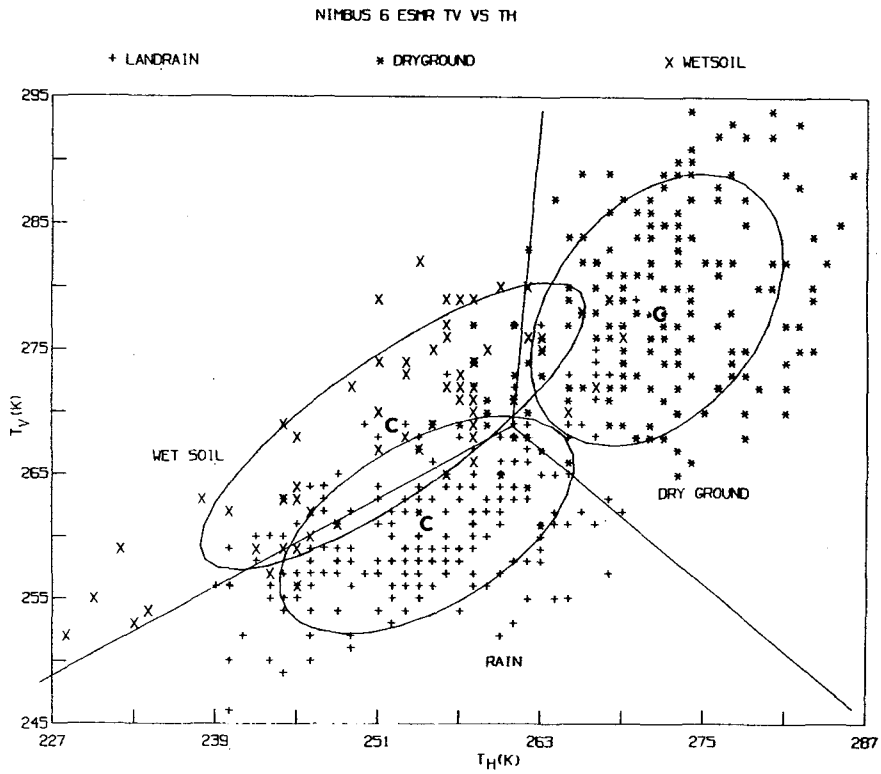

FIG. 3. Vertically polarized vs horizontally polarized EMSR 6 T_B for each sampled category (rain over land, and wet and dry land surfaces)

standard deviation) of the data within the population. The ellipses reveal the extent of scattering of data from each population, the correlation between the dual polarization T_B's (T_H and T_V) within each population (the higher the correlation the larger the eccentricity of the ellipse), and the extent of overlap among the populations. The three concurrent lines drawn in this figure are the Fisher (1938) linear discriminant lines which separate two-by-two the rain over land area (S_R), the dry land surface (S_D), and the wet land surface (S_W) populations represented by the T_B pairs (T_H,T_V).

It can be seen from Fig. 3 and Table 2, that T_B's from rain areas over land are colder than those T_B's from dry land surface areas. Further, the difference between the mean horizontally and vertically polarized T_B's from rain areas over land (6.45 K) is much smaller than that for wet land surfaces (16.81 K). This is in accordance with theoretical findings that microwave radiation emerging from hydrometeors is essentially unpolarized (Weinman and Guetter, 1977) whereas radiation emanating from wet land surfaces is polarized. It is also seen from Fig. 3 that the largest overlap occurs between the data obtained from rainfall areas and wet land surfaces. The reason for this is that sometimes in sampling rain over land the total upwelling radiance received by the radiometer contains a direct surface contribution. This may occur when an IFOV of the ESMR 6 measurement is partially filled with moderate to heavy rain or when it is completely filled with light rain (background being wet land surface). Consequently, the T_B's for each category are somewhat similar, thus producing the overlap between rain over land and wet land surface classes.

Since the surface emission is given by ϵT_S, where ϵ is the surface emissivity and T_S the surface thermodynamic temperature, there is an influence of T_S on ESMR 6 measured dry land surface T_B. A decrease in T_S results in a decrease in T_B from dry ground and consequently, the T_B contrast between dry land surfaces and rain over land will also decrease. These effects can be observed in Figs. 4 and 5 and in statistical Tables 3 and 4. The figures and tables are identical to Fig. 3 and Table 2, respectively, except that Fig. 4 and Table 3 correspond to sample cases where the surface thermodynamic temperatures were above 15°C, while Fig. 5 and Table 4 correspond to cases where the surface thermodynamic temperatures were between 5° and 15°C. It is clear from Fig. 5 and Table 4 that rain over land is difficult to delineate from dry land surfaces when the surface thermodynamic temperature is below 15°C. Since the populations in Fig. 5 cannot be separated, the Fisher linear discriminant lines are not drawn.

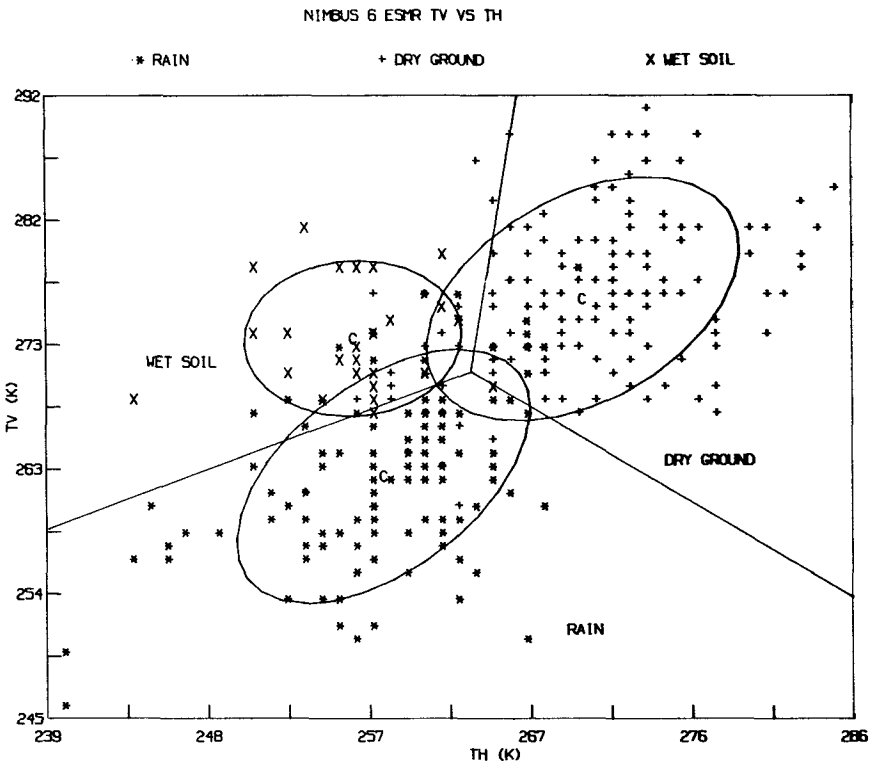

FIG. 4. As in Fig. 3 except for surfaces whose thermodynamic temperatures are greater than 15°C.

