AVHRR Pixel Level Clear-Sky Classification Using Dynamic Thresholds (CLAVR-3)

S. VEMURY*
Scientific Management and Applied Research Technologies, Inc., Landover, Maryland

L. L. STOWE
NOAA/NESDIS, Camp Springs, Maryland

V. R. ANNE
I.M. Systems Group, Inc., Kensington, Maryland

(Manuscript received 26 January 1999, in final form 24 April 2000)

ABSTRACT

Clear-sky classifications from the Clouds from the Advanced Very High Resolution Radiometer (AVHRR)-Phase 1 (CLAVR-1) program are used to create an 8-day rotating clear-sky radiation dataset. This dataset is used to create satellite zenith angle dependent, dynamic, cloud/no-cloud albedo and temperature thresholds for ocean and six different vegetation index groups in 10° latitude intervals over the globe. Individual pixels from ambiguously classified 2 × 2 pixel arrays from CLAVR-1 (MIXED and RESTORED-CLEAR) are reexamined at the individual pixel level, using these dynamic thresholds, and reclassified as CLEAR, MIXED, or CLOUDY. This methodology is referred to as the CLAVR-3 algorithm. It is found that many of the MIXED (partially cloudy or mixed overcast) pixels from CLAVR-1 are cloud free after using the dynamic thresholds. A smaller number of RESTORED-CLEAR pixels are found to be CLEAR, while many are classified MIXED. Poleward of about 80°N and 60°S, the CLAVR-1 algorithm does not provide sufficient detection of unambiguous CLEAR pixels for the creation of the necessary albedo and temperature angular distribution models. Other techniques or datasets will have to be employed at these latitudes to provide the necessary dynamic thresholds for reclassification of CLAVR-1 ambiguous pixels.

The CLAVR-3 reclassification leads to a 75% increase in CLEAR pixel population, globally, during the ascending segment (mostly daytime portion) of orbits and a 95% increase during the descending segment (mostly nighttime). Maps of differences in albedo and brightness temperature for grid cells containing clear pixels before and after application of CLAVR-3 are used, as well as histogram analyses, to demonstrate the quality of clear pixels derived from the CLAVR-3 algorithm. These analyses show that the quality of CLEAR pixels from CLAVR-3 may be slightly lower compared with CLAVR-1, particularly over oceans, which is most likely the result of cloud contamination. However, this may be an acceptable consequence in exchange for the resulting dramatic increase in CLEAR pixel population. Each CLEAR pixel application (e.g., sea surface temperature) will have to evaluate how CLAVR-3 impacts their products. It is quite possible that by making minor adjustments and modifications to the threshold tests currently used in CLAVR-3, an optimum increase in CLEAR pixels can be achieved with acceptable levels of cloud contamination. Based on the results presented, the authors conclude that the CLAVR-3 algorithm concept has merit and can be used to enhance the spatial coverage of daily operational land, ocean, and atmospheric parameters that depend on clear-sky radiance observations from AVHRR.

1. Introduction

Cloud detection, specification, and screening using satellite measurements are critical in providing observations useful for studies of climate and weather forecasting. The purpose of cloud screening is to identify the presence of cloud in an instrument’s field of view (pixel), so as to eliminate it from clear pixel processing. These automated, clear pixel processing systems derive surface (land or ocean) and cloud-free atmospheric properties. The output of these systems is very useful in specifying land cover types, particularly vegetation (DeFries et al. 1995); determining sea surface temperatures (SSTs; McMillin and Crosby 1984; McClain et al. 1985; Antoine et al. 1992) and land surface temperatures (LSTs; Gutman et al. 1995); temperature/moisture profiles (McMillin and Dean 1982); and aerosol optical thickness (Stowe et al. 1997).

* Deceased.

Corresponding author address: Dr. Larry L. Stowe, NOAA/NESDIS/ORA, NSC (WWB), 5200 Auth Rd., Camp Springs, MD 20746. E-mail: lstowe@nesdis.noaa.gov
The most commonly used cloud screening techniques apply separate visible (albedo) or infrared (temperature) thresholds, bispectral thresholds (visible and infrared simultaneously), or multispectral thresholds to each pixel’s measured radiometric observations. These thresholds should be dependent on view and illumination angle, as well as surface type and season of the year. The effects of geometry on albedo and temperature have been reported in the literature, and are usually specific to a particular instrument, geographical location, or time of the year (e.g., Hutchison and Hardy 1995; Wu and Cihlar 1995; Taylor and Stowe 1984; Suttles et al. 1988a,b).

In this paper, these dynamic, regional and angularly dependent cloud/no-cloud thresholds are developed from data from the Advanced Very High Resolution Radiometer (AVHRR) using clear-sky classifications from the Phase 1 algorithm of the Clouds from AVHRR (CLAVR-1) project in the National Oceanic and Atmospheric Administration/National Environmental Satellite, Data, and Information Service (NOAA/NESDIS), described in Stowe et al. (1999). CLAVR has four phases with the objectives of developing a unified technique for cloud screening (Phases 1 and 3) and for cloud typing and optical property retrieval (Phases 2 and 4) using all five channels of the AVHRR on NOAA polar orbiting satellites. The CLAVR-1 cloud screening algorithm is a 2 × 2 pixel array classification scheme for AVHRR global area coverage data to separate clear from partially and fully cloudy arrays. The classifications are based on a sequence of threshold tests applied to “equivalent-isotropic” albedo \( \pi L/FD\mu_o \), where \( L \) is observed radiance, \( F \) is solar irradiance, \( D \) is earth–sun distance normalization factor, and \( \mu_o \) is cosine solar zenith angle and “brightness” temperature values, their spectral differences, and their spatial uniformity using all five spectral channels of data. Pixel arrays that could not be unambiguously classified as CLEAR or CLOUDY, were put either into a MIXED (partially cloudy or mixed overcast) or a RESTORED-CLEAR category (subsequently termed “uncertain” pixels).

