Automated Cloud Classification of Global AVHRR Data Using a Fuzzy Logic Approach

BRYAN A. BAUM

Atmospheric Sciences Division, NASA Langley Research Center, Hampton, Virginia

VASANTH TOVINKERE AND JAY TITLOW

Science Applications International Corporation, Hampton, Virginia

RONALD M. WELCH

Institute of Atmospheric Sciences, South Dakota School of Mines and Technology, Rapid City, South Dakota

ABSTRACT

A fuzzy logic classification (FLC) methodology is proposed to achieve the two goals of this paper: 1) to discriminate between clear sky and clouds in a 32×32 pixel array, or sample, of 1.1-km Advanced Very High Resolution Radiometer (AVHRR) data, and 2) if clouds are present, to discriminate between single-layered and multilayered clouds within the sample. To achieve these goals, eight FLC modules are derived that are based broadly on airmass type and surface type (land or water): equatorial over land, marine tropical over land, marine tropical/equatorial over water, continental tropical over land, marine polar over land, marine polar over water, continental polar over land, and continental polar/arctic over water. Derivation of airmass type is performed using gridded analyses provided by the National Centers for Environmental Prediction.

The training and testing data used by the FLC are collected from more than 150 daytime AVHRR local area coverage scenes recorded between 1991 and 1994 over all seasons and over all continents and oceans. A total of 190 textural and spectral features are computed from the AVHRR data. A forward feature selection method is implemented to reduce the number of features used to discriminate between classes in each FLC module. The number of features selected ranges from 13 (marine tropical over land) to 24 (marine tropical/equatorial over water). An estimate of the classifier accuracy is determined using the hold-one-out method in which the classifier is trained with all but one of the data samples; the classifier is applied subsequently to the remaining sample.

The overall accuracies of the eight classification modules are calculated by dividing the number of correctly classified samples by the total number of manually labeled samples of clear-sky and single-layer clouds. Individual module classification accuracies are as follows: equatorial over land (86.2%), marine tropical over land (85.6%), marine tropical/equatorial over water (88.6%), continental tropical over land (87.4%), marine polar over land (86.8%), marine polar over water (84.8%), continental polar over land (91.1%), and continental polar/arctic over water (89.8%). Single-level cloud samples misclassified as multilayered clouds range between 0.5% (continental polar over land) and 3.4% (marine polar over land) for the eight airmass modules.

Classification accuracies for a set of labeled multilayered cloud samples range between 64% and 81% for six of the eight airmass modules (excluded are the continental polar over land and continental polar/arctic over water modules, for which multilayered cloud samples are difficult to find). The results indicate that the FLC has an encouraging ability to distinguish between single-level and multilayered clouds.

1. Introduction

The motivation for this work is to develop new techniques for the detection of multilayered clouds from satellite imagery in preparation for the upcoming launches of the Clouds and the Earth's Radiant Energy System (CERES) radiation budget instrument (Wielicki

et al. 1995). CERES is slated for launches on the Tropical Rainfall Measuring Mission (TRMM) platform (Simpson et al. 1988) and subsequently on the Earth Observing System platforms (EOS-AM and EOS-PM). As part of its data processing, the CERES team will incorporate narrowband radiometric data from the TRMM visible and infrared scanner (VIRS, 2-km resolution at nadir) and the EOS-AM and EOS-PM moderate resolution imaging spectrometer (MODIS; 0.25-, 0.5-, or 1-km resolution at nadir) instruments. The VIRS and MODIS satellite imagery will be used to determine cloud and clear-sky properties. Once determined, the

Corresponding author address: Dr. Bryan A. Baum, Atmospheric Sciences Division, NASA Langley Research Center, MS 420, Hampton, VA 23681-0001.
E-mail: b.a.baum@larc.nasa.gov

clear-sky and cloud information will be convolved with the much larger CERES field of view (FOV, ~10-km resolution at nadir on TRMM, ~20-km resolution at nadir on EOS-AM and EOS-PM). In preparation for the CERES production code development, the CERES team is implementing new algorithms and testing their performance using AVHRR and ERBE (Earth Radiation Budget Experiment) data. This paper specifically focuses on one aspect of that process—the development and implementation aspects of an automated fuzzy logic classification system appropriate for use with global daytime AVHRR (and in the future, VIRS and MODIS) imagery. The classifier is used to infer the presence of clear-sky, single-layer clouds, or multilayered clouds within an array of AVHRR pixels.

The pervasive occurrence of multiple cloud layers is described by global cloud climatologies derived from surface-based synoptic observations (Hahn et al. 1982; 1984). Observations of multiple cloud layers are discussed in Warren et al. (1985) and Tian and Curry (1989). These studies have found that mid- and high-level cirrus clouds tend to co-occur with lower-level clouds. If multiple cloud layers are present but are assumed a priori to be single level, the satellite-retrieved single-level cloud height will be somewhere between the upper and lower cloud layers. The magnitude of the error depends on the optical depth and cloud fraction of the upper cloud layer (e.g., Menzel et al. 1992; Baum and Wielicki 1994). Errors in retrieved cloud heights cause errors in the microphysical cloud property retrievals.

Part of the problem is that one needs to know where multiple cloud layers exist in the imagery before an algorithm can be developed to infer the macrophysical and microphysical properties of each of the cloud layers present. Pankiewicz (1995) and Welch et al. (1992) present reviews of pattern recognition techniques for the identification of clouds and cloud systems. In recent years, textural and spectral signatures have been used to classify cloud and surface types from high-resolution Landsat multispectral scanner imagery (e.g., Chen et al. 1989; Lee et al. 1990; Welch et al. 1988, 1989) or from lower resolution imagers such as AVHRR (e.g., Tovinkere et al. 1993). The classification schemes presented in these studies have the limitation that they are designed for a specific area, such as for maritime regions (e.g., Bankert 1994), tropical regimes (e.g., Shenk et al. 1976; Inoue 1987), or polar regions (e.g., Ebert 1987, 1989; Key et al. 1989; Key 1990; Welch et al. 1990, 1992; Rabindra et al. 1992; Tovinkere et al. 1993). Peak and Tag (1992) report on the development of an automated cloud classification system using GOES data for use in the interpretation of synoptic-scale events as an aid in forecasting on U.S. Navy ships. The focus of these studies is primarily that of distinguishing between various cloud types or, in some cases, between different surface types in clear-sky conditions, but not of distinguishing whether there are one or more cloud layers.

In the context of the CERES experiment, the goals of the classification process are twofold:

1) To discriminate imager pixel arrays that contain clear land or ocean from those that contain clouds.
2) If a pixel array contains clouds, to discriminate between single and multiple cloud layers.

While automated cloud-layer classification has the potential for providing useful information for the cloud retrieval process, it has been demonstrated only for a case study analysis (Baum et al. 1995). No studies exist to our knowledge that examine global classification of single and multiple cloud layers. In this study, we modify the fuzzy logic classifier (FLC) proposed by Tovinkere et al. (1993) and Baum et al. (1995) for use with global daytime AVHRR data.

Section 2 discusses the satellite and ancillary data used in this study, section 3 describes the classes, sample labeling process, and spectral and textural features. Section 4 outlines the classification methodology, results are shown in section 5, and conclusions are given in section 6.

2. Data

a. Satellite data

The FLC is trained and tested using *NOAA-11* AVHRR satellite imagery collected for the time period between 1991 and 1994 from all continents and oceans. The *NOAA-11* polar-orbiting platform has equatorial crossing times of 0140 (descending node) and 1340 (ascending node) local time (LT). Spatial resolution of the local area coverage (LAC) data is 1.1-km at nadir. The spectral data consists of AVHRR channels 1 (0.55–0.68 μm), 2 (0.725–1.1 μm), 3 (3.55–3.93 μm), 4 (10.5–11.5 μm), and 5 (11.5–12.5 μm), which include visible (channel 1), near-infrared (channels 2 and 3), and infrared (channels 4 and 5) wavelengths. Conversion of channel 1 and channel 2 raw counts to radiances is performed using the calibration detailed by Rao and Chen (1994). The channel 1 radiances are converted to reflectances ρ_1 by

$$\rho_1 = \frac{\pi I_1}{\mu_o F_1},\tag{1}$$

where I_1 is the shortwave spectral radiance in AVHRR channel 1, F_1 is the incoming solar spectral flux for channel 1, and μ_o is the cosine of the solar zenith angle. Conversion of channel 2 radiances to bidirectional reflectances is performed similarly. The near-infrared (NIR) and infrared (IR) radiances are calculated from the nominal calibration provided in the NOAA level 1-B data stream (Kidwell 1997). The IR channels include corrections due to the nonlinear response of channels 4 and 5 reported by Brown et al. (1993).

TABLE 1. Thresholds for tropopause pressure, precipitable water (pw), and surface temperature T_s.

Airmass type	Tropopause pressure (mb)	Precipitable water (cm)	Surface temperature		
			Daytime	Twilight	Nighttime
Equatorial	$P < 135$	$pw > 6$	$T_s > 32$	$T_s > 31$	$T_s > 30$
Tropical	$135 < P < 225$	$3 < pw < 6$	$22 < T_s < 32$	$18 < T_s < 31$	$14 < T_s < 30$
Polar	$225 < P < 300$	$1 < pw < 3$	$-16 < T_s < 22$	$-18 < T_s < 18$	$-20 < T_s < 14$
Arctic	$P > 300$	$pw < 1$	$T_s < -16$	$T_s < -18$	$T_s < -20$

b. NCEP gridded profiles

National Centers for Environmental Prediction (NCEP) gridded profiles of temperature, humidity, and horizontal winds at 0000 and 1200 UTC are used as the source of meteorological data for the AVHRR image analyses. The NCEP gridded profiles are provided at 14 pressure levels including the surface and are derived from many sources including rawinsonde data and the TIROS Operational Vertical Sounder (TOVS) instruments on the NOAA platforms. The quality of the gridded profiles is more suspect in data sparse regions such as the Southern Hemisphere.

Cloud classification on a global scale involves training on a wide variety of cloud systems, surface types, and climatological regimes. A classifier trained and validated using data from any limited region can be expected to have only limited success globally. To accurately classify clouds in a given region one should use a classifier based upon a training set that is appropriate for the local surface type(s) and ambient climatological conditions. That is, one expects clouds to have similar radiometric features in regions with similar surface characteristics and with similar meteorological conditions. Indeed, there is a strong correlation between the spectral signature of a cloud and the air mass in which it lies. Our approach is to use the meteorological gridded profiles to determine the most likely air mass for any local region in an image. The decision tree for the designation of airmass type is based on a simple set of thresholds as described below. Airmass types include arctic (A), maritime polar (mP), continental polar (cP), maritime tropical (mT), continental tropical (cT), and equatorial (E). Each of the pixel arrays collected for training purposes is tagged with the appropriate air mass using the NCEP gridded data.

The following meteorological parameters are used to determine the air mass: profile mean wind speed, vertical wind shear, tropopause height, precipitable water, boundary layer precipitable water, mean profile relative humidity, K stability index, boundary layer lapse rate, and surface dewpoint depression. In addition, since several of these parameters display diurnal fluctuations, local time is included for each analysis region. Local time is used subsequently to fit the analysis region into one of three diurnal regimes: day, night, and twilight. First, a set of threshold tests (Table 1) is used to determine one of four basic airmass types: arctic, polar, tropical, or equatorial. A cumulative ''score'' value is kept during the tests that ranges from 8 (equatorial) to -8 (arctic). Surface air temperature is doubly weighted because of its relative importance in determining the air mass. Scores greater than 5 are classified as equatorial, between 5 and 0 as tropical, between 0 and -5 as polar, and less than -5 as arctic.