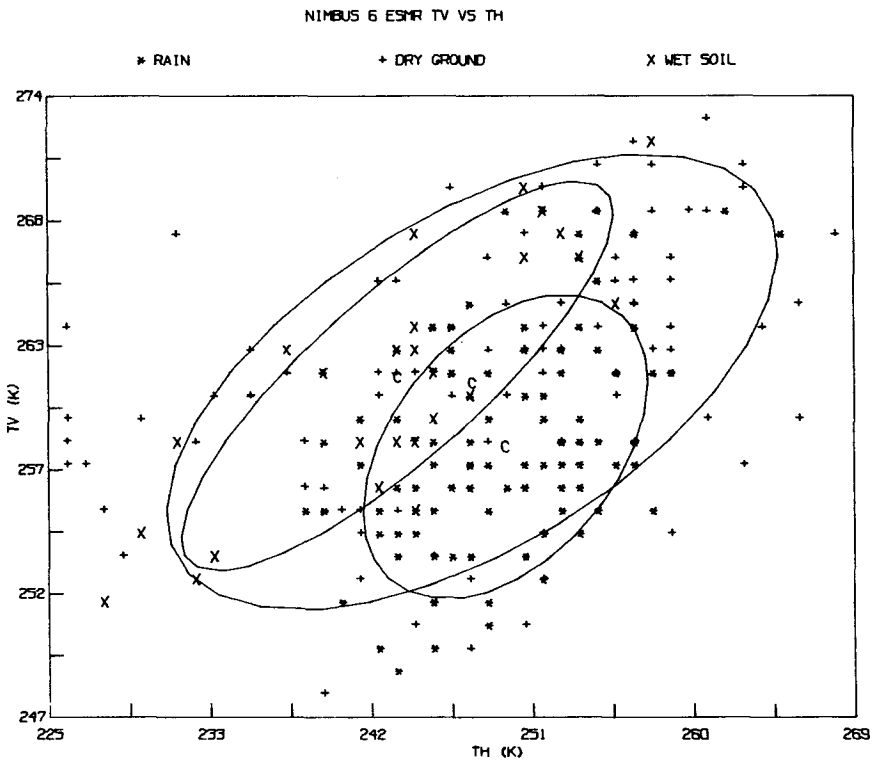

FIG. 5. As in Fig. 3 except for surfaces whose thermodynamic temperatures are between 5° and 15°C.

TABLE 3. Elementary statistics of sampled data (surface temperature >15°C).

Sample size (N)	Rain area 112		Dry ground 145		Wet soil 26	
	T_{HR}	T_{VR}	T_{HD}	T_{VD}	T_{HW}	T_{VW}
Mean (μ)	258.57	263.22	270.22	276.61	256.81	273.62
Mean brightness temperature difference	4.65		6.39		16.81	
Standard deviation (d)	5.70	6.36	6.07	6.16	4.22	3.90
Sample correlation coefficient between T_H and T_V (ρ)	0.53		0.42		0.05	

Fig. 6 displays the marginal densities (histograms) of the sampled horizontally and vertically polarized T_B's from the three populations. Table 5 presents the results of the chi-square test (Cochran, 1952) performed to validate the normal distribution of the data. Since each observed chi-square value in Table 5 is comparable to the corresponding critical (table) value at 1%, it is assumed that each marginal distribution of the data is Gaussian. Therefore, it is reasonable to assume that the data from each of the populations S_R, S_D or S_W satisfy the bivariate Gaussian density distribution

$$f(\mathbf{x}) = \frac{1}{2\pi|c|^{\frac{1}{2}}} \exp\left[-\tfrac{1}{2}(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}} c^{-1}(\mathbf{x}-\boldsymbol{\mu})\right], \qquad (1)$$

where $\mathbf{x}$ is the two-dimensional column vector (T_H, T_V), $\boldsymbol{\mu}$ is the mean of $\mathbf{x}$, c is the covariance matrix of the population, c^{-1} is the inverse of c, $|c|$ is the determinant of c, and $(\mathbf{x}-\boldsymbol{\mu})^{\mathrm{T}}$ is the transpose of $(\mathbf{x}-\boldsymbol{\mu})$. The $\boldsymbol{\mu}$ and $\mathbf{c}$ are estimated using the sampled data from each class. Then $\boldsymbol{\mu}$'s are provided by Table 2 and c, c^{-1} and $|c|$ by Table 6.

Prior to employing the data in Table 2 for the purpose of developing classification algorithms, the data were examined to verify whether the three populations were statistically distinguishable from one another. To accomplish this, an F (variance ratio) test, in terms of Hotelling's T^2 and Mahalanobis's D^2 (Kshirsagar, 1972), was performed to determine the

TABLE 4. Elementary statistics of sampled data (surface temperature 5–15°C).

Sample size (N)	Rain area 104		Dry ground 98		Wet soil 28	
	T_{HR}	T_{VR}	T_{HD}	T_{VD}	T_{HW}	T_{VW}
Mean (μ)	249.92	258.78	248.11	261.57	244.04	261.82
Mean brightness temperature difference	8.86		13.46		17.78	
Standard deviation (d)	5.16	4.40	11.08	6.58	7.87	5.59
Sample correlation coefficient between T_H and T_V (ρ)	0.43		0.58		0.84	

TABLE 5. Chi-square test for normality.