The objective of Phase 3 of CLAVR is to increase the quantity of CLEAR pixels from CLAVR-1 while maintaining their quality. It involves construction of clear-sky albedo and temperature fields from the CLEAR pixels of CLAVR-1, from which regional, dynamic (time varying) angular dependence models (ADMs), first reported in Stowe et al. (1993), can be constructed. Spanning over a sequence of prior days in the repeat cycle of the NOAA satellite, these ADM statistics yield estimates of clear-sky albedo and temperature that can be used to predict the dynamic, cloud/no-cloud thresholds for the next day in the sequence. These are used to test for additional CLEAR pixels from within the “uncertain” set of pixels from CLAVR-1 during that next day.

Section 2 defines the problem and describes the characteristics of CLEAR and uncertain pixels from the 2 × 2 pixel array classification scheme used in CLAVR-1. Details on the construction of the clear-sky radiation datasets (CRDS) and methods to build regional angular dependence models for albedo and temperature from them are presented in section 3. Section 4 describes the dynamic, cloud/no-cloud thresholds developed from the regional ADMs and focuses on the algorithm used to recategorize uncertain pixels from CLAVR-1. The improvements obtained using the present recategorization methodology are presented in section 5. A brief description of future enhancements that may be implemented in the classification methodology under CLAVR-3 is included in section 6 together with conclusions.

2. CLAVR-1 classification and implications for CLAVR-3

There are two nonclear pixel array classifications from CLAVR-1 that may frequently contain clear pixels. These are (i) pixels with a MIXED classification and (ii) pixels with a RESTORED-CLEAR classification (Stowe et al. 1999). These two categories of pixel arrays are hereafter referred to as uncertain pixels.

The focus in CLAVR-1 was on the reliable detection of perfectly CLEAR pixels, both over land and over ocean. This led to a very conservative classification scheme, where all four pixels in a 2 × 2 array must fail as many as seven cloud tests to be classified as CLEAR. As a consequence, many CLEAR pixels were classified as MIXED because of their proximity to cloud contaminated pixels within the same 2 × 2 array. Also, a “restoral-to-clear” process was incorporated into CLAVR-1, based on the realization that for ocean sun glint, bright and hot land, or for snow/ice covered surfaces, the CLAVR-1 tests could lead to a MIXED or CLOUD classification when a pixel was actually CLEAR.

Recovering CLEAR pixels from these uncertain ones is the objective of the CLAVR-3 algorithm, accomplished by testing the above two classifications of pixels from CLAVR-1 with dynamic, cloud/no-cloud thresholds, which are viewing angle, surface type, and latitude (illumination angle) dependent.

a. CLEAR pixels in CLAVR-1

A complete month of AVHRR data for September 1989 from the NOAA-11 satellite has been classified by CLAVR-1 for testing of the CLAVR-3 methodology. One day [9 September 1989 (day 89 252)], processed using the present pixel level and dynamic threshold scheme (CLAVR-3) has been used in the analyses presented here.

As an example of the global distribution of CLEAR pixels from CLAVR-1, a clear albedo map for the ascending (mostly daytime) segment of day 89 252 is...
shown in Fig. 1. It is derived only from CLEAR pixel arrays mapped on a grid of 110 km × 110 km equal area cells and averaged to produce a mean clear albedo or temperature for each of the channels of AVHRR. The white areas, which cover much of the oceans and the poles, indicates that no CLEAR pixels were available from CLAVR-1. To be consistent with the albedo threshold tests used in the CLAVR-1 algorithm, channel-2 (0.7–1.1 μm) albedo over ocean and channel-1 (0.58–0.68 μm) albedo over land are used in the figure. Albedo values generally range from 0% to 50% for clear-sky cases, with the higher albedos occurring primarily over deserts. Albedos over the oceans are generally in the 5%–10% range, highest in the direction of sun glint (darker bands between 30°N and 10°S latitudes).

A clear-sky brightness temperature map derived from mean channel-4 (10.3–11.3 μm) radiances for the same day is shown in Fig. 2. Mean channel-4 CLEAR temperatures (T4) over oceans vary between 271 K and about 300 K, while over land, they range between 250 and 320 K. Global desert areas are clearly discernable from their high temperatures. There is no evidence of cloud contamination (irregular bright or cold patterns) in either figure. For land applications, Gutman and Ignatov (1996) have clearly demonstrated the usefulness of CLAVR-1 clear-sky flags for enhancing the quality of the AVHRR Land Pathfinder dataset.

