1. INTRODUCTION

The National Aeronautics and Space Administration (NASA) Earth Observing System (EOS) is a global observation system that launched the satellites Terra and Aqua to study climate and climate change (Parkinson 2003). The sensor suites on these satellites have other applications, as well, including retrievals for diagnostics and weather forecasting. In this research, the Advanced Infrared Sounder (AIRS) flown on the Aqua satellite is used to develop a cloud mask to detect the presence of clear and cloudy pixels.

There are several key factors toward the motivation of the development of an AIRS cloud mask. One key factor is that the method of cloud detection is critical to a number of applications that require knowledge of clear/cloudy pixels in AIRS data, such as the development of accurate profiles of atmospheric variables and data assimilation. ‘Clear’ radiances are needed to obtain accurate retrievals, whereas if a ‘cloudy’ radiance is used, unreasonable profiles may be produced. A second factor is that the AIRS science team does not currently produce an AIRS cloud mask; it primarily works on cloud clearing to produce cloud-cleared radiances, which are used in retrievals of temperature and moisture (Susskind et. al. 2003).

2. BACKGROUND

2.1 Previous Cloud Detection Approaches

Several cloud masks are routinely produced from GOES and MODIS satellite data (Jedlovec and Laws 2003; Ackerman et. al. 2002; Haines et. al. 2004). MODIS is unique in its spectral and spatial resolutions. The GOES Imager has limited spectral and spatial resolutions, consisting of only five spectral bands with varying spatial resolutions between 1 km and 4 km (Menzel and Purdom 1994). The cloud mask algorithm developed for GOES at the Global Hydrology and Climate Center (GHCC) consists of a combination of spectral and spatial tests. Before implementing any of the tests, an hourly difference image is produced from the 10.7 and 3.9-µm channels. This difference image is then integrated into four different tests, the first of which is an adjacent pixel test to detect cloud edges. Then the one-dimensional spatial variability of the pixels is analyzed to determine the ‘body’ of the cloud. Two other tests are then used if the first two tests fail, a minimum difference test and an IR threshold test, in order to determine whether a pixel is deemed cloudy (Jedlovec and Laws 2003; Haines et. al. 2004).

MODIS data consists of 36 bands with a spatial resolution of 250 m in two of the visible bands, 500 m resolution in five visible and near-IR bands, and 1 km resolution in the remaining bands (Ackerman et. al. 2002). Due to the larger amount of spectral information available from MODIS, as compared to GOES, the MODIS cloud mask algorithm produced by the EOS Science Team consists primarily of spectral tests using 19 of the visible and infrared bands (Ackerman et. al. 2002). These tests range from very simple threshold tests, such as the BT11 (where BT is brightness temperature) threshold test where a brightness temperature around 11 µm above or below a given threshold would be deemed clear or cloudy, to tests on the differences between brightness temperatures of two different bands. An example of the latter is the BT11-BT3.9 test, which can be used to detect partial or thin clouds at night based on longwave emission or low-level water clouds during the day based on reflection of solar energy. Another difference test used by the EOS science team is the BT7.3-BT11 test used primarily at night over land to detect mid- to high-level clouds in MODIS data. Under clear-sky conditions BT7.3 is sensitive to temperature and moisture in middle levels of the atmosphere while BT11 measures radiation mainly from warmer surfaces. The tests mentioned above, and others, are used to classify pixels as one of four types: cloudy, uncertain clear, probably clear, and clear. The tests used to develop the cloud mask combined with high spatial resolution should make the EOS MODIS cloud mask algorithm a very robust global product.

2.2 AIRS

AIRS uses advances in infrared technology to provide a much higher spectral resolution than previous instruments. It operates over the range of 3.7 to 15.4 µm and provides 2378 spectral bands with a nominal spectral resolution of $\lambda/\Delta\lambda = 1200$. The spatial resolution of each pixel is approximately 15 km (Aumann et. al. 2003). Figure 1 shows a single channel image (11.1 µm brightness temperature) from 12 March 2003 as an example of the resolution of AIRS and coverage for a five-minute portion of the orbit. This image conveys brightness temperatures present in the window region channel, where the effect of the atmosphere is minimal. Therefore, without the presence of clouds, the brightness temperatures are relatively warm approximating the land surface temperature and sea surface temperature values.

