Rapid Daytime Estimation of Cloud Properties over a Large Area from Radiance Distributions

ANDREW K. HEIDINGER

NOAA/NESDIS Office of Research and Applications, Madison, Wisconsin

(Manuscript received 26 August 2002, in final form 3 March 2003)

ABSTRACT

An algorithm is developed to rapidly estimate cloud properties for a large area from daytime imager data. In this context, a large area refers to a grid cell composed of many imager pixels. The algorithm assumes a gamma distribution to model the subgrid variability in the optical depth and estimates both the mean and the width of the horizontal distribution of optical depth. Optical depth in this study refers to a vertically integrated value at 0.63 μm. Mean values of the cloud-top effective particle radius and cloud-top temperature are also estimated. Retrievals were performed separately for ice and water cloud layers within a grid cell. Applications of this approach to data from NOAA’s Advanced Very High Resolution Radiometer (AVHRR) are presented. Simulations indicate that this method performs well for all retrieved parameters except for thin clouds with very broad distributions of optical depth. Comparison of this approach versus rigorous pixel-level retrieval results for an actual scene with multiple cloud layers indicate that comparable performance is achieved with a two to three orders of magnitude increase in computational efficiency. This approach is being implemented into the Clouds from AVHRR (CLAVR) suite of cloud algorithms at NOAA. The computational efficiency of this approach will allow for efficient reprocessing of the entire data record of the AVHRR.

1. Introduction

In this paper, an algorithm is presented that derives the statistics of the horizontal distribution of the cloud optical depth over a large region (a grid cell) based on the statistics of the horizontal distribution of the visible reflectance and the mean values of other channel observations. In this context, a large region is an area consisting of many pixels such as grid cells used for numerical weather prediction (NWP) models that have resolutions ranging from 20 to 100 km. In addition to the statistics of the horizontal distribution of optical depth, this algorithm also estimates the mean cloud-top effective particle radius and cloud-top temperature within each grid cell.

There are three main reasons why this type of algorithm is appealing. The first reason is computational efficiency. The production of real-time pixel-level global cloud data from imagers with resolutions of 1–4 km is challenging. The reprocessing of imager datasets at full resolution for climate studies presents even more computational challenges. The computational efficiency of this algorithm is achieved by requiring only one retrieval per cloud layer per grid cell. In contrast, pixel-level algorithms perform a retrieval for each cloudy pixel. For a 50-km grid cell composed of 1–4-km pixels, this results in potentially a two to three order of magnitude reduction in required retrievals. Simulations will demonstrate that this reduction in computations preserves a meaningful representation of the cloudiness.

The second reason is that the radiance distributions required for this algorithm already exist and continue to be compiled as standard products. For example, a standard International Satellite Cloud Climatology Project (ISCCP) product includes the mean and standard deviation of the cloudy radiances within each grid cell. In addition, the Advanced Very High Resolution Radiometer (AVHRR) Pathfinder Atmospheres (PATMOS) also provides this information from the afternoon polar orbiter AVHRR data from 1981 to 2001 (Stowe et al. 2002). The advantage of this approach is that once the radiance distributions are made, there is no need to reprocess the full-resolution data. The reprocessing capability is crucial to cloud climate studies because improvements in the radiative modeling of clouds are ongoing. For example, as new and more accurate models of cirrus scattering properties become available (Key et al. 2002), this algorithm would allow for relatively rapid recomputation of cloud climatologies consistent with these improvements.

Last, the product of this algorithm, the horizontal distribution of cloud optical depth, has been shown to be sufficient for accurate modeling of the solar radiative heating within a grid cell. Barker et al. (1996) dem-
onstrated this using three-dimensional radiative transfer modeling applied to high-resolution cloud fields. Oreopoulos and Barker (1999) developed a computationally efficient two-stream radiative transfer solution that directly incorporates the grid cell distribution of optical depth. The capability of remotely sensing the optical depth distribution within a grid cell may therefore be a critical tool for the efficient assimilation of relevant cloud properties by numerical weather prediction models.

In summary, the goal of this paper is to present an algorithmic approach to the daytime estimation of cloud properties that operates on the radiance distributions from cloud layers over a large area or grid cell. The cloud properties provided by this approach are the statistics of the horizontal distribution of the vertically integrated optical depth, the mean cloud-top particle effective radius, and the mean cloud-top temperature. The purpose of this paper is to demonstrate that the assumptions necessary for this algorithm are sufficiently valid and robust to allow for meaningful estimation of grid cell cloud properties. This demonstration will be conducted using both simulations and actual satellite observations.

The algorithm presented in this paper is currently being applied to data from the Advanced Very High Resolution Radiometer and serves as the daytime cloud property retrieval module in the Clouds from AVHRR (CLAVR) system. The CLAVR system will serve as the National Oceanic and Atmospheric Administration’s (NOAA) operational AVHRR cloud-processing system. The CLAVR products are generated at a resolution of 50 km and will be produced for the AVHRR data from the morning, midmorning, and afternoon polar-orbiting satellites. A separate but conceptually similar approach is applied for the nighttime generation of cloud properties. In addition to the cloud retrieval algorithms, the CLAVR system is composed of pixel-level cloud masking (Stowe et al. 1999; Heidinger et al. 2002) and cloud phase determination as well as the cloud-layering algorithm used to derive the radiance distributions used for the cloud property retrievals. The evolution of this algorithm can be traced in the conference proceedings of Heidinger and Stowe (1999) and Heidinger and Liu (2001).