During the descending portion of an orbit, only brightness temperatures can be used in CLAVR-3, as most of this portion of an orbit is viewed without solar illumination (nighttime). The global distribution of CLEAR T4 from CLAVR-1 during this portion is shown in Fig. 3. The temperature scale used is the same as the one used for the ascending part. There is no noticeable difference from the temperature range in Fig. 2 over the oceans, but over the land, the highest T4 values are clearly lower at night (~300 K), as expected due to no solar heating and due to the lower heat capacity of land compared to that of water. Due to the latter, some land areas appear cooler than the ocean. These three figures provide visual evidence of the quality of CLAVR-1 CLEAR classifications over land and oceans.

Sample sizes of CLEAR, RESTORED-CLEAR, MIXED, and CLOUDY categories from CLAVR-1 for
day 89 252 are consistent with the results presented in Stowe et al. (1999, cf. their Tables 1 and 2). At the pixel level, only 11.6% of the ascending pixels and 14.5% of the descending pixels are classified as CLEAR. The numbers for ascending are consistent with values obtained by Prince and Goward (1996) in their analysis using CLAVR flags in the Land Pathfinder dataset. The uncertain group, consisting primarily of MIXED pixels (RESTORED-CLEAR pixels have low populations except over Greenland and Antarctica), constitutes a major portion of the sample; 49.4% and 40.8% for ascending and descending portions, respectively.

b. Uncertain pixels in CLAVR-1

The distribution of mean albedo over the globe from the RESTORED-CLEAR pixels for this day is shown in Fig. 4. These pixels come from nonclear grid cells (grid cells with zero CLEAR pixel population) as well as clear grid cells. Over the oceans, these RESTORED-CLEAR populations lie along the western side of each orbital swath due to the equator crossing time of the NOAA-11 (afternoon) satellite. However, besides the elevated albedos resulting from sun glint, there is apparently some other high albedo surface being restored to clear. This is probably low stratus cloud which occurs most frequently in regions west of continents (Warren et al. 1988). Stratus clouds in other ocean regions are apparently also being similarly classified, as indicated by RESTORED-CLEAR albedos in excess of 40%. This happens because stratus clouds are commonly thermally uniform and bright, conditions required of RESTORED-CLEAR pixels in CLAVR-1 when viewing a clear ocean in the sun glint direction. Also, sea ice surfaces poleward of 50°, which are bright in channels 1 and 2 but dark in channel 3 (3.5–3.9 μm; cf. Stowe et al. 1999), may be RESTORED-CLEAR, although on this day there does not appear to be many classifications of this type. Exceptions are a few grid cells just off the coast of Antarctica and at the very highest latitudes in the Arctic.

Over land, three regions with RESTORED-CLEAR pixels are observed: those associated with bright, hot deserts; those associated with hot vegetated land surfaces; and those associated with snow cover. Except for snow cover over Antarctica, the locations and magnitudes of the grid cell mean albedos of RESTORED-CLEAR pixels are consistent with CLEAR albedos in
Fig. 3. Same as Fig. 2 except for descending (mostly nighttime) parts of orbits.

Fig. 1. Figure 4 shows that these pixels add information about the Earth’s surface beyond what is provided from CLEAR pixels alone.

Overall, the RESTORED-CLEAR process in CLAVR-1 appears to be working well for land, with the exception of Greenland, where on this day the CLAVR-1 algorithm classifies it as either MIXED (see Fig. 5) or CLOUDY. The same can be concluded for oceans, with the exception of RESTORED-CLEAR pixels contaminated by stratus cloud.

In general, grid cell average albedos of MIXED pixels (Fig. 5) lie between 10% and 25% over ocean and between 40% and 70% over land. This suggests, when compared with the magnitudes and spatial distribution of albedo for CLEAR from Fig. 1, that many of these MIXED pixels may be cloud-free. Unlike Fig. 4, grid cells with MIXED pixels are more randomly distributed over the globe.

Since MIXED and RESTORED-CLEAR classifications represent about 40%–50% of all AVHRR pixels on given day, a reclassification of them to identify additional CLEAR pixels is highly desirable. The remainder of this paper describes the CLAVR-3 algorithm, where regional, dynamic thresholds, derived from CLEAR observations from CLAVR-1, are used to select additional CLEAR pixels from these two uncertain ones.

3. Developing angular reflection and emission models for use in reclassification of uncertain pixels

This section discusses the development of clear-sky radiation datasets (CRDS), which are the basic ingredients in the construction of dynamic cloud/no-cloud radiometric thresholds necessary for reclassification. When cloud decisions are based upon a single visible and/or IR channel, for them to be effective, the cloud detection thresholds cannot be global in nature but need to be adjusted for local conditions. This is important because albedo and temperature fields are region specific. In particular, clear-sky temperatures vary significantly from the tropics to the poles. The thresholds also need to be dynamic because the fields vary daily, and even hourly.

NOAA polar orbiting satellites have a repeat cycle which varies between 6 and 10 days, depending on the stability of the orbit over the duration of the satellite mission. For the case of NOAA-11, a region viewed on
day 1 at a particular satellite viewing zenith angle (VZA) is seen at that same angle after 8 days. During those 8 days, the region is viewed by the radiometer from different VZAs until it comes back on the ninth day to be “in phase” with the first day (Gutman 1989). This easily permits the inclusion of VZA variation into the dynamic thresholds.