2.3 AIRS Data Format and Processing

AIRS Level 1B radiance data (Version 3.0.10) was obtained in HDF format. The HDF data files consist of various bits of information including

---

A CLOUD MASK FOR AIRS

Nicole J. Brubaker¹, Gary Jedlovec²

¹ University of Alabama in Huntsville, Huntsville, AL
² Global Hydrology and Climate Center, NASA/MSFC, Huntsville, AL

---
calibrated radiances assigned to each wave number in the spectrum, geolocation information, etc. Calibrated radiances were converted to brightness temperature with the inverse of the Planck function and bad channels were eliminated by way of information contained in the HDF file. Further information about AIRS specifications and calibration can be found in Pagano et. al. (2003).

2.4 Spectral Signatures of Clouds

After converting the radiance data to the binary format and eliminating bad data, spectral plots corresponding to clear and cloudy regions were created to better understand the spectral signatures of clouds in AIRS (Figures 2 and 3). From the spectral plots, several key characteristics can be seen about the infrared spectral signatures of clouds. When comparing clear and cloudy spectra, as in Figures 2(a & b) and 3(a & b), one will notice that brightness temperatures in the 10-12 µm window region are much warmer in a clear spectrum (~290 K) than in a cloudy spectrum (~240-260 K). Because temperature decreases with height in the troposphere, it’s only natural that clouds would be colder than the Earth’s surface for most cases. Also, little to no water vapor absorption is present above the clouds in the cloudy spectrum. Water vapor concentration is highest near the Earth’s surface, which would indicate why a clear spectrum would have much more water vapor absorption (spikes in the radiances in the 10-12 µm region) than a cloudy spectrum.

Another feature of the AIRS spectra in the figures is the difference in brightness temperature between the 11 µm and 3.9 µm spectral regions. When comparing a cloudy and clear spectrum, the 11 µm brightness temperature is much lower than the 3.9 µm brightness temperature in the cloudy spectrum (Figure 2b), while the clear spectrum (Figure 2a) holds both values closer to one another. Therefore, a more negative difference would indicate the presence of cloud, while a less negative/positive value would indicate clear sky regions.

A key feature that stands out in cloudy AIRS spectra is the increased slope in the window region of cloudy spectra due to varying cloud emissivities, as well as select other regions such as the 3.7-4.1 µm region (c and d in Figures 2 and 3). Although the 10-12 µm region seems to have a linear feature over the region, the 3.7-4.1 µm region seems to have several different slopes (i.e. 3.93-3.96 µm, 3.96-3.98 µm, and 3.85-3.88 µm). This feature is the premise for the slope test approach to the AIRS cloud mask algorithm discussed later in this paper. This slope would not be as distinguishable (if at all) in a MODIS or GOES spectrum due to the limited number of bands available to work with. Because only AIRS possesses this feature, it may prove to be a powerful tool in many future studies.

3. AIRS CLOUD DETECTION METHODOLOGY

3.1 Tests of AIRS Cloud Mask Algorithm

The 15 km spatial resolution of AIRS pales in comparison to the spatial resolutions of GOES and MODIS, therefore some of the spatial techniques used by GOES and MODIS would not be preferential for the AIRS cloud mask algorithm. The use of difference imaging combined with an adjacent pixel test, as in the GOES cloud mask algorithm (Jedlovec and Laws 2003), would not be worthwhile for AIRS due to its coarser spatial resolution; cloud edges would not be as well defined by AIRS as they are by GOES. The AIRS cloud mask algorithm developed here will expand upon the MODIS tests discussed in section 2.1 (BT 11, BT11-BT3.9, BT7.3-BT11) using Level 1B calibrated radiances. With MODIS, the BT7.3-BT11 test is only used at night, but the AIRS cloud mask algorithm will attempt to use it for both day and night.