	Rain		Dry		Wet	
	T_H	T_V	T_H	T_V	T_H	T_V
Number of cells	8	8	8	9	6	5
Degrees of freedom	5	5	5	6	3	2
Table value of χ^2 at 0.01	15.09	15.09	15.09	16.81	11.34	9.21
Observed value of χ^2	14.57	28.17	18.99	11.59	8.93	10.33

significance of the differences between the means of any two classes. Then the simultaneous confidence intervals were estimated for these differences by Scheffé's procedure (Scheffé, 1959; Bennett, 1951).

Table 7 displays D^2 and T^2 as well as the computed and table (critical) values of F. The difference between the means of any two classes is highly significant since the observed value of F is much higher for each pair of classes than the corresponding critical (table) value of F at the 1% confidence level, i.e., the probability that the mean vectors of any two populations are identical is less than 1 in 100.

Table 8 shows the estimated confidence intervals. It can be seen that only the interval for the differences between the wet land surface and rainfall over land mean horizontal polarization T_B's contains zero. Therefore, the three populations are distinguishable from one another when the dual polarization information is taken into consideration. However, the lower bounds of the mean differences between rainfall over land and wet land surface T_B's is smaller than those of the other two pairs. This indicates that it will be more difficult to distinguish an area of rain over land from wet land surfaces.

TABLE 6. Covariance matrices of sampled data.

Rain area

	Matrix		Inverse	
	T_{HR}	T_{VR}	T_{HR}	T_{VR}
T_{HR}	52.23	23.02	0.0273	−0.0185
T_{VR}	23.02	33.93	−0.0185	0.0420

Covariance matrix determinant: 1242.4240

Dry ground

	Matrix		Inverse	
	T_{HD}	T_{VD}	T_{HD}	T_{VD}
T_{HD}	38.36	16.51	0.030	−0.010
T_{VD}	16.51	52.14	−0.010	0.022

Covariance matrix determinant: 1727.7496

Wet soil

	Matrix		Inverse	
	T_{HW}	T_{VW}	T_{HW}	T_{VW}
T_{HW}	90.39	59.73	0.034	−0.035
T_{VW}	59.73	58.28	−0.035	0.053

Covariance matrix determinant: 1700.3525

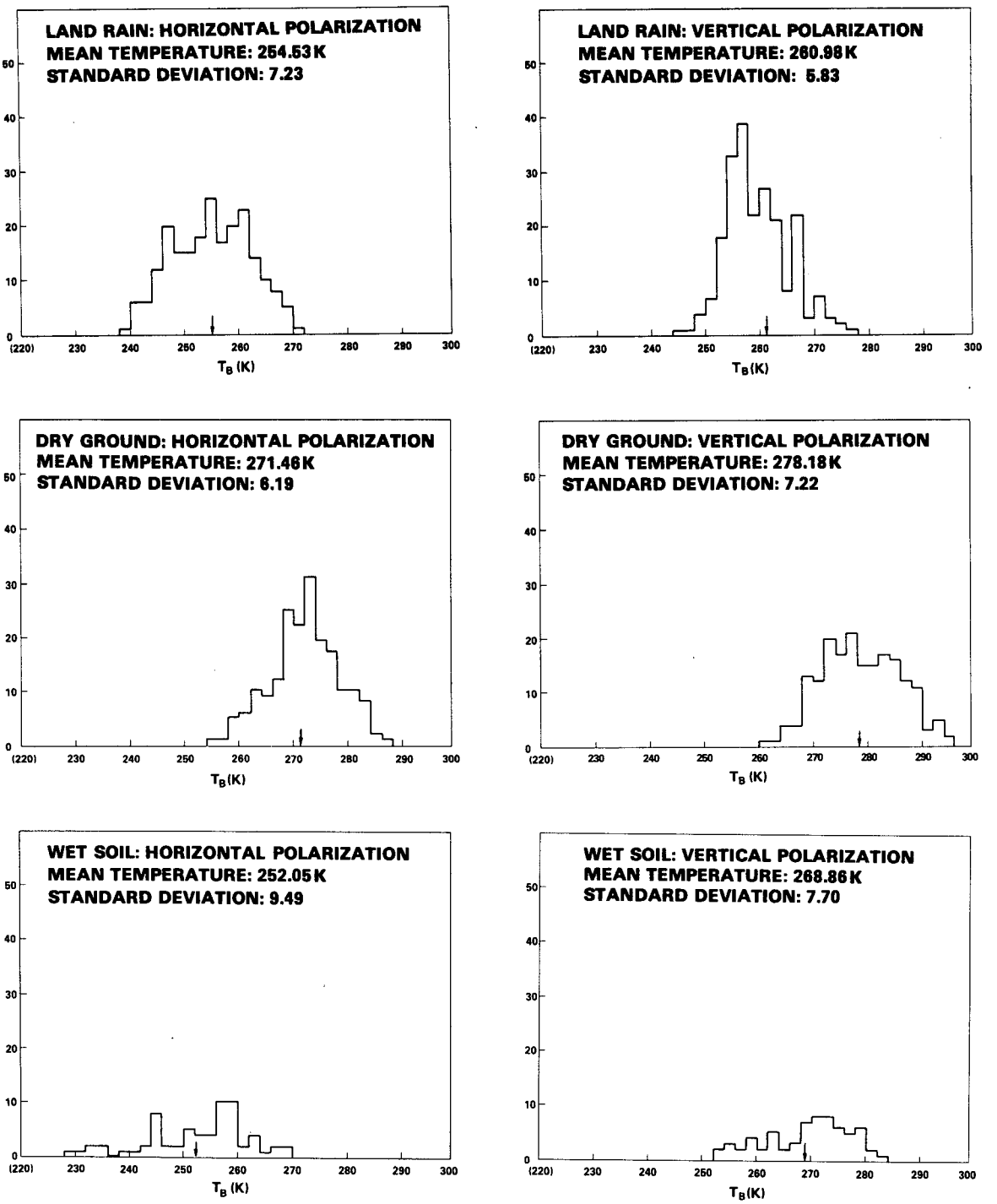

FIG. 6. Marginal densities (histograms) of the sampled horizontally and vertically polarized T_B's from the three populations.

5. Classification algorithms

Since the populations were found to be statistically distinguishable and satisfied the Gaussian frequency distribution, three classification techniques were considered with the purpose of developing an efficient and effective classification algorithm to detect and delineate active rainfall over land from dry and wet land surfaces. The three techniques are the Bayesian classifier, the Fisher linear discriminant classifier and a non-parametric linear discriminant classifier.