In view of this, albedo and temperature grid cell statistics for selected AVHRR channel pixels classified as CLEAR are computed for each day in the 8-day period to produce clear-sky radiation datasets (CRDS8). These consist of radiance means and standard deviations along with a mean Normalized Difference Vegetation Index (NDVI, computed from channel-1 and -2 clear-sky albedos), a surface characterization (land, ocean, and coastal) and mean viewing zenith (VZA), solar zenith, and relative azimuth angles for each of 41 252 equal area (EA) grid cells covering the earth. Separate CRDS8 files are used for the ascending and descending parts of each orbit. During the ascending part, most of the grid cells have albedo and temperature statistics from all channels, while during the descending portion, they mostly have temperatures from the infrared channels. At higher latitudes, some grid cells may be viewed during several consecutive orbits, thus providing several means and standard deviations. In such situations, CRDS8 chooses statistics from the orbit with the smallest VZA (i.e., most nadir view). During the 8-day interval, means and standard deviations in albedo and temperature vary primarily with changes in VZA, with lesser changes due to solar zenith or azimuth angles. This is a result of the AVHRR being on a sun-synchronous satellite, so that the relationship between solar zenith and relative azimuth with satellite view zenith repeats with the repeat cycle of the satellite, and only gradually changes as the solar declination moves with the seasons. Thus, the albedo and temperature dependence on solar zenith angle is accounted for by binning the 8 days of observations only by viewing zenith angle.

Each day’s mean values are used to construct ADMs from which localized cloud/no-cloud thresholds on albedo and temperature are derived. These thresholds differ substantially from the universal thresholds used in CLAVR-1. Another feature of this approach is that CRDS8 is updated as each additional day is processed, and the most recent 8-day statistics are used to construct the ADMs and the associated thresholds for the next
day. This process makes the ADMs regional, dynamic, and angle dependent.

It may eventually become necessary to compare dynamic thresholds from the 8-day composite ADMs with concurrent measurements of snow/ice cover from microwave instruments and surface skin temperature from weather prediction models, to account for day-to-day variability not present in the 8-day composites due to weather changes.

a. Building clear-sky angular dependence models

In principle, the CRDS8 file permits development of ADMs for each grid cell. But for many grid cells, particularly at high latitudes, they are frequently cloud covered, and thus clear-sky samples may not exist during most or all of the 8 days. Grid cell statistics for the first 8-day period in September 1989 indicate that nearly 7200 of the 42,152 grid cells do not have clear-sky observations for the entire 8-day period. Only 1500 grid cells have clear-sky observations for all 8 days. We have therefore adopted a methodology for compositing grid cell statistics to provide ADMs as a function of latitude and surface type.

We currently use 10° latitude zones for the accumulation of grid cell statistics. This interval is large enough to provide an adequate number of clear grid cells for creating ADMs for most surface types, while small enough to retain characteristic differences in ADMs due to solar zenith angle changes with latitude. The grid cells in each 10° latitude zone are classified into separate vegetation type (VT) categories. All grid cells over ocean are put into category 1 (VT = 1). Over land, each grid cell is categorized from an analysis of its monthly mean Normalized Difference Vegetation Index (NDVI), computed from the daily mean NDVI estimates on CRDS8. This parameter has been shown to be useful for land cover classification (DeFries and Townshend 1994).

The map of monthly mean NDVI for September 1989 is shown in Fig. 6 (scaled by the factor 10). Comparisons with climatological atlases established that the global distribution of vegetation is well represented by this analysis. Generally, as vegetation cover becomes more dense and healthy, the NDVI grid cell value becomes more positive (darker shades). The white areas represent grid cells where no clear-sky albedos were observed during any day of the month. As an interesting aside,
FIG. 6. Global distribution of grid cell monthly mean NDVI derived from daily CLEAR pixels from CLAVR-1 for Sep 1989. The data shown correspond to NDVI $\times 10$.

it is also possible to see aerosols over the oceans from the NDVI map (e.g., dark shaded area west of the Sahara). NDVI is negative over oceans because of the spectral redistribution of sunlight by molecular scattering, but it is less negative in regions of high aerosol concentrations because of the more spectrally uniform scattering by the larger aerosol particles. These NDVI patterns over oceans are very consistent with satellite derived distributions of aerosol optical thickness over oceans (Husar et al. 1997).

In order to separate these NDVI values into clusters representative of a few vegetation types, a frequency distribution of grid cell NDVI monthly mean values over land was constructed for September 1989, as shown in Fig. 7. Six vegetation types (clusters) have been identified from visual analysis of the modes in the distribution (separated by vertical lines): NDVI less than 0.0 (VT = 2); NDVI between 0.0 and 0.1 (VT = 3), 0.1 and 0.18 (VT = 4), 0.18 and 0.25 (VT = 5), 0.25 and 0.34 (VT = 6); and greater than 0.34 (VT = 7). With ocean considered as VT = 1, there are a total of seven distinct VT categories.

Thus, each latitude zone consists of grid cells which are classified as either water (VT = 1) or one of the six land types (VT = 2 through VT = 7). To construct ADMs from this collection of grid cells, daily clear-sky albedo and temperature grid cell means are sorted by latitude zone, VT, and VZA. They are then combined over an 8-day period in $10^\circ$ VZA bins, and albedo and temperature means and standard deviations are computed for each latitude zone, VT, and VZA bin. These means and standard deviations then provide ADMs representing the variation of albedo and temperature with VZA for the VT within any latitude zone, over the 8-day period.