2. Modeling $h(\tau)$ with a gamma distribution

Because this retrieval methodology uses radiance distributions from many pixels within a grid cell, some assumption about the subgrid variability in cloudiness is required. In this paper, it is assumed that the horizontal distribution function (hdf) of vertically integrated optical depth, $h(\tau)$, for each cloud layer in a grid cell can be modeled using a gamma distribution. With this assumption, the $h(\tau)$ can be expressed as

$$h(\tau) = \frac{1}{\Gamma(\nu)} \left(\frac{\nu}{\tau}\right)^{\nu-1} \exp\left(-\nu\tau\right),$$

where $\tau$ is the optical depth, $\bar{\tau}$ is the mean optical depth, and $\nu$ is the width parameter. The standard deviation of $\tau$, $\sigma_\tau$, is given by $\bar{\tau}/\sqrt{\nu}$. In this paper, $h(\tau)$ refers to the probability of occurrence within a cloud layer of a vertically integrated optical depth of value $\tau$.

To illustrate the behavior of the gamma distribution, Fig. 1 shows plots of $h(\tau)$ for $\bar{\tau} = 10$ and $\nu = 1$ and 10. As the gamma distribution width parameter, $\nu$, increases, the distribution narrows. For values of $\nu \ll 1$, the gamma distribution takes on the shape of an exponential distribution. For values of $\nu > 10$, the gamma distribution approaches the form of a Gaussian distribution. The distributions of the 0.63-μm reflectances corresponding to the optical depth distributions shown in Fig. 1 are given in Fig. 2. These reflectances were computed assuming overhead illumination, an oceanic surface, and a cloud composed of water droplets with an effective radius of 10 μm. Because of the asymptotic behavior of the variation of reflectance with optical depth, the distributions of reflectance tend to be more narrow than the distributions of optical depth. If reflectance varied linearly with optical depth, the two distributions would have the same shape. This retrieval does not assume any particular shape of the reflectance distribution. The only assumption is that knowledge of the statistics of the reflectance distribution is sufficient to uniquely retrieve $h(\tau)$. Results shown later will support this assumption.

Modeling $h(\tau)$ with a gamma distribution has been proposed by Barker et al. (1996). In this study, the distributions of optical fields from marine stratocumulus
were derived from high-resolution (28 m) Landsat data. The results indicated a gamma distribution accurately modeled the observed \( h(\tau) \). To test these assumptions for a wider range of cloudiness and for observations with the resolution of AVHRR GAC (4 km), an orbital segment of NOAA-16 AVHRR GAC data was analyzed. The left panel in Fig. 3 shows the 0.63-µm reflectance of an orbital segment of NOAA-16 AVHRR GAC that stretches from the equator to roughly 35°N. The coast of California is evident in the upper-right corner of this image. This scene was chosen for analysis because it contains multiple types of cloudiness. Present within this image are stratus, convective, and cirrus clouds. The radiance distributions are compiled separately for water and ice cloud pixels in each grid cell. The decision as to which pixels enter into the radiance distribution is critically dependent on the cloud mask and cloud phase algorithms. The center panel in Fig. 3 shows the CLAVR cloud mask for this scene. The CLAVR cloud mask discretely classifies each pixel as being clear (0), partly clear (1), partly cloudy (2), or cloudy (3). The current version of CLAVR allows for one ice and one liquid water cloud layer in each grid cell. The right panel in Fig. 3 shows the cloud type or phase derived for this scene with the dark gray pixels being the noncloudy, the gray pixels representing the liquid water clouds, and the white pixels representing the ice phase clouds. Both the cloud mask and cloud typing algorithms use a multispectral approach.

To analyze the pixel-scale distribution of cloud properties, a pixel-scale retrieval of \( \tau \) and \( e \) was used. This retrieval uses only 0.63- and 1.6-µm reflectances and does not retrieve the cloud-top temperature \( T_c \). Values of the 11-µm brightness temperature are used for surrogates of \( T_c \) for this analysis. The forward models used in the retrieval for this section are identical to the ones described in the next section except that each pixel is modeled as a plane-parallel (horizontally uniform) cloud. Once the pixel values of \( \tau \) and cloud-top effective radius, \( r_e \), were computed for this image, the pixels were grouped into grid cells with each grid cell being a 20 × 20 array of pixels. The analysis was found to be insensitive to the size of the grid cells; therefore, the results apply to a wide range of grid resolutions. The mean and standard deviation computed from observed \( h(\tau) \) were used to define a gamma distribution for each grid cell. A reduced \( \chi^2 \) test (Taylor 1982) was performed to determine how well the gamma distribution fit the observed \( h(\tau) \). If the reduced \( \chi^2 \) value is of the order

![Graph showing horizontal distribution functions for \( R_1 \) for \( \tau = 10 \) and \( \nu = 1 \) and \( \nu = 10 \). In legend, \( \langle \tau \rangle \) is equivalent to \( \tau \).](image1)

![Image showing reflectance of an orbital segment of NOAA-16 AVHRR GAC data, CLAVR cloud mask, and cloud type or phase derivation](image2)
or less than unity, the gamma distribution is a good model of the observed distribution. The bin size used for computing \( h(\tau) \) was 1 and the allowable range was 0 to 100. For most grid-cell-reduced \( \chi^2 \) computations, the number of degrees of freedom was much larger than 10. Therefore, this analysis should be able to test the validity of the gamma distribution for modeling \( h(\tau) \).

The distributions of the values of reduced \( \chi^2 \) for this scene are given in Fig. 4 and are scaled to integrate to unity. Distributions were constructed for both ice and water cloudy pixels as well as for all cloudy pixels combined. Because the cloud mask can classify a pixel as either cloudy or partly cloudy, the distribution for cloudy and partly cloudy pixels combined is also shown. The values in parentheses in the legend of Fig. 4 are the mean values of reduced \( \chi^2 \) for each distribution. These results indicate that the \( h(\tau) \)'s derived from cloudy pixels with and without cloud typing are well modeled by the gamma distribution. The inclusion of the partly cloudy pixels does raise the mean reduced \( \chi^2 \) but not to the point of indicating a failure of the gamma distribution assumption. Even without cloud typing, the gamma distribution appears to be valid for most grid cells in this scene. Only for less than 1% of the grid cells, does reduced \( \chi^2 \) exceed 1.5 indicating an inability of the gamma distribution to accurately model \( h(\tau) \).

