JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 102, NO. D2, PAGES 1739-1756, JANUARY 27, 1997

Cloud fraction errors caused by finite resolution measurements

Larry Di Girolamo[1]

Department of Atmospheric and Oceanic Sciences, McGill University, Montreal, Quebec, Canada

Roger Davies

Institute of Atmospheric Physics, University of Arizona, Tucson

Abstract. The standard method of deriving cloud fraction from space simply finds the fraction of total image pixels that contains some cloud. The sensitivity of this method to sensor resolution has been examined by *Shenk and Salomonson* [1972], assuming the cloudy pixels are detected perfectly. Their experiment is reexamined to show that the sensitivity is much more complex than they predicted. We derive the upper and lower bounds of the true cloud fraction, A_t, given the standard method estimate, A_e, to show that the range of possible values for A_t can be, in general, very wide. By including the fraction of apparent cloud edge and cloud interior pixels, the bounds can be reduced and improvements to the standard method can be obtained. However, this improvement is also resolution limited by the misidentification of partially cloudy pixels as cloud interior rather than cloud edge. A potentially better technique for estimating the true cloud fraction is therefore explored using a pattern recognition approach. A nearest neighbor classification rule is used in two sets of experiments: one using 684 simulated cloud fields as a training set, the other using 370 cloud fields based on Advanced Very High Resolution Radiometer (AVHRR) measurements. Given the underlying distribution of A_t, A_e overestimates A_t with an overall average bias of 32% and standard error of 11% for the simulated training set, and a bias of 35% and a standard error of 3% for the AVHRR training set. The pattern recognition estimate, A_p, is essentially unbiased and has a standard error of 12% for both training sets. The relevance of these training sets to new scenes and the importance of imperfect cloud detection have yet to be investigated, but the pattern recognition technique shows considerable potential advantage over the standard technique in providing unbiased estimates of cloud fraction that are less sensitive to the effects of sensor resolution.

1. Introduction

Clouds have long been recognized for their important role in the Earth's climate system. Yet our understanding of their microphysical and macrophysical properties, their interaction with the environment, and their influence on a changing climate remains incomplete [e.g., *Arakawa*, 1975; *Cess et al.*, 1990; *Arking*, 1991]. To improve our understanding of cloud and cloud processes, further monitoring is necessary. To this end, programs such as the International Satellite Cloud Climatology Project (ISCCP) [*Schiffer and Rossow*, 1983; *Rossow and Schiffer*, 1991] have been extracting numerous cloud variables through the interpretation of satellite measurements. One of these variables is cloud fraction.

The importance of cloud fraction as a variable in the physics and parameterizations of cloud processes has been demonstrated in the literature at least as far back as *London* [1957]. Since then, cloud fraction has been shown to be important in, for example, the parameterization of cloud field albedo [e.g., *Welch and Wielicki*, 1989; *Kobayashi*, 1988] and in the derivation of the surface radiation budget [e.g., *Frouin et al.*, 1988]. Indeed, cloud fraction is recognized as a critical variable required to accurately assess cloud feedback effects in climate [e.g., *Harshvardhan*, 1982; *Arking*, 1991]. It is no wonder that accurate measurements of cloud fraction have been put on the ISCCP high-priority list [*Schiffer and Rossow*, 1983].

Unfortunately, obtaining good estimates of cloud fraction from satellite measurements has long been a problem [*Rossow*, 1989; *Wielicki and Parker*, 1992] (hereinafter WP92). The problem can be coarsely broken into four parts: cloud definition, cloud fraction definition, resolution effects, and threshold effects. The definitions of cloud and cloud fraction are discussed

[1]Now at Institute of Atmospheric Physics, The University of Arizona, Tucson.

Paper number 96JD02663.
0148-0227/97/96JD-02663$09.00

in section 2. Resolution effects refer to the problems caused by the measuring instrument's finite field of view. Threshold effects include such factors as instrument signal-to-noise ratio, the amount of contrast between cloud and the background, and the logic used to determine the threshold of an observed quantity that separates clear from cloudy. Note that these problems are not necessarily independent from each other (e.g., the threshold effect depends on the definition of cloud and on the resolution of the measurements). For this reason, it has been difficult to unravel any one particular effect; a combination of the four is usually what is under investigation [e.g., *Wielicki and Welch*, 1986; WP92; *Key*, 1994].

This study focuses on isolating and examining the resolution effect and its associated errors on measurements of cloud fraction. Section 2 reviews the definition of cloud and cloud fraction and defines the "perfect cloud detector," which allows the resolution effect to become independent of the other three problems mentioned above. A similar approach was taken by *Shenk and Salomonson* [1972] (hereinafter SS72), and in section 3 we revisit their experiment to shed additional light on this problem. We discover errors in their experimental results concerning the effect of sensor resolution and cloud size on estimates of cloud fraction. Section 4 develops a mathematical framework to set upper and lower limits on our estimate of the true cloud fraction and derives several equations that describe the dependence of measured cloud fraction on resolution. The derivation of these equations helps to reveal why the analysis of SS72 breaks down. The outcome of sections 3 and 4 stresses the importance of cloud pattern in determining the resolution effects. Therefore section 5 takes a pattern recognition approach to estimating the true cloud fraction and compares the errors with the "standard method" (defined in section 2) and with (17) derived in section 4. Section 6 summarizes the results of this study.

2. A Matter of Definition

There are numerous ways in which satellite data can be used for cloud detection (for a review, see *Rossow et al.* [1985], *Goodman and Henderson-Sellers* [1988], and *Rossow* [1989]). However, a long-standing problem with satellite cloud detection algorithms is obtaining a proper method of validation. This is due to the lack of a "truth" data set with which to compare the derived cloud cover [e.g., *Rossow et al.*, 1985; *Rossow*, 1989]. One of the underlying problems in obtaining a "truth" data set lies in how we currently define "cloud."

Most attempts at defining cloud have been measurement-based [*Rossow*, 1989]. For example, all cloud detection techniques have some defined threshold in that some value of an observed quantity divides the population of the measured quantity into either clear or cloudy categories. Although this may prove useful in many cloud research areas, the definition of cloud is based on instrument and algorithm performance. This contributes to a large variance in estimated cloud fraction by different instruments and algorithms [e.g., *Hughes*, 1984; *Di Girolamo*, 1992; WP92]. In practice, we may forever have to use such a definition; however, the problem is more fundamental in that a precise quantitative definition of cloud does not presently exist.

Clouds are conventionally defined as "a mass of water or ice particles suspended in air." This definition lacks quantitative precision. For example, do two suspended water droplets 10 μm in diameter, placed 200 m apart, constitute a cloud? If so, present state-of-the-art satellite cloud detection methods will need some refinements! If not, then the first step in establishing any sort of "truth" is to establish a precise definition of cloud. Note that any definition of cloud must include a minimum scale over which the cloud exists.

Cloud fraction is defined as the fractional area coverage of cloud over a particular background domain [e.g., *Henderson-Sellers and McGuffie*, 1990]. A common approach to estimating cloud fraction from satellite images is to compute the fraction of total image pixels that contains some cloud and call this the "standard method" for estimating cloud fraction. Note that pixels labeled as "cloudy" may be only partially cloud covered with reference to some established cloud definition. Several methods have been proposed to explicitly account for partially cloud-filled pixels [e.g., *Coakley and Bretherton*, 1982; *Arking and Childs*, 1985; *Lovejoy et al.*, 1987; *Stowe et al.*, 1988]. However, these methods all rely on further assumptions about the cloud field, assumptions that are open for debate [e.g., WP92]. For the remainder of this study, only the standard method in estimating cloud fraction is examined, and it is compared to (17) derived in section 4 and to a pattern recognition approach for estimating cloud fraction in section 5.

Since we are interested in isolating the resolution effects from threshold and definition effects, we will assume that measurements are made using a "perfect cloud detector." The perfect cloud detector offers perfect cloud detection in that any amount of cloud, however defined, within an image pixel flags that pixel as "cloudy." This approach eliminates threshold effects. It also eliminates definition effects, except for the minimum scale over which the cloud is defined. Thus the standard method estimate of cloud fraction is examined in this study assuming a cloud-segmented image obtained from a perfect cloud detector. This approach to studying resolution effects on derived cloud fraction was first introduced by SS72, except that the definition effects were not discussed. We now reexamine the work done by SS72 to gain better insight into the problem.

3. Paper Cloud Experiment Revisited

The paper cloud experiment of SS72 gave the first demonstration of how finite sensor resolution measurements can affect the interpretation of cloud fraction.

Their experimental apparatus consisted of paper clouds placed on a black background with a grid placed over top, the resolution of which they could vary. An observer then assessed the state of cloudiness of each pixel, for more than 2 million pixels! Since SS72 also assumed perfect detection as defined in section 2, the observer included pixels that contained even the smallest amount of cloud as "cloudy" (their Method 1). Their experiment further assumed the standard method estimate of cloud fraction, A_e, to be a function of only the true cloud fraction, A_t, and a parameter, R, that is equal to the ratio of the true average areal cloud area, λ, to sensor resolution area, r^2; that is,

$$A_e = A_e(R, A_t). \tag{1}$$

Their experimental results showed this to be true for all cloud patterns so long as there were no large clear areas within the image (i.e., for what they called "uniformly distributed clouds"). The results are summarized in Figure 2 of SS72 and show that considerable overestimates of the true cloud fraction occur for small R.

SS72 concluded that if R can be estimated, then the empirically derived equation (1) can be used to obtain a better estimate of the true cloud fraction compared to A_e alone. They suggested that λ (hence R) can be estimated from the original coarse resolution image, even though λ will be overestimated due to coarse resolution measurements. However, R can be better estimated by measuring A_e at two different resolutions. From (1) we can write

$$dA_e = \frac{\partial A_e}{\partial R} dR + \frac{\partial A_e}{\partial A_t} dA_t. \tag{2}$$

For any given scene, the average cloud size, λ, is constant, so that

$$\frac{\partial A_e}{\partial R} = -\frac{r}{2R} \frac{\partial A_e}{\partial r}; \tag{3}$$

thus

$$dA_e = -\frac{r}{2R} \frac{\partial A_e}{\partial r} dR + \frac{\partial A_e}{\partial A_t} dA_t. \tag{4}$$

Consider a cloud field under analysis. The true cloud fraction does not change; so, $dA_t = 0$. Measurements of $\partial A_e/\partial r$ are made by degrading the spatial resolution of the sensor. Equation (1) is obtained empirically, and from this $(dA_e/dR)(A_t)$ is obtained. This leaves two equations [(1) and (4)] with two unknowns; hence, by measuring A_e and $\partial A_e/\partial r$, R and A_t can in principle be solved, given the behavior of Figure 2 of SS72.

This approach requires considerable accuracy in (1). We thus recreate the experiment of SS72, except that we make use of the modern computer that allows us to use many more pixels in our image with an enormous saving in time compared to 1972 technology. However, by discretizing our clouds, we essentially define our smallest element as the true cloud scale. This remains consistent with our discussion in section 2, in that any established definition of cloud must include a scale over which the definition holds true. Further discussion on this point is given in section 4. Also, SS72 considered three different cloud patterns in their experiment. We consider only their pattern 1 (regularly spaced clouds having the same size), since detailed information to recreate their patterns 2 and 3 were not provided. This should not pose a problem since they concluded that the three types of cloud pattern produced the same results.

We begin with a grid that has 2048 pixels per side and define the pixel scale as our true cloud scale. A pixel at this resolution is either fully clear or fully cloudy. We then overlay regularly spaced columns and rows of circular clouds, all having the same size (the clouds are not exactly circular due to the discrete nature of computer processing). We then degrade the resolution by a factor of 2 and count up the number of coarse resolution elements that contain some cloud. We continue to degrade by factors of 2 until $A_e = 1$.

Our experimental results revealed one important difference compared to those reported in SS72: as R became smaller, significant deviation in A_e existed among different cloud sizes (or, equivalently, a different number of cloud elements) for a given A_t. This is demonstrated in Figure 1 for two values of A_t. As the number of cloud elements increases for a particular A_t, A_e increases more slowly for a decreasing R. This behavior was observed for the range of A_t [0.05, 0.78] used in our experiment ($A_t = 0.78$ is the upper limit for regularly spaced circular clouds of the same size), as well as for different grid sizes ($512^2 - 2400^2$) and different degradation methods (i.e., degrading by the prime numbers of the image size versus by a factor of 2).