TABLE 7. Significance between means (F test).

	Mahalanobis's distance squared D^2	Hotelling's T^2	Observed variance ratio F	Table value of F at 1%
Rain vs dry	9.13	920.35	459.04	3.83
Dry vs wet	6.03	295.13	146.28	3.87
Rain vs wet	4.00	202.06	100.67	3.86

The Bayesian classifier is a parametric classifier (i.e., it assumes the functional form of the relevant density function). The non-parametric linear discriminant classifier does not assume a density function (Bond and Atkinson, 1972). The Fisher linear discriminant classifier may be either parametric or non-parametric (Fisher, 1938). All three methods are termed supervised in the sense that it is necessary to use known sample data for the various classes to train the algorithms. Algorithms were developed using all three classifiers and tested using independent data. It was found that the results from the Bayesian classifier were superior to the other two methods. Hence, only the Bayesian classification technique will be described in the following.

The Bayesian classifier is a Gaussian parametric maximum likelihood quadratic classifier which requires the knowledge of the *a priori* probabilities for the occurrence of each class (Duda and Hart, 1973; Fu *et al.*, 1969). It minimizes the average loss due to misclassification while assuming that each misclassification is equally costly.

It minimizes the conditional average loss

$$L(\mathbf{x},S_k)=\sum_{i=1}^{3}\lambda(S_k|S_i)P(S_i|\mathbf{x}), \tag{2}$$

where $\lambda(S_k|S_i)$ is the loss incurred when a measurement $\mathbf{x}=(T_H,T_V)$ actually belonging to class S_k is placed in class S_i and $P(S_i|\mathbf{x})$ is the *a priori* probability of the class S_i occurring having observed $\mathbf{x}$. The symmetrical loss function $\lambda(S_k|S_i)$ is given by

$$\lambda(S_k|S_i)=\begin{cases}0, & i=k\\ 1, & i\neq k\end{cases} \quad i,k=1,2,3. \tag{3}$$

Hence, all misclassifications are equally costly and Eq. (2) now reduces to

$$L(\mathbf{x},S_k)=1-P(S_k|\mathbf{x}), \tag{4}$$

TABLE 8. Simultaneous confidence intervals for differences between mean brightness temperatures representing rain (R), dry (D) and wet (W) areas.

Polarization	Rain vs dry	Dry vs wet	Wet vs rain
Horizontal	$14.88\leqslant\mu_{HD}-\mu_{HR}\leqslant 18.98$	$15.89\leqslant\mu_{HD}-\mu_{HW}\leqslant 22.93$	$-1.15\leqslant\mu_{HR}-\mu_{HW}\leqslant 6.11$
Vertical	$15.00\leqslant\mu_{VD}-\mu_{VR}\leqslant 19.20$	$6.04\leqslant\mu_{VD}-\mu_{VW}\leqslant 12.60$	$4.93\leqslant\mu_{VW}-\mu_{VR}\leqslant 10.83$

where $P(S_k|\mathbf{x})$ is the conditional probability that the class S_k to which $\mathbf{x}$ is assigned is correct.

The likelihood function $P(S_k|\mathbf{x})$ is given by Bayes' rule

$$P(S_k|\mathbf{x})=P(\mathbf{x}|S_k)P(S_k)/\sum_{j=1}^{3}P(\mathbf{x}|S_j)P(S_j), \tag{5}$$

where $P(\mathbf{x}|S_j)$ is the bivariate Gaussian probability density function of $\mathbf{x}$ given that $\mathbf{x}$ is in S_j, and $P(S_j)$ is the *a priori* probability of the class S_j occurring. Sample data sizes given in Table 2 provide the values of $P(S_j)$. They are 0.459, 0.401 and 0.140 for the classes S_R, S_D and S_W, respectively.

Since the loss given by Eq. (4) is to be minimized, the quadratic discriminant functions are

$$g_i(\mathbf{x})=P(S_i)P(\mathbf{x}|S_i), \quad i=1,2,3. \tag{6}$$

These functions, considering the relationship in Eq. (1), lead to the following decision rule. The measurement $\mathbf{x}$ belongs to the class S_k if

$$2\ln P(S_k)-\ln|c_k|-(\mathbf{x}-\boldsymbol{\mu}_k)^{\mathrm{T}}c_k^{-1}(\mathbf{x}-\boldsymbol{\mu}_k) > 2\ln P(S_i)-\ln|c_i|-(\mathbf{x}-\boldsymbol{\mu}_i)^{\mathrm{T}}c_i^{-1}(\mathbf{x}-\boldsymbol{\mu}_i) \tag{7}$$

for all $i\neq k$, where c_j and $\boldsymbol{\mu}_j$ are the covariance matrix and the mean vector of the class S_j. Then by substituting the relevant values into Eq. (7), one arrives at the following Bayesian algorithm. The pixel corresponding to the given vector (T_H,T_V) is rainfall over land, dry land surface or wet land surface, respectively, depending on which of the following values is the largest:

$$P_R(T_H,T_V)=-0.027T_H^2+0.038T_HT_V-0.042T_V^2+3.826T_H+12.250T_V-2094.097, \tag{8}$$

$$P_D(T_H,T_V)=-0.030T_H^2+0.020T_HT_V-0.022T_V^2+10.720T_H+6.811T_V-2412.165, \tag{9}$$

$$P_W(T_H,T_V)=-0.034T_H^2+0.070T_HT_V-0.053T_V^2-1.678T_H+10.846T_V-1261.721. \tag{10}$$

The quadratic function (Mahalanobis's distance squared)

$$Q_k(\mathbf{x})=(\mathbf{x}-\boldsymbol{\mu}_k)^{\mathrm{T}}c_k^{-1}(\mathbf{x}-\boldsymbol{\mu}_k) \tag{11}$$

is a measure of distance in probabilistic terms between $\boldsymbol{\mu}_k$ and $\mathbf{x}$ and has a chi-square distribution with two degrees of freedom (Scheffe, 1959). Therefore, a con-

TABLE 9. Probabilities of misclassification: Theoretical computation.

Known	Classified Rain	Dry	Wet
Rain	77.15	6.66	16.19
Dry	6.67	82.08	11.25
Wet	16.28	11.29	72.43

Average accuracy: 77.22%.

TABLE 10. Bayesian classification error matrix determined from sampled data.