The other assumption in this retrieval is that it is sufficient to estimate only the mean values of \( r_c \) and \( T_c \), for each layer in each grid cell. To test this assumption, the distributions of \( r_c \) and \( T_c \) were also computed. The bin sizes used were 0.5 \( \mu m \) for \( r_c \) and 0.5 K for \( T_c \). The mean value of the width parameter for \( h(\tau) \) was 4. In corresponding mean width, parameters for the distribution of \( r_c \) and \( T_c \) were 10 and 30. The narrow distributions observed for \( r_c \) and \( T_c \) lend support to the assumption that only the mean values of these parameters need to be retrieved. This work does not imply that there is no information in these distributions but only that a meaningful representation of cloudiness for this application is provided by their mean values.

3. Retrieval algorithm

To retrieve the cloud parameters, a one-dimensional variational (1DVAR) technique is used. In this retrieval, four parameters are being estimated from five observations. A 1DVAR approach was selected because more traditional iterative approaches were found to offer no advantage in speed or in convergence: A 1DVAR approach also offers the benefits of ensuring the retrievals are consistent with the uncertainties in both the forward models and the measurements.

a. Required observations

As mentioned previously, the data used for this algorithm comes from NOAA’s AVHRR. The observations used are the channel 1 (0.63 \( \mu m \)) reflectance \( R_1 \), the channel 4 (11 \( \mu m \)) brightness temperature \( T_4 \), and the channel 5 (12 \( \mu m \)) brightness temperature \( T_5 \). The near-infrared reflectance offered by the AVHRR is obtained from either channel 3a (1.6 \( \mu m \)), \( R_{3a} \), or channel 3b (3.75 \( \mu m \)), \( R_{3b} \). The AVHRR instruments on NOAA-15 and later can only record channel 3a or channel 3b. Before NOAA-15, all AVHRR’s provided only \( ch3b \). Because the current operational afternoon-orbiting AVHRR (NOAA-16) records channel 3a during daylight operation, results for retrievals using \( R_{3a} \) are emphasized. Simulations will be presented that offer insight into the differences in retrieval performance for both channel configurations. In this algorithm, both the solar and thermal components of channel 3b are combined and expressed as a reflectance. Separate solar and thermal models, described later, are combined to serve as the forward model of \( R_{3b} \).

The required observations to this algorithm are the mean and standard deviation of \( R_1 \) and the mean values of \( R_{3a+b} \), \( T_4 \), and \( T_4 - T_5 \). As stated before, these observations are compiled separately for ice and water cloud layers within each grid cell. The use of \( R_1 \) and \( R_{3a+b} \) to simultaneously derive cloud optical depth and cloud-top effective radius has been well documented (i.e., Nakajima and King 1990). In addition, many algorithms have incorporated \( T_4 \) to estimate cloud-top temperature. The use of \( T_4 - T_5 \) is not so common in cloud retrievals. Figure 5 shows the modeled variation of \( T_4 - T_5 \) as a function of \( \tau \) and \( r_c \) for a cirrus cloud viewed at nadir with \( T_c = 240 \) K over a surface with a temperature of 300 K. The width parameter \( \nu \) was set to 20 for these simulations and these clouds are therefore relatively spatially uniform in optical depth or emissivity.
ity. As this figure shows, for thin clouds (τ < 4), values of $T_3 - T_x$ constrain the allowable values of $r_e$ and $τ$. For thick clouds, $T_4 - T_x$ provides little information except for a small dependence on $r_e$. In addition, use of $T_4 - T_3$ aids in achieving continuity between the cloud products with the nighttime algorithm. As Parol et al. (1991) demonstrated, $T_4 - T_3$ can exhibit a strong sensitivity to particle shape. Since a fixed particle shape is assumed here, this potential error will be included in the uncertainty assigned to estimates of these observations.

b. 1D-VAR retrieval

As described below, this algorithm uses five observations to derive four parameters. The traditional direct iteration techniques used by algorithms retrieving two or three parameters, such as the one described by Nakajima and King (1990), become difficult to implement for this algorithm. Instead, an optimal estimation or 1DVAR retrieval methodology is employed. A 1DVAR approach offers a simultaneous retrieval of all parameters based on minimizing a cost function composed of the difference in the observations and the estimates from a forward model. The form of the 1D-var retrieval methodology used here is taken from Rodgers (1996). This approach has also been used previously for cloud retrievals by Heidinger and Stephens (2000) as well as Miller et al. (2000). Using Rodger’s notation, the vector of measurements, $y$, for this algorithm is given by

$$y = (R_1, \sigma_1, R_{3a}, T_4, T_4 - T_3),$$

where $\sigma_1$ is the standard deviation of $R_1$. The bar superscript denotes the mean for all cloudy pixels of the same phase within a grid cell.

The vector of estimated parameters, $x$, is defined as

$$x = [\log(\tau), \log(r_e), T_x, \log(\nu)],$$

where $\tau$ is the mean optical depth, $r_e$ is the mean effective radius, $T_x$ is the mean cloud-top temperature, and $\nu$ is the width parameter of $b(\tau)$. Logarithmic values of $\tau$, $r_e$ and $T_x$ are retrieved to improve the convergence characteristics.

In addition, this retrieval method requires the definition of a priori estimates of the retrieved parameters, $x_{ap}$, that are derived based on representative values for each cloud phase. The final solution is an optimal estimate of $x$ that balances the uncertainties in the measurements, the forward models, and the a priori parameters.