Conceptually, this should result in angular models for 126 combinations of surface type and latitude (18 latitude intervals by 7 vegetation types). However, clear-sky radiance statistics typically only exist for grid cells from slightly more than half the globe. For September 1989, most of these grid cells are located at low and midlatitudes (cf. Figs. 2, Fig. 3), so CLAVR-3 can only be applied to those regions. Other sources of ADM data (e.g., theoretical models modified to match weather prediction model output, empirical statistics collected over longer time intervals) will have to be used to overcome this deficiency in CLAVR-3.

During the construction of ADMs for the descending
portion of orbits, the grid cells maintain their daytime derived VT classifications. It is important to maintain this separation in surface type, since each has a different emissivity and heat capacity that affects the angular dependence and magnitude of the observed brightness temperatures.

b. Characteristics of the clear-sky models

As an initial test of CLAVR-3, ADMs have been developed for channel-2 albedo ($A_2$) and channel-4 brightness temperature ($T_4$). Other channels or channel combinations could also be used. Examples for the first 8-day period, from 1–8 September 1989, are shown in Fig. 8. Figure 8a shows the behavior of the mean (boxes) and standard deviation (asterisk, plotted at the mean + 1σ for albedo, and mean − 1σ for temperature; these become the dynamic thresholds, as discussed in section 4) of $A_2$ and $T_4$ for different VT values for latitude interval 8 (10$^\circ$–20$^\circ$S) for ascending data. The three horizontal panels display the VZA dependence of sample size, albedo and temperature, respectively, for each of the seven VTs. Negative VZA corresponds to viewing the side of the orbit away from the sun. The ocean (VT = 1) has the largest sample size, and is best behaved in its statistical description of the angular dependence of the radiative quantities, that is, smooth angular variations with small standard deviations.

Over land in this latitude interval, all VT groups are represented, but the dominant population size comes from VT = 4 and VT = 5. The sample size (number of EA grid cells in an 8-day period) generally is smallest near nadir (VZA = 0$^\circ$) and largest at the limb (VZA ~ 60$^\circ$) due to grid cells having their largest cross-track angular width near nadir ($\Delta$VZA ~ 12$^\circ$), and smallest angular width at the limb ($\Delta$VZA ~ 4$^\circ$). Thus, up to three times as many EA grid cells may be viewed at the limb than at nadir, since AVHRR samples approximately in equal intervals of VZA.

Over ocean, the mean channel-2 albedos are in the range of 2% to 4%, with small standard deviations. A weak specular reflection effect is noticeable at VZA of 25$^\circ$ for this latitude in September, as is the expected limb brightening due to increased atmospheric scattering along the slant viewing path away from the sun (cf. Wagener et al. 1997). The channel-2 albedos are much larger over land. They initially decrease with increasing VT (increasing vegetation cover or greenness), but then tend to increase. Due to the very small sample sizes in VT = 2 (NDVI less than 0), variation with VZA is not well behaved. All other land VTs exhibit the expected bi-directional reflectance distribution function pattern with maximum albedo in the backscattered direction (Gutman et al. 1989), further confirming that the CLAVR-1 cloud mask is working well.

Ocean $T_4$ means are typically lower than land, and all except for the poorly sampled VT = 2 case exhibit the expected limb darkening (Suttles et al. 1988b), due
to water vapor and continuum absorption in the 10–12 micron region. Over land, mean temperatures tend to drop with increasing VT, consistent with progressively increasing density of vegetation that lowers canopy temperatures through evapotranspiration processes (Nemani and Running 1989).

For the descending part, Fig. 8b, the middle panel’s albedo values are zero because this latitude band is viewed in darkness. For every land VT type, the mean clear-sky temperatures are lower at night than in daytime (Fig. 8a), and this diurnal difference is largest for the least vegetated types (lowest VT), consistent with Gutman and Ignatov (1996). Also, limb darkening is less at night than in the day for land types due to the effects of the nonlinear dependence of the Planck emission function on temperature and how that affects the magnitude of radiation transmitted from the surface. The standard deviations are somewhat larger at night, which might indicate that the CLAVR-1 cloud mask is not working quite as well at night as in daytime (three infrared channels at night, without the benefit of the two daytime albedo channels).

4. Dynamic thresholds and the reclassification algorithm

The ADMs described in the above section provide a basis for creating dynamic cloud/no-cloud thresholds for the ninth day following the 8-day period from which the models were derived. These thresholds represent AVHRR observations of the most recent clear conditions available and are thus extremely useful for testing whether uncertain pixels from CLAVR-1 are CLEAR.

a. Dynamic cloud/no-cloud thresholds

Dynamic cloud/no-cloud thresholds developed here provide a statistical description of the behavior of $A_2$ and $T_4$ for varying VZA conditions for a given latitude interval and known vegetation type for September 1989. The albedo threshold is derived from the ADMs by adding one standard deviation to the mean channel-2 albedo ($\langle A_2 \rangle$) at each mean VZA value ($[A_2 = \langle A_2 \rangle + \sigma_{A_2}]$) (asterisks in Fig. 8). The channel-4 standard deviation is subtracted from the mean temperature ($[T_d = \langle T_4 \rangle - \sigma_{T_4}]$)
$\langle T_4 \rangle - \sigma_T$] to obtain the temperature threshold. This is done to allow for expected variation in these two radiative quantities due to the changing conditions observed over the 8-day period from which they were computed. The sign is chosen to allow that variation to be on the side of cloud contamination, but only by 1 standard deviation. A multiplier different from unity for the standard deviation may be chosen in future applications (e.g., SST retrieval) based on evidence of cloud contamination from validation against ground-truth data (e.g., buoy SSTs) or high resolution imagery.