The deviation of our results from SS72 can easily be explained. In order for SS72 to obtain a complete curve

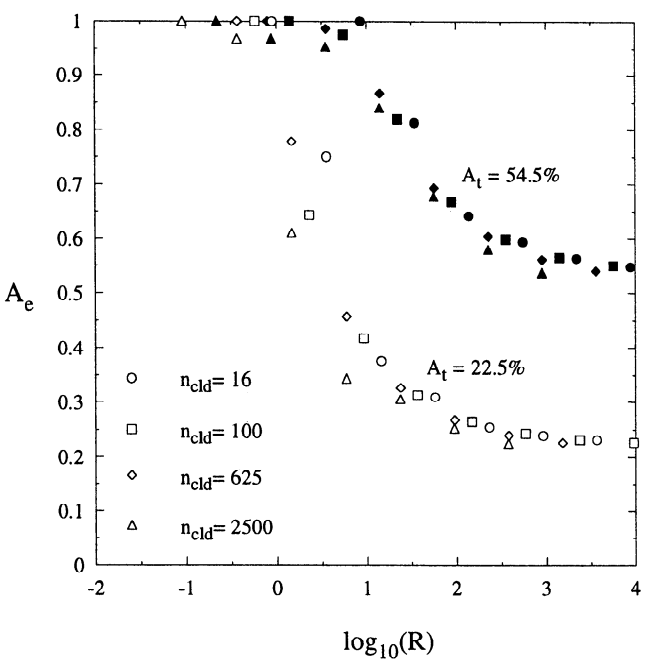

Figure 1. The behavior of A_e as a function of R (average cloud area/pixel area) for $A_t = 0.545$ and $A_t = 0.225$ for several regular cloud arrays, each having a different number of cloud elements generated from our experiment.

for a particular true cloud fraction, they had to use a range of cloud sizes (hence a range in number of cloud elements) in order to obtain a complete range of R (e.g., see Figure 1 of SS72). For example, small, medium, and large clouds were used to get small, medium, and large values of R, respectively. This was necessary due to the limitation of their experimental apparatus, a problem that our experimental apparatus does not possess. Thus we conclude from our experiment that a unique empirical relationship in the form of (1) that is applicable to cloud patterns having clouds "uniformly distributed" over the image does not exist, in contradiction to the conclusions of SS72.

Accordingly, A_e as revealed by our "paper" cloud experiment not only depends on R and A_t as stated in (1), but it also depends on the number of cloud elements, n_{cld}, within the image. One may suspect that all we need to continue is to determine an empirical relationship of the form $A_e = A_e(R, A_t, n_{\text{cld}})$ and check for its uniqueness. Unfortunately, with real satellite images, n_{cld} and λ are not always easy to define, even for perfect detection, due to their dependence on the resolution of the measurements, the irregular shape of clouds, and our assumption of the connectivity of the cloudy pixels that make up individual clouds. Moreover, we have so far assumed that the clouds are "uniformly distributed" over the scene; in reality, this is often not the case. For cloud scenes having clouds clustering within the image, the results presented in SS72 tend to overestimate the effects of resolution [SS72]. *Wielicki and Welch* [1986] came up with the same conclusion, but their results also included threshold effects, making direct comparisons to the results presented by SS72 inappropriate. Thus it appears that the above approach at characterizing the effects of sensor resolution on measurements of cloud fraction is inadequate. The next section presents a more thorough analysis of the problem, showing where (1) fails, how it is affected by subpixel cloudiness, and what the real uncertainties are in cloud fraction under various assumptions.

4. Mathematical Look at the Paper Cloud Experiment

This section takes a formal look at the paper cloud experiment in order to obtain a better understanding of the behavior of A_e with resolution. We begin by deriving the upper and lower bounds of A_t given a measurement A_e over a general cloud scene. To do so, we follow closely the mathematical methodology of *Raffy* [1993]. We also solve for A_t under various assumptions and cast it in the form of (1). The limitations of the derivation are shown, which point to where and why the assumptions used in (1) fail.

Let the resolution length of our measurement be equal to r_i for a square pixel. Consider discrete scale variations according to

$$r_i = p^{-i},\ i = 0, 1,\ 2,\ \ldots,\ t, \tag{5}$$

where p is an integer greater than 1, and r_t is the scale at which the cloud truth is (arbitrarily) defined. Note that $r_0 = 1$ defines the image scale and that the total number of pixels of scale r_i within the image is equal to p^{2i}. The experimental setup used in section 3 had $p = 2$ and $t = 11$.

Let us label each pixel of scale size r_i within the image by the index $\omega_j(r_i)$, where $j = 1, \ldots, p^{2i}$. When the resolution is increased from r_i to r_{i+1}, the pixel $\omega_j(r_i)$ gets subdivided into p^2 intervals, denoted by $\omega_j^m(r_{i+1})$ with the index $m = 1, \ldots p^2$, since (from (5))

$$r_{i+1} = \frac{r_i}{p}. \tag{6}$$

Let f be the image function whose value, measured at a resolution r_i and located at $\omega_k(r_i)$, is equal to $f_{r_i}(\omega_k)$. *Raffy* [1993] takes $f_{r_i}(\omega_k)$ to be the mean value of f over the smaller subintervals, ω_k^m, that belong to ω_k as described by *Raffy* [1993, Eq. (31)], and further uses a constant threshold that is independent of scale. However, in the paper cloud experiment, no such regularization takes place; $f_{r_i}(\omega_k)$ remains binary at all levels of r_i.

Lemma 1

(1) If the pixel $\omega_j(r_i)$ is cloudy, then $f_{r_i}(\omega_j) = 1$ and at least one $f_{r_{i+1}}\left(\omega_j^m\right) = 1,\ m \in \{1,\ \ldots,\ p^2\}$.

(2) If the pixel $\omega_j(r_i)$ is clear, then $f_{r_i}(\omega_j) = 0$ and $f_{r_{i+1}}\left(\omega_j^m\right) = 0\ \forall m,\ m \in \{1,\ \ldots,\ p^2\}$.

Let $N(r_i)$ be the total number of pixels within the image containing resolution elements of scale r_i that have $f_{r_i} = 1$. That is,

$$N(r_i) = \sum_{j=1}^{p^{2i}} f_{r_i}(\omega_j). \tag{7}$$

Using the standard method, our estimate of cloud fraction is given by

$$A_e(r_i) = \frac{N(r_i)}{p^{2i}}. \tag{8}$$

From Lemma 1, condition (1), we can immediately write

$$N(r_i) \le N(r_{i+1}) \le p^2 N(r_i),$$

which becomes

$$\frac{A_e(r_i)}{p^2} \le A_e(r_{i+1}) \le A_e(r_i) \tag{9}$$

with the substitution of (8). Repeated application of (9) to determine $A_e(r_{i+d})$ yields

$$\frac{A_e(r_i)}{p^{2d}} \le A_e(r_{i+d}) \le A_e(r_i) \qquad d \in \{1,\ \ldots,\ t-i\}. \tag{10}$$

To determine the bounds on $A_e(r_t) \equiv A_t$, we set $d = t - i$ and combine (5) and (10) to yield

$$\left(\frac{r_t}{r_i}\right)^2 A_e(r_i) \le A_t \le A_e(r_i). \quad (11)$$

As an example, say we are using satellite data having a resolution of 1 km, and we define our true cloud scale to be 10 m. If the instrument and detection are perfect, then for an observed cloud fraction of 80%, (11) sets the bounds on the true cloud fraction as $0.008\% \le A_t \le 80\%$. Physically, the upper bound occurs when all pixels detected as cloudy are fully cloud covered; the lower bound occurs when each pixel detected as cloudy contains only one cloud having a size equal to the smallest defined cloud scale (10 m in the above example). Although the lower bound situation may be rare, the derivation is general. Thus, even with perfect detection, the standard method to derive cloud fraction can leave us with a wide range of possibilities in the true cloud fraction when using coarse resolution measurements.

To narrow the bounds on A_t, we consider the case when only cloud edge pixels carry the error in $A_e(r_i)$. Therefore cloud interior pixels are assumed to be fully cloudy (i.e., our working assumption omits the possibility of partial cloudiness within a cloud interior pixel). This seems to be a reasonable assumption given the observations of *Wielicki and Welch* [1986]. To be labeled a cloud edge pixel, the pixel must be cloudy and at least one of its eight neighboring pixels must be clear. Lemma 1, condition (1), stands for cloud edge pixels. A cloud interior pixel must have all eight neighbors as cloudy.

Lemma 2

> If the pixel $\omega_j(r_i)$ is an interior cloud pixel, then $f_{r_i}(\omega_j) = 1$ and $f_{r_{i+1}}(\omega_j^m) = 1 \ \forall \ m, \ m \in \{1, \ldots, p^2\}$.

Let $N_{\text{int}}(r_i)$ and $N_{\text{edge}}(r_i)$ be the number of interior cloud pixels and cloud edge pixels, respectively, within an image having resolution r_i. Let $A_{\text{int}}(r_i)$ and $A_{\text{edge}}(r_i)$ be their respective fractional coverage with the same correspondence as (8). Note that $N(r_i) = N_{\text{int}}(r_i) + N_{\text{edge}}(r_i)$. From Lemma 2 we can immediately write

$$p^2 N_{\text{int}}(r_i) + N_{\text{edge}}(r_i) \le N(r_{i+1}) \le p^2 N(r_i). \quad (12)$$

Note that from Lemma 2,

$$N_{\text{int}}(r_{i+d}) = p^{2d} N_{\text{int}}(r_i);$$

and from Lemma 1,

$$N_{\text{edge}}(r_i) \le N_{\text{edge}}(r_{i+d}).$$

Thus repeated application of (12) to determine $N(r_{i+d})$ yields

$$p^{2d} N_{\text{int}}(r_i) + N_{\text{edge}}(r_i) \le N(r_{i+d}) \le p^{2d} N(r_i). \quad (13)$$

From our definition of fractional coverage in (8), (13) can be rewritten as

$$A_{\text{int}}(r_i) + \frac{A_{\text{edge}}(r_i)}{p^{2d}} \le A_e(r_{i+d}) \le A_e(r_i). \quad (14)$$

To determine the bounds on $A_e(r_t) \equiv A_t$, we set $d = t - i$ and combine (5) and (14) to yield:

$$A_{\text{int}}(r_i) + \left(\frac{r_t}{r_i}\right)^2 A_{\text{edge}}(r_i) \le A_t \le A_e(r_i). \quad (15)$$

As an example, say we are again using satellite data having a resolution of 1 km, and we define our true cloud scale to be at 10 m. If the detection is perfect, then for an observed internal cloud fraction of 60% and an observed edge cloud fraction of 20%, (15) sets the bounds on the true cloud fraction as $60.002\% \le A_t \le 80\%$. Thus, by assuming that only the cloud edge pixels (as defined in our coarse resolution measurements) can be partially cloudy and that cloud interior pixels are fully cloudy, we have narrowed the bounds on A_t as compared to (11).

Equations (11) and (15) show that for a finite defined true cloud scale, the bounds on A_t narrow as the resolution of the measurements increases. If the measurements are made at the true cloud scale (i.e., if $r_i = r_t$), then the standard method estimate of cloud fraction is equal to the true cloud fraction under the assumption of perfect detection.

We can proceed one step further under the assumptions used to derive (15) to obtain an exact solution of A_t using coarse resolution measurements. Say we have a large number, $N_{\text{edge}}(r_i)$, of cloud edge pixels in our image of resolution r_i. From Lemma 1, each cloud edge pixel must contain at least one cloud edge pixel of resolution r_{i+d} and at most p^{2d}. If the image has at least $N_{\text{edge}}(r_i)$ cloud edge pixels at the r_{i+d} scale, and the rest of the pixels $[(p^{2d} - 1)N_{\text{edge}}(r_i)$ of them] are equally likely to be cloudy, then we can expect $[(p^{2d} - 1)/2]N_{\text{edge}}(r_i)$ of these remaining pixels to be cloudy; therefore

$$N(r_{i+d}) = p^{2d} N_{\text{int}}(r_i) + N_{\text{edge}}(r_i) + \left(\frac{p^{2d} - 1}{2}\right) N_{\text{edge}}(r_i). \quad (16)$$

From our definition of fractional coverage in (8) and by setting $d = t - i$, (16) becomes

$$A_t = A_{\text{int}}(r_i) + \left[1 + \left(\frac{r_t}{r_i}\right)^2\right] \frac{A_{\text{edge}}(r_i)}{2}, \quad (17)$$

which is the exact solution of the true cloud fraction from coarse resolution measurements under the assumption of perfect detection, the assumption that only the cloud edge pixels (as defined in our coarse resolution measurements) can be partially cloudy, and the assumption that we have a large number of cloud edge pixels having an average of nearly 50% cloud cover. If these assumptions are met, then (17) can be used to obtain

A_t by simply measuring $A_{\text{int}}(r_i)$ and $A_{\text{edge}}(r_i)$ within the image under analysis.