One conceptual difficulty in applying these techniques for the retrieval of cloud properties is the specification of $x_{ap}$. In many applications, model or climatologic values can be used but for this application, this data is generally not available for these properties. The a priori values of $\tau$, $r_e$, and $T_x$ are taken from simple relationships derived from analysis of simulations of clouds over the ocean using the forward models described in a following section. These relationships are meant to offer only rough estimates of $\tau$ and $r_e$. The expression used to derive the a priori of $\tau$ is given by

$$\tau_{ap} = 0.7(R_1)^{0.9},$$

where $R_1$ is a reflectance expressed as a percentage. To derive a priori estimates of $r_e$, separate expressions for ice and water clouds are necessary. When using $R_{3a}$, the ratio of $R_{3a}$ to $R_1$ was found to be the dominant driver in the retrieval of $r_e$. For water clouds, the a priori value of $r_e$ is computed as

$$r_{e, ap} = 8 + 16.25 \left( \frac{R_{3a}}{R_1} - 1.2 \right)$$

and for ice cloud the a priori value of $r_e$ is computed as

$$r_{e, ap} = 15 + 15 \left( \frac{R_{3a}}{R_1} - 1.0 \right).$$

When $R_{3b}$ is used, the a priori values of $r_e$ are set to 10 μm for water clouds and to 20 μm for ice clouds. To estimate the a priori value of $\nu$, the width parameter derived from the mean and standard deviation of $R_1$ is used (see Figs. 1 and 2).

$$\nu_{ap} = \frac{\nu \bar{\tau}}{2}. \quad (5)$$

The reduction of the reflectance width parameter in half accounts for the tendency for optical depth distributions to be wider than reflectance distributions (see Figs. 1 and 2). For the a priori value of $T_x$, the value of $T_3$ is used. The above values can be inaccurate estimates under many conditions. The proper setting of the uncertainties of these estimates will prevent them from overconstraining the retrieval.

Another requirement of the 1DVAR retrieval not typ-
ically levied on other cloud algorithms is the requirement to specify the uncertainties of $x_{ap}$, the measurements $y$, and in the forward model's estimate of the measurements $f$. Because little information is available to refine the value of $x_{ap}$, their uncertainties are set to high values to prevent improper constraints on the solution. For example, the uncertainty for the a priori value of $\log T$ is assumed to be half its value plus 0.2. The uncertainty in $\log \nu$ and $\log r_s$ are both set to 0.5. In general, these values imply greater than 100% relative uncertainty in terms of the actual values of $\tau$, $r_s$, and $\nu$. The uncertainty in a priori $T_c$ is set to 20 K. These values are admittedly not based on rigorous analysis. The attempt here is to include all relevant errors and to adjust these uncertainties when knowledge warrants it.

As given by Rodgers (1996), once the errors in the measurements, forward model, and a priori parameters are specified, the 1DVAR retrieval allows for estimation of some diagnostic terms. One diagnostic term is a measure of the relative weight of the retrieval on the observations compared to the a priori parameters. These weights are given by matrix $A$, referred to as the measurement resolution matrix (Menke 1989), which can be expressed as

$$A = \frac{K^T S^{-1} K}{S^{-1}}$$

where $S_x$ is the error covariance matrix of $x$ and is given by

$$S_x^{-1} = S_f^{-1} + K S_c^{-1} K^T$$

and $K$ is the kernel matrix composed of the Jacobian of $f$ to $x$ (i.e., $df/dx$). In (9), $S_x$ is the error covariance matrix of the a priori parameters and $S_c$ is the error covariance of the measurements and the forward model. Here $S_x$ and $S_c$ are treated as diagonal matrices with each diagonal term being the square of the uncertainties given above. In an ideal observing system, $A$ would be the identity matrix indicating no reliance of the retrieval on the a priori parameters. As will be shown later, the diagonal terms of $A$ are useful for predicting conditions where the retrieval relies significantly on the a priori information.

c. Forward modeling

The goal of the forward models put forth here is to simulate the behavior of cloud fields within grid cells well enough to allow for meaningful estimation of their properties. Since global, faster than real time processing is required of this algorithm, true radiative transfer models are approximated by lookup tables. The forward models described here are similar to the forward models used in pixel-scale retrievals except that the width of the $\tau$ distribution for a grid cell, $\nu$, is an additional free parameter. The forward models described below provide the vector $f$ and the kernel matrix $K$, where each element contains the terms of the Jacobean matrix ($\partial f/\partial x$). To compute the scattering properties of clouds, the cloud particles were assumed to be spherical droplets. While this is a well accepted practice for water clouds, for ice clouds it can lead to large uncertainties. Finding methods for proper specification of ice crystal scattering properties is an area of ongoing research (Baran et al. 1999; Yang et al. 2000). In future validation studies, more appropriate ice crystal models will be implemented. The use of spherical ice particles does not directly impact the relative findings of this study.
1) Modeling solar reflectance

To model the observed solar reflectance, a lookup table is computed using an adding/doubling model for the reflectance of a plane parallel cloud above a dark surface with no other atmospheric effects, \( R_{\text{cc}} \). Once the lookup tables for single-layer plane-parallel clouds are made, lookup tables with the additional dimensions of \( \nu \) and the Lambertian surface reflectance, \( a_s \), are made. To compute the top of atmosphere reflectance for a plane-parallel cloud, \( R_{\text{pp}} \), above a Lambertian reflecting surface, the following expression taken from Chandrasekhar (1960) is used:

\[
R_{\text{pp}}(\tau) = t_{\text{sc}}^\nu \left[ R_c + \frac{T(\mu) a_s' T(\mu_c)}{1 - a_s' \alpha_{\text{sph}}} \right],
\]  