b. Reclassification algorithm

Only RESTORED-CLEAR (excluding snow/ice) and MIXED pixels from CLAVR-1 are tested using these dynamic thresholds. Given the VZA value for an uncertain pixel from CLAVR-1, located in a grid cell with known VT and latitude interval, the prediction models permit determination of cloud/no-cloud thresholds for channel-2 albedo ($A_2$) and channel-4 temperature ($T_4$). The observed ($A_2$) and/or temperature ($T_4$) values for the uncertain pixels are compared with the dynamic threshold values ($A_d$ and $T_d$), and a decision is made as to whether the pixel is CLEAR. The value of the threshold at the observed VZA is computed by linearly interpolating between two neighboring VZA bins of the prediction model. Use of polynomial functions to fit the points has been considered and tested, but linear interpolation seems adequate.

CLAVR-1 uses cloud codes in its classification to identify the path followed by the pixel array through the algorithm (Stowe et al. 1999). There are over 40 codes, but they can be grouped into four primary classifications: CLEAR, RESTORED-CLEAR, MIXED, and CLOUDY. When CLAVR-1 is not able to make a classification, a cloud code of zero is assigned [note: RESTORED-CLEAR SNOW/ICE classification is assumed to be CLEAR in this initial test of CLAVR-3 since it is not possible to develop reliable prediction models for this classification because of the difficulty of separating snow and ice from some types of clouds in CLAVR-1 (cf. Stowe et al. 1999)].

The reclassification algorithm is outlined in Table 1. If during daytime the uncertain pixel albedo $A_2$ is less than the corresponding dynamic threshold $A_d$, and the pixel temperature $T_4$ is greater than $T_d$—that is, the pixel satisfies both threshold tests—the pixel is reassigned to CLEAR. If either one of the two tests or both tests are failed, the pixel is either retained as MIXED or reclas-
Dynamic cloud/no-cloud thresholds

\[ A(\text{LA}, \text{VT}, \text{VZA}) = < A_2 > + \sigma_i \]

\[ T(\text{LA}, \text{VT}, \text{VZA}) = < T_4 > - \sigma_i \]

\( \text{LA} = \) lat index

\( \text{VT} = \) vegetation type

\( \text{VZA} = \) view zenith angle

RESTORED-CLEAR (cc = 13-16; 30)—daytime

If \( A_2 \) and \( T_4 \), then CLEAR, else MIXED (cc = 14-16) or CLOUDY (cc = 13)

RESTORED-CLEAR (cc = 60-69)—nighttime

If \( T_4 \), then CLEAR, else MIXED

MIXED (cc = 17-27)—daytime

If \( A_2 \) and \( T_4 \), then CLEAR, else MIXED

MIXED (cc = 17-26; 90-99)—nighttime

If \( T_4 \), then CLEAR, else MIXED

No ADM: RESTORED-CLEAR \( \rightarrow \) CLEAR MIXED \( \rightarrow \) MIXED

5. Evaluation of CLAVR-3 performance

The following two subsections describe various qualitative and quantitative tests applied to the output of the CLAVR-3 algorithm, first for data from ascending parts of orbits and second for descending parts.

a. Ascending (daytime) data

1) Quantity of CLEAR pixels after CLAVR-3 (ascending)

Nearly 50% of pixels classified by CLAVR-1 from the ascending orbits on day 89 252 were in the uncertain (MIXED and RESTORED-CLEAR) category. Of the 18.2 million uncertain pixels, albedo and temperature threshold tests were both satisfied for 17.8% of these pixels, that is, they were reclassified as CLEAR. About 73.7% failed either one or both tests and were not reclassified as CLEAR. There were 8.5% that could not be reclassified because a threshold could not be determined due to lack of an ADM. A total of 3.2 million pixels were added to the CLEAR population after reclassification with CLAVR-3, an increase of 75% over the CLAVR-1 CLEAR sample size.

2) Quantity of CLEAR pixels in grid cells after CLAVR-3 (ascending)

The change in CLEAR pixel population within equal area grid cells due to the addition of CLEAR pixels from CLAVR-3 is shown globally in Fig. 9. The increase is as large as 600 and the largest increase is along the satellite subtrack due to pixel area enlargement with increasing viewing angle (~6 times larger at the limb than at nadir). Since uncertain pixels constitute about 50% of the total, and most of these are MIXED, there is no dependence on position across the orbit, that is, dependence on solar geometry, as would be the case if RESTORED-CLEAR dominated. Higher populations generally fall in the fair weather regions and in desert areas, consistent with these regions typically having fewer clouds.

The white area (no CLEAR pixels in grid cell) in Fig. 9 is noticeably smaller than in Fig. 1, indicating that grid cells containing CLEAR observations have been recovered with CLAVR-3 and added to those available from CLAVR-1. Specifically, over oceans, CLEAR grid cells increased by 23.8%, while over land only by 12.5%. This again is consistent with the predominate uncertain classification being MIXED, as these pixels tend to be more frequently detected over oceans than over land (cf. Stowe et al. 1999). Globally, CLAVR-3 results in an increase in CLEAR grid cells of 19.3% (4340 CLEAR grid cells added). This increased coverage will result in better daily characterization of remotely sensed Earth surface parameters, that is, SST, Aerosol Optical Thickness (AOT), and NDVI, if derived from mean CLEAR grid cell radiances. This is the approach used in the AVHRR Pathfinder Atmosphere project (Stowe and Jacobowitz 1997).