A further set of threshold tests are performed if either a Polar or tropical air mass is chosen to determine whether the air mass is maritime or continental in nature. Three threshold tests (Table 2) are applied: precipitable water, surface dewpoint depression, and boundary layer lapse rate. The boundary layer lapse rate test is weighted half as much as the other two tests.

3. Classes, region labeling, and features

The fuzzy logic classifier uses a supervised learning approach in which a set of labeled samples for each class is required to train the classifier. A class can be a specific type of cloud, such as cirrus, or a surface type, such as ocean. Samples, defined for this study as 32-pixel $\times$ 32-pixel arrays (approximately 35 km $\times$ 35 km at nadir), are collected manually for each defined class from 1.1-km AVHRR satellite imagery. The number of samples required for training depends on the number of classes. A rule of thumb is to label a minimum of approximately (15 $\times$ number of classes) samples per class to ensure an adequate representation for each class.

TABLE 2. Thresholds used to differentiate between continental and maritime tropical and polar air masses. The tests include precipitable water pw, surface dewpoint depression $T - T_d$, and boundary layer lapse rate.

Airmass type	Precipitable water (cm)	$T - T_d$ (°C)			Lapse rate (°C km^{-1})		
		Day	Twilight	Night	Day	Twilight	Night
Continental tropical	<1.8	>22	>16	>8	>7.5	>6	>4.5
Maritime tropical	>1.8	<22	<16	<8	<7.5	<6	<4.5
Continental polar	<0.8	>12	>7	>8	>4	>3	>2
Maritime polar	>0.8	<12	<7	<8	<4	<3	<2

Without an adequate set of samples, the classifier will not be robust. This requirement proved to be a limitation in our technique because in several air masses, certain cloud classes, primarily midlevel clouds, were extremely difficult to populate. This is discussed further in section 5.

a. Classes

Our goal is to design and build a robust global cloud classification system that is independent of season, but exclusive of strong sunglint conditions and surfaces that are covered by snow or ice. For a classifier to be effective, one must first define a set of classes that are well separated by a set of features derived from the multispectral channel radiometric data. The choice of classes is not always straightforward and may depend upon the desired applications. For instance, some investigators choose a set of standard cloud types such as cirrostratus, altocumulus, or cumulus (e.g., Garand 1988; Ebert 1987; Bankert 1994). However, for inferring cloud properties from satellite data, more useful criteria are cloud height (low, mid, or high), cloud fraction (broken or uniform), and visual opacity (thin or thick). Few if any cloud property retrieval algorithms depend on the label attached to a cloud type, that is, whether a low cloud is stratocumulus or cumulus; what matters is the effective cloud amount and cloud height.

For high-level clouds, a separate class is defined as "uniform and thick" because optically thick, extremely cold clouds such as cumulonimbus have different spectral and textural properties than other high-level clouds such as cirrus. The set of classes used in this study are as follows:

1) Clear sky
2) Broken low-level cloud
3) Uniform low-level cloud
4) Broken midlevel cloud
5) Uniform midlevel cloud
6) Broken high-level cloud
7) Uniform high-level cloud
8) Uniform, thick high-level cloud

Low-level clouds consist of cumulus (Cu), stratocumulus (Sc), stratus (St), fog, and fractostratus (Fs). Midlevel clouds consist of altocumulus (Ac), altostratus (As), and towering cumulus. High-level clouds are cirrus (Ci), cirrostratus (Cs), cirrocumulus (Cc), and cumulonimbus (Cb). The cloud classes are separated by cloud fraction and cloud height. Note that only single-layered cloud classes are defined. The reason for this is that the fuzzy logic method uses the concept of class membership to determine whether more than one cloud layer exists. If more than one cloud layer exists, the sample should theoretically exhibit high class memberships for more than one cloud layer. Therefore, the classifier is not trained with samples containing multilayered clouds. Further discussion on this point is provided in section 4.

b. Region labeling

Since the labeled samples are used during both the training and testing phases, any inaccuracies introduced during the labeling process will result in lower classification accuracies. The ultimate ability of the FLC depends on a judicious selection of class samples. Sample labeling is not performed by an analyst solely on the basis of the imagery. While imagery provides contextual information about a scene, there are many situations when additional information is required to make a more accurate judgment. For instance, thin cirrus can be extremely difficult to recognize if it overlies uniform stratus.

The manual collection of samples is performed independently by two analysts. A cloud sample is placed into a height category based on the cloud-top temperature. Midlatitude clouds are placed into the "low" category if their cloud-top temperature places them at heights below approximately 2 km, into the midlevel category if the height is between 2 and 5 km, and into the high cloud category if the height is above approximately 5 km. The height range is expanded for tropical and equatorial air masses (low: <3 km; mid: 3–6 km; high: >6 km).

Cloud fraction in a sample is determined by comparing the individual pixel values within the 32×32 array with the clear-sky brightness temperatures or reflectances. For each of these cloud types, a further distinction is made based on cloud fraction within the pixel array of broken (not completely cloud covered) or uniform (completely cloud covered). Note that the term "uniform" applies only to cloud fraction; it does not have any connotation for the optical depth of the particular sample. A separate field is available for the analyst to list subjective notes about the sample, such as that the scene appeared hazy or that the cirrus was chosen from a hurricane's outflow. This proves beneficial when investigating samples that are misclassified, such as when a clear-sky sample contains thick smoke from biomass burning and the classifier returns an answer of low-level cloud.

The sample labeling process is strengthened through the use of ancillary datasets including NCEP gridded analyses of temperature and humidity profiles. Regions of enhanced relative humidity or a surface temperature inversion may indicate the presence of a cloud layer. Sample selection is made using the Satellite Imagery Visualization System (SIVIS; Baum et al. 1995). The SIVIS software allows the user to analyze an image interactively and view ancillary data related to each sample. While some error is inescapable in the sample collection process, this software provides a wide variety of analysis tools, which strongly limits labeling errors. Each collected sample is stored in a database along with

the scaled radiometric data, latitude, longitude, time of observation, analyst comments, viewing geometry (solar zenith, viewing zenith, relative azimuth, and scattering angles), ancillary data providing surface characteristics such as ecosystem and elevation, and finally meteorological parameters calculated from the NCEP gridded profiles such as tropopause height, airmass type, and boundary layer lapse rate.

c. Features

A variety of textural and spectral features have been explored in the literature and are briefly mentioned here. As noted in Welch et al. (1992), texture refers to a set of statistical measures of the spatial distribution of gray levels in an image. The textural features are derived from the average spatial relationships of gray-level values of pairs of pixels across the pixel array (Haralick et al. 1973). The gray-level values are based on an absolute scale in which reflectance values of 0 and 1 correspond to gray-level values of 0 and 255, respectively. Brightness temperature values of 180 K and 330 K correspond to gray-level values of 0 and 255, respectively. Reflectance gray levels are calculated for AVHRR channels 1 and 2, and brightness temperature gray levels are calculated for AVHRR channels 3, 4, and 5. In addition to the textural features, a set of spectral features is computed that is based on reflectances, brightness temperatures, or a combination of both.

1) TEXTURAL FEATURES

Many textural features are computed using the gray-level difference method (GLDV) (Haralick et al. 1973; Weszka et al. 1976; Chen et al. 1989). The GLDV method is based on the absolute differences between pairs of pixels having gray levels I and J separated by a distance d at angle ϕ with a fixed direction. The GLDV probability density function $P_{\text{GLDV}}(m)$ is defined for $m = I - J$, where I and J are the corresponding gray levels having a value between 0 and 255. The function $P_{\text{GLDV}}(m)$, which depends on d and ϕ, is obtained by normalizing the gray-level frequencies of occurrence by the total number of frequencies. Once $P_{\text{GLDV}}(m)$ has been formed, the calculation of the textural features listed in Table 3 is straightforward. Further details on the textural features may be found in Chen et al. (1989).

Some useful information may be lost in the use of difference vectors that may be inherent in the original imagery. For example, a gray-level difference of zero may be derived from a pair of pixels in which both pixels are cloud free. The same result could be obtained if both pixels are completely covered by thick cloud. One way of supplementing the information provided by the GLDV method is to use the gray levels of the original image instead of the gray-level differences. This approach will be referred to henceforth as the gray-level vector (GLV) method. For each data sample, a gray-

TABLE 3. Texture features derived using the gray-level difference vector (GLDV); m is the absolute value of the difference between gray-level differences one pixel apart in the horizontal direction ($d = 1$, $\phi = 0°$). Here, $P(m)$ is the difference vector probability density function obtained by normalizing the gray-level difference frequencies of occurrence by the total number of frequencies.

Mean	$\mu = \sum_m mP(m)$
Standard deviation	$\sigma = \left[\sum_m (m - u)^2 P(m) \right]^{1/2}$
Angular second moment (ASM)	$\sum_m [P(m)]^2$
Entropy (ENT)	$\sum_m P(m) \log P(m)$
Local homogeneity (HOM)	$\sum_m \dfrac{P(m)}{(1 + m^2)}$
Contrast (CON)	$\sum_m m^2 P(m)$
Cluster shade (SHADE)	$\dfrac{\sum_m (m - \mu)^3 P(m)}{\sigma^3}$
Cluster prominence (PROM)	$\dfrac{\sum_m (m - \mu)^4 P(m)}{\sigma^4} - 3.0$

level histogram is calculated and normalized by the total number of points. The normalized histogram becomes the GLV-based density function $P_{\text{GLV}}(m)$. The features defined in Table 3 are calculated using both the GLV and GLDV methods.

2) SPECTRAL FEATURES

A number of spectral features may be formed by exploiting the thermal or reflectance characteristics of various cloud and surface types (e.g., Garand 1988; Ebert 1987; Bankert 1994). The features are formed from the gray-level representation of the bidirectional reflectances for AVHRR channels 1 and 2 and from the gray-level representation of brightness temperatures for the NIR and IR channels.

Spectral features are formed from individual AVHRR channels and combinations of the five AVHRR channels. The spectral features are calculated for a single channel quantity X (where a quantity is a gray-level representation of either a reflectance or a brightness temperature), for two different quantities X and Y, or for three quantities X, Y, and Z, as follows.

1) Mean X: This spectral feature is the mean gray-level value of either reflectance or brightness temperature calculated from the array.
2) Band difference ($X - Y$): This spectral feature is the difference of the gray-level means of two channels.
3) Band ratio (X/Y): This feature is formed from taking the ratio of mean gray-level values between two channels.

4) Overlay (*X, Y, Z*): This spectral feature forms a packed integer from the mean gray-level value of three quantities *X, Y,* and *Z.* It is similar in nature to the idea of using 24-bit color graphics to form false-color imagery. With the proper channel combination, warm reflective stratus clouds would have different values than cold, thin, less reflective cirrus. The overlay of *X, Y,* and *Z* is calculated from Overlay $= Z \times 2^{16} + Y \times 2^8 + X$. This particular feature is useful in separating the clear-sky land, low-level cloud, midlevel cloud, and high-level cloud classes.

5) Low *X*: This feature is the GL mean of the lowest 2% of the values within the array.

6) High *X*: This feature is the GL mean of the highest 2% of the values within the array.

7) Spatial coherence: For a given array, means and standard deviations are calculated for all local (2 × 2 or 4 × 4) pixel groups within the 32 × 32 array. The spatial coherence feature is the mean of those local pixel groups that have a standard deviation less than 2.5.