(10)

where \( \alpha_{\text{sph}} \) is the spherical albedo of the cloud layer, \( t_{\text{sc}} \) is the nadir transmission from the top of atmosphere to cloud top, \( m \) is the airmass factor \((1/\mu + 1/\mu_c)\), \( \mu_c \) is the cosine of the solar zenith angle, and \( \mu \) is the cosine of the viewing zenith angle. The dependencies on the solar/viewing geometries and other parameters are not included in (10) for simplicity. The terms \( T(\mu) \) and \( T(\mu_c) \) are the flux transmissions through the cloud layer (direct and diffuse) for a solar beam incident at zenith angles defined by \( \mu \) and \( \mu_c \). The modified surface albedo, \( a_s' \), accounts for absorption between the cloud and the surface and is approximated as

\[
a_s' = t_{\text{sc}}^\nu a_s,
\]

where \( t_{\text{sc}} \) is the nadir transmission from the cloud to the surface. Because this study deals with clouds over an oceanic surface, \( a_s \) was set to 3% for all reflectance computations. In the global application, \( a_s \) is determined for each channel as a function of the surface type and the solar zenith angle.

The final mean reflectance for a distribution of optical depth \( \bar{R}(\tau, \nu) \) is found by integrating (10) over the assumed \( h(\tau) \) to give

\[
\bar{R}(\tau, \nu) = \int_0^\infty R_{\text{pp}}(\tau) h(\tau) \, d\tau
\]

(11)

with the standard deviation \( \sigma \) being computed as

\[
\sigma(\tau, \nu) = \sqrt{\int_0^\infty [\bar{R}(\tau, \nu) - R_{\text{pp}}(\tau)]^2 h(\tau, \nu, \tau) \, d\tau.}
\]

(12)

The result of these computations are lookup tables for the mean and the standard deviation in reflectance as a function of the cloud properties, \( \tau, r_c, \nu \), the surface reflectance \( a_s \); the viewing geometry defined by \( \mu_c, \mu \); and the relative solar azimuth angle \((\phi_c - \phi)\). As mentioned earlier, the dimensions of the reflectance lookup tables are composed of equally spaced vectors of \( \log(\tau) \), \( \log(r_c) \), and \( \log(\nu) \).

2) Modeling thermal emission

In a similar manner to the modeling of solar reflectance, an adding/doubling model is used to compute the cloud emissivity and cloud transmissivity. The emissivity, \( \epsilon_c \), was computed as the ratio of the radiance emanating out of the top of cloud divided by the radiance emanating from a blackbody at the same temperature with no contribution from radiance below the cloud. The transmissivity, \( t_c \), was computed by turning off all cloud emission and computing the ratio of the radiance emanating out of the top of the cloud to that incident at the base of the cloud. These computations therefore result in nearly temperature-independent quantities that include effects of scattering. Values of the mean emissivity \( \bar{\epsilon} \), and mean transmissivity \( \bar{t} \), are computed using the plane-parallel lookup tables for all values of \( \tau \) and an integration over \( h(\tau) \) as shown below:

\[
\bar{\epsilon} = \int_0^\infty \epsilon(\tau) h(\tau) \, d\tau.
\]

Once the lookup tables of emissivity and transmissivity are made as functions of \( \log(\tau), \log(\nu), \log(r_c) \), and \( \mu \), the cloud layer is embedded in a nonscattering atmosphere over a non-scattering surface to model the top-of-atmosphere radiance using the following relation:

\[
E = E_{\text{cc}} (1 - E_o) + t_{\text{cc}}^\nu \bar{\epsilon} B(T_o) + \bar{t}, E_{\text{clear}},
\]

(13)

where \( E_{\text{clear}} \) is the clear-sky radiance, \( E_{\text{cc}} \) is the emitted radiance from the layer above the cloud, and \( m \) is air mass for emitted radiation \((1/\mu)\). To account for the lapse rate within the cloud, the cloud emissivity is redefined. Using a linear-in-depth variation of the emission through the cloud, a modified mean emissivity of the cloud, \( \bar{\epsilon}' \), can be computed to account for temperature variations using the following relation:

\[
\bar{\epsilon}' = \bar{\epsilon} \left( 1 - \frac{B_n}{B_o} e^{-\tau_{\text{top}}} \right) + \left( 1 - e^{-\tau_{\text{base}}} \right) \frac{(B_n - B_o)}{\tau_{\text{base}}},
\]

(14)

where \( B_n \) and \( B_o \) are the blackbody-emitted radiances at the cloud top and cloud base and \( \tau \) in this expression is defined as \(-\ln(t_{\text{cc}})\).

4. Simulated performance of retrievals

In this section, simulations are used to illustrate the performance of the retrieval over a wide range of conditions. These simulations focus only on the retrieval over an ocean surface. The results for retrievals over bright surfaces show larger errors but are qualitatively similar. The goal here is to diagnose which grid cell distributions of optical depth, \( h(\tau) \), will present difficulties to this retrieval approach. Because significant differences exist, results using both \( R_{3b} \) and \( R_{3a} \) are shown.

In these simulations, \( h(\tau) \) is modeled with a gamma distribution to compute the reflectance and radiance dis-
Fig. 6. Simulated retrieval results for AVHRR using $R_{3a}$ for water clouds. The $x$ axis is the domain-averaged optical depth $\tau$, and the $y$ axis is optical depth width parameter $\nu$. (a), (c), (e), (g) Errors in $\tau$, $\nu$, $r_e$, and $T_c$, respectively. (b), (d), (f), (h) Observation weight for $\tau$, $\nu$, $r_e$, and $T_c$, respectively. No value is less than 0.95.