Known	Classified Rain	Dry	Wet
Rain	89.35	6.02	4.63
Dry	7.41	91.53	1.06
Wet	27.27	15.15	57.58

Average accuracy: 79.49%.

fidence value F can be associated with each classified pixel. F is given by

$$F(\mathbf{x})=255\left\{1-\frac{[Q_K(\mathbf{x})]^{\frac{1}{2}}}{n_\sigma}\right\}, \tag{12}$$

where n_σ is the distance, in terms of standard deviation from the mean, to which zero confidence value is assigned.

This maximum likelihood decision rule selects one class from a set of predetermined classes (S_R, S_D, S_W in the present context) to which a pixel represented by $\mathbf{x}\equiv(T_H,T_V)$ most likely belongs. The associated confidence value, $F(\mathbf{x})$, measuring the distance in probabilistic terms of $\mathbf{x}$ from the mean point of the selected class, is calculated according to Eq. (12). The hypothesis that a pixel actually belongs to a selected class may be accepted or rejected based on this confidence value. If the confidence value is greater than a predetermined value (153.0, 178.5 and 204.0, respectively, in the case of 60, 70 and 80% confidence levels), the hypothesis is accepted and the pixel is put in the selected class; otherwise, the hypothesis is rejected and the pixel is put in the unknown class. The problem, of course, is to assign the appropriate values for thresholding the confidence value parameter.

6. Error analysis

An error estimate was made in order to evaluate quantitatively the performance of the Bayesian classification algorithm. The error rates were computed according to the asymptotic formulas given by Okamoto (1963), assuming that the populations satisfy the Gaussian distribution, have different means, and have the same covariance matrices. The results are shown in Table 9. Virtually all of the misclassification probability in each case was accounted for by the first term of the asymptotic expansion

$$\Phi(-\tfrac{1}{2}\Delta)=\frac{1}{(2\pi)^{\frac{1}{2}}}\int_{-\infty}^{-\frac{1}{2}\Delta}\exp\left[\frac{-z^2}{2}\right]dz, \tag{13}$$

where

$$\Delta^2=\frac{N_1+N_2-5}{N_1+N_2-2}D^2-\frac{2(N_1+N_2)}{N_1N_2}. \tag{14}$$

D is Mahalanobis's distance and N_1, N_2 are sample sizes of the populations under consideration. Only a small fraction is contributed by the rest of the terms. From Table 9 it is clear that the chance of incorrectly classifying wet land surfaces or dry land surfaces as rain over land is nearly 23%. But when a given pixel is classified as a raining area and each of the eight contiguous pixels that cluster around it is also classified as rain over land, then the chance of misclassification of that central pixel is reduced to 7.7×10^{-6}% assuming each pixel is independently classified.

Table 10 displays the actual probabilities that the Bayesian algorithm classify the sampled training data into the various populations is as indicated. The average accuracy is the mean of the diagonal elements of the corresponding error matrix, and these averages compare well with the estimated average.

Tables 11 and 12 show the estimated error matrices corresponding to data which came from land areas where the surface thermodynamic temperature was less than or greater than 15°C, respectively. It is apparent from the tables that the classifications are not definitive when the surface thermodynamic temperature is between 5° and 15°C.

7. Algorithm evaluation

A case not previously used in sampling was tested to verify qualitatively the performance of the Bayesian classification algorithm. This case consisted of a synoptic-scale rain pattern over the southeastern United States (14 September 1976) which was observed by the ESMR 6 sensor (surface thermodynamic temperature $\geqslant$15°C). Fig. 7 shows the rainfall area delineated by the WSR 57 radars and hourly rainfall reporting stations. The approximate time of the radar PPI images was 1630 GMT (within 5 min of the Nimbus 6 pass). The reporting times of the hourly

TABLE 11. Probabilities of misclassification: Theoretical computation (<15°C).

Known	Classified Rain	Dry	Wet
Rain	45.99	36.23	17.78
Dry	36.26	23.99	39.75
Wet	18.03	41.86	40.11

Average accuracy: 36.70%.

TABLE 12. Probabilities of misclassification: Theoretical computation (>15°C).

Known	Classified Rain	Dry	Wet
Rain	74.83	11.76	13.41
Dry	11.75	77.66	10.59
Wet	13.59	10.72	75.69

Average accuracy: 76.06%.