8) Normalized difference indices (NDIs) are computed from the gray levels (GLs) of a pair of channels *i* and *j*:

$$NDI_{ij} = \frac{GL_i - GL_j}{GL_i + GL_j}. \qquad (2)$$

The NDI features are computed from the gray-level values of any two channel combinations. If AVHRR channels 1 and 2 are used, the NDI would be similar to the normalized vegetation index that is widely used in remote sensing investigations of surface cover.

Additional information may be gleaned from pseudochannels computed from a combination of two or more AVHRR channels. For instance, false color imagery often provides more insight as to what is in a scene as compared to gray-scale imagery. As an example, three channels may be used to provide information to the red (*R*), green (*G*), and blue (*B*) color cube. The RGB values may be transformed into values of hue *H,* saturation *S,* and intensity *I.* The HSI transform (Foley and Van Dam 1984) is used to cluster points within a three-dimensional subspace in which the color intensity of each channel is scaled according to the gray-level value. One way of perceiving the HSI features is that they provide a quantitative method of clustering data according to the color scale for a specific class in a false color image. With a particular selection of overlay channels, clouds may appear white, the vegetated surface as green, and the ocean as dark blue. The HSI transform separates classes by quantifying the colors as HSI values. A set of additional HSI features are computed for the following two RGB channel combinations: [1, 2, (3–4)] and [4, 5, 1].

4. The fuzzy logic approach

The fuzzy logic approach uses a different conceptual model to classify objects than that used by the traditional classification approaches such as the maximum likelihood estimator (MLE) or by neural network schemes. The theory of fuzzy logic is based on approximate reasoning. That is, fuzzy sets display a strength of membership for all potential classes; the degree of membership to any particular class is provided by a membership function. Stronger membership values reveal the potential for a particular class to be present in the sample. Further discussion of the fuzzy logic method may be found in Tovinkere et al. (1993), Giarratano and Riley (1989), or Penaloza and Welch (1996).

With neural network or MLE schemes, the final classification decision is a single class. If the sample being classified contains more than one class, only the most likely class will be chosen. The strength of the fuzzy logic method is that more than one class may be chosen depending on the strength of membership of the sample to the chosen classes (Tovinkere et al. 1993; Baum et al. 1995). Instead of having explicitly defined classes that represent a combination of cloud types, including many multilayered cloud combinations, only a core set of single-layered cloud classes are necessary. Note that a set of eight classes are defined in section 3a as opposed to the 18 and 20 classes described in Ebert (1987) and Garand (1988), respectively. Note that the classes must be populated with samples for each of the airmass types for training and testing the classifier.

The generation of the FLC for each air mass follows the scheme shown in Fig. 1. There are five steps to the process: choose a set of classes (discussed in section 3a), collection of samples for each class (sets of labeled 32 × 32 pixel arrays discussed in section 3b), selection of features that discriminate between classes (discussed in section 4b), using the feature set to derive membership functions (discussed further in sections 4a and 4c), and finally saving the selected feature set and the parameters that describe the fuzzy membership functions. Further details on the generation of the classifier are provided in the following sections.

a. Fuzzy membership functions

The fuzzy logic classifier is trained using the sets of labeled samples collected for each airmass type. Since a labeled dataset is used in the training phase, the classifier is said to be trained with supervision. Means (μ) and standard deviations (σ) are calculated for the complete training dataset for each feature and for each class. The *S* membership function shown in Fig. 2a is formed from mean and standard deviation values as follows:
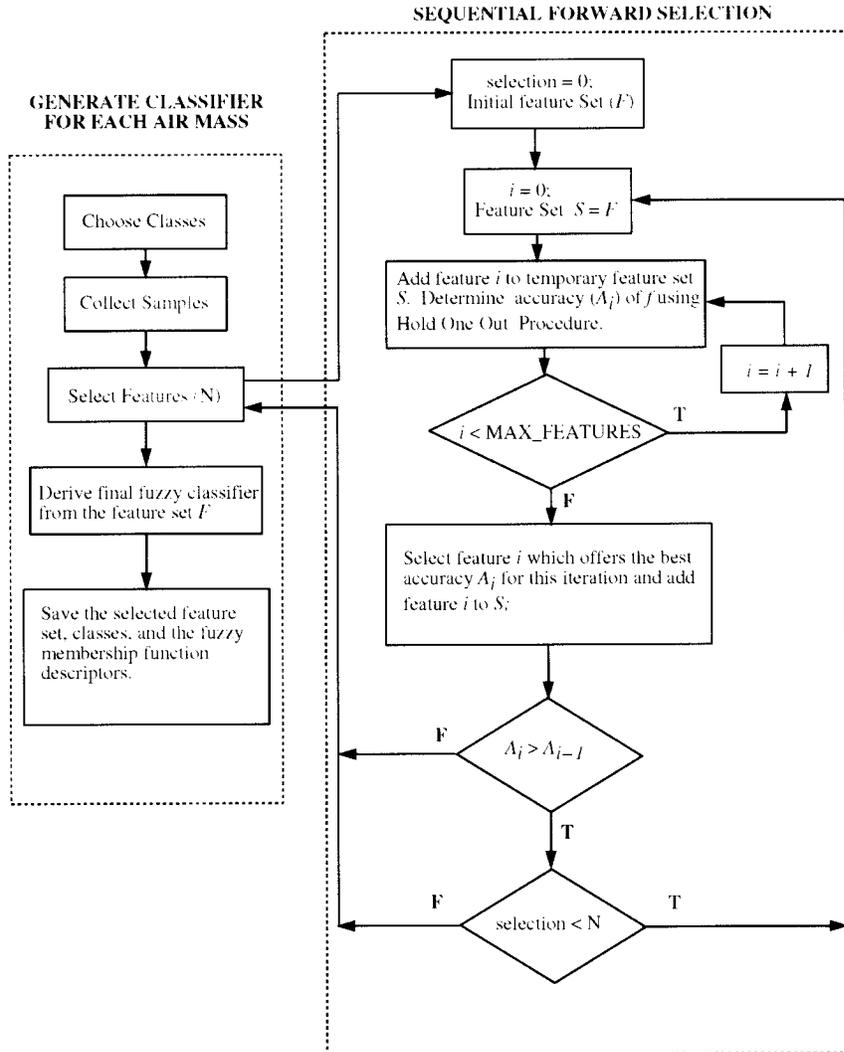
FIG. 1. Flowchart of the methodology used to generate the fuzzy logic classifier. There are five steps to the process of generating a classifier: choose a set of classes (discussed in section 3a), collection of samples for each class (i.e., sets of labeled 32×32 pixel arrays discussed in section 3b), selection of features that discriminate between classes (discussed in section 4b), use of the feature set to derive membership functions (discussed further in sections 4a and 4c), and finally saving the selected feature set and the parameters that describe the fuzzy membership functions.

$$S(x, \alpha, \beta, \gamma) \to \begin{cases} 0, & \text{for } x \leq \alpha, \\ 2\left(\dfrac{x - \alpha}{\gamma - \alpha}\right)^2, & \text{for } \alpha \leq x \leq \beta, \\ 1 - 2\left(\dfrac{x - \alpha}{\gamma - \alpha}\right)^2, & \text{for } \beta \leq x \leq \gamma, \\ 1 & \text{for } x \geq \gamma. \end{cases} \tag{3}$$

For fuzzy sets that may be represented by Gaussian distributions, the Π function (Fig. 2b) is used:

$$\Pi(x, \beta, \gamma)$$
$$\to \begin{cases} S\left(x, \gamma - \beta, \gamma - \dfrac{\beta}{2}, \gamma\right), & \text{for } x \leq \gamma, \\ 1 - S\left(x, \gamma, \gamma + \dfrac{\beta}{2}, \gamma + \beta\right), & \text{for } x \geq \gamma. \end{cases} \tag{4}$$

In Tovinkere et al. (1993), the Π function approximated the Gaussian distribution such that $\beta = 3\sigma$ and γ was the mean value μ. This resulted in a number of samples going unclassified by the expert system. By increasing

(a) S–function

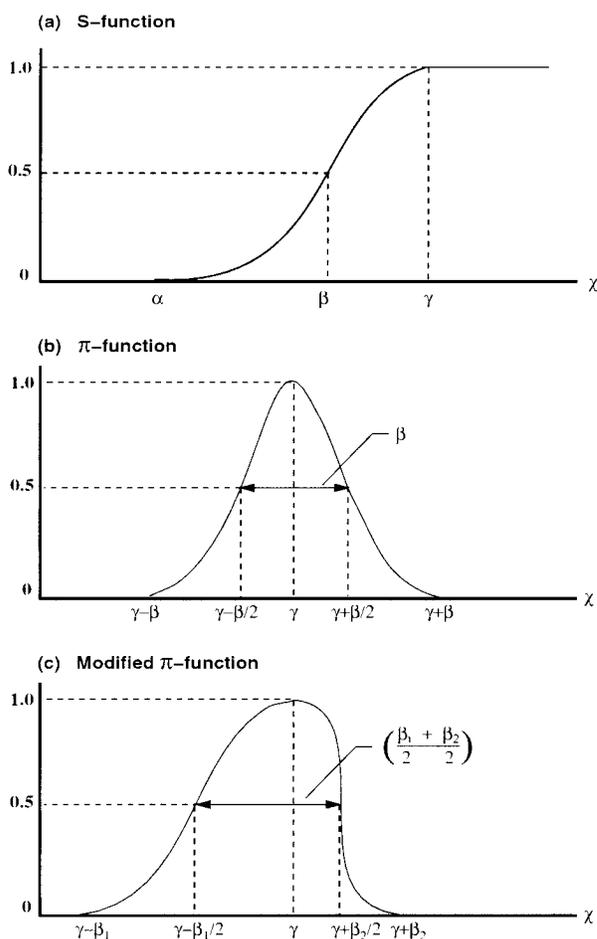
(b) π–function

(c) Modified π–function

FIG. 2. (a) The traditional S function used to define fuzzy membership functions. (b) The Π function traditionally used to represent Gaussian distributions of data. (c). The modified Π function used in this study.

the spread of the fuzzy sets (i.e., increasing β), the number of unclassified samples is decreased. The value of $\beta = 5\sigma$ was eventually chosen to allow greater overlap of the fuzzy sets, thereby minimizing the number of samples that could not be classified.

The derivation of fuzzy membership functions has been revisited in the current study because many fuzzy sets form non-Gaussian distributions. One way of accounting for the non-Gaussian distribution is to simply

TABLE 4. Values of β_1 and β_2 used to define the modified-Π function.