Results will therefore not directly test the validity of using a gamma distribution to model $h(\tau)$. The analysis of the previous section indicates this is not warranted. These results do, however, test the assumption of using the mean values of $R_{3b}$, $T_a$, and $T_s$ by allowing $r_e$ and $T_c$ to vary with $\tau$ when computing the simulated observations. Based on the analysis of marine stratocumulus clouds by Szczodrak et al. (2001), $r_e$ is assumed to vary linearly with $\tau^{0.2}$. The actual form used was

$$r_e = 6.31\tau^{0.2}, \quad (15)$$

which gives a value of $r_e = 10 \mu m$ for $\tau = 10$. After Minnis et al. (1993), the variation of cloud geometrical thickness in kilometers, $H$, is given by

$$H = 0.045\tau^{0.3}. \quad (16)$$

While these relations are not meant to be globally valid, they do serve to introduce subgrid-cell variation in $r_e$ and $T_c$ that is correlated with $\tau$, as would be expected for most cloud fields. Other specified parameters include an oceanic surface, a cloud base of 500 m, and an atmosphere modeled using the standard midlatitude summer profile. The solar zenith angle was 30°, the satellite zenith angle was 30°, and the relative azimuth was 150°. These simulations are meant only to illustrate the performance of the retrieval under realistic calibration errors and the error incurred by assuming uniform values of $r_e$ and $T_c$. Retrievals were run on the simulated data using the above relations and the retrieval parameters as described above. Random errors of 5% were applied to $R_a$ and $R_{3b}$, 2% to $R_{3a}$, and 0.5 K to $T_a$ and $T_s$. These errors are conservative estimates of the calibration errors of the AVHRR. To simulate uncertainties in the knowledge of the atmosphere and surface, the skin temperature of the ocean surface was randomly varied by 2 K while the total precipitable water amount was randomly varied by 30%.

Figure 6 shows the results for simulated retrievals for an AVHRR using $R_{3a}$ as function $\tau$ and $\nu$ for a water.
cloud with the properties described above. The results in Fig. 6 are the mean values computed from 100 simulations with errors described above randomly varied. The 1DVAR retrieval approach automatically gives estimates of the error in the retrievals, $S_x$, and the relative weight of the observations in the retrieval, $A$. To show the performance of these simulated retrievals, the values of $S_x$ were not used in favor of the actual relative differences between the retrieved and true parameters. The contours of $S_x$ and $A$ were found to be highly correlated and the presentation of the actual errors along with the observation weights was found to provide a more insightful diagnosis of the retrieval performance.

The plots in the left column are the relative errors in the retrieved parameters relative to the true values used to compute the radiance distributions. The right column contains the contour plots of the weight of the retrieval of each parameter on the observation taken from the diagonal elements of $A$. In general, the errors are largest where the reliance on the observations is least. The interpretation of these simulated results depends ultimately on the required accuracy of these retrievals. For the interpretations given here, only when errors in $\tau$, $\nu$, $r_s$, or $\nu$ exceed 20% or the errors in $T_c$ exceed 1 K will a potential weakness with the retrieval performance be noted.

The results for optical depth (Figs. 6a and 6b) indicate good performance and high observation weight over most values of $\tau$ and $\nu$. Only for large values of $\tau$, do optical depth errors exceed 20%. The increase in optical depth error for large values of $\tau$ is due to the decrease in sensitivity of reflectance to optical depth and the corresponding increase in the weight of the a priori value of $\tau$.

Figures 6c and 6d give the results for the retrieval of $\nu$. Errors in $\nu$ generally exceed 10% for all simulations. However, errors in $\nu$ exceed 20% for all cloud layers with exponential distributions ($\nu < 0.5$). The largest errors (over 100%) occur for optically thick layers with small values of $\nu$. The values of $A$ in Fig. 6d are con-

Fig. 7. Simulated retrieval results with AVHRR using $R_{\nu}$ for ice clouds. The $x$ axis is the domain-averaged optical depth $\tau$, and the $y$ axis is optical depth width parameter $\nu$. (a), (c), (e), (g) Errors in $\tau$, $\nu$, $r_s$, and $T_c$, respectively. (b), (d), (f), (h) Observation weight for $\tau$, $\nu$, $r_s$, and $T_c$, respectively.
F I G . 8. Simulated results with AVHRR using $R_{3b}$ for $r_e$. The $x$ axis is the domain-averaged optical depth $\tau$, and the $y$ axis is optical depth width parameter $\nu$. (a) The errors and (b) the observation weights.

The performance of the retrieval of $r_e$ can be taken from Figs. 6e and 6f. Errors in $r_e$ are greater than 20% for most $\nu$ when $\tau < 10$. Errors for $r_e$ exceed 50% for $\tau < 20$ and $\nu < 2$. The corresponding observation weights depart significantly below unity for these large error regions. The reason for these large errors is the lack of significant absorption of $R_{3a}$ by water droplets for optically thin clouds. For a given value of $\tau$, as $\nu$ decreases, more and more of the cloud layer optical depth is proportionally distributed among smaller values of optical depth. Because small values of optical depth imply less absorption and less sensitivity to particle size, this contributes to the increase in error for small values of $\nu$.

The simulated results for the retrieval of $T_c$ are shown in Figs. 6g and 6h. Only for optically thin clouds ($\tau < 4$) with exponential distributions do errors in $T_c$ exceed 1 K and do the observation weights drop slightly below unity.

Analogous simulations to the ones described above were applied to ice clouds over an oceanic surface and are presented in Fig. 7. The results are qualitatively similar. The main difference in the ice cloud performance is the increase in the uncertainty of the forward model and the increase in the relevance of $T_4 - T_5$ for optically thin ice clouds. The results for the retrieval of $\tau$ and $\nu$ show no dramatic difference. The errors in the retrieved $\nu$ are generally larger for optically thin ice cloud than water cloud. However, the largest errors in $\nu$ still occur for layers with exponential distributions ($\nu < 1$). Due to an increase in the sensitivity in $R_{3a}$ to particle size in the ice cloud simulations, the errors in the retrieved $r_e$ are correspondingly less than for the water cloud simulations. Because ice clouds are much colder than the surface, there is greater potential for error in $T_c$ for semitransparent clouds. Correspondingly, the errors in $T_c$ approach 4 K and the observation weights depart significantly from unity for $\tau < 4$ and $\nu < 1$.