3) Quality of CLAVR-3 grid cell CLEAR pixel mean albedos and temperatures (ascending)

The increased pixel and grid cell populations resulting from CLAVR-3 are only useful if the resulting CLEAR
FEBRUARY 2001

FIG. 9. Difference in grid cell CLEAR pixel population between CLAVR-3 and CLAVR-1, shown in lat/long projection for ascending orbits, day 89 252.

radiances (albedos and temperatures) are of comparable quality to those available from CLAVR-1. To visualize the impact of CLAVR-3 on the quality of CLEAR albedo, histograms of the number of grid cells whose mean CLEAR pixel albedo ranges from 0% and 49% in 1% intervals, both before and after the use of dynamic thresholds are computed separately for ocean and land (Fig. 10). CLEAR ocean grid cell mean albedos lie in the range of 1%–13%. The distribution appears to be slightly shifted to higher albedos as a result of CLAVR-3, whereas one would ideally expect to see the distribution change uniformly with respect to albedo. This ideal is expected since the CLAVR-3 albedo dynamic thresholds (cf. Fig. 8) are well below, and the temperature thresholds well above the values used in CLAVR-1 at most latitudes for these respective tests [CLAVR-1 ocean cloud/no-cloud channel-2 albedo threshold is 30%; ocean channel-4 temperature threshold is 271 K]. This greatly limits the bright (cold) range over which the additional CLAVR-3 CLEAR albedos (temperatures) are added to the grid cell mean values from CLAVR-1. Nothing in the CLAVR-3 algorithm should make it any less likely for reclassified uncertain pixels to be relatively dark (warm) with respect to the mean value observed from CLAVR-1. The observed shift in albedo would be expected if a majority of the reclassified CLEAR pixels were the result of specular reflection (RESTORED-CLEAR) pixels from CLAVR-1. However, since most uncertain pixels are MIXED, it is not likely to be the principal cause of this shift. Rather, it is more likely that MIXED pixels are being reclassified to CLEAR with slightly elevated albedos, either due to their adjacency with pixels containing cloud cover or due to their containing some partial cloud cover themselves. To resolve which of these possibilities is the cause will require further study with high resolution satellite cloud images or applying the CLEAR pixel (or grid cell average) radiances to the retrieval of parameters verifiable with ground truth (e.g., SST, channel-1 aerosol optical thickness).

The distribution of grid cell mean CLEAR albedo over land indicates that CLAVR-3 is yielding CLEAR pixel albedo ranges that are very similar to those from CLAVR-1, CLAVR-3 providing about the same number of additional grid cells at each albedo interval. The most frequently occurring albedo over land is about 12%. The distributions are nearly identical for albedos above 25%.
Thus, CLAVR-3 does not seem to be introducing noticeable cloud contamination of land in daytime.

Histograms of grid cell mean CLEAR temperature (from channel 4 in 1 K intervals) for ascending (daytime) orbital segments from CLAVR-1 and CLAVR-3 on day 89 252 are shown separately for land and ocean in Fig. 11. The grid cell average CLEAR temperatures from CLAVR-1 and -3 are distributed similarly to the CLEAR albedo distributions in Fig. 10, but with the temperature redistribution over ocean skewed towards the cold side, whereas albedo was skewed towards the bright side. This is consistent with the possibility that CLAVR-3 may be introducing some cloud contamination over oceans.

Over some land areas (particularly mountainous regions), other analyses have shown that CLAVR-3 increases both grid cell mean CLEAR temperature and albedo, which supports the conclusion that CLEAR pixels from CLAVR-3 are not noticeably contaminated by cloud over land. This behavior is consistent with the interpretation that pixels in these grid cells were CLEAR but were classified as MIXED by CLAVR-1, due to horizontal inhomogeneities in surface albedo and/or temperature. More horizontally homogenous terrain exhibit small changes in temperature between the two algorithms.

b. Descending (nighttime) data

1) Quantity of CLEAR pixels after CLAVR-3 (descending)

During the descending portion of orbits, the dynamic thresholds are primarily derived from the temperature ADMs, as only a small fraction of uncertain pixels are observed during daytime conditions in summer polar latitudes. Thus, misclassifications may be more likely at night, as only one channel is being used for the determination of CLEAR from the uncertain pixels of CLAVR-1. Of the 14.5 million uncertain pixels during the descending orbit segments of 89 252, 4.9 million pixels (33.5%) were reclassified as CLEAR. This is almost double the percent reclassified as CLEAR for ascending data (17.8%). The number of descending CLEAR pixels after CLAVR-3 now represents 28.2% of the total pixel population, whereas from CLAVR-1, CLEAR pixels only comprised 14.5% of the total. This corresponds to a 94.9% increase over the CLAVR-1 CLEAR sample size for descending data. About the same number of uncertain pixels as occurred in the ascending case could not be reclassified.