Value of β_1 or β_2	Operation performed on the membership function
0–1	dilation of power 3
1–2.5	dilation of power 2
2.5–5.5	concentration of power 1
5.5–7.5	concentration of power 2
>7.5	concentration of power 3

increase β, but this tends to result in an increased number of incorrect classifications. For fuzzy sets that may have non-Gaussian distributions, a modified Π function (Π^{mod}) is used (see Fig. 2c):

$$\Pi^{mod}(x, \beta, \gamma)$$

$$\rightarrow \begin{cases} S\left(x, \gamma - \beta_1, \gamma - \dfrac{\beta_1}{2}, \gamma\right), & \text{for } x \leq \gamma, \\ 1 - S\left(x, \gamma, \gamma + \dfrac{\beta_2}{2}, \gamma + \beta_2\right), & \text{for } x \geq \gamma, \end{cases}$$

$$(5)$$

where $\beta_1 = (\gamma - \min)/\sigma$ and $\beta_2 = (\max - \gamma)/\sigma$, σ is the standard deviation of the feature, and min and max represent the lower and upper numerical limits of the feature for a class as derived from the training set. The variables β_1 and β_2 represent the integer number of standard deviations used to generate the modified-Π membership functions. To better adjust the function shape to the data provided by the set of labeled samples, dilation and concentration operations (Giarratano and Riley 1989) are performed on the membership functions according to the values for β_1 or β_2 as shown in Table 4.

b. Feature selection

The primary purpose of the feature selection process is to find a subset of the 190 features that minimizes computational resources and maximizes classifer accuracy. The selection of features depends on the classes desired; a change in specified classes usually results in the selection of a different feature set. As shown in Fig. 1, a modified sequential forward selection (SFS) scheme (Devijver and Kittler 1982) is implemented for the purpose of selecting an optimal feature set for each air mass. The SFS is a simple search procedure in which one feature at a time is added to the developing feature set (denoted by S in Fig. 1). A feature may be selected more than once in our scheme. For each feature selected and added to the feature set S, a criterion function is evaluated to determine how well the feature set performs. The criterion function used to assess the discriminatory power of each feature set is the fuzzy classifier built from the current feature set. The overall accuracy (A_i, where i is the iteration number) obtained from application of the classifier to the various sets of labeled samples is used as the criterion for feature selection. A probabilistic distance measure such as the Bhattacharya distance is not used as the criterion function in our approach. The Bhattacharya distance measure is an overall measure of class separability in the dimensional space of the feature vector as determined from the total set of "N" chosen features. That is, if the number of features in the current feature set is 10, the distance measure is the class separability in the 10th dimension. In contrast, the FLC is an additive one-dimensional scheme where

the class separabilities are determined by the degree of membership. Thus, the features selected by use of the Bhattacharya distance may not necessarily yield the best feature set for the FLC.

It has been our experience that a small amount of noise may be encountered during the training phase due to the similarity of, for example, broken low, mid, and high clouds using certain textural features. The noise is evidenced by very small residual values for each class. If a labeled sample belongs to a class, one would expect to see the membership for that class to be equal to 1. Residual values are membership values obtained in classes other than the actual class assigned to the sample by the analyst. A running average of the residual noise is maintained for each class. The residual values from the classification output are normalized to get the relative strength of the residuals for each class. The residue values are used to set the final membership thresholds used in the classifier. For each of the features selected and for each air mass, the parameters describing the membership function (β_1, β_2, and σ) and the residual values are saved in the knowledge base.

c. Fuzzy logic expert system

The fuzzy logic expert system comprises an initialization phase and a decision and result phase. The details of the classifier are provided in Tovinkere et al. (1993) and are briefly outlined here.

1) INITIALIZATION PHASE

For each class, membership function definitions for each feature are loaded from the knowledge base, and the membership functions are constructed. The class membership functions for each feature represents the fuzzy set for that feature.

2) DECISION AND RESULT PHASE

The membership values for each feature and for each class are computed from the data contained within the sample being tested. An average membership value is computed for each class. The average membership values are subsequently normalized to indicate the relative memberships of each class. The residual values computed during the feature selection and training phase are used to compute threshold values for the various membership functions. For a given data sample, class membership is said to exist if the membership strength exceeds the thresholds determined during the training phase. If membership values exceed the threshold for more than one cloud class, the sample is classified as containing multiple cloud layers. Once a sample has been classified according to its membership values, the feature values computed for the sample being tested are cleared and the system is ready to run another case.

An example of the process may be illustrative. Figure 3 shows a representative set of fuzzy functions for a single feature constructed from the mean gray-level values of channel 4 (MeanGL4). The functions are derived using the sets of data samples collected during the training process. We now assume an unclassified sample has a MeanGL4 value of 86. The actual membership values are 0.0 for the land, broken low cloud, uniform low cloud, and broken midlevel cloud classes; 0.03 for the uniform midlevel cloud class, 0.46 for the broken high-level cloud class; and 0.06 for the uniform high-level cloud class. The normalized membership values for the nonzero-valued classes are 0.05 for the uniform midlevel cloud class, 0.84 for the broken high-level cloud class, and 0.11 for the uniform high-level cloud class. The same process is followed for each of the features (based on airmass type) selected for the given data sample. Once the feature values have been calculated and normalized for each of the selected features, an average class membership is derived for each class (i.e., the sum of the normalized class memberships divided by the number of features). For this example, we assume final average membership values are as follows: 0.0 for the land, broken low cloud, uniform low cloud, and broken midlevel cloud classes; 0.08 for the uniform midlevel cloud class; 0.74 for the broken high-level cloud class; and 0.21 for the uniform high-level cloud class. If a value of 0.3 is required for the sample to membership in any of the classes (the actual threshold is derived during the training/testing phase), the sample would be classified as containing broken high-level clouds. As a final note, for a sample to be classified as containing multilayered clouds, the membership of two classes (in separate height categories) would have to exceed the threshold.

5. Results

The features chosen by the feature selection routine are shown in the appendix. The number of features selected for use in the classifier depends on the airmass type and the number of classes in each air mass. The number of features chosen for each air mass range from 13 (mT over land) to 24 (mT/E over water). Some features are chosen more than once for a given air mass by the feature selection scheme.

a. Clear-sky and single-level clouds

Table 5 outlines the total number of labeled samples collected by the analysts for the eight classification modules. In selecting clear-sky samples, especially for the warmer air masses over land (E, mT, and cT), it became apparent that if there was clear-sky or broken cloud cover, there was some likelihood of finding smoke and fires, especially over northern South America and central Africa. An effort was made initially during the labeling process to obtain clear land samples that did not contain fires or smoke. Since biomass burning is so
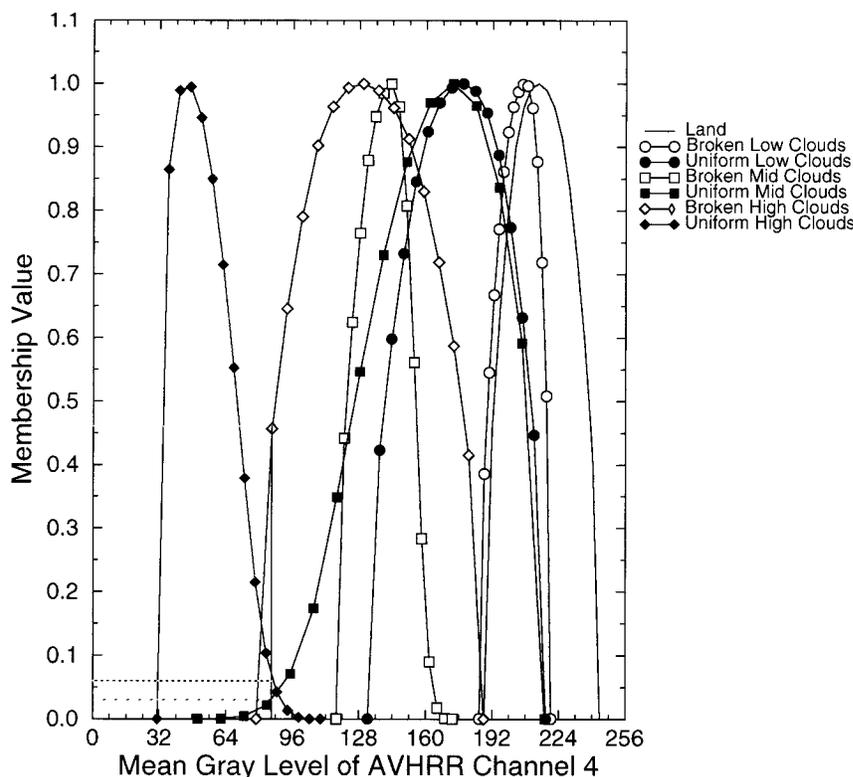
FIG. 3. Example of normalized membership values for a single feature representing the mean gray-level value of AVHRR channel 4 (see section 3c). The membership values for each class are derived during the training phase (see section 4).

prevalent, however, we quickly abandoned this effort and simply collected samples regardless of whether fires or smoke were present. The implications of including fire and smoke samples is discussed later in this section.

Two methods, the hold-one-out method (Bankert 1994) and the modified bootstrap method (Chatterjee and Chatterjee 1983; Bankert 1994), were used to estimate the classification accuracy of the fuzzy logic method for each of the airmass modules. With the hold-one-out method, the classifier is trained with all but one of the labeled samples; the classifier is applied subsequently to the remaining sample. New membership functions are derived for each run of the hold-one-out

test. The procedure is repeated for each of the samples. After the labeled samples have been classified, the results are normalized by the number of labeled samples in each class. The modified bootstrap method (Chatterjee and Chatterjee 1983; Bankert 1994) involves choosing a random selection of samples for each class. A sample may be selected more than once during the random selection process. However, the classifier is tested with only those samples that are not used during the training process. About 80% of the number of samples in each class are randomly selected; the classifier is tested on the remaining samples. The random selection and accuracy testing process was repeated 25 times to

TABLE 5. Number of manually labeled clear-sky and single-level cloud samples in each class for each airmass module. The "broken" designation refers to a cloud sample having a cloud fraction less than 1, and the "uniform" designation refers to a cloud fraction of 1.

	E (land)	mT (land)	cT (land)	mT/E (water)	mP (land)	mP (water)	cP (land)	cP/A (water)
Clear-sky	166	225	262	197	212	163	311	137
Broken low	175	147	280	300	241	184	178	238
Uniform low	—	160		126	133	212	164	195
Broken mid	141	216	192	119	162	155	—	—
Uniform mid	113	124	—	125	178	154	—	—
Broken high	122	188	234	124	160	166	168	170
Uniform high	125	138	164	173	174	155	276	190
Uniform, thick high	178	174	99	121	—	—	—	—
Total	1020	1372	1231	1285	1260	1189	1097	930

TABLE 6. Matrix of classifications for hold-one-out accuracy test (%) for cloud-free and single cloud layer samples for the E over land module. Fuzzy logic classification is presented in columns; expert classification in rows. Overall accuracy for the E over land module, determined by dividing the number of correctly classified samples by the total number of samples (1020), is 86.2%. The "broken" designation refers to a cloud sample (32 × 32 pixel array) having a cloud fraction less than 1, and the "uniform" designation refers to a cloud fraction of 1.

Classes	Clear	Broken low	Uniform low	Broken mid	Uniform mid	Broken high	Uniform high	Thick high	Multiple layers
Clear	94.6	5.4	—	0	0	0	0	0	0
Broken low	2.3	91.4	—	2.3	0	0	0	0	4.0
Uniform low	—	—	—	—	—	—	—	—	—
Broken mid	0	2.1	—	86.5	6.4	1.4	0	0	3.5
Uniform mid	0	0	—	14.2	81.4	1.8	0.9	0.9	0.9
Broken high	0	0	—	7.4	0	78.7	9.8	1.6	2.5
Uniform high	0	0.8	—	2.4	0	13.6	74.4	8.8	0
Uniform thick high	0	0	—	0	0	0	10.7	89.3	0

create 25 different bootstrap sample sets and 25 sets of results. A final set of results was compiled from the average of the individual iterations. The comparison between the results for the two methods was quite similar with overall accuracies (i.e., all classes within an airmass type) within 3%. The bootstrap results for individual classes within each air mass were most stable when the available pool of samples was much larger than that required (about 100 samples per class). Because some of the classes had insufficient sampling, the results from the hold-one-out method are deemed to be more reliable and are shown below.

1) EQUATORIAL AIR MASS OVER LAND (E OVER LAND)

The overall accuracy of the E over land classifier module for clear-sky and single-layer cloud samples is 86.2% as calculated by taking the ratio of correctly classified samples to the total number of samples. Detailed results of this module are presented in Table 6. Because so few samples of uniform (nonbroken) low-level cloud were found in the equatorial air mass using the 32 × 32 imager pixel array size, the uniform low-level cloud class was disabled in the classifier.