Assuming an average of about 50% cloud cover for partially cloudy pixels seems to be reasonable [*Molnar and Coakley*, 1985]. However, we can expect that the assumption of only cloud edge pixels being partially cloudy and cloud interior pixels being fully cloudy holds true for scales smaller than some cutoff scale r_c; that is, $\forall\, r_i,\ i \in \{c, \ldots, t\}$. For $r_i > r_c$, it is possible that a pixel $\omega_j(r_i)$ be partially cloudy, yet still be observed as an interior cloud pixel. This situation is depicted in Figure 2. Thus, for r_i having $i \in \{1, \ldots, c-1\}$, (17) may not be valid.

The validity of (17) can be tested using our "paper" cloud experiment. To do so, Appendix A shows how (17) can be transformed into (1), yielding

$$A_e = A_t + \frac{b^2}{2} + \frac{b}{2}\left(b^2 + 4A_t\right)^{1/2}, \qquad (18)$$

where $b > 0$ and is equal to

$$b = \left(\frac{\pi A_t}{R}\right)^{1/2} - n_{\text{cld}} r_t^2 \left(\frac{\pi R}{A_t}\right)^{1/2}$$

and n_{cld} is the number of cloud elements in the image. Figure 3 shows a plot of A_e versus R for various values of A_t along with the results from our paper cloud experiment for a value of $n_{\text{cld}} = 100$. We observe that (18) agrees with the experiment for large values of R; however, as R decreases, the deviation between (18) and the experiment increases. The deviation begins below a cutoff R that appears to be a function of A_t. Indeed, the cutoff is a function of A_t as shown in Appendix B,

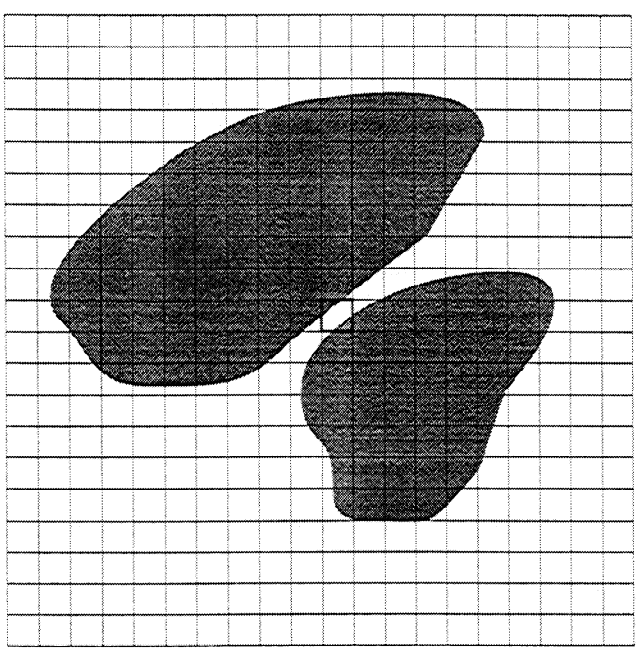

Figure 2. The highlighted pixel is an example of a partially cloud-covered pixel being labeled as a cloud interior pixel rather than a cloud edge pixel.

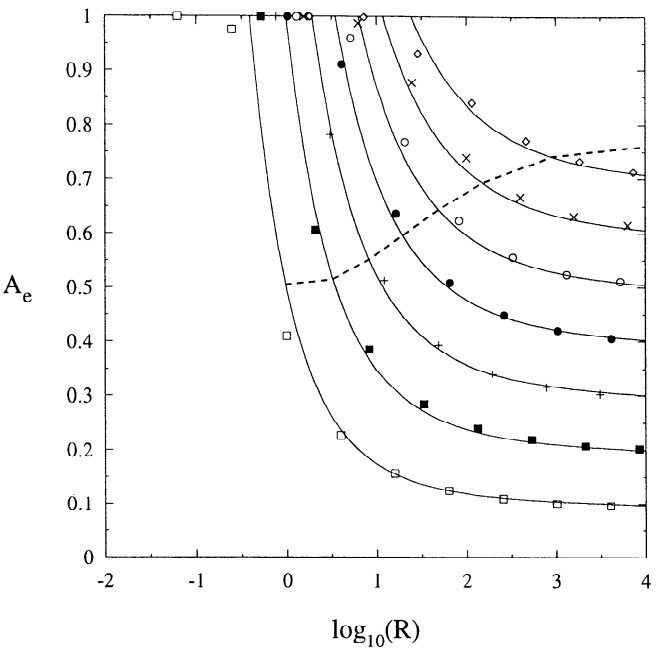

Figure 3. The behavior of A_e as a function of R for values of A_t ranging from 0.1–0.7 in steps of 0.1 as derived from (18) (solid lines) and from our experiment (symbols) for $n_{\text{cld}} = 100$. The dashed line represents the cutoff of (18), above which the validity is not guaranteed.

which derives the cutoff value of R beyond which (18) is no longer valid for an array of regularly spaced circular clouds of the same size. The cutoff is also shown on Figure 3. Note that the cutoff tends to be conservative and that (18) does reasonably well beyond the cutoff. This is expected since the population of cloud interior pixels that are partially cloudy is still small just beyond the cutoff; only when the population becomes sufficiently large does the deviation between (18) and the experiment become significant.

Figure 4 shows an example of the weak dependence (18) has on n_{cld}. The dependence of A_e on n_{cld} becomes smaller with decreasing R. This is in contrast with the experimental results presented in Figure 1, where n_{cld} has the largest effect for small R; however, this effect lies beyond the cutoff region for which (18) is valid. We can thus conclude that the empirical behavior of A_e with R, for small R, depends on the degree to which $\omega_j(r_i)$ is partially covered by more than one cloud element (see Figure 2). This, in turn, depends on the distribution of clear areas within the image.

Unfortunately, observations from meteorological satellites tend to be in the small R regime where (18), hence (17), may no longer be valid. However, in examining (18) we have gained a better insight into the reasons behind the behavior of A_e with resolution. It would appear that A_e depends on the pattern the clouds have over the image. We can hypothesize that characteristic cloud patterns result in similar values of A_e. Evidence for this hypothesis already exists in the literature. For

example, WP92 reported that when individual cloud scenes used to generate their Figure 2 (which depicts the error in cloud fraction as a function of resolution for several popular cloud detection algorithms applied to many cloud fields) were grouped in terms of cloud type, a more systematic dependence of cloud fraction error became evident. Since each cloud type tends to cluster about a certain range of patterns that are characterized by a set of feature descriptors [e.g., *Welch et al.*, 1988], support for our hypothesis exists. Thus the next logical approach is to develop a pattern recognition technique that identifies a certain cloud pattern and attempts to give a better estimate of A_t as compared to the standard estimate, A_e.

5. Pattern Recognition Approach to Estimate Cloud Fraction

Our goal in this section is to extract cloud pattern information (i.e., features) from a coarse resolution cloud mask image and to use this information to obtain a new estimate, A_p, of A_t. We will show that A_t is estimated more accurately using A_p than using the standard estimate, A_e, or estimates using (17), A_{17}. We proceed by describing the list of features used to characterize the pattern within the cloud mask image, the classification rule that gives A_p, the training set used to form the classification rule, the allocation rate estimator appropriate for the classification rule, the optimum feature vector, and finally, the results.

5.1. Feature List

Let $\mathbf{X}$ be a n-dimensional feature vector $(x_1, \ldots, x_n)$, where each x_k $(k = 1, \ldots, n)$ is a measurable feature from our image. Ideally, we would like x_k to be measured at the true cloud scale, r_t, because any measure we use to characterize the pattern will depend on resolution [e.g., *Welch et al.*, 1989]. From an implementation point of view, we have only measurements taken at r_i; therefore the features used to characterize the pattern are derived from the measurements taken at the scale of r_i. This limits the classification rule derived from a training set to be applicable only to data of the same resolution as the training set.

There are many features we can use to characterize pattern in our cloud mask image. From section 4, obvious features include $A_e(r_i)$ and $A_{\text{edge}}(r_i)$. Other popular features for the purpose of pattern recognition can be found in the literature [e.g., *Gu et al.*, 1989; *Teague*, 1980; *Haralick*, 1979]. We choose two common approaches: gray level difference statistics (GLDS) [*Weszka et al.*, 1976] and moment invariants [*Hu*, 1962]. GLDS has been very useful in describing texture within satellite images for the purpose of classifying clouds [e.g., *Gu et al.*, 1989; *Chen et al.*, 1989]. Basically, GLDS is based on statistics gathered from absolute differences between pairs of gray levels found in the image. The statistics are summarized by such parameters as

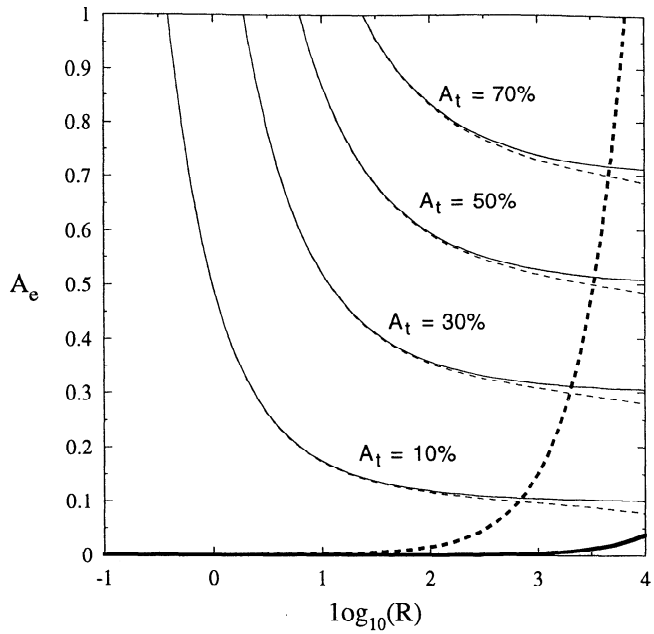

Figure 4. The behavior of A_e as a function of R for values of A_t equal to 0.1, 0.3, 0.5 and 0.7, and n_{cld} equal to 16 (solid lines) and 625 (dashed lines) derived from (18). The thick lines represent the behavior of (18) when $b = 0$; equation (18) is not valid below this line.

the mean, variance, contrast, entropy, local homogeneity, and angular second moment to name only some of the more popular ones [e.g., *Chen et al.*, 1989]. For pattern recognition, these parameters define a feature space used by a classification rule. However, many of these parameters have been empirically shown to be correlated, making some of the information content within the feature space redundant [*Baraldi and Parmiggiani*, 1995]. Indeed, it is easy to show mathematically that many of the parameters are correlated for a binary image; the mean becomes equivalent to contrast and local homogeneity, and the variance becomes equivalent to the angular second moment. Thus we choose as features the mean, variance, and entropy. A one-pixel difference is used along the x and y axes of the image to compute the GLDS statistical parameters. They are then added vectorially to produce the features that belong to the feature vector.

Moments are often classified as a shape descriptor rather than a texture descriptor [*Haralick and Shapiro*, 1991]. The moments that we use are the moment invariants of *Hu* [1961, 1962] who has derived seven moment invariants that remain invariant under image translation, rotation, and scaling. These moment invariants have since proven useful in pattern recognition problems [e.g., *Hu*, 1962; *Dudani et al.*, 1977; *Simpson et al.*, 1991]. In this study we add to the feature space the two lowest moment invariants as these contain the most information.

From (4) it is reasonable to assume that information at a degraded scale is also desirable. As a result, we double the dimension of our feature space by including

all the features mentioned above measured at a scale r_{i-1} (for $p = 2$). Thus the dimension of our feature space is $n = 14$.

5.2. Nearest Neighbor Classification Rule

Say we have N images to train our classifier. From each of the N images, we can extract the training feature vectors $\{\mathbf{X_1}, \mathbf{X_2}, \ldots, \mathbf{X_N}\}$ as well as the true cloud fractions $\{A_{t1}, A_{t2}, \ldots, A_{tN}\}$, which belong to the set of cloud fraction classes $\{C_1, C_2, \ldots, C_z\}$. In this study we have $z = 19$ true cloud fraction classes, from 0.05 to 0.95 in steps of 0.05.