2) Quantity of CLEAR pixels in grid cells after CLAVR-3 (descending)

Figure 12 shows grid cell CLEAR pixel population differences between CLAVR-3 and CLAVR-1 during
the descending portion of orbits on day 89-252. As one would expect, coverage by CLEAR grid cells is more extensive after CLAVR-3, as can be seen by comparison with Fig. 3, which shows the spatial extent of grid cell mean CLEAR temperatures from CLAVR-1. The largest increases in grid cells occur in the Southern Hemisphere, particularly over oceans, which is different from the ascending part, where the largest increases occurred in the Northern Hemisphere. The pixel population gains within grid cells containing CLEAR pixels in CLAVR-1 appears to have a similar pattern to ascending, with maximum increases occurring along the subsatellite track, for the same reasons (that most uncertain pixels are MIXED and therefore should have distributions peaking where the grid cells are largest relative to the size of the pixels). There are some relatively large increases in the Arctic region for descending that were not present for ascending. This perhaps results from smaller temperature differences between cloud and clear surfaces in this region, which also raises concerns about possible cloud contamination being introduced by CLAVR-3.

3) Quality of CLAVR-3 Grid Cell CLEAR Pixel Mean Temperatures (Descending)

As for the ascending case, histograms of grid cell average CLEAR pixel temperature in 1K intervals are shown for ocean and land in Fig. 13 for the descending data, separately for CLAVR-1 and CLAVR-3. As expected, there are more CLEAR grid cells in CLAVR-3 at most temperatures. The histograms are more similar between ocean and land for descending (nighttime) orbits as compared to ascending orbits, where elevated daytime land temperatures were evident. The change in the ocean histogram due to CLAVR-3 is similar to the ascending change but shows a more pronounced shift to lower temperatures. Over land, contrary to the behavior of histograms for ascending data, CLAVR-3 appears to noticeably shift temperatures to lower values, again consistent with cloud contamination. Further research is required to determine how much of this shift is indeed cloud contamination, and for those grid cells where it is a problem, to properly adjust and perhaps change the thresholds used in the CLAVR-3 tests to remove it.

6. Conclusions and future work

The Phase-3 Clouds from AVHRR (CLAVR-3) algorithm uses angular distribution models (ADMs) for albedo and temperature, derived from CLEAR pixel classifications from CLAVR-1 over an 8-day period, to search for additional CLEAR pixels from the “uncertain” (RESTORED-CLEAR and MIXED) pixel classifications in CLAVR-1 for the next day after the 8-day period. These models provide dynamic, surface type, and angularly dependent albedo and temperature cloud/
no-cloud thresholds. Use of these thresholds in CLAVR-3 replaces the universal, static thresholds used in CLAVR-1. This work demonstrates that the current CLAVR-3 algorithm’s dynamic thresholds enhance the clear-sky sample sizes both over land and ocean.

For one sample day, 9 September 1989, dynamic thresholds resulted in a 75% increase in CLEAR pixels during the ascending (mostly daytime) part and a 95% increase during the descending (mostly nighttime) part of that day’s orbits. This increase in CLEAR pixels is particularly important for ocean and land surface applications, where reliable identification of clear-sky pixels is necessary for deriving surface parameters like sea surface temperature (SST), Normalized Difference Vegetation Index (NDVI), and aerosol optical thickness (AOT) on a daily basis. As a result of CLAVR-3, 20.4% of ascending and 28.2% of descending pixels on this sample day are being classified as CLEAR compared to 11.6% and 14.5% from CLAVR-1, respectively.

Histogram analysis suggests that part of this enhancement may be the result of CLAVR-3 allowing some cloud contaminated pixels to be classified as CLEAR. If further analysis supports this hypothesis, remedies could be found through choosing AVHRR channels other than channels 2 and 4 for the threshold tests or other scaling factors for the standard deviation about the mean ADMS used in computing the thresholds. The ultimate test of the need to correct for possible cloud contamination will depend upon the application of the CLAVR-3 CLEAR pixels to derivation of surface and atmospheric parameters with ground truth (e.g., SST and AOT), to determine if the amount of cloud contamination significantly effects the accuracy of these derived parameters.

The current dynamic thresholds have discontinuities at 10° latitudinal boundaries. Artificial effects caused by these discontinuities can be removed through the use of smoothly changing dynamic thresholds from one latitude interval to the next. Angular dependence models cannot be constructed in some latitude intervals and for some vegetation types with an 8-day composite CLEAR dataset. A process of accumulating data over longer periods or utilizing auxiliary data, such as output from weather prediction models, needs to be evaluated to supplement the current procedure. Additionally, the value of using surface type information derived independently from AVHRR mean NDVI values should be assessed. Finally, independent snow/ice datasets should be used to remove the ambiguity between cloud and
CLEAR snow/ice in CLAVR-1, because only RESTORED-CLEAR pixels emerge from CLAVR-1 for this condition, from which angular models for snow/ice cannot be reliably constructed.

Acknowledgments. The authors acknowledge helpful discussions on the methods and analyses used in this work with members of the CLAVR science team. Particular recognition is due to Gang Luo for graphical support, to Charlie Dean for a careful review of the revised manuscript, and to Paul Davis, Ken Campana, Dan Tarpley, and Bill Pichel for critical reviews of earlier versions. We also thank Doug Love for coding support during the course of this project and Sreedevi Kumar for her patience in typing multiple revisions of the manuscript. We appreciate the reviewers and the editor of this journal, who made many valuable suggestions which have been incorporated into this manuscript. This work was supported with funds from the NOAA Climate and Global Change Program administered by Drs. William Murray and Rex Fleming.

REFERENCES