It is interesting to note that analysts found fire or smoke in about 30% of the clear-sky samples, 6% of the low broken cloud samples, 3% of the midlevel broken cloud samples, and less than 1% of the broken and uniform high-level cloud samples. The clear-sky results indicate that 5.4% (nine samples) are misclassified as containing broken low cloud. Of the nine misclassified clear-sky samples, five contain sizeable amounts of smoke; the other four samples are taken from mountain ranges with strong terrain shadows.

The results for the single-level cloud samples indicate that the classifier does very well at placing clouds within the defined height category, but it has more of a problem in determining the cloud fraction within the height category. Approximately 1.6% of the 1020 total samples are misclassified as multilayered clouds. Further inspection into the classification process for high clouds shows that samples that contain extremely thin cirrus, such as thin cirrostratus with little if any texture, may be misclassified as low- or midlevel cloud.

2) MARITIME TROPICAL AIR MASS OVER LAND (mT OVER LAND)

The overall accuracy of the mT over land module is 85.6%, with detailed results shown in Table 7. For this module, the most difficult part of the labeling process was to obtain a significant number of uniform (nonbroken) midlevel cloud samples to build the uniform mid-

TABLE 7. Matrix of classifications for hold-one-out accuracy test (%) for cloud-free and single cloud layer samples for the mT over land module. Fuzzy logic classification is presented in columns; expert classification in rows. Overall accuracy for the mT over land module, determined by dividing the number of correctly classified samples by the total number of samples (1372), is 85.6%. The "broken" designation refers to a cloud sample (32 × 32 pixel array) having a cloud fraction less than 1, and the "uniform" designation refers to a cloud fraction of 1.

Classes	Clear	Broken low	Uniform low	Broken mid	Uniform mid	Broken high	Uniform high	Thick high	Multiple layers
Clear	99.6	0.4	0	0	0	0	0	0	0
Broken low	4.8	87.1	0.7	2.0	0	0	0	0	5.4
Uniform low	0	7.5	90.6	0	0	0	0	0	1.9
Broken mid	0	1.4	0	84.3	6.0	1.4	0	0	6.9
Uniform mid	0	0	1.6	8.9	82.3	6.5	0	0	0.8
Broken high	0.5	2.1	0	2.1	0	72.9	17.6	1.1	3.7
Uniform high	0.7	0	0	0	0.7	14.5	73.2	8.0	2.9
Uniform thick high	0	0	0	0	0	0	10.3	89.7	0

TABLE 8. Matrix of classifications for hold-one-out accuracy test (%) for cloud-free and single cloud layer samples for the cT over land module. Fuzzy logic classification is presented in columns; expert classification in rows. Overall accuracy for the cT over land module, determined by dividing the number of correctly classified samples by the total number of samples (1231), is 87.4%. The "broken" designation refers to a cloud sample (32 × 32 pixel array) having a cloud fraction less than 1, and the "uniform" designation refers to a cloud fraction of 1.

Classes	Clear	Broken low	Uniform low	Broken mid	Uniform mid	Broken high	Uniform high	Thick high	Multiple layers
Clear	91.6	8.4	—	0	—	0	0	0	0
Broken low	1.8	90.7	—	4.6	—	0.4	0	0.4	2.1
Uniform low	—	—	—	—	—	—	—	—	—
Broken mid	0	3.6	—	84.4	—	2.1	0	0	9.9
Uniform mid	—	—	—	—	—	—	—	—	—
Broken high	0.4	4.3	—	1.3	—	80.8	9.8	0	3.4
Uniform high	0	0.6	—	0	—	12.8	82.3	3.1	1.2
Thick uniform high	0	0	—	0	—	0	3.0	97.0	0

level cloud class. Analysts noted biomass burning in 26% of the clear-sky samples, 5% of the broken low-level cloud samples, 2% of the broken midlevel cloud samples, and 3% of the broken and uniform high-level cloud samples. However, only one of the clear-sky samples is misclassified as containing low cloud; in actuality, thick smoke almost completely covers the misclassified sample.

As with the E over land module, the overall results indicate that the classifier does very well at placing clouds within the defined height category, but has more of a problem in determining the cloud fraction within the height category. Of the 1372 total samples collected, 2.8% of the single-level cloud samples are misclassified as multilayered clouds. The misclassification of single-level cloud samples as containing multiple clouds seems to occur more often with the broken cloud classes. Further investigation indicates that the misclassified samples contained well-defined cloud shadows within the array.

3) CONTINENTAL TROPICAL AIR MASS OVER LAND (cT OVER LAND)

The cT over land module (Table 8) has an overall accuracy for clear-sky and single-layer samples of 87.4%. So few samples of uniform low-level cloud and uniform midlevel cloud were found in imagery designated as having a cT airmass type that these classes were disabled in the classifier. The air masses denoted as being continental in origin are much drier than the marine air masses. The majority of the broken midlevel cloud samples have the appearance of being convective Cu and have very distinctive shadows. Scenes having a cT airmass designation tend to have small-celled Cu with overlying Ci at the time of the *NOAA-11* afternoon overpass, when the prolonged solar insolation has heated the surface sufficiently to cause convection.

Biomass burning is noted by analysts in 16% of the clear-sky samples and 5% of the broken low-level cloud samples. Upon further inspection of the misclassified clear-sky samples, we find that approximately 3% of the

samples are more than half covered by thick smoke and that 5% of the samples are taken from mountains with strong terrain shadows. Evidently, the thermal contrast between the mountaintops and valleys combined with the strong reflectance change caused by the shadows causes the classifier some confusion. These results indicate the necessity for further study of cloud cover over mountains.

The results for the high-level cloud samples show that the main problem with the classifier is in distinguishing between broken and uniform samples. Part of this problem stems from the observation that cirrus and alto-stratus often do not have distinct cloud boundaries and may be present even if at an extremely low optical depth (<0.1). Even with all the ancillary data available to the analyst, the location of the cloud boundary is often unclear. This uncertainty shows up in the classifier results. Since there is no uniform low- or uniform midlevel cloud class, we can make no judgment on how well the classifier would do regarding the cloud fraction issue. Of the total 1231 samples collected, approximately 3.6% of the single-layer cloud samples are misclassified as multilayered clouds, with the highest percentage coming from the broken midlevel cloud class.

4) MARITIME TROPICAL/EQUATORIAL AIR MASS OVER WATER (mT/E OVER WATER)

The mT/E over water module (Table 9) has an overall accuracy for clear-sky and single-layer cloud samples of 88.6%. The analysts noted that it was difficult in practice to obtain broken and uniform midlevel cloud samples from these air masses. While every effort was made to accurately label samples for the training phase, the results for the uniform midlevel and broken high-level cloud classes are below 80%. Generally, the cloud samples are classified as belonging to the correct height level for low-, mid-, and high-level cloud samples. As with the other tropical airmass modules, there seems to be some confusion between the broken and uniform high-level classes. Only 25 of 1285 total samples (less than 2%) are misclassified as multilayered clouds.

TABLE 9. Matrix of classifications for hold-one-out accuracy test (%) for cloud-free and single cloud layer samples for the mT/E over water module. Fuzzy logic classification is presented in columns; expert classification in rows. Overall accuracy for the mT/E over water module, determined by dividing the number of correctly classified samples by the total number of samples (1285), is 88.6%. The "broken" designation refers to a cloud sample (32 × 32 pixel array) having a cloud fraction less than 1, and the "uniform" designation refers to a cloud fraction of 1.

Classes	Clear	Broken low	Uniform low	Broken mid	Uniform mid	Broken high	Uniform high	Thick high	Multiple layers
Clear	98.0	1.0	1.0	0	0	0	0	0	0
Broken low	0	96.0	2.3	0.7	0	0	0	0	1.0
Uniform low	0	7.1	92.9	0	0	0	0	0	0
Broken mid	0	0	0	79.0	10.9	1.7	5.0	0	3.4
Uniform mid	0	0	0	8.0	77.6	0	7.2	0	7.2
Broken high	0	0	0	1.6	3.2	73.4	20.2	0	1.6
Uniform high	0	0	0	0	6.9	4.0	81.5	3.5	4.0
Thick, uniform high	0	0	0	0	0	0	2.5	97.5	0

5) MARINE POLAR (mP) AIR MASS OVER LAND OR WATER

The mP over land module (Table 10) has an overall accuracy of 86.8%. Fire or smoke are noted in 30 of the clear-sky samples and in one each of the broken low- and broken midlevel cloud samples. The majority of the samples are collected in transitional seasons (spring and fall) from the midlatitudes in both hemispheres. The underlying surface types include a variety of ecosystem types such as forests and grassland. Further, 88 of the 212 clear-sky samples contain lakes or rivers. Quite often in transitional seasons, the lakes and rivers can have a temperature much different than that of the surrounding land. One concern is that temperature differences between the water and land surfaces may cause some confusion to the classifier with the result that clouds might be deemed present in the sample, when in fact the sample may be cloud free.

In general, classification results are above 80% for all but the uniform midlevel cloud class. Nine of the clear-sky samples are misclassified as containing low cloud. Five of these samples can be attributed to the presence of smoke, two misclassifications may be due to strong terrain shadows, and two misclassifications are given to samples containing large lakes that have a much lower temperature than the land. In response to our previously stated concern, the classifier does not tend to confuse clear-sky land samples with lakes or rivers as containing cloud. With the cloud samples, the classifier performs well at placing the samples into the correct height class but has trouble in distinguishing whether the sample is broken or uniform. Approximately 3.4% of the 1260 samples collected for the mP over land air mass are misclassified as containing multilayered clouds.

Since results for the mP over water module are so similar to that of mP over land, the results are not shown. The mP over water classifier has an overall accuracy of 84.4%, with the lowest accuracies in the broken mid- and broken high-level cloud classes (both less than 80%). It is sometimes difficult to determine which height class should be assigned to a sample since at higher latitudes, the troposphere has less depth than near the equator. Approximately 2.8% of the samples collected for the mP over water air mass are misclassified as containing multilayered clouds, with the midlevel cloud samples providing the most difficulty for the classifier.

As a general rule for the mP over land and mP over water modules, the analysts note that the midlevel cloud classes are difficult to populate with samples and attribute the problem to the difficulty of finding a sample of midlevel clouds with no other higher or lower clouds present. That is, midlevel clouds tend to co-occur with other cloud types.

TABLE 10. Matrix of classifications for hold-one-out accuracy test (%) for cloud-free and single cloud layer samples for the mP over land module. Fuzzy logic classification is presented in columns; expert classification in rows. Overall accuracy for the mP over land module, determined by dividing the number of correctly classified samples by the total number of samples (1260), is 86.8%. The "broken" designation refers to a cloud sample (32 × 32 pixel array) having a cloud fraction less than 1, and the "uniform" designation refers to a cloud fraction of 1.

Classes	Clear	Broken low	Uniform low	Broken mid	Uniform mid	Broken high	Uniform high	Multiple layers
Clear	95.8	4.2	0	0	0	0	0	0
Broken low	2.9	93.4	2.5	0	0	0	0	1.2
Uniform low	0.8	2.2	88.0	0	5.3	0	0	3.7
Broken mid	0	0	0	84.6	8.0	3.7	0.6	3.1
Uniform mid	0	0	0.6	10.1	77.0	1.1	1.1	10.1
Broken high	0	1.3	0.6	4.4	0	83.1	5.6	5.0
Uniform high	0.6	0.6	1.1	0	2.3	12.1	81.6	1.7

TABLE 11. Matrix of classifications for hold-one-out accuracy test (%) for cloud-free and single cloud layer samples for the cP over land module. Fuzzy logic classification is presented in columns; expert classification in rows. Overall accuracy for the cP over land module, determined by dividing the number of correctly classified samples by the total number of samples (1097), is 91.1%. The ''broken'' designation refers to a cloud sample (32 × 32 pixel array) having a cloud fraction less than 1, and the ''uniform'' designation refers to a cloud fraction of 1.