Given a new feature vector $\mathbf{X_0}$, the nearest neighbor classification rule assigns $\mathbf{X_0}$ to a class C_v as follows: Let $\mathbf{X}_j$ satisfy $\min \{\|\mathbf{X_0} - \mathbf{X}_i\|\} \; \forall \; i \in \{1, \ldots, N\}$, where $\| \cdot \|$ denotes the Euclidean metric. Assign $\mathbf{X_0}$ to the class C_v that $\mathbf{X}_j$ belongs to. If more than one $\mathbf{X}_j$ is a nearest neighbor to $\mathbf{X_0}$, the ties are resolved randomly.

5.3. Training Set

The training set should consist of N images from which we know $\{\mathbf{X}_i, A_{ti}\}_{i=1}^{N}$, where (ideally) N is sufficiently large to capture the rich variety of cloudy systems found over the globe. In order to know A_{ti}, the training scenes must be measured at the resolution of the true cloud scale r_t. Since r_t is at present arbitrarily defined at a reasonably small scale, say less than 100 m, difficulties exist in obtaining an appropriate data set, given the current network of Earth-observing satellite instruments. We have considered Landsat data, but unfortunately we do not have the resources to obtain a large volume of these types of data to derive a wide variety of cloud patterns to properly train our classification rule.

Since this study deals with perfect cloud detectors, this work must be viewed as pedagogical in that any application to real satellite data must take into account the effects of thresholds in determining clear versus cloudy. Moreover, the aim of this section is to compare the errors between A_e, A_{17}, and A_p. For this purpose, any set of binary images will do so long as $r_t \ll r_0$. Naturally, we would like the images to resemble clouds as much as possible. With this in mind, we identify two sources for the training set.

5.3.1. Simulated data. We generated stochastic cloud field geometries using the model of *Várnai* [1996], which is based on the model of *Barker and Davies* [1992]; the main difference is the ability to include a break in the scaling. Basically, the model generates random Fourier coefficients according to the desired scaling and then performs an inverse Fourier transform to obtain two-dimensional (2-D) variations in the cloud structure. The required inputs are image size (in terms of number of pixels), the scale break, the scaling before and after the scale break, and the cloud fraction. We chose an image size of 1024 × 1024 pixels. To produce a wide variety of cloud fields, we varied the scale break from 1 to 13, the scaling before the break from 0 to -1, and the scaling after the break from -3 to -4. We generated 684 cloud fields, with cloud fraction varying from 0.05 to 0.95 in steps of 0.05. To produce the cloud mask image, we simply look at where the cloud does and does not exist. The simulated cloud fields do look realistic (examples are given by *Loeb et al.* [1996] and *Várnai* [1996]); however, they are confined to a single layer.

5.3.2. AVHRR data. The source of the second training set comes from the Advanced Very High Resolution Radiometer (AVHRR). We used AVHRR data taken in the local area coverage mode, which has a subsatellite ground track resolution of 1.1 km. We have acquired data taken over the North Atlantic, spanning the period of August 1 to October 31, 1993. We have extracted 370 scenes (maximum number of usable data) that were 1024 × 1024 pixels in size. The scenes are taken about the center of the swath in order to minimize distortions in the cloud fields due to pixel expansion. Appendix C describes the cloud mask applied to the 370 scenes. The distribution in cloud fractions taken from the AVHRR data is shown in Figure 5. Note the characteristic bell shape of the distribution as predicted by *Falls* [1974], given the large image scale under analysis.

The two training sets define the true cloud scale to be at the pixel scale of the original cloud mask image. The true cloud fraction A_t for each scene is derived from this scale. The images are degraded by a factor of 32 to produce binary images that are 32 × 32 pixels in size. Since we are assuming perfect cloud detection, a coarse pixel is clear if all high-resolution pixels within

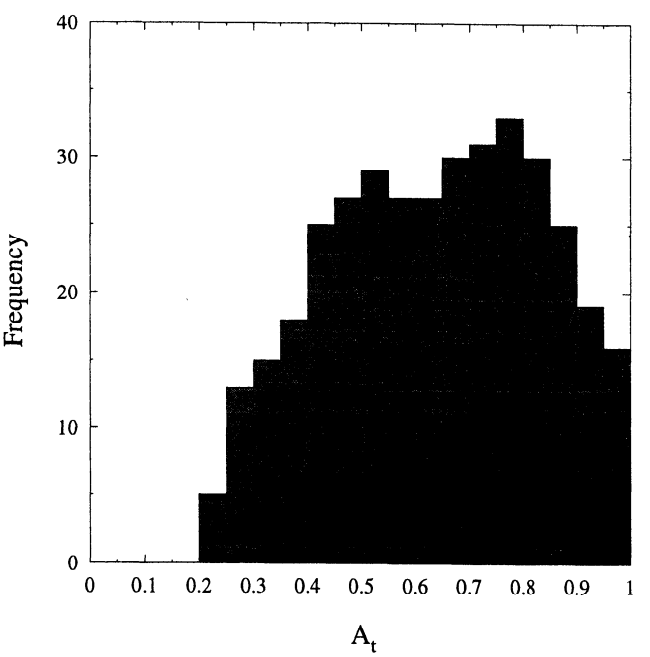

Figure 5. The distribution of "true" cloud fractions derived from the AVHRR training set.

it are clear; otherwise, it is cloudy (see Lemma 1). The feature vectors are extracted from the coarse resolution images.

5.4. Allocation Rates

A common approach to evaluating the performance of a classifier is to determine its allocation rate [*McLachlan*, 1992]. For convenience, let $\mathbf{X_T}$ represent the training set, i.e., $\mathbf{X_T} = \{(\mathbf{X}_1, A_{t1}), \ldots, (\mathbf{X_N}, A_{tN})\}$. The allocation rate, e_{ij}, for a given classification rule, $\eta(\mathbf{X}, \mathbf{X_T})$, is simply the probability that a randomly chosen scene (with feature vector $\mathbf{X}$) belonging to C_i is allocated to C_j, that is,

$$e_{ij}(\eta) = pr\{\eta(\mathbf{X}, \mathbf{X_T}) = C_j \,|\mathbf{X} \in C_i\}. \tag{19}$$

As in so many situations, the only data available to test the classifier are the data at hand (i.e., the training set $\mathbf{X_T}$), and they usually do not contain all of the possible situations. Most procedures used to estimate the allocation rates give rise to estimates that are biased with respect to the true allocation rates, as well as having large variances (for a review, see *Toussaint* [1974]). In recent years, however, bootstrap estimates [*Efron*, 1979] have gained widespread acceptance as the standard for allocation rate estimation because of their small variance and reduced bias (a very good introduction to bootstrap methods is given by *Efron and Tibshirani* [1986]). The so-called EO bootstrap estimator [*Jain et al.*, 1987] is used in this study, as it is particularly appropriate for estimating the allocation rate of nearest neighbor classifiers. A total of 200 bootstrap samples is used in the EO estimate, in accordance with *Efron* [1983].

5.5. Optimizing the Feature Vector

It has been well documented that too many features in a feature vector can result in a poor classifier [e.g., *McLachlan*, 1992]. For this reason, feature selection procedures are usually carried out in order to optimize the performance of a classifier. We make use of an error rate-based procedure. This involves choosing the set of features that minimizes $e(\eta)$, the sum of $e_{ij}(\eta), i \neq j$. Typically, as in our case, the initial number of features is large enough to prohibit the testing of all possible combinations of feature vectors, because the computational load is too great. To reduce the computational load, we use the following procedure: each of the 14 features alone can be considered as 14 one-dimensional (1-D) feature vectors. Each 1-D vector is used in the calculation of $e(\eta)$. Let $x_{\min}$ be the feature that has the minimum $e(\eta)$. The remaining features are then combined with $x_{\min}$ to form 13 2-D feature vectors. Each of these 2-D vectors is used in the calculation of $e(\eta)$. The feature vector that has the minimum $e(\eta)$ is retained and combined with the remaining 12 features to form 12 three-dimensional feature vectors. This procedure is continued until we have built up a fourteen-dimensional feature vector, whose feature entries are ranked in order of importance. The optimum feature vector contains the first m features, where m satisfies $\{EO_m\}_{m=1}^{14}$.

Because the EO bootstrap estimator of $e(\eta)$ has some variance associated with it (i.e., each EO realization is not exactly the same), it is possible for a different rank and a different optimum feature vector to be selected by the selection procedure. For this reason, the selection procedure was carried out 10 times using the simulated and AVHRR training sets. The optimum feature vector that stood out was six-dimensional and, in order of preference, equal to $(\mathrm{VAR}(r_i), A_e(r_i), A_{\mathrm{edge}}(r_i), M_1(r_i), \mathrm{MEAN}(r_i), \mathrm{ENT}(r_i))$, where VAR, MEAN, and ENT are the GLDS variance, mean, and entropy, respectively, and M_1 is the first moment invariant. The first two features reduce the error rates substantially, with only a small improvement with the addition of the last four. Note, the most important feature is not $A_e(r_i)$ but $\mathrm{VAR}(r_i)$. This tells us that if we had to use a single feature to estimate A_t, we would do better using $\mathrm{VAR}(r_i)$ rather than $A_e(r_i)$. Also, the information at the degraded scale r_{i-1} is not required, thereby substantially reducing the computational load. The optimum feature vector is used in the next section to estimate A_p for the two training sets.

5.6. Results

Figures 6 and 7 show the allocation rates, as defined in (19), for A_e, A_{17} and A_p for the simulated and AVHRR training sets, respectively. Although the calculations, and those that follow, are performed using cloud fraction classification, we have converted them to cloud fraction (percent) for the reader's convenience. The percentage cloud fractions that are reported represent the lower bound of the cloud fraction classes. Note the absence of data in Figure 7 for $A_t < 15\%$; this is consistent with Figure 5. As expected, a strong bias for A_e to overestimate A_t exists for both training sets. (We also note that the results of A_e presented in Figures 6a and 7a are reminiscent of those in Figure 7a of *Chang and Coakley* [1993]. The differences are likely due to our assumption of perfect cloud detection). The results are slightly improved for A_{17}; however, A_{17} also always equals or overestimates A_t. The allocation rates for A_p are scattered about the diagonal of the figure, with the highest allocation rates clustering near the diagonal of the figure.

For the three estimators, the bias and variance are a function of A_t. We define the average bias in the cloud fraction class estimate for a given A_t as

$$\beta(A_{ti}) = \sum_{j=1}^{19} (A_j - A_{ti}) e_{ij} \tag{20}$$

and the variance as

$$\sigma^2(A_{ti}) = \sum_{j=1}^{19} [A_j - A_{ti} - \beta(A_{ti})]^2 e_{ij}, \tag{21}$$

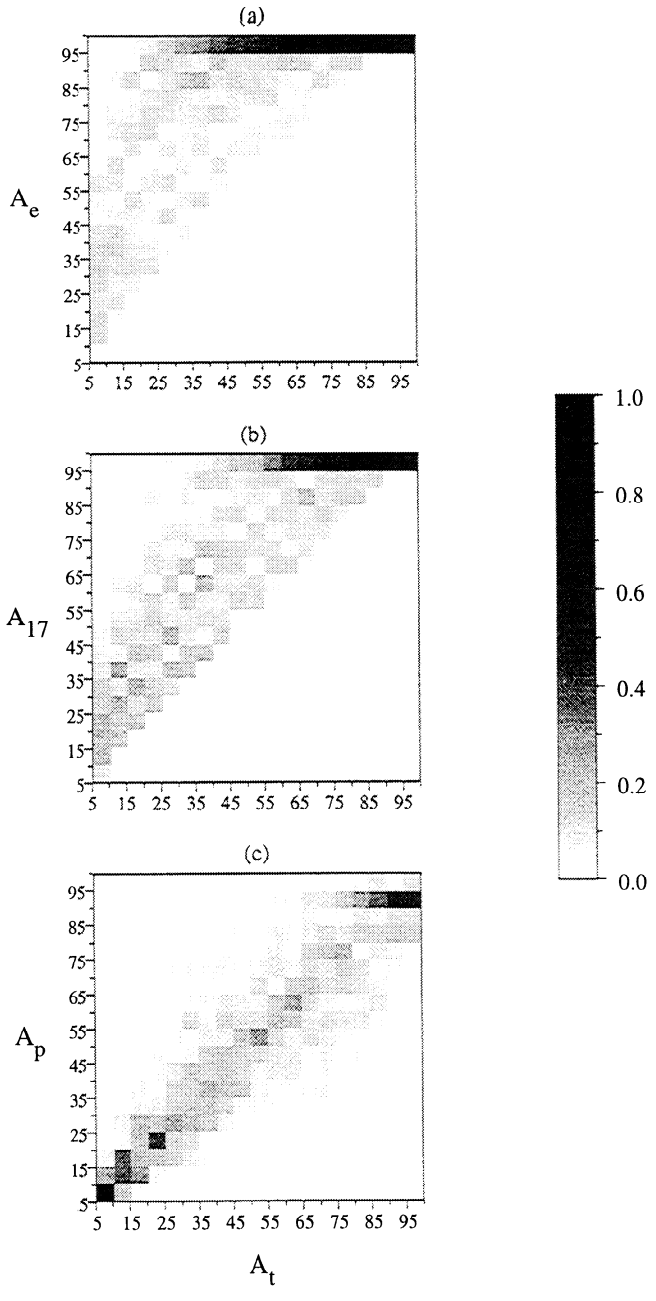

Figure 6. Allocation rates, as defined by (19), for (a) A_e, (b) A_{17}, and (c) A_p derived from the simulated training set. The axes are converted to percentage cloud fraction.

where A_j is the cloud fraction class estimate given by either A_e, A_{17}, or A_p, and $i \in \{1, \ldots, 19\}$.