As mentioned earlier, current AVHRR instruments have the option to measure $R_{3b}$ rather than $R_{3a}$. To explore the consequences of this in the context of these simulations, the water cloud simulations described above were repeated substituting $R_{3b}$ for $R_{3a}$ and the
results are given in Fig. 8. The results for the retrieval of $\tau$, $\nu$, and $T_c$ were little affected by the channel choice and are not shown. Dramatic differences were evident in the performance for the retrieval of $r_e$. Where the results using $R_{3a}$ generally showed errors exceeding 10%, the results using $R_{3b}$ show errors of less than 10% for most values for $\tau$ and $\nu$. Only when $\nu < 1$ and $\tau < 3$, do the simulated errors exceed 20%. This performance improvement is reflected in the observation weights that exceed 0.90 for most values and never go below 0.75. The reason for this improvement is the presence of more water droplet absorption at 3.75 $\mu$m than at 1.6 $\mu$m. In addition, differences in retrieved $r_e$ using $R_{3a}$ or $R_{3b}$ can occur due to differing sensitivities at different levels in the cloud between the two channels (Platnick 2000). These effects do indicate a potential discontinuity in the data record of $r_e$ derived from AVHRR.

The results of these simulations are necessarily dependent on the choices made for the uncertainties used in computing $S_e$ and $S_a$, and in the prescription of the a priori parameters, $x_\text{ap}$. For example, if the measurements were assumed perfect, the errors in the $r_e$ retrieval using $R_{3a}$ or $R_{3b}$ can occur due to differing sensitivities at different levels in the cloud between the two channels (Platnick 2000). These effects do indicate a potential discontinuity in the data record of $r_e$ derived from AVHRR.

The results of these simulations are necessarily dependent on the choices made for the uncertainties used in computing $S_e$ and $S_a$, and in the prescription of the a priori parameters, $x_\text{ap}$. For example, if the measurements were assumed perfect, the errors in the $r_e$ retrieval using $R_{3a}$ or $R_{3b}$ can occur due to differing sensitivities at different levels in the cloud between the two channels (Platnick 2000). These effects do indicate a potential discontinuity in the data record of $r_e$ derived from AVHRR.

The results of these simulations are necessarily dependent on the choices made for the uncertainties used in computing $S_e$ and $S_a$, and in the prescription of the a priori parameters, $x_\text{ap}$. For example, if the measurements were assumed perfect, the errors in the $r_e$ retrieval using $R_{3a}$ or $R_{3b}$ can occur due to differing sensitivities at different levels in the cloud between the two channels (Platnick 2000). These effects do indicate a potential discontinuity in the data record of $r_e$ derived from AVHRR.

The results of these simulations are necessarily dependent on the choices made for the uncertainties used in computing $S_e$ and $S_a$, and in the prescription of the a priori parameters, $x_\text{ap}$. For example, if the measurements were assumed perfect, the errors in the $r_e$ retrieval using $R_{3a}$ or $R_{3b}$ can occur due to differing sensitivities at different levels in the cloud between the two channels (Platnick 2000). These effects do indicate a potential discontinuity in the data record of $r_e$ derived from AVHRR.

The results of these simulations are necessarily dependent on the choices made for the uncertainties used in computing $S_e$ and $S_a$, and in the prescription of the a priori parameters, $x_\text{ap}$. For example, if the measurements were assumed perfect, the errors in the $r_e$ retrieval using $R_{3a}$ or $R_{3b}$ can occur due to differing sensitivities at different levels in the cloud between the two channels (Platnick 2000). These effects do indicate a potential discontinuity in the data record of $r_e$ derived from AVHRR.

5. Comparison to pixel-level retrievals

Without independent cloudiness observations, the results of this retrieval cannot be verified. A relative validation is possible from the comparison of the grid cell retrievals to the results from composited pixel-scale retrievals. As the simulations showed above, the results from this approach should not be significantly worse than the standard pixel-scale retrievals for many cloudiness conditions.

To perform this comparison with the pixel-level retrievals, the data from the image in Fig. 3 is used. The CLAVR system was used to perform the cloud mask and cloud phasing shown in Fig. 3 and to generate the statistics of the AVHRR channel observations for ice and water cloud layers. Any one grid cell can report separate ice and water cloud layers. This analysis uses 0.5° equal area grid with the grid cells having a corresponding spatial resolution of approximately 55 km. This grid corresponds to the CLAVR implementation at NOAA.

The results of the grid cell retrieval for $\tau$ and $\nu$ are shown in Fig. 9 and for $r_e$ and $T_c$ in Fig. 10. Though the algorithm retrieves values for ice and water clouds separately, the results in Figs. 9 and 10 show the cloud-fraction weighted average of the ice and water results. The retrieval results are generally consistent with the expected properties for oceanic cloud systems. The water phase clouds have optical depths generally ranging from 4 to 20 and with effective radii ranging from 6 to
Fig. 11. Comparison of grid cell results with pixel-scale retrievals for grid cells containing water clouds. (a) Results for \( \tau \), (b) results for \( r_e \), and (c) results for \( \nu \).

Fig. 12. Comparison of grid cell results with pixel-scale retrievals for grid cells containing ice clouds. (a) Results for \( \tau \), (b) results for \( r_e \), and (c) results for \( \nu \).