Classes	Clear	Broken low	Uniform low	Broken mid	Uniform mid	Broken high	Uniform high	Multiple layers
Clear	99.4	0.3	0	—	—	0	0.3	0
Broken low	2.8	90.4	1.7	—	—	0	3.4	1.7
Uniform low	1.8	5.5	92.1	—	—	0	0.6	0
Broken mid	—	—	—	—	—	—	—	—
Uniform mid	—	—	—	—	—	—	—	—
Broken high	0	2.4	1.8	—	—	80.4	14.9	0.6
Uniform high	0.4	1.1	0.4	—	—	9.8	88.0	0.4

6) CONTINENTAL POLAR AIR MASS OVER LAND (cP OVER LAND) AND cP/A OVER WATER

In general, cloud samples collected from cP and A air masses, over both land and water, tend to be stratiform in nature. While striation may be present within an individual cloud layer (i.e., multiple layers spaced closely together), the appearance of a striated cloud layer tends to look like a single layer from the satellite's perspective. If there is little thermal signature between cloud layers spaced closely together, it is difficult for an analyst to be absolutely sure that only a single cloud layer exists in a sample. Given this limitation, the overall accuracy for the cP over land module is 91.1% and for the cP/A module is 89.8%. Since the results are similar for both airmass modules, detailed results are shown in Table 11 only for the cP over land module. For both the cP over land and cP/A over water modules, the broken and uniform midlevel cloud classes are disabled because so few samples could be obtained from the imagery. Smoke or fire are noted in 15 of the 311 clear-sky samples collected over land.

For both the cP over land and cP/A over water modules, less than 1% of the 32 × 32 pixel arrays is classified as containing multilayered clouds. The higher overall accuracies of these modules as compared to the

TABLE 12. Matrix of classifications for hold-one-out accuracy test (%) for multilayered cloud samples for six airmass modules. The multilayer cloud test was not performed for the cP over land and cP/A over water airmass modules because of the difficulty in collecting a significant number of multilayered cloud samples. Fuzzy logic classification is presented in columns; expert classification in rows. The overall accuracy, determined by dividing the number of correctly classified samples by the total number of multilayered samples collected, is 71.3%.

Classes	Total # samples	Clear-sky	Low-level clouds	Mid-level clouds	High-level clouds	Multiple layers
E (land)	187	0	7.5	7.5	7.5	77.5
cT (land)	155	0	1.9	12.3	4.5	81.3
mT (land)	233	0	2.6	7.7	17.2	72.5
mT/E (water)	340	0	4.7	19.4	11.5	64.4
mP (land)	235	0	1.7	12.8	12.8	72.7
mP (water)	211	0	5.2	17.5	10.9	66.4

other modules may be due in part to the disabling of the midlevel cloud classes. While the overall accuracy of the cP over land classifier is slightly higher than that for cP/A over water, the ability of the classifier to discriminate clear-sky samples from cloud samples is better over water than over land. The misclassifications of the uniform high-level cloud samples occur for very thin cirrus.

b. Multiple-layered clouds

The previous section presented the facility of the FLC to discriminate between clear-sky and single-layer cloud samples. In this section we address the second goal of this study, specifically to investigate the ability of the classifier to discriminate between samples that contain one cloud layer from those that contain more than one cloud layer. Note that we do not train the classifier in any way with samples containing multilayered clouds but only use a collection of multilayered cloud samples to test the classifier performance.

One way of gaining insight as to the performance of the various cloud classification modules is to test each classification module with a set of collected multilayered cloud samples. The samples are labeled using the procedure in section 3b. The multilayered samples contain only cloud layers separated by at least 2 km in height. If cloud layers are very close together, we anticipate that the classifier will have little facility in discerning more than one layer. The result of this exercise is shown in Table 12. There are no entries for the cP over land and cP/A over water classification modules because too few multilayered cloud samples could be collected by the analysts to draw any conclusions for the module performance. The samples contain a variety of multilayered cloud conditions, such as thin cirrus completely or partially covering Sc, Cu, St, Ac, or As. Note that our definition of multilayered clouds includes both adjacent and overlapping clouds. Other samples contain Cu under conditions where convection is taking place, so that each individual cloud cell may have a different height. The overall accuracy is 71.3% for the set of collected multilayered cloud samples. Further
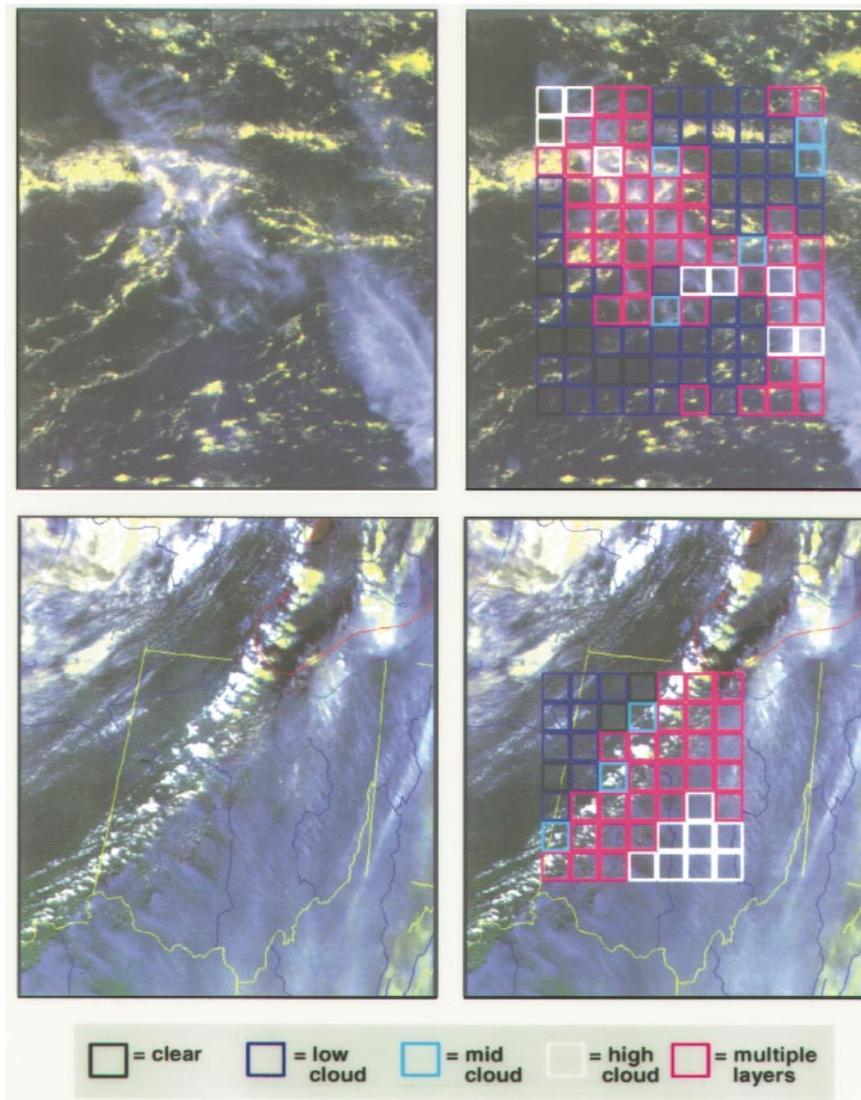
FIG. 4. Cloud classification results are provided for two scenes. Upper left: A *NOAA-11* AVHRR image recorded over the Pacific Ocean (centered at 14°S, 163°E) on 31 January 1993. The low clouds are clearly visible through much of the cirrus. False color imagery is derived from a combination of AVHRR channel 1 in the red channel, AVHRR channel 2 in the green channel, and the brightness temperature differences between AVHRR channels 3 and 4 in the blue channel. Water-covered surfaces appear dark, land surfaces range from green to brown, low clouds are bright white/yellow, and high-level clouds appear white/blue in color. Upper right: Classification results using the mT/E over water classification module. The classified results for individual 32 × 32 pixel arrays are denoted as white for high clouds, yellow for midlevel clouds, blue for low-level clouds, black for clear sky, and red for multilayered clouds. Lower left: A *NOAA-14* AVHRR image recorded at approximately 1925 UTC over the north central United States (centered over Ohio) on 20 April 1996. A cold front is passing through the region denoted in the image, with low-level clouds behind the front to the northwest, convective clouds near the front, and a cirrus shield ahead of the front that covers much of Ohio, West Virginia, and Kentucky. Lower right: Classification results are derived for 32 × 32 pixel arrays using the mP over land module.

analysis of the results provides some insight as to why the classifier fails to identify more of the samples as containing more than one cloud layer. One problem is with a certain multilayered cloud scenario—extremely thin cirrus over boundary layer cloud. It is often the case where the cirrus is so thin that it has no texture, that is, the cirrus is noticeable in the context of an image, but is uniform in texture across the data sample. When this happens, the classifier tends to label the sample as a midlevel cloud.

Figure 4 shows a graphic demonstration of how the FLC algorithm works for two different scenes. Neither scene was used in the generating the classifiers. The first scene, denoted in the upper left-hand panel, was recorded by *NOAA-11* at approximately 0400 UTC 31 January 1993 over the central Pacific Ocean. The false color image is derived from a three-channel overlay, with AVHRR channel 1 reflectance in the red channel, channel 2 reflectance in the green channel, and the brightness temperature difference between channels 3 and 4 (i.e., [3–4]) in the blue channel. Clouds having a high reflectance and a relatively low [3–4] brightness temperature difference (e.g., Cu) tend to be white/yellow, while clouds having a relatively high [3–4] brightness temperature difference (e.g., optically thin Ci, Cs, and in some cases Sc and St) will tend to have more of a bluish cast. The image depicts primarily Cu clouds with a Ci veil over much of the scene. The Ci cloud appears white/blue as compared to the bright white/yellow shades of the cumulus, and the ocean's surface appears dark. The classification results appear as colored boxes, where black represents clear-sky, blue represents low cloud, yellow represents midlevel cloud, white depicts high cloud, and red represents multilayered clouds. Each box represents a 32 × 32 pixel array of 1-km AVHRR data (approximately 35 km × 35 km square). The classifier appears to clearly recognize the arrays that contain clear-sky, low-level broken clouds, and high-level cirrus. Additionally, the classifier correctly identifies the arrays containing visually obvious multilayered clouds, although two arrays are classified (probably incorrectly) as containing midlevel cloud. When the Ci is extremely thin and has little texture, the classification is less accurate.

The lower left panel in Fig. 4 shows a *NOAA-14* image from approximately 1925 UTC 20 April 1996 over Ohio in the north central United States. The false color imagery is created using the same channel combination that is used with the previous image. Water surfaces (specifically Lake Erie in the upper-right corner of the image) are dark; land has varying shades of green and brown. State and lake boundaries are outlined in black and rivers are blue. In this scene, a weak cold front is passing over Ohio. The notable features in this figure are the Ci shield ahead of the front, the line of convective clouds near the frontal boundary, the dry slot behind the line of convection, and low-level clouds well behind the front to the northwest. The classifier results are shown in the bottom right panel. The FLC results show a steady progression from low clouds behind the front to the northwest, midlevel or multilayered clouds in the convective region near the front, and finally high-level clouds in the southwest portion of the analysis region. This analysis provides an indication of the robust nature of the classifier because the classifier was developed using only *NOAA-11* data and was not modified for application to *NOAA-14* AVHRR imagery.