Figures 8 and 9 show the bias and standard deviation (σ) in the cloud fraction estimates for the simulated and AVHRR training sets, respectively. The bias in A_e peaks at an $A_t = 25\%$, with a bias of 46% in cloud fraction for the simulated cloud fields and 63% for the AVHRR training set. Thus cloud fields having a true cloud fraction of 25% tend to give rise to the largest bias in A_e. The bias in A_{17} peaks at an $A_t = 35\%$, with a bias of 30% in cloud fraction for the simulated training set and 53% for the AVHRR data set. The bias in A_p peaks at an $A_t = 95\%$, with a bias of -9% in cloud fraction for the simulated training set and -14% for the AVHRR training set.

The behavior in bias for A_e and A_{17} helps to explain some of the behavior in their variance. First, the bias and variance equals zero when $A_t = 95\%$ cloud fraction for both training sets. This is expected since $A_t = 95\%$ is the largest cloud fraction bin and these cloud fraction estimators always equal or overestimate A_t. Second, since the estimators always equal or overestimate A_t, the average variance is smaller for A_e as compared to A_{17} because the overall bias is larger for A_e than it is for A_{17}. This also explains the lower variances observed for the AVHRR training set.

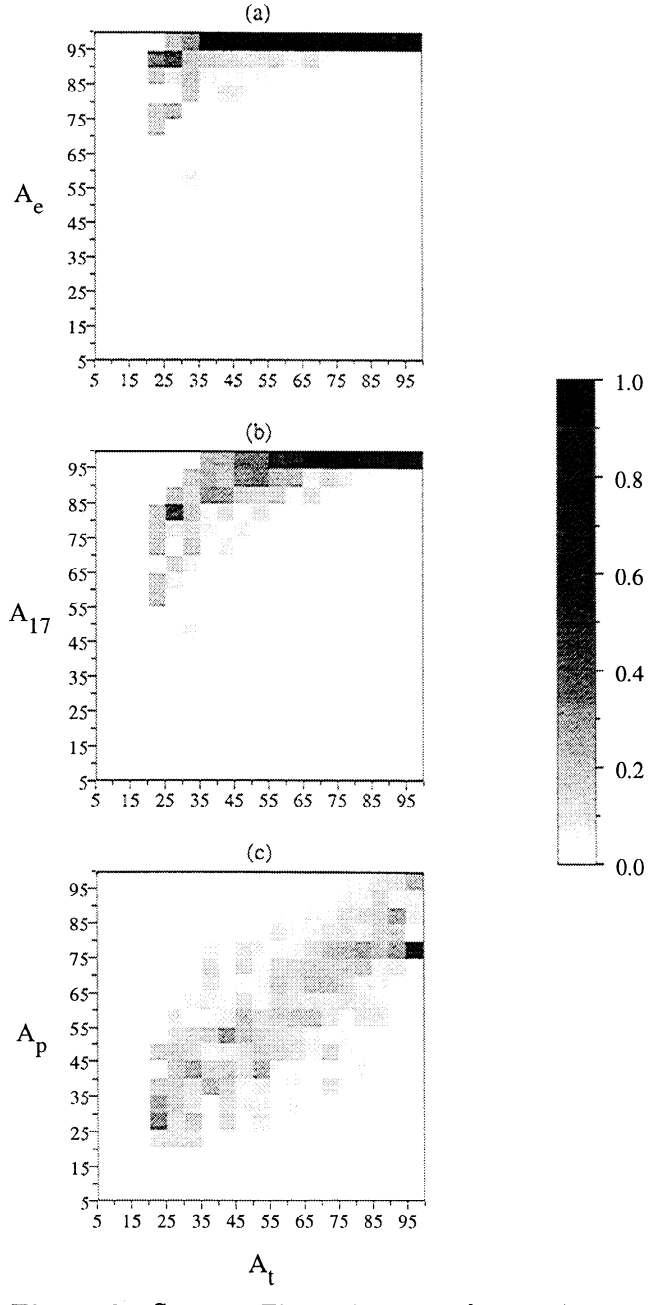

Figure 7. Same as Figure 6, except for the AVHRR training set.

As for A_p, the lowest variances tend to be with the smaller A_t classes, indicating that cloud pattern information is well characterized at small values of A_t. The variability is largest for an A_t of about 25% to 85%. In this region, a lot of overlap in cloud pattern information can exist. This is because the cloud pattern information has been derived at the scale of the coarse resolution measurements and not at the true cloud scale. As a result, cloud fields having a true cloud class of 25% can appear in the coarse resolution measurements as having much larger cloud fractions (which is where the peak bias in A_e occurs) and cloud patterns resembling those of larger true cloud fractions. Moreover, the variability in cloud pattern information is reduced as A_e gets larger. The overall average bias and standard deviation for the three estimators derived from the two training sets is tabulated in Table 1.

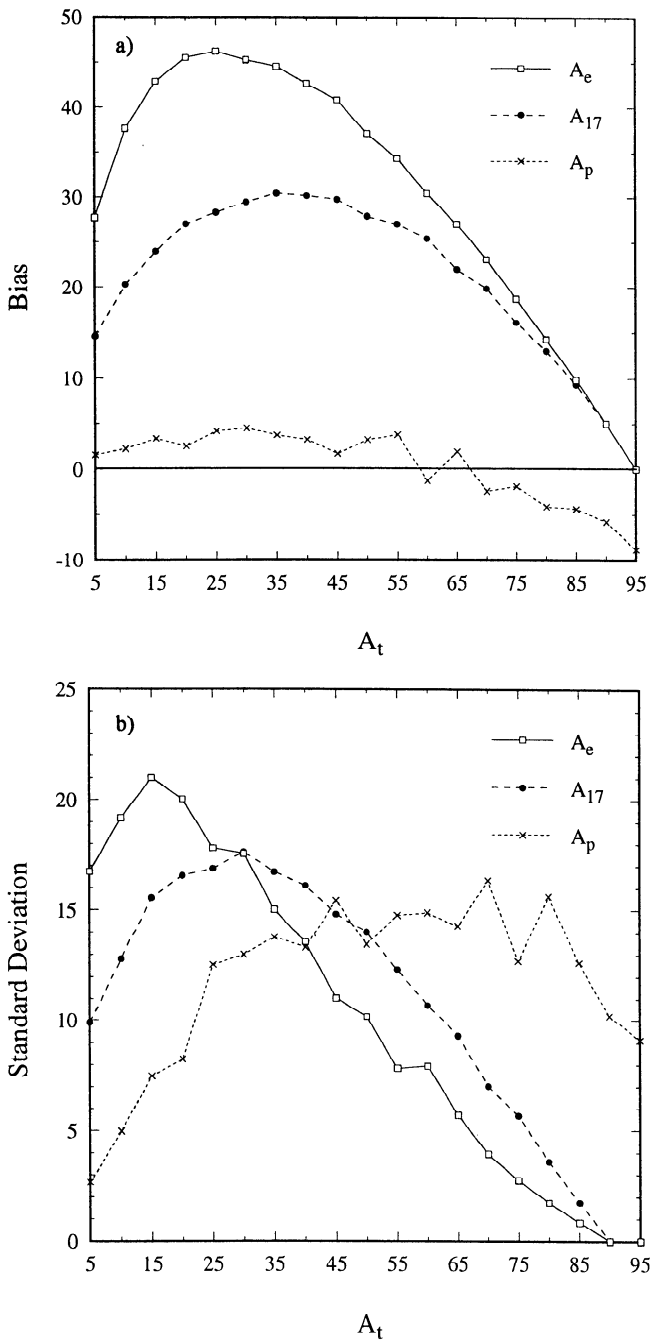

Figure 8. (a) The bias and (b) the standard deviation in A_e, A_{17}, and A_p for the allocation rates defined by (19) applied to the simulated training set and converted to percentage cloud fraction.

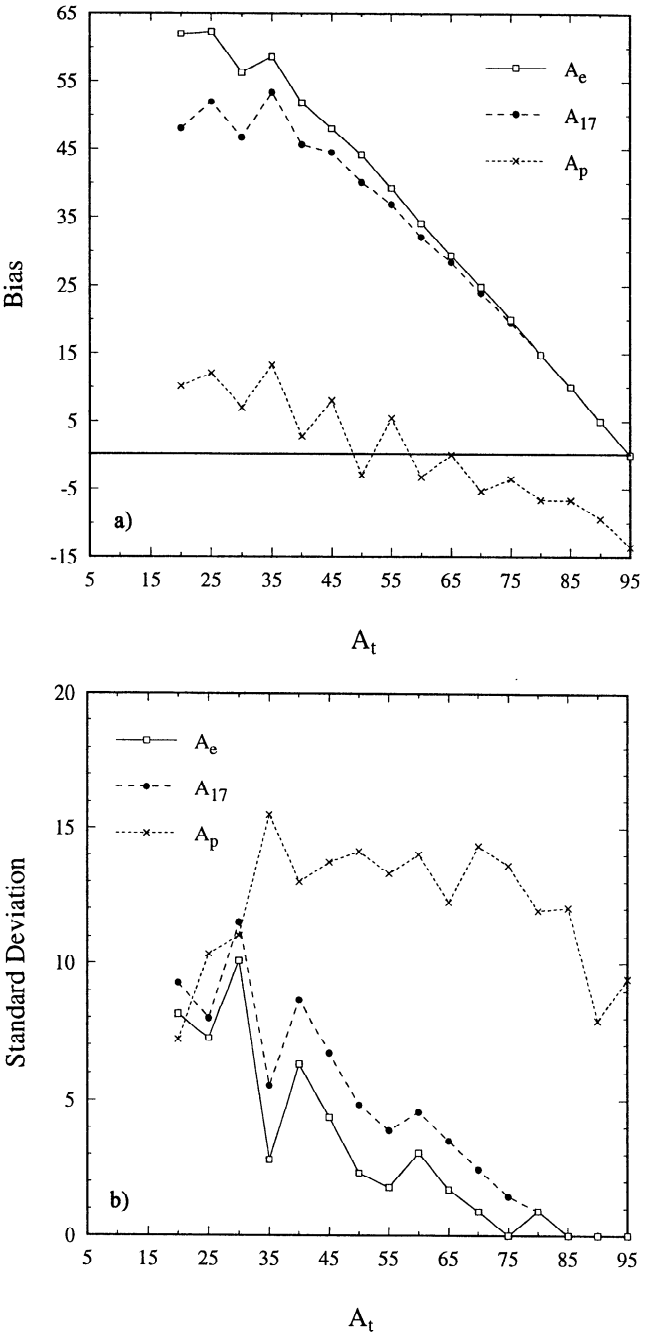

Figure 9. Same as Figure 8, except for the AVHRR training set.