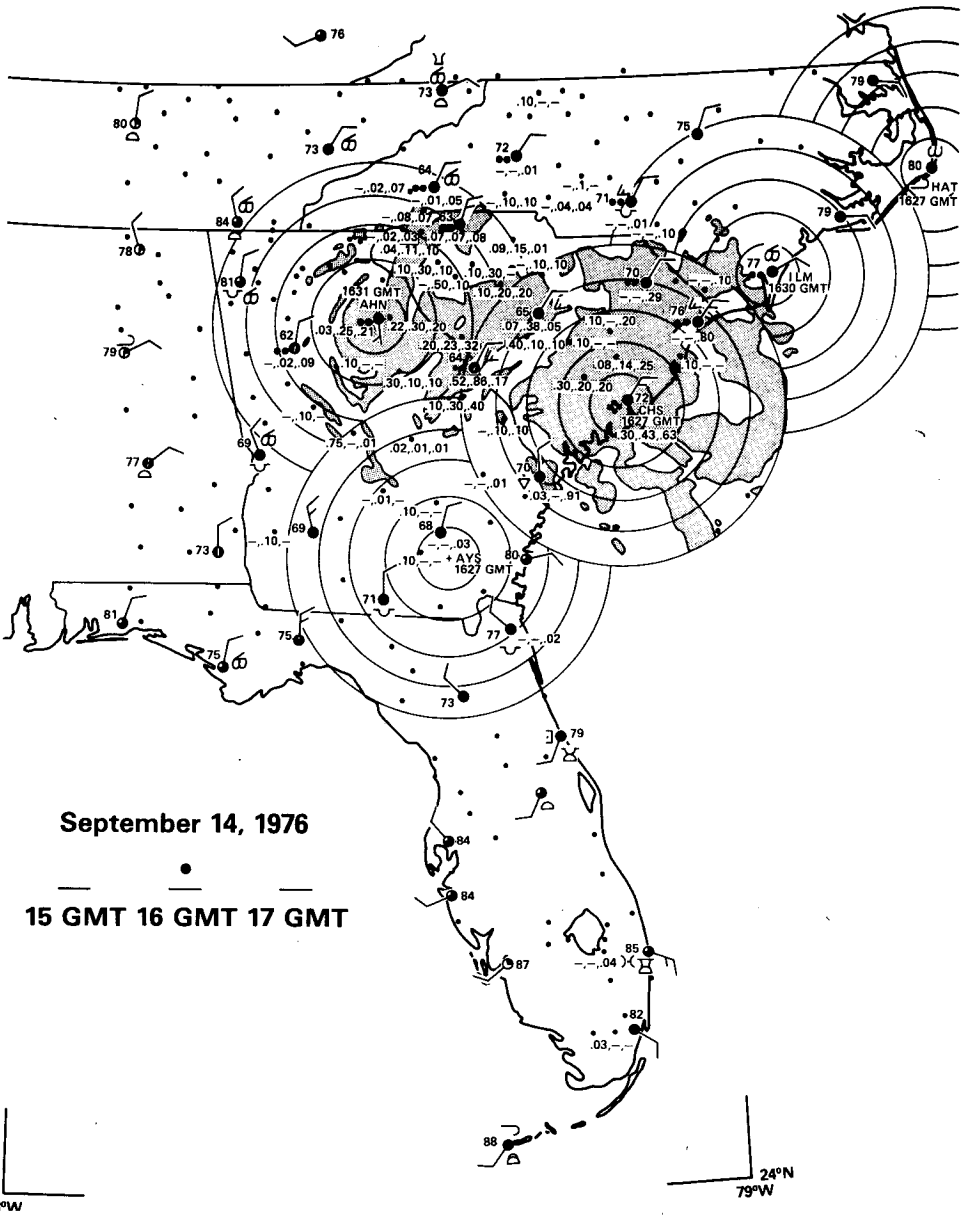

FIG. 7. Rainfall over the southeast United States as delineated by the WSR 57 radar and hourly rainfall reporting stations. Time of the data is approximately 1630 GMT 14 September 1976.

precipitation amounts were 1500, 1600 and 1700 GMT. The shaded area within the WSR 57 radar PPI range (232 km) is rainfall area with rain rates >2.5 mm h^{-1}. The radars were located at Waycross and Macon, Georgia; Charleston, South Carolina; and Wilmington and Cape Hatteras, North Carolina. Surface station data (present weather, temperatures, cloud type and amount, precipitation amount in three hours, and wind velocity and direction) were taken at 1800 GMT. Hourly rainfall is also shown. (See model in Fig. 7.)

The Bayesian (70 and 80% confidence) classification maps are seen in Figs. 8 and 9, respectively. Areas of clouds most likely producing rain are delineated by the Nimbus 6 THIR 11.5 μm channel where equivalent blackbody temperatures $T_{BB} \leqslant 270$ K (Shenk *et al.*, 1976). Rain areas in the absence of rain producing clouds are considered misclassifications. Regions only covered by clusters of contiguous pixels classified into a single individual class are shown, since the probability of misclassifying clusters is much less than that of a single pixel.

It is seen by comparing the two Bayesian classification maps at 70 and 80% confidence level (Figs. 8 and 9, respectively) with the map delineating observed rain (Fig. 7) that they agree well, particularly at the 80% confidence level. No attempt was made to verify

wet land surfaces. The 80% confidence Bayesian classification, however, did not delineate rain over eastern South Carolina as well as the 70% confidence classification. This is because the requirement of 80% confidence level is obviously too stringent.

The main discrepancies found between the ESMR 6 observed rainfall and ground observed rainfall is seen over North Carolina and southwestern Georgia. The rainfall indicated by ESMR 6 over North Carolina may be suspended liquid water in the clouds and/or virga ahead of the rain area (the area of rain was moving northeastward toward North Carolina). The ESMR 6 delineated rain over southwestern Georgia, which was upstream from the rain area, may be due to wet land surfaces produced by the rain that fell a few hours prior to the Nimbus 6 pass.

The Bayesian classification algorithm was applied to another test case (1645 GMT 27 August 1976, surface thermodynamic temperatures were ≥ 15°C) over the same geographical area as the previous case in order to determine whether the surface characteristics (vegetation, soil moisture and surface roughness) had influenced the classification performed in the previous case. During these periods the area was under the influence of a Bermuda high. Also there was a squall line located in southwestern Virginia and extending southwestward into Tennessee. Rainfall

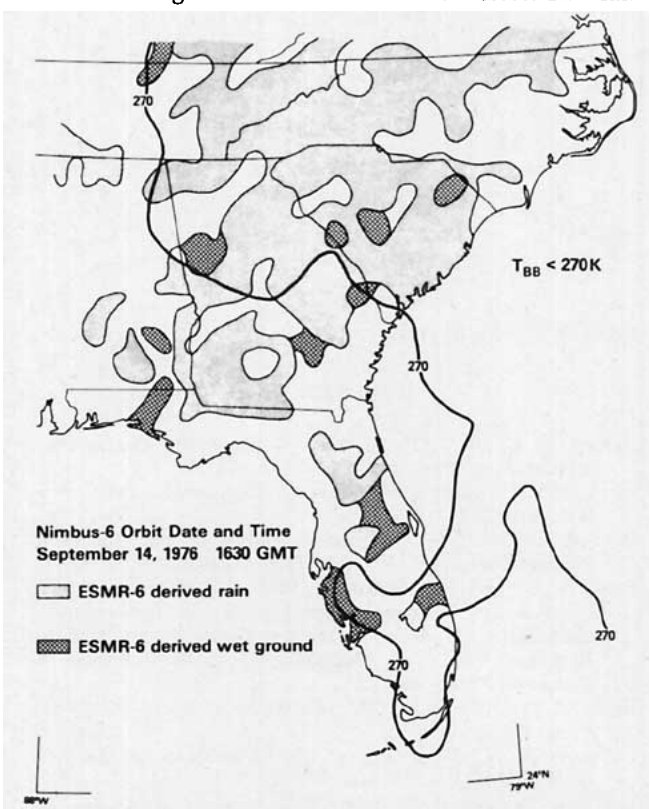

FIG. 8. ESMR 6 derived rainfall distribution using the Bayesian classifier with a confidence level of 70%. Time of Nimbus 6 pass was 1630 GMT 14 September 1976.