A second drawback of the classifiers is that they tend to be overly sensitive to the presence of Ci. This problem stems from the feature selection process, where features are chosen to maximize the overall accuracy using the labeled sample set of clear-sky and single-layer cloud samples. Since Ci clouds are usually the most difficult to classify because of the rapidly changing microphysical and optical properties, the features chosen tend to increase the sensitivity of the classifier to the presence of Ci.

6. Summary

A fuzzy logic classification methodology is proposed to achieve the two goals of this paper: 1) to discriminate between clear-sky and clouds in a 32 × 32 pixel array, or sample, of global 1.1-km Advanced Very High Resolution Radiometer data, and 2) if clouds are present, to discriminate between single-layered and multilayered clouds within the sample. To achieve these goals, a set of eight classification modules are derived based broadly on airmass type: equatorial over land, marine tropical over land, marine tropical/equatorial over water, continental tropical over land, marine polar over land, marine polar over water, continental polar over land, and continental polar/arctic over water. The classifiers are not designed to operate over snow- or ice-covered surfaces or in extreme sunglint conditions. Derivation of airmass type is performed using gridded analyses provided by the National Center for Environmental Prediction. The fuzzy logic classifier modules are developed using sets of manually labeled samples of daytime AVHRR 1-km data. The samples are collected from 150 daytime scenes of *NOAA-11* data collected between 1991 and 1994, covering all seasons over every continent and ocean. These scenes contain a wide variety of both clear and cloudy conditions For example, clear-sky samples collected over land often contain lakes or rivers in which the water has a different temperature than the surrounding land. By using clear-sky samples with complex surface types, the classifiers are more robust. Further, no effort was made to filter out smoke or fires from clear-sky samples collected over land. If a sample contained a significant amount of smoke or fire, it was labeled as a clear-sky sample. Future research may dictate separating clear-sky samples from those that contain fire and smoke.

Spectral and textural features are computed from all five AVHRR channels to help distinguish up to eight classes: clear sky, broken low-, mid-, and high-level cloud, and uniform low-, mid-, and high-level cloud, and finally uniform, thick high-level cloud. Textural features are computed using the gray-level difference vector method and also a modification of the GLDV where the gray-level values of the original image are used [the so-called gray-level vector method].

Interestingly, in several of the modules, so few samples could be found for one or more of the cloud classes that some classes had to be disabled, such as the uniform

low-level cloud class in the E over land module, mid-level clouds in the cP over land, and cP/A over water modules. Generally, we find that midlevel cloud samples are extremely difficult to collect because of the high likelihood of finding another cloud layer present.

The "hold-one-out" method is used to estimate the theoretical accuracy of the classifier for clear-sky and single-layered cloud samples. The premise of the "hold-one-out" estimation method is that one sample is held out and the classifier is trained on the remaining samples; the remaining sample is subsequently classified. This procedure is repeated for each of the samples within an airmass module. Overall accuracies for the clear-sky and single-layer cloud classification modules range from 84% to 91%. The classification scoring matrices (Tables 6–11) show that the classifiers perform well at placing a cloud sample into the "correct" height category, but may err in determining when the sample is broken or uniform in cloud fraction. Further, only 201 samples out of a total of 9384 clear-sky and single-level cloud samples (about 2%) are misclassified as containing multilayered clouds. From these results we draw the following conclusions. First, the classifier does not tend to misclassify clear-sky samples as containing cloud if rivers, small lakes, fires, or small amounts of smoke are contained within the sample. Second, the classifier does not tend to misclassify samples containing a single cloud layer as containing multilayered clouds.

For single-layer cloud samples, three areas of difficulty are noted. First, the classifier has difficulty with midlevel cloud samples, sometimes placing the sample into a low-level or high-level cloud category. This result, in hindsight, is not unexpected since the World Meteorological Organization definition of a midlevel cloud provides a range of cloud-base temperatures as a guideline. The inclusion of midlevel clouds increases the complexity of the problem since the midlevel cloud temperature range depends on the air mass within which the clouds appear. Second, the classifier sometimes misclassifies samples containing terrain shadows, strong cloud shadows, or thick smoke. Third, if thin cirrus with little or no identifiable texture (meaning varying microphysical or optical properties) is present over a lower-level cloud layer, the sample may be misclassified as containing a midlevel cloud.

When the resulting FLC modules are tested against sets of multilayered cloud samples collected for the various airmass modules (except for cP over land and cP/A over water, where multilayered cloud samples are difficult to find), the classifier accuracies range from 64% to 81%. While the accuracies are generally lower for multilayered cloud samples than for the single-layered cloud samples, we can be reasonably confident that a sample classified as containing more than one cloud layer will actually contain more than one cloud layer. Given the complexity of global cloud classification, we are encouraged by the results for clear-sky, single-layer cloud, and multilayered cloud classification.

Because of the difficulty in finding single-layer midlevel cloud samples for training purposes and the difficulty in distinguishing between the low-, mid-, and high-level cloud classes, it may be more useful to simply define cloud classes as low or high, thereby leaving out the midlevel cloud classes. After all, the primary purpose of the FLC is to distinguish multiple from single cloud layers. Introduction of the midlevel cloud classes, while interesting from a meteorological point of view, actually tends to add to the classifier's difficulty in distinguishing between cloud heights. The potential gains in multilayered cloud classification by neglecting midlevel clouds is an aspect we intend to pursue in future research.

APPENDIX

Textural (T) and spectral (S) features selected to discriminate the classes for the eight airmass modules. (The check marks refer to the number of times the particular feature is chosen for each airmass module. Note that a feature may be chosen more than once for a particular air mass.)

Feature	E land	mT/E water	mT land	cT land	mP land	mP water	cP land	cP/A water	Feature type	Feature description
GLDV Con 3	√	√							T	Contrast computed from the GLDV of channel 3
GLDV Con 4							√		T	Contrast computed from the GLDV of channel 4
GLDV Con 5	√√	√		√	√		√		T	Contrast computed from the GLDV of channel 5
GLDV Mean 3							√		T	Mean computed from the GLDV of channel 3
GLDV Hom 1					√	√			T	Homogeneity computed from the GLDV of channel 1
GLDV Ent 1					√				T	Entropy computed from the GLDV of channel 1
GLDV Ent 2								√	T	Entropy computed from the GLDV of channel 2
Difference [3–4]						√			S	Difference of the GL means of channels 3 and 4
Difference [4–5]						√			S	Difference of the GL means of channels 3 and 4
GLDV Shade 1				√					T	Shade computed from the GLDV of channel 1
GLDV Shade 3	√		√			√			T	Shade computed from the GLDV of channel 3
GLDV Shade 4	√							√	T	Shade computed from the GLDV of channel 4
GLDV Prom 1			√		√			√	T	Prominence computed from the GLDV of channel 1
GLDV Prom 4	√				√				T	Prominence computed from the GLDV of channel 4
NDI 3,4			√						S	Difference index computed for the GL means of channels 3 and 4
NDI 3,5				√					S	Difference index computed for the GL means of channels 3 and 5
NDI 4,5			√	√				√	S	Difference index computed for the GL means of channels 4 and 5
GLV Mean 5	√	√	√	√	√	√	√	√	T	Mean computed from the GLV of channel 5
GLV Con 2							√		T	Contrast computed from the GLV of channel 2
GLV Con 3		√			√				T	Contrast computed from the GLV of channel 3
GLV Hom 1				√					T	Homogeneity computed from the GLV of channel 1
GLV Hom 2	√								T	Homogeneity computed from the GLV of channel 2
GLV Ent 1			√						T	Entropy computed from the GLV of channel 1
GLV Ent 2								√	T	Entropy computed from the GLV of channel 2
GLV Ent 3						√			T	Entropy computed from the GLV of channel 3
GLV Ent 4		√							T	Entropy computed from the GLV of channel 4
GLV Ent 5					√		√	√	T	Entropy computed from the GLV of channel 5
GLV Asm 1					√	√			T	Angular second moment computed from the GLV of channel 1
GLV Asm 2			√						T	Angular second moment computed from the GLV of channel 2
GLV Asm 4								√	T	Angular second moment computed from the GLV of channel 4
GLV Shade 3		√							T	Shade computed from the GLV of channel 3

APPENDIX. (*Continued*).

Feature	E land	mT/E water	mT land	cT land	mP land	mP water	cP land	cP/A water	Feature type	Feature description
GLV Prom 1	√						√		T	Prominence computed from the GLV of channel 1
GLV Prom 3				√			√√		T	Prominence computed from the GLV of channel 3
GLV Prom 4								√	T	Prominence computed from the GLV of channel 4
Std. dev. 1			√						T	Standard deviation computed from GLDV of channel 1
Std. dev. 2								√	T	Standard deviation computed from GLDV of channel 2
GLV Std. dev. 1		√							T	Standard deviation computed from GLV of channel 1
GLV Std. dev. 2							√		T	Standard deviation computed from GLV of channel 2
GLV Std. dev. 3								√	T	Standard deviation computed from GLV of channel 3
GLV Std. dev. 4			√						T	Standard deviation computed from GLV of channel 4
Mean spatial coherence 2, 4 × 4		√							S	Mean of all the 4 × 4 pixel groups of channel 2 with $\sigma < 2.5$
Std. dev. spatial coherence 2, 2 × 2	√√								S	Standard deviation of the 2 × 2 pixel groups of channel 2 with $\sigma < 2.5$
Std. dev. spatial coherence 2, 4 × 4					√	√√			S	Standard deviation of the 4 × 4 pixel groups of channel 2 with $\sigma < 2.5$
Std. dev. spatial coherence 3, 4 × 4					√				S	Standard deviation of the 4×4 groups of channel 3 with $\sigma < 2.5$
Std. dev. spatial coherence 4, 4 × 4							√		S	Standard deviation of the 4×4 pixel groups of channel 4 with $\sigma < 2.5$
High GL 1				√					S	GL mean of the highest 2% values of channel 1
High GL 2								√	S	GL mean of the highest 2% values of channel 2
Low GL 3					√	√			S	GL mean of the lowest 2% values of channel 3
Low GL 4				√					S	GL mean of the lowest 2% values of channel 4
Low GL 5	√√√ √√	√√	√√	√	√√	√√	√√	√	S	GL mean of the lowest 2% values of channel 5
[High–Low] GL 1				√	√	√√√			S	Difference of the mean of the top 2% to that of the mean of the bottom 2% of channel 1 GL values
Run length 2	√								S	The absolute difference of the maximum GL runs in each direction (x and y) for channel 2
Run length 3								√	S	The absolute difference of the maximum GL runs in each direction (x and y) for channel 3
Ratio [3–4]								√	S	Ratio of the top 2% to the bottom 2% of [3–4] GL values.
Top 2% OVL 14		√							S	Packed integer formed from the mean GL of the top 2 percent of the pixels from channels 1 and 4
Top 1% OVL 14			√			√			S	Packed integer formed from the mean GL of the top 1 percent of the pixels from channels 1 and 4
Top 1% OVL 134					√√				S	Packed integer formed from the mean GL of the top 1 percent of the pixels from channels 1, 3, and 4

APPENDIX. (*Continued*).