We have thus far examined the behavior of the three cloud fraction estimators as they depend on the true cloud fraction, A_t. However, from a remote sensing point of view, we do not know A_t. To study the errors in our cloud fraction retrievals, the allocation rate from

Table 1. Overall Average Bias and Standard Deviation of A_e, A_{17}, and A_p, Computed From the Allocation Rates Defined by (19) and Converted to Percentage Cloud Fraction

Estimators	Simulated Training Set		AVHRR Training Set	
	$\bar{\beta}$	$\bar{\sigma}$	$\bar{\beta}$	$\bar{\sigma}$
A_e	31.9	11.4	35.1	3.3
A_{17}	22.2	11.9	31.3	4.5
A_p	0.4	11.9	0.5	12.1

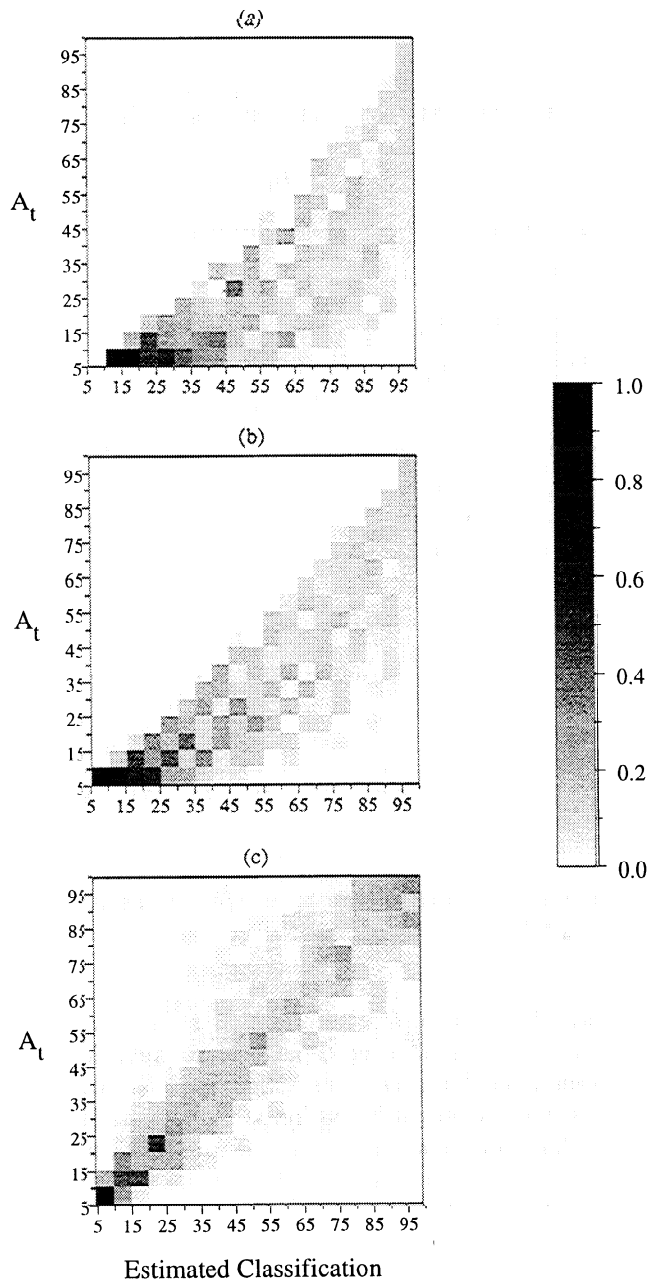

Figure 10. Allocation rates, as defined by (22), for the estimators (a) A_e, (b) A_{17}, and (c) A_p derived from the simulated training set.

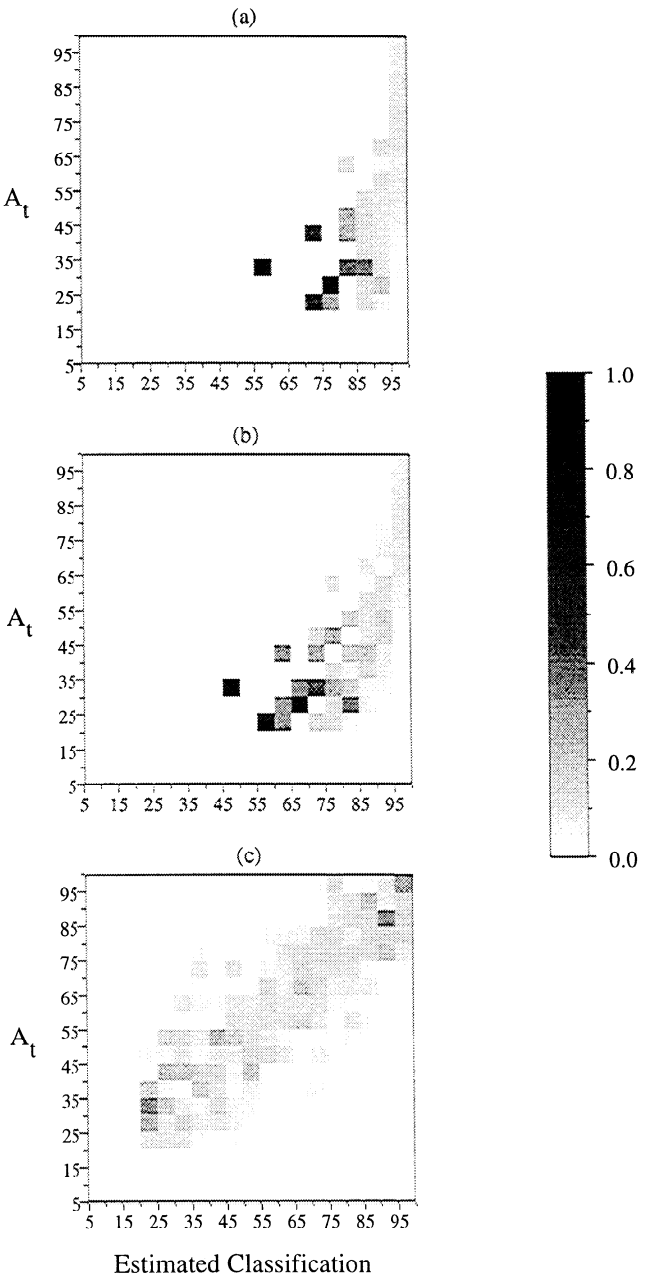

Figure 11. Same as Figure 10, except for the AVHRR training set.

a remote sensing point of view needs to be computed. This is simply the probability that a randomly chosen scene (with feature vector $\mathbf{X}$) having an estimate C_i belongs to the true class, C_j, that is,

$$e_{ij}(\eta) = pr\{\mathbf{X} \in C_j \,|\, \eta(\mathbf{X}, \mathbf{X_T}) = C_i\} \qquad (22)$$

Figures 10 and 11 show the allocation rates, as defined in (22), for A_e, A_{17} and A_p for the two training sets. These results are consistent with the behavior of the allocation rates of Figures 6 and 7: true cloud fractions will always be equal to or lie below the estimates given by A_e and A_{17}, and they will cluster about the diagonal given cloud fraction estimates from A_p.

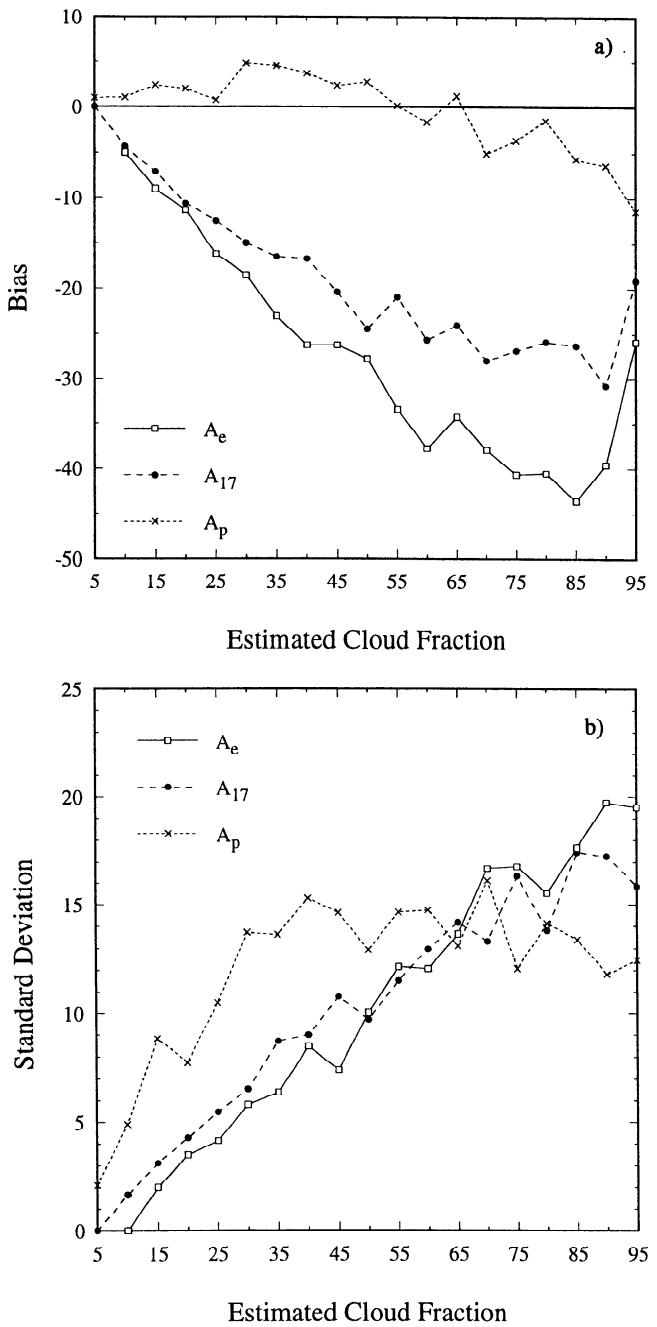

Figure 12. Same as Figure 8, except for the allocation rates defined by (22).

Figures 12 and 13 show the bias and standard deviation in the true cloud fractions as they depend on the estimators cloud fraction classification for the simulated and AVHRR training sets, respectively. Note the lack of data below $A_e = 55\%$ and $A_{17} = 45\%$ in Figure 13. This is simply due to the lack of estimates below these values for the AVHRR training set. Similarly, there are very few A_e and A_{17} estimates below a class estimate of 90% for the AVHRR training set; therefore the bias and standard deviation estimates should be interpreted with caution. The bias for the A_e estimator peaks at $A_e = 85\%$, with a value of -44% in cloud fraction for the simulated training set, and $A_e = 75\%$ with a value of -51% in cloud fraction for the AVHRR training set. Thus, given A_e estimates of cloud fraction between 75 and 90%, the true cloud fraction will lie on average 44 – 51% below the estimate. The results improve slightly for A_{17}. For A_p, the bias estimates are much lower than A_e and A_{17} except for the case $A_{17} = 5\%$ with the simulated data (note that no $A_e = 5\%$ data exist). The variance is lowest when the estimator classification is low for all three estimators; this is where the biases are also low. Thus, given a low estimate of cloud fraction from either of the three estimators, we can assume the estimate to be good. The overall average bias and stan-

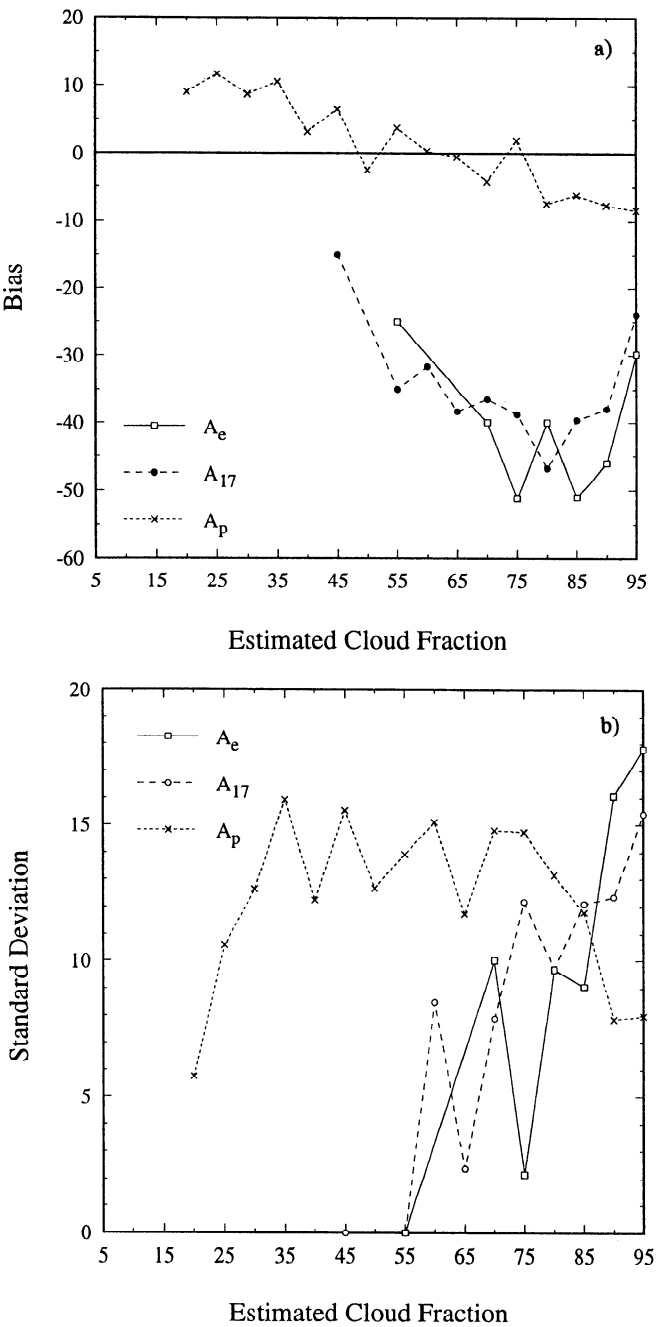

Figure 13. Same as Figure 9, except for the allocation rates defined by equation (22).

dard deviation for the three estimators derived from the two training sets is tabulated in Table 2.