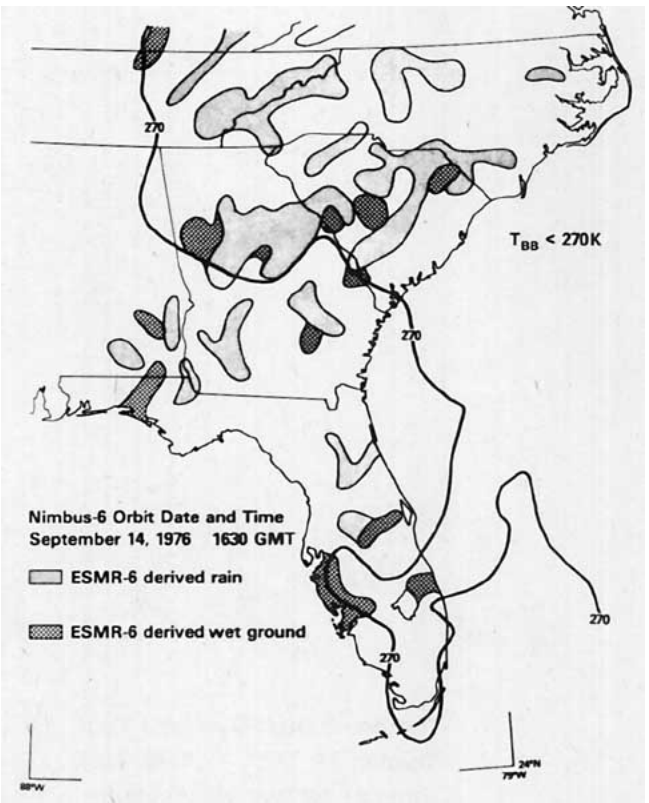

FIG. 9. As in Fig. 8 except with a confidence level of 80%.

was associated with this squal line and along the Gulf States. Fig. 10 shows the 80% confidence level Bayesian classification map superimposed over the surface station models. The reporting time for these stations was 1800 GMT. The figure shows that the only areas classified as rain over land were along the Gulf Coast and in eastern Tennessee. The regions in the previous case where the algorithm showed rainfall were classified as dry land surfaces. Hence, there were no influences of extraneous surface characteristics on the outcome of the previous case study.

However, contradicting results occurred when the Bayesian classification algorithm was applied to a nighttime Nimbus 6 pass over the same geographical area (0525 GMT, 13 September 1976) where surface thermodynamic temperatures were >15°C and there was no synoptic-scale rainfall. Almost all pixels were classified by the algorithm as rain over land. An examination of the ESMR 6 vertically polarized T_B's showed that the temperatures were below 0°C. Since calibration of the nighttime data is better than that of daytime data (Wilheit, 1978), this anomaly may be attributed to the changes in the surface emissivity caused by the presence of dew on the vegetation. The 0600 GMT National Weather Service map indicated that the conditions were ideal for the formation of dew. A large anticyclone centered over Virginia produced clear skies, winds less than 5 kt and dew

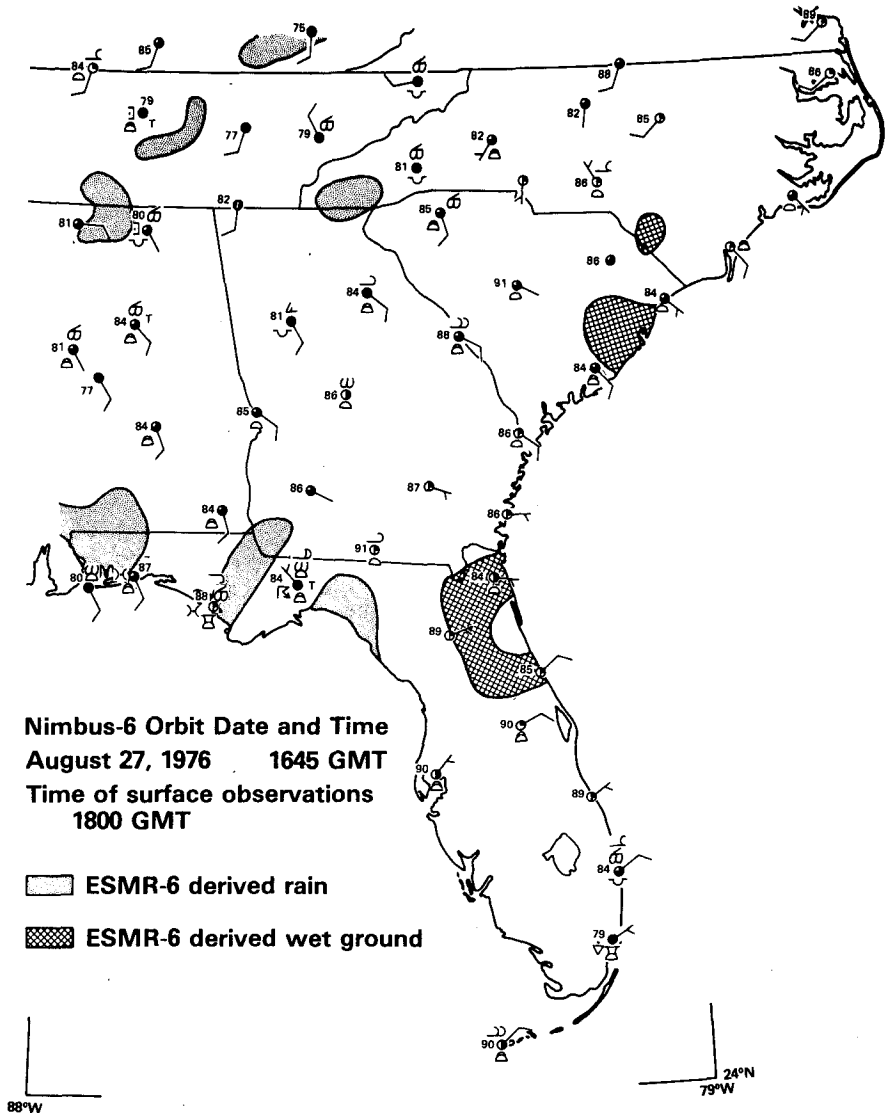

FIG. 10. As in Fig. 9 except at 1645 GMT 27 August 1976.

point temperature differences of less than 3°C over the majority of the reporting stations in the southeast United States. Therefore, the classification algorithm trained by data sampled from Nimbus 6 daytime passes can be employed only when dew is absent.