| Feature | Airmass type | | | | | | | | Feature type | Feature description |
	E land	mT/E water	mT land	cT land	mP land	mP water	cP land	cP/A water		
Con [3–4]		√			√				T	Contrast computed from the GLDV of the channel created from the differences of channels 3 and 4
Con [2/1]	√								T	Contrast computed from the GLDV of the channel created from the ratio of channels 2 and 1
Con H [1, 2, 3–4]					√				T	Contrast computed form the GLDV of the channel created from the hue values of [1, 2, 3–4]
Con H [4, 5, 1]					√	√		√	T	Contrast computed from the GLDV of the channel created from the hue values of [4, 5, 1]
Con I [4, 5, 1]	√					√√	√		T	Contrast computed from the GLDV of the channel created from the intensity values of [4, 5,1]
Mean [1–2]		√							T	Mean computed from the GLDV of the channel created from [1–2]
Mean H [1, 2, 3–4]			√				√		T	Mean computed from the GLDV of the channel created from the hue values of [1, 2, 3–4]
Mean H [4, 5, 1]				√				√	T	Mean computed from the GLDV of the channel created from the hue values of [4, 5, 1]
Mean I [4, 5, 1]								√	T	Mean computed from the GLDV of the channel created from the intensity values of [4, 5, 1]
Mean S [1, 2, 3–4]		√√							T	Mean computed from the GLDV of the channel created from the saturation values of [1, 2, 3–4]
Mean S [4, 5, 1]							√		T	Mean computed from the GLDV of the channel created from the saturation values of [4, 5, 1]
Hom H [4, 5, 1]	√								T	Homogeneity computed from the GLDV of the channel created from the hue values of [4, 5, 1]
Hom I [1, 2, 3–4]		√							T	Homogeneity computed from the GLDV of the channel created from the intensity values of [1, 2, 3–4]
Hom S [4, 5, 1]		√				√			T	Homogeneity computed from the GLDV of the channel created from the saturation values of [4, 5, 1]
Ent [4–5]				√					T	Entropy computed from the GLDV of the channel created from the differences of channels 4 and 5
Ent I [1, 2, 3–4]		√							T	Entropy computed from the GLDV of the channel created from the intensity values of [1, 2, 3–4]
Ent I [4, 5, 1]		√							T	Entropy computed from the GLDV of the channel created from the intensity values of [4, 5,1]
Ent S [1, 2, 3–4]						√			T	Entropy computed from the GLDV of the channel created from the saturation values of [1, 2, 3–4]
Ent H [4, 5, 1]							√√		T	Entropy computed from the GLDV of the channel created from the hue values of [4, 5, 1]
Asm H [1, 2, 3–4]	√								T	Angular second moment computed from the GLDV of the channel created from the hue values of [1, 2, 3–4]
Asm H [4, 5, 1]						√			T	Angular second moment computed from the GLDV of the channel created from the hue values of [4, 5, 1]
Asm I [4, 5, 1]		√							T	Angular second moment computed from the GLDV of the channel created from the intensity values of [4, 5, 1]

APPENDIX. (*Continued*).

Feature	E land	mT/E water	mT land	cT land	mP land	mP water	cP land	cP/A water	Feature type	Feature description
GLV Mean [1–2]							√		S	GLV mean of the channel created from the differences of channels 1 and 2
GLV Mean [2/1]	√								S	GLV mean of the channel created from the ratios of channels 2 and 1
GLV Mean [4–5]		√							S	GLV mean of the channel created from the differences of channels 4 and 5
GLV Mean I [1, 2, 3–4]					√				S	GLV mean of the channel created from the intensity values obtained from using channels [1, 2, 3–4]
GLV Mean S [1, 2, 3–4]							√		S	GLV mean of the channel created from the saturation values of [1, 2, 3–4]
S [1, 2, 3–4]		√							S	GL mean of the channel created from the saturation values of [1, 2, 3–4]
H [1, 3, 4]		√		√					S	GL mean of the channel created from the hue values of [1, 3, 4]
H [1, 2, 3–4]						√			S	Gray level mean of the channel created from the hue values of [1, 2, 3–4]
I [1, 3, 4]							√		S	GL mean of the channel created from the intensity values of [1, 3, 4]

REFERENCES

Bankert, R. L., 1994: Cloud classification of AVHRR imagery in maritime regions using a probabilistic neural network. *J. Appl. Meteor.,* **33,** 909–918.

Baum, B. A., and B. A. Wielicki, 1994: Cirrus cloud retrieval using infrared sounding data: Multilevel cloud errors. *J. Appl. Meteor.,* **33,** 107–117.

——, and Coauthors, 1995: Satellite remote sensing of multiple cloud layers. *J. Atmos. Sci.,* **52,** 4210–4230.

Brown, J. W., O. B. Brown, and R. H. Evans, 1993: Calibration of Advanced Very High Resolution Radiometer infrared channels: A new approach to nonlinear correction. *J. Geophys. Res.,* **98,** 18 257–18 268.

Chatterjee, S., and S. Chatterjee, 1983: Estimation of misclassification probabilities by bootstrap methods. *Commun. Stat. Simul. Comp.,* **12,** 645–656.

Chen, D. W., S. K. Sengupta, and R. M. Welch, 1989: Cloud field classification based upon high spatial resolution textural features, Part 2, Simplified vector approaches. *J. Geophys. Res.,* **94,** 14 749–14 765.

Coakley, J. A., Jr., and F. P. Bretherton, 1982: Cloud cover from high resolution scanner data: Detecting and allowing for partially filled fields of view. *J. Geophys. Res.,* **87,** 4917–4932.

Devijer, P. A., and J. Kittler, 1982: *Pattern Recognition: A Statistical Approach.* Prentice Hall, 448 pp.

Ebert, E., 1987: A pattern recognition technique for distinguishing surface and cloud types in polar regions. *J. Climate Appl. Meter.,* **26,** 1412–1427.

——, 1989: Analysis of polar clouds from satellite imagery using pattern recognition and a statistical cloud analysis scheme. *J. Appl. Meteor.,* **28,** 382–399.

Foley, J. D., and A. VanDam, 1984: *Fundamentals of Interactive Computer Graphics.* Addison-Wesley, 664 pp.

Garand, L., 1988: Automated recognition of oceanic cloud patterns. Part I: Methodology and application to cloud climatology. *J. Climate,* **1,** 20–39.

Giarratano, J., and G. Riley, 1989: *Expert Systems: Principles and Programming.* PWS-Kent, 632 pp.

Hahn, C. J., S. G. Warren, J. London, R. M. Cherrin, and R. Jenne, 1982: Atlas of simultaneous occurrence of different cloud types over the ocean. NCAR Tech. Note TN-201+STR, 211 pp.

[Available from National Center for Atmospheric Research, Boulder, CO 80307.]

——, ——, ——, ——, and ——, 1984: Atlas of simultaneous occurrence of different cloud types over land. NCAR Tech. Note TN-241+STR, 209 pp. [Available from National Center for Atmospheric Research, Boulder, CO 80307.]

Haralick, R., M., K. S. Shanmugam, and I. Dinstein, 1973: Textural features for image classification. *IEEE Trans. Syst. Man. Cybern.,* **SMC-3** (6), 610–621.

Inoue, T., 1987: A cloud type classification with NOAA 7 split-window measurements. *J. Geophys. Res.,* **92,** 3991–4000.

Key, J., 1990: Cloud cover analysis with arctic AVHRR data, Part II: Classification with spectral and textural measures. *J. Geophys. Res.,* **95,** 7661–7675.

——, J. A. Maslanik, and A. J. Schweiger, 1989: Classification of merged AVHRR and SMMR Arctic data with neural networks. *Photogramm. Eng. Remote Sens.,* **55,** 1331–1338.

Kidwell, K. B., 1997: NOAA polar orbiter data users guide. NOAA National Climatic Data Center, Satellite Data Services Division, 412 pp. [Available from National Climatic Data Center, 151 Patton Ave., Room 120, Asheville, NC 28801-5001.]

Lee, J., R. Weger, S. K. Sengupta, and R. M. Welch, 1990: A neural network approach to cloud classification. *IEEE Trans. Geosci. Remote Sens.,* **28,** 846–855.

Menzel, W. P., D. P. Wylie, and K. I. Strabala, 1992: Seasonal and diurnal changes in cirrus clouds as seen in four years of observations with the VAS. *J. Appl. Meteor.,* **31,** 370–385.

Pankiewicz, G. S., 1995: Pattern recognition techniques for the identification of cloud and cloud systems. *Meteor. Appl.,* **2,** 257–271.

Peak, J. E., and P. M. Tag, 1992: Toward automated interpretation of satellite imagery for navy shipboard applications. *Bull. Amer. Meteor. Soc.,* **73,** 995–1008.

Penaloza, M. A., and R. M. Welch, 1996: Feature selection for classification of polar regions using a fuzzy expert system. *Remote Sens. Environ.,* **58,** 81–100.

Rabindra, P., S. K. Sengupta, and R. M. Welch, 1992: An interactive hybrid expert system for polar cloud and surface classification. *Environmetrics,* **3,** 121–148.

Rao, C. R. N., and J. Chen, 1994: Post-launch calibration of the visible and near-infrared channels of the Advanced Very High Resolution Radiometer on NOAA-7, -9, and -11 spacecraft. NOAA

Tech. Rep. NESDIS 78, 22 pp. [Available from National Technical Information Service, U.S. Dept. of Commerce, Sills Bldg., 5285 Port Royal Road, Springfield, VA 22161.]

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.

Simpson, J., R. F. Adler, and G. R. North, 1988: A proposed Tropical Rainfall Measuring Mission (TRMM) satellite. *Bull. Amer. Meteor. Soc.,* **69,** 278–295.

Tian, L., and J. A. Curry, 1989: Cloud overlap statistics. *J. Geophys. Res.,* **94,** 9925–9935.

Tovinkere, V. R., M. Penaloza, A. Logar, J. Lee, R. C. Weger, T. A. Berendes, and R. M. Welch, 1993: An intercomparison of artificial intelligence approaches for polar scene identification. *J. Geophys. Res.,* **98,** 5001–5016.

Warren, S. G., C. J. Hahn, and J. London, 1985: Simultaneous occurrence of different cloud types. *J. Climate Appl. Meteor.,* **24,** 658–667.

Welch, R. M., S. K. Gupta, and K. S. Kuo, 1988: Marine stratocumulus cloud fields off the coast of southern California observed using LANDSAT imagery. Part II: Textural analysis. *J. Appl. Meteor.,* **27,** 363–378.

——, M. S. Navar, and S. K. Sengupta, 1989: The effect of spatial resolution upon texture-based cloud field classifications. *J. Geophys. Res.,* **94,** 14 767–14 781.

——, K. S. Kuo, and S. K. Sengupta, 1990: Cloud and surface textural features in polar regions. *IEEE Trans. Geosci. Remote Sens.,* **28,** 520–528.

——, S. K. Sengupta, A. K. Goroch, P. Rabindra, N. Rangaraj, and M. S. Navar, 1992: Polar cloud and surface classification using AVHRR imagery: An intercomparison of methods. *J. Appl. Meteor.,* **31,** 405–420.

Weszka, J. S., C. R. Dyer, and A. Rosenfeld, 1976: A comparative study of texture measures for terrain classification. *IEEE Trans. Syst. Man. Cybern.,* **SMC-6** (4), 2269–2285.

Wielicki, B. A., R. D. Cess, M. D. King, D. A. Randall, and E. F. Harrison, 1995: Mission to planet earth: role of clouds and radiation in climate. *Bull. Amer. Meteor. Soc.,* **76,** 2125–2153.