Overall, the results are slightly better for the simulated training set compared to the AVHRR training set. Moreover, Figures 8, 9, 12, and 13 show the results for the AVHRR training set to be more "noisy" than those for the simulated training set. This may be due to the richer variety of cloud patterns that exist in nature, especially when one considers multilayered systems, and/or to the smaller amount of data used in the AVHRR training set as compared to the simulated training set. Regardless of the small differences that exist between the two training sets, the results clearly demonstrate that a pattern recognition approach in estimating A_t is for the most part significantly better than simply using the standard method or the correction method of (17). Moreover, even though A_{17} did produce better results than A_e, the fact that the A_{17} results are still poor suggests that we are applying (17) to situations where the equation is no longer valid. Therefore from the derivation of (17), we can conclude that most of the coarse resolution cloud scenes used in this section had many cloud interior pixels that were partially cloudy.

Finally, we have presented results for the case when $r_t/r_i = 1/32$, which is a reasonable ratio to consider for present-day observations (e.g., Landsat/AVHRR). For different ratios, different quantitative results are obtained; however, the results remain qualitatively consistent, and better results are obtained for all three estimators as r_t/r_i gets larger. This remains true up to $r_i = r_t$, where all three estimators give the correct A_t. This is clearly the case for A_e and A_{17}, as expected from section 4. It is also true for A_p since optimizing the feature vector at this resolution results in a 1-D optimum feature vector whose entry is equal to A_e. At the other extreme, when $r_i = r_0$, all training scenes appear 100% cloudy. Given A_t, the overall average bias for A_e and A_{17} is 45% with a standard error of 0%. For A_p, the choice of A_e is random (as expected from the nearest neighbor classification rule); therefore, the overall average bias is 0% with a standard error of 27%.

6. Summary and Conclusions

The standard method in deriving cloud fraction is to compute the fraction of total image pixels that contains some cloud. We have examined the errors associated with this method due to the effects of finite measurement resolution. The concept of perfect cloud detection was employed in order to isolate resolution effects from threshold and cloud definition effects, as did SS72. We have revisited the paper cloud experiment of SS72 and demonstrated that their results are not unique for uniformly distributed clouds. This conclusion was further supported on theoretical grounds that dismiss the assumptions made in SS72 which state that the standard method estimate, A_e, of cloud fraction depends only on the true cloud fraction, A_t, and the ratio of the average cloud size to pixel size (see (1)).

Table 2. Overall Average Bias and Standard Deviation of A_e, A_{17}, and A_p, Computed From the Allocation Rates Defined by (22) and Converted to Percentage Cloud Fraction

Estimators	Simulated Training Set		AVHRR Training Set	
	$\bar{\beta}$	$\bar{\sigma}$	$\bar{\beta}$	$\bar{\sigma}$
A_e	-27.6	10.6	-40.5	9.3
A_{17}	-19.8	10.1	-34.4	8.1
A_p	-0.5	12.0	1.2	12.3

To understand the uncertainties in using A_e as the estimate for A_t, we have derived the upper and lower limits of A_t given A_e over a general cloud field under the assumption of perfect detection. The derivation is general and shows A_e to be the upper bound on A_t, while the lower bound is essentially zero for measurements having resolutions typical of current meteorological satellites (> 1 km) and for a reasonably small defined true cloud scale [see (11)]. Further assumptions were used to significantly reduce the bounds on A_t and to obtain an exact relationship (see (17)); however, the assumptions all become questionable for coarse resolution measurements. This is principally due to partially cloud-covered pixels being observed as cloud interior pixels rather than cloud edge pixels within coarse resolution imagery (see Lemma 2). Although we have not examined radiance-based methods to estimate subpixel cloud fraction in this study, we note that the statistical approach by *Molnar and Coakley* [1985], which uses an empirical relationship similar to (17), shows promising results.

We have also demonstrated that a pattern recognition estimate, A_p, can be used as an estimate of A_t. Comparisons were made between A_p, A_e, and the estimate from (17), A_{17}, as estimates for A_t. Overall, A_p leads to nearly unbiased estimates of A_t, unlike A_e and A_{17}, leaving us with unbiased estimates of global cloudiness if such an approach were adapted using perfect cloud detectors. Of course, an appropriate training set for the estimate A_p is required. The two training sets we have used are for demonstration purposes only.

Applying the findings of this study to present satellite data will be difficult given that perfect cloud detection does not yet exist. Any real application requires us to consider the effect of thresholds. True errors are difficult to assess due to the interdependence of resolution and threshold, and the lack of a precise quantitative definition of what constitutes a cloud. Once these issues are resolved, further efforts at measuring cloud fraction using the information within the measured radiances may be pursued. This includes using the radiances rather than the binary image to characterize the image pattern.

As a final note, much of what has been presented in this study can apply to the fractional area coverage of many geophysical parameters (e.g., vegetation, ice, ice leads, etc.). However, caution should still be observed in defining the relevant geophysical parameter and in deriving an appropriate training set for a pattern recognition approach.

Appendix A: Derivation of (18)

In this appendix we transform (17) into the form of (1) for a regular array of circular clouds all having the same size. We begin by rewriting (17) as

$$A_e(r_i) = 2A_t - A_{\text{int}}(r_i) - \left(\frac{r_t}{r_i}\right)^2 A_{\text{edge}}(r_i). \quad (A1)$$

Let n_{cld} be the total number of circular clouds in our scene. We assume that an image pixel does not contain partial coverage by more than one cloud, which may cause cloud edge pixels to be observed as cloud interior pixels, as depicted in Figure 2. With this assumption, $A_{\text{edge}}(r_i)$ and $A_{\text{int}}(r_i)$ are given by

$$A_{\text{edge}}(r_i) = 2\pi n_{\text{rad}}(r_i) r_i^2 n_{\text{cld}} \quad (A2)$$

$$A_{\text{int}}(r_i) = \pi n_{\text{rad}}(r_i)^2 r_i^2 n_{\text{cld}} - 2\pi n_{\text{rad}}(r_i) r_i^2 n_{\text{cld}} \quad (A3)$$

where $n_{\text{rad}}(r_i)$ is the number of pixels that make up the cloud radius at a scale r_i, with

$$A_e(r_i) = \pi n_{\text{rad}}(r_i)^2 r_i^2 n_{\text{cld}}. \quad (A4)$$

Substituting (A2) and (A3) into (A1) yields

$$A_e(r_i) = 2A_t - \pi n_{\text{rad}}(r_i)^2 r_i^2 n_{\text{cld}} + 2\pi n_{\text{rad}}(r_i) n_{\text{cld}} (r_i^2 - r_t^2). \quad (A5)$$

R is defined as the ratio of the true average area of a cloudy element (A_t/n_{cld} in this example) to the pixel area ($r_i \times r_i$), that is,

$$R = \frac{A_t}{n_{\text{cld}} r_i^2}. \quad (A6)$$

Combining (A6) with (A4), we can rearrange the terms to give

$$n_{\text{rad}}(r_i) = \left(\frac{R A_e(r_i)}{\pi A_t}\right)^{1/2}. \quad (A7)$$

Substituting (A6) and (A7) into (A5) yields

$$A_e(r_i) = A_t + \left[\left(\frac{\pi A_t}{R}\right)^{1/2} - n_{\text{cld}} r_t^2 \left(\frac{\pi R}{A_t}\right)^{1/2}\right] A_e(r_i)^{1/2}. \quad (A8)$$

Equation (A8) is quadratic in $A_e(r_i)^{1/2}$, the solution of which gives (18) in section 4.

Appendix B: Limitations of (18)

Here we derive the cutoff value for R in which (18) is no longer valid for a regular array of circular clouds all of the same size. The validity of (18) breaks when partially cloud covered pixels are observed as cloud interior pixels rather than cloud edge pixels (see Figure 2). This can occur when the pixel size is greater than the minimum distance between the edges of neighboring clouds. The true distance, D_c, between nearest neighbor cloud centers is given by

$$D_c = \left(\frac{1}{n_{\text{cld}} r_t^2}\right)^{1/2} \quad (B1)$$

and the minimum distance between the edges of neighboring clouds (corresponding to r_c in section 4) is given by

$$r_c = \frac{D_c - 2n_{\text{rad}}(r_t)}{2}. \quad (B2)$$

Then with (B1), (B2), (A6), and (A7), the condition $r_i > r_c$ given in section 4 transforms to

$$R < \frac{A_t}{\left[\frac{1}{2} - \left(\frac{A_t}{\pi}\right)^{1/2}\right]^2}. \quad (B3)$$

Equation (B3) describes the region for which (18) is no longer valid due to the limitations of the assumption that only cloud edge pixels can be partially cloudy.

Appendix C: The AVHRR Cloud Mask Algorithm

In this appendix we describe the cloud mask that was used to extract one of the training sets from the AVHRR data. The data come from the AVHRR instrument flown aboard the NOAA 11 spacecraft. The instrument has five spectral channels centered at about 0.67, 0.86, 3.7, 11, and 12 μm, which are labeled channels 1 through 5, respectively. The channel 1 and 2 counts are converted to reflectances (R_1 and R_2, respectively) using the calibration coefficients shown in Table C1a [*P. M. Teillet*, personal communication, 1995] with their meanings described by *Teillet and Holben* [1994]. The channel 4 and 5 counts are converted to brightness temperatures (T_4 and T_5, respectively) using the procedure described by *Wooster et al.* [1995]. Nonlinearity corrections in calibrating the infrared channels are followed using the procedure described by *Weinreb et al.* [1990].

The cloud mask algorithm that we employed is depicted in Figure C1. The list of observables employed by the cloud mask, the references that describe their rationale, and the thresholds are shown in Table C1b. The OWE14 data set [*NOAA-EPA Global Ecosystems*

Table C1a. Calibration Coefficients for AVHRR Channels 1 and 2

Channel	Gain		Offset	
	Slope	Intercept	Slope	Intercept
1	2.783×10^{-5}	1.654	3.832×10^{-4}	39.26
2	-4.630×10^{-5}	2.534	8.914×10^{-4}	38.28

Table C1b. Cloud Detection Observables, Description, and Thresholds Used by the Algorithm Depicted in Figure C1

Observable	Threshold	Description
R_1	$R_{1t} = 0.6$	gross reflectance test to filter out very bright clouds
R_2	$PCST$[a]	see Di Girolamo and Davies [1995]
D[b]	$D_t = 1.9$	see Di Girolamo and Davies [1995]
$DSVI$[c]	$DSVI_t = 0.004$	see Di Girolamo and Davies [1994]
T_5	$T_{5t} = 250K$	gross brightness temperature test to filter out very high, cold clouds
σ[d]	$\sigma_t = 0.5K$	see Saunders and Kriebel [1988]
$T_4 - T_5$	T_{dt}[e]	see Saunders and Kriebel [1988]

[a] $PCST$ (predetermined clear sky threshold) is a function of the Sun view geometry.

[b] $D = \|NDVI\|^b / R_1^2$, where $NDVI = (R_2 - R_1)/(R_2 + R_1)$ and b is chosen to maximize the separation between clear and cloudy D values. A constant b value of 0.6 was used in our study since all land classes were predominantly vegetation covered.

[c] $DSVI$ is the spatial variability index of D, defined as the absolute difference between D and the mean value of its surrounding 3×3 matrix.

[d] Variable σ is the standard deviation of T_4.

[e] T_{dt} is a function of T_4 and the view angle.