8. Conclusion

Statistical analyses were performed on the sampled ESMR 6 data for the purpose of delineating rainfall areas over land from dry and wet land surfaces. It was found that synoptic-scale rainfall over land, where surface thermodynamic temperatures were greater than 15°C and where the vegetation was not covered with dew, could be delineated despite the large ESMR 6 IFOV. However, there was some ambiguity in distinguishing between rainfall over land and wet land surfaces.

REFERENCES

Barrett, E. C., 1970: The estimation of monthly rainfall from satellite data. *Mon. Wea. Rev.*, **98**, 322–327.

——, 1973: Forecasting daily rainfall from satellite data. *Mon. Wea. Rev.*, **101**, 215–222.

Bennett, B. M., 1951: Note on a solution of the generalized Behrens-Fisher problem. *Ann. Inst. Statist. Math.*, **2**, 87–90.

Bond, A. D., and R. J. Atkinson, 1972: An integrated feature selection and supervised learning scheme for fast computer classification of multispectral data. *Remote Sensing of Earth Resources*, Vol. 1, F. Shahrokhi, Ed., The University of Tennessee Press, 645-672.

Born, M., and E. Wolf, 1975: *Principles of Optics*. Pergamon Press, 182 pp.

Cochran, W. G., 1952: The χ^2 test of goodness of fit. *Ann. Math. Statist.*, **23**, 315–345.

Duda, R. O., and P. E. Hart, 1973: *Pattern Classification and Scene Analysis*. Wiley, 482 pp.

Fisher, R. A., 1938: The statistical utilization of multiple measurements. *Ann. Eugenics*, **8**, 376–386.

Follansbee, W. A., and V. J. Oliver, 1975: A comparison of infrared imagery and video pictures in the estimation of daily rainfall from satellite data. NOAA Tech. Memo. NESS 62, 14 pp.

Fu, K. S., D. A. Landgrebe and T. L. Phillips, 1969: Information processing of remotely sensed agricultural data. *Proc. IEEE*, **57**, 631–653.

Griffith, G. C., W. L. Woodley, S. Browner, J. Teijeiro, M. Maier, D. W. Martin and D. N. Sikdar, 1978: Rain estimation from geosynchronous satellite imagery—visible and infrared studies. *Mon. Wea. Rev.*, **106**, 1153–1171.

Gunn, K. L. S., and T. U. R. East, 1954: The microwave properties of precipitation particles. *Quart. J. Roy. Meteor. Soc.*, **80**, 522–545.

Hall, C. D., R. Davies and J. A. Weinman, 1978: The distribution of precipitation derived from Nimbus-6 data. Paper presented at 18th Conf. Radar Meteorology, Atlanta, Amer. Meteor. Soc.

Idso, S. B., R. D. Jackson and R. J. Reginato, 1975: Detection of soil moisture by remote surveillance. *Amer. Sci.*, **63**, 549–557.

Kshirsagar, A. M., 1972: *Multivariate Analysis*. Marcel Dekker, 534 pp.

Martin, D. W., and W. D. Scherer, 1973: Review of satellite rainfall estimation methods. *Bull. Amer. Meteor. Soc.*, **54**, 661–674.

Martin, D. W., J. Stout and D. N. Sikdar, 1975: GATE area rainfall estimation from satellite images. Rep. NOAA Grant 04-5-158-47, University of Wisconsin, 28 pp.

McFarland, M. J., and B. J. Blanchard, 1977: Temporal correlation of antecedent precipitation with Nimbus-5 ESMR brightness temperature. *Preprints 2nd Conf. Hydrometeorology*, Toronto, Amer. Meteor. Soc., 311–315.

Meeks, M. L., and A. Lilley, 1963: The microwave spectrum of oxygen in the earth's atmosphere. *J. Geophys. Res.*, **68**, 1683–1703.

Meneely, J. M., 1975: Application of the Nimbus-5 ESMR to rainfall detection over land surfaces. Contract Rep. NAS 5-20878, 48 pp.

Mie, G., 1908: Beiträge zur Optik trübes media, Speziell kolloidaler Metallösungen. *Ann. Phys.*, **26**, 597–614.

Okamoto, M., 1963: An asymptotic expansion for the distribution of the linear discriminant function. *Ann. Math. Statist.*, **34**, 1286–1301.

Savage, R. C., and J. A. Weinman, 1975: Preliminary calculations of the upwelling radiance from rain clouds at 37.0 and 19.35 GHz. *Bull. Amer. Meteor. Soc.*, **56**, 1272–1274.

——, P. J. Guetter and J. A. Weinman, 1976: The observation of rain clouds over land in Nimbus-6 electrically scanned microwave radiometer (ESMR-6) data. *Preprints 7th Conf. Aerospace and Aeronautical Meteorology and Symp. on Remote Sensing from Satellite*, Melbourne, Amer. Meteor. Soc., 131–136.

Scheffé, H., 1959: *The Analysis of Variance*. Wiley, 477 pp.

Scofield, R. A., and V. J. Oliver, 1977: A scheme for estimating convective rainfall from satellite imagery. NOAA Tech. Memo. NESS 86, 47 pp.

Shenk, W. E., R. J. Holub and R. A. Neff, 1976: A multispectral cloud type identification method developed for tropical ocean areas with Nimbus-3 MRIR measurements. *Mon. Wea. Rev.*, **104**, 284–291.

Staelin, D. A., 1966: Measurements and interpretation of the microwave spectrum of the terrestrial atmosphere near one centimeter wavelength. *J. Geophys. Res.*, **71**, 2875–2881.

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

Wilheit, T. T., 1975: The electrically scanning microwave radiometer (ESMR) experiment. *Nimbus-6 Users Guide*, NASA Goddard Space Flight Center, 87–108.

——, 1978: The effect of wind on the microwave emission from the ocean's surface at 37 GHz. NASA/GSFC Tech. Memo. 79588, 24 pp. (To be published in *J. Geophys. Res.*)

——, A. T. C. Chang, M. S. V. Rao, E. B. Rodgers and J. S. Theon, 1977: A satellite technique for quantitatively mapping rainfall rates over the oceans. *J. Appl. Meteor.*, **16**, 551–560.