20 \( \mu \)m. The ice clouds have greater dynamic range in \( \tau \) and generally larger values of \( r_e \). The cloud-top temperature also varies as expected. The retrieved values of \( \nu \) also capture the variation in roughness of the \( R_i \) image shown in Fig. 3 with high values of \( \nu \) being retrieved in uniform stratus region and lower values being retrieved in the more visually rough ice clouds.

To compare the grid cell results to pixel-level results, the same mapping routine used to make the AVHRR pixels was used to map the pixel-level retrievals into the appropriate grid cells. The mean and standard deviation of pixel-level retrievals of \( \tau \) and \( r_e \) were computed separately for the ice and water phase cloudy pixels. The value of \( \nu \) was derived from the pixel-level results from the mean and standard deviation of \( \tau \). Comparisons were only done when the cloud fraction for ice

<table>
<thead>
<tr>
<th>( \tau &gt; 0 )</th>
<th>( \tau &gt; 5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Std dev</td>
</tr>
<tr>
<td>( \tau )</td>
<td>-3.8</td>
</tr>
<tr>
<td>( r_e )</td>
<td>-6.7</td>
</tr>
<tr>
<td>( \nu )</td>
<td>-9.6</td>
</tr>
</tbody>
</table>

Table 2. Mean and standard deviation of the relative difference (%) between the grid cell and pixel-scale retrievals for grid cells containing ice clouds.

<table>
<thead>
<tr>
<th>( \tau &gt; 0 )</th>
<th>( \tau &gt; 5 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>Std dev</td>
</tr>
<tr>
<td>( \tau )</td>
<td>-19</td>
</tr>
<tr>
<td>( r_e )</td>
<td>-8.4</td>
</tr>
<tr>
<td>( \nu )</td>
<td>-12</td>
</tr>
</tbody>
</table>
or water exceeds 5% of the grid cell. To avoid any effects of navigation errors on the retrievals, any grid cell with one or more pixels being classified as land was not used. The effect of these constraints is to ensure that the same pixels that were used to derive the radiance distributions were used to construct the pixel-level retrievals. The resulting grid cells used for this analysis numbered approximately 1600 for the water cloud and 900 for the ice clouds.

The results of the comparison of the grid cell and the pixel-level water cloud values of $\tau$, $v$, and $r_e$ are shown in Fig. 11. There are no results for $T_c$ because the pixel-level retrieval does not estimate it. The results in Fig. 11 are plotted separately for $\tau > 5$ and for $\tau < 5$. As the simulations showed, the least reliance on the observations and largest errors occurred for $\tau < 5$. The mean and standard deviations computed for the water cloud results are given in Table 1 expressed as a percentage relative to the pixel-level value. The errors in Table 1 are shown separately for $\tau > 0$ and for $\tau > 5$. The results show excellent agreement for $\tau > 5$ with the mean and standard deviation of the differences for $\tau$ and $r_e$ being generally much less than 10%. Only the standard deviation of $r_e$ for $\tau > 0$ values exceeds 10%. For optically thick clouds, the mean difference in $v$ is small while the inclusion of optically thin clouds causes the mean difference to approach $-10\%$. The standard deviation of $v$ is roughly 20%.

The corresponding results for the comparisons of the pixel-level and grid cell retrievals for the ice clouds are shown in Fig. 12 and Table 2. The results are generally similar to those for water clouds. The most notable change is the mean difference in the retrieval of $\tau$ being roughly 20%. The other notable difference between the ice and water results is the significantly larger standard deviation of both the pixel-level and grid cell results for retrieval of $v$.

As noted above, calculations shown in Fig. 12 and Table 2 indicate a systematic difference in the mean optical depth results from the two methods for the optically thick ice cloud retrievals. The reason for this difference is unclear but is allowable given the large uncertainties in the retrieval of $\tau$ for optically thick clouds. In Fig. 7, errors in $\tau$ exceeded 20% for optically thick clouds. Given that both the plane-parallel and the grid cell retrievals suffer from this increase in error for optically thick cloud, it is unclear which retrieval is closer to the truth. A tuning exercise could have been carried out where the a priori estimates and prescribed uncertainties were varied to give better agreement between the two methods for large values of optical depth. However, because the differences are within the expected errors based on the ice cloud simulations, this exercise was not warranted nor is it clear it would increase the absolute accuracy of the grid cell results. It is notable that for the water cloud results, where the assumed uncertainties are less, this behavior is not evident.

6. Conclusions

A retrieval approach has been developed to rapidly estimate the cloud properties from grid cells composed of many imager pixels. The properties estimated are the mean and width of the horizontal distribution of optical depth and the mean value of the cloud-top effective radius and temperature for each cloud layer in the grid cell. Previous modeling studies of Barker et al. (1996) and Oreopoulos and Barker (1999) indicate that the cloud properties retrieved here are sufficient to accurately compute the mean solar radiative heating within the cloud layer. Simulations indicate that this method should perform well except for very optically thin cloud layers with very broad distributions of optical depth. It
is not clear that even pixel-level retrievals perform well for these cloud fields. The differences between the comparisons to the pixel-level comparisons are generally less than the errors predicted by simulations. This finding indicates that there is no significant penalty incurred in the accuracy of the cloud properties when this rapid approach is used.

In the future, this algorithm will be applied globally as part of NOAA’s operational CLAVR cloud product system. The current resolution of this product is 0.5°. To present a qualitative look at the application of this algorithm globally, Fig. 13 shows the results of this algorithm applied to all the ascending data (13:30 local time) from NOAA-16 on day 230 of 2001. The image does not show latitudes poleward of 60° because there is no solar reflectance in this region at this time. Work is ongoing to apply this technique to the radiance statistics of ISCCP or PATMOS to derive new cloud climatologies.

Acknowledgment. The author appreciates the critical review provided by Sharon Nebuda formerly of NASA’s Data Assimilation Office.

REFERENCES