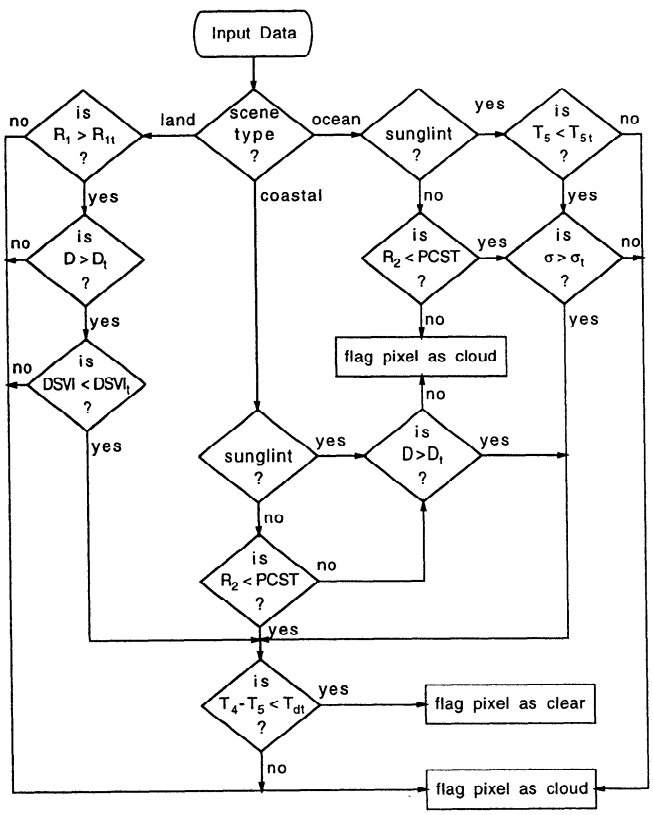

Figure C1. The cloud detection algorithm employed in our study.

Database Project, 1992] was used to classify the pixels as ocean, land or coastal (coastal being a pixel that has land and ocean amongst its eight neighboring pixels). Of the 370 scenes that were extracted, approximately 93% of the total pixels were classified as ocean. Also, all land pixels were classified as having vegetation cover. With this in mind, we simply used constant thresholds for the observables D and DSVI, since small errors in cloud detection are not important in our study. However, all thresholds employed by the cloud mask lean toward being clear sky conservative.

Acknowledgments. Partial support from the Jet Propulsion Laboratory of the California Institute of Technology under contract 959085 and from the Natural Sciences and Engineering Research Council of Canada is gratefully acknowledged. The AVHRR data used in this study were provided by the U. S. Geological Survey EROS Data Center DAAC. We thank Jeff Eidenshink and Eugene Clothiaux for the timely and efficient manner in which the data were distributed. Thanks are also extended to Bryan Baum and Lazaros Oreopoulos for the AVHRR unpacking code, Phil Teillet for the AVHRR calibration coefficient, and Tamás Várnai for his cloud model code.

References

Arakawa, A., Modeling clouds and cloud processes for use in cloud models, *GARP Publ. Ser.*, *16*, 183–197, 1975.

Arking, A., The radiative effects of clouds and their impact on climate, *Bull. Am. Meteorol. Soc.*, *72*, 795–814, 1991.

Arking, A., and J. D. Childs, Retrieval of cloud cover parameters from multispectral satellite measurements, *J. Clim. Appl. Meteorol., 24,* 322–333, 1985.

Baraldi, A., and F. Parmiggiani, An investigation of the textural characteristics associated with gray level cooccurrence matrix statistical parameters, *IEEE Trans. Geosci. Remote Sens., 33,* 293–304, 1995.

Barker, H. W., and J. A. Davies, Solar radiative fluxes for stochastic, scale-invariant broken cloud fields, *J. Atmos. Sci., 49,* 1115–1126, 1992.

Cess, R. D., *et al.*, Intercomparison and interpretation of climate feedback processes in 19 atmospheric general circulation models, *J. Geophys. Res., 95,* 16,601–16,615, 1990.

Chang, F.-L., and J. A. Coakley, Estimating errors in fractional cloud cover obtained with infrared threshold methods, *J. Geophys. Res., 98,* 8825–8839, 1993.

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

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

Di Girolamo, L., On the detection of cirrus clouds from satellite measurements, M. Sc. thesis, McGill Univ., Montréal, Qué., 1992.

Di Girolamo, L., and R. Davies, Optimizing the use of 0.67 μm and 0.86 μm radiometric data for cloud detection, paper presented at 8th Conference on Atmospheric Radiation, pp. 466–467, Amer. Meteorol. Soc., Nashville, Tenn., Jan. 23–28, 1994.

Di Girolamo, L., and R. Davies, The image navigation cloud mask for the Multi-angle Imaging SpectroRadiometer (MISR), *J. Atmos. Oceanic Technol., 12,* 1215–1228, 1995.

Dudani, S. A., K. J. Breeding, and R. B. McGhee, Aircraft identification by moment invariants, *IEEE Trans. Comput., C-26,* 39–46, 1977.

Efron, B., Bootstrap methods: Another look at the jackknife, *Ann. Stat., 7,* 1–26, 1979.

Efron, B., Estimating the error rate of a prediction rule: Improvement on cross-validation, *J. Am. Stat. Assoc., 78,* 316–331, 1983.

Efron, B., and R. Tibshirani, Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy (with discussion), *Stat. Sci., 1,* 54–77, 1986.

Falls, L. W., The Beta distribution: A statistical model for world cloud cover, *J. Geophys. Res., 79,* 1261–1264, 1974.

Frouin, R., C. Gautier, and J. J. Morcrette, Downward longwave irradiance at the ocean surface from satellite data: Methodology and in situ validation, *J. Geophys. Res., 93,* 597–619, 1988.

Goodman, A. H., and A. Henderson-Sellers, Cloud detection analysis: A review of recent progress, *Atmos. Res., 21,* 203–228, 1988.

Gu, Z. Q., C. N. Duncan, E. Renshaw, M. A. Mufflestone, C. F. N. Cowan, and P. M. Grant, Comparison of techniques for measuring cloud texture in remotely sensed satellite meteorological image data, *IEE Proc., 136,* 236–248, 1989.

Haralick, R. M., Statistical and structural approaches to texture, *Proc. IEEE, 67,* 786–804, 1979.

Haralick, R. M., and L. G. Shapiro, Glossary of computer vision terms, *Pattern Recognit., 24,* 69–93, 1991.

Harshvardhan, The effect of brokenness on cloud-climate sensitivity, *J. Atmos. Sci., 39,* 1853–1861, 1982.

Henderson-Sellers, A., and K. McGuffie, Are cloud amounts estimated from satellite sensor and conventional surface-based observations related?, *Int. J. Remote Sens., 11,* 543–550, 1990.

Hu, M. K., Pattern recognition by moment invariants, *Proc. IRE, 49,* 1428, 1961.

Hu, M. K., Visual pattern recognition by moment invariants, *IEEE Trans. Inf. Theory, IT-8,* 179–187, 1962.

Hughes, N. A., Global cloud climatologies: A historical review, *J. Clim. Appl. Meteorol., 23,* 724–751, 1984.

Jain, A. K., R. C. Dubes, and C.-C. Chen, Bootstrap techniques for error estimation, *IEEE Trans. Pattern Anal. Mach. Intel., PAMI-9,* 628–633, 1987.

Key, J. R., The area coverage of geophysical fields as a function of sensor field-of-view, *Remote Sens. Environ., 48,* 339–346, 1994.

Kobayashi, T., Parameterization of reflectivity for broken cloud field, *J. Atmos. Sci., 45,* 3034–3045, 1988.

Loeb, N. G., T. Várnai, and R. Davies, The effects of cloud inhomogeneities on the solar zenith angle dependence of nadir reflectance, *J. Geophys. Res.,* in press, 1996.

London, J., A study of the atmospheric heat balance, *Final Rep. AFC-TR-57-287, OTSB129551,* College of Eng., N. Y. Univ., 99 pp., 1957.

Lovejoy, S., D. Schertzer, and A. A. Tsonis, Functional box counting and multiple elliptical dimensions in rain, *Science, 235,* 1036–1038, 1987.

McLachlan, G. F., *Discriminant Analysis and Statistical Pattern Recognition,* 526 pp., John Wiley, New York, 1992.

Molnar, G., and J. A. Coakley, The retrieval of cloud cover from satellite imagery data: A statistical approach, *J. Geophys. Res., 90,* 12,960–12,970, 1985.

NOAA-EPA Global Ecosystems Database Project, *Global Ecosystems Database Version 1.0, User's Guide: Documentation, Reprints, and Digital Data on CD-ROM.* US DOC/NOAA National Geophys. Data Cent., Boulder, Colo., 1992.

Raffy, M., Remotely-sensed quantification of covered areas and spatial resolution, *Int. J. Remote Sens., 14,* 135–160, 1993.

Rossow, W. B., et al., ISCCP cloud algorithm intercomparison, *J. Clim. Appl. Meteorol., 24,* 877–903, 1985.

Rossow, W. B., Measuring cloud properties from space: A review, *J. Clim., 2,* 201–213, 1989.

Rossow, W. B., and R. A. Schiffer, ISCCP cloud data products, *Bull. Am. Meteorol. Soc., 72,* 2–20, 1991.

Saunders, R. W., and K. T. Kriebel, An improved method for detecting clear sky and cloudy radiances from AVHRR data, *Int. J. Remote Sens., 9,* 123–150, 1988.

Schiffer, R. A., and W. B. Rossow, The International Satellite Cloud Climatology Project (ISCCP): The first project of the World Climate Research Program, *Bull. Am. Meteorol. Soc., 64,* 779–784, 1983.

Shenk, W. E., and V. V. Salomonson, A simulation study exploring the effects of sensor spatial resolution on estimates of cloud cover from satellites, *J. Appl. Meteorol., 11,* 214–220, 1972.

Simpson, M. L., R. L. Schmoyer, and M. A. Hunt, Moment invariants for automated inspection of printed material, *Opt. Eng., 30,* 424-430, 1991.

Stowe, L. L., C. G. Wellemeyer, T. F. Eck, and H. Y. M. Yeh, Nimbus-7 global cloud climatology, 1, Algorithm and validation, *J. Clim., 1,* 445-470, 1988.

Teague, M. R., Image analysis via the general theory of moments, *J. Opt. Soc. Am., 70,* 920-930, 1980.

Teillet, P. M., and B. N. Holben, Towards operational radiometric calibration of NOAA AVHRR imagery in the visible and near-infrared channels, *Can. J. Remote Sens., 20,* 1-10, 1994.

Toussaint, G. T., Bibliography on estimation of misclassification, *IEEE Trans. Inf. Theory, IT-20,* 472-479, 1974.

Várnai, T., Reflection of solar radiation by inhomogeneous clouds, Ph.D. thesis, McGill Univ., Montréal, Qué., 1996.

Weinreb, M. P., G. Hamilton, S. Brown, and R. J. Koczor, Nonlinearity corrections in calibration of Advanced Very High Resolution Radiometer infrared channels, *J. Geophys. Res., 95,* 7381-7388, 1990.

Welch, R. M., and B. A. Wielicki, Reflected fluxes for broken clouds over a Lambertian surface, *J. Atmos. Sci., 46,* 1384-1395, 1989.

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

Welch, R. M., S. K. Sengupta, and D. W. Chen, Cloud field classification based upon high spatial resolution textural features: 1. Gray level co-occurrence matrix approach, *J. Geophys. Res., 93,* 12,663-12,681, 1988.

Weszka, J. S., C. R. Dyer, and A. Rosenfeld, A comparative study of texture measures for terrain classification, *IEEE Trans. Syst. Man Cybern., SMC-6,* 269-285, 1976.

Wielicki, B. A., and L. Parker, On the determination of cloud cover from satellite sensors: The effect of sensor resolution, *J. Geophys. Res., 97,* 12,799-12,823, 1992.

Wielicki, B. A., and R. M. Welch, Cumulus cloud properties derived using Landsat satellite data, *J. Clim. Appl. Meteorol., 25,* 261-276, 1986.

Wooster, M. J., T. S. Richards, and K. Kidwell, NOAA-11 AVHRR/2 – thermal channel calibration update, *Int. J. Remote Sens., 16,* 359-363, 1995.

R. Davies and L. Di Girolamo, Institute of Atmospheric Physics, The University of Arizona, 1118 E. Fourth St., Rm. 542, P. O. Box 210081, Tucson, AZ 85721-0081. (e-mail: davies@air.atmo.arizona.edu; larry@air.atmo.arizona.edu)

(Received March 21, 1996; revised August 8, 1996; accepted August 8, 1996.)