For the BT11 threshold test in the AIRS cloud mask algorithm, the BT11 value is computed by taking the average brightness temperature over the range of approximately 10.95 to 11.08 µm. This averaged value is then compared to a predetermined threshold value. High clouds with large emissivities will have temperatures colder than this threshold; although clouds with lower emissivities with brightness temperatures similar to land surface temperatures may be missed by the BT11 test.

For the BT11-BT3.9 test, the average brightness temperature at 11 µm and the average brightness temperature between 3.85 and 3.95 µm are used, and the same is used for the BT7.3-BT11 test, where BT7.3 is determined by calculating the average brightness temperature between 7.25 and 7.35 µm.

Due to the unique enhanced spectral resolution of AIRS and information obtained from cloudy and clear spectra (as discussed in section 2.4), a slope test approach was developed to take advantage of the spectral resolution. A simple linear regression analysis was used to determine the slope of a particular wavelength region. The emphasis was on parts of the 3-4 µm region because various subsections of this region, as seen in Figure 4, hold unique slope
Figure 2. AIRS spectra for clear (a) and cloudy (b) regions during daytime pass for 23 April, 2003 at approximately 19:30 UTC. (c) and (d) are expanded regions (3.7-4.1 µm) of (a) and (b).

Figure 3. AIRS spectra for clear (a) and cloudy (b) regions during nighttime pass for 23 April, 2003 at approximately 8:30 UTC. (c) and (d) are expanded regions (3.7-4.1 µm) of (a) and (b).
characteristics and because the 10-12 $\mu$m window region would be complicated by the presence of water vapor absorption. The subsections included in this investigation were the regions of 3.96-3.98 $\mu$m, 3.93-3.96 $\mu$m, and 3.85-3.88 $\mu$m. Preliminary tests indicated that the 3.85-3.88 $\mu$m region provided the most consistent slopes to differentiate cloudy from clear pixels. The 3.96-3.98 $\mu$m and 3.93-3.96 $\mu$m slopes did not seem to be the optimal ranges for the slope test at this time largely due to great overdetection along the edges of the AIRS pass, although they did perform quite well in detecting smaller scale clouds. Various 'subregions' of the 10-12 $\mu$m window region were also investigated due to the obvious change in slope as seen in different AIRS spectra (see Figures 2 and 3), but results at this time do not seem favorable. Based on results from the various slope tests, the 3.85-3.88 $\mu$m slope test was incorporated into the AIRS cloud mask algorithm. The effect of sunglint on these slope values may be problematic.

3.2 Determination of Thresholds

Thresholds are determined based on scatter plots of the test data and preliminary cloud information from the MODIS cloud mask (Ackerman et. al. 2002). Clustering was made apparent in the majority of scatter plots, which were then examined to determine whether the clusters represented clear or cloudy regions. Labeling the region as ‘clear’ or ‘cloudy’ was based on the previous discussion of the behavior of the AIRS spectra in section 2.4. Finding a median value between the two clusters indicated the possibility of a useful threshold. These thresholds were then used in the AIRS cloud mask algorithm to produce a preliminary cloud mask for several training data sets for each season (spring, summer, winter, fall). Initially, only one test was used at a time to produce the cloud mask, which was then compared quantitatively to a “ground truth” (MODIS cloud mask) image to obtain statistics that would point to the ‘optimal’ threshold of each test for each training data set. Different ‘optimal’ thresholds were determined for daytime passes. An average ‘optimal’ threshold was then calculated for each season to determine a threshold closer to what would be used in the AIRS cloud mask algorithm.

At this point, all cloud mask tests were then used to create an AIRS cloud mask using the thresholds determined previously. To understand the performance of the various tests, colors were assigned to pixels in the cloud mask image to provide information on which test dominated per season, which test ‘overdetected’ or ‘underdetected’ more often when compared to the MODIS-derived AIRS mask, and to show what type of cloud each test was most likely to detect (Figure 5). In the figure, white represents what pixels were solely identified by the BT$_{11}$ test, while red represents the BT$_{11}$-BT$_{3.9}$ test, blue represents the BT$_{7.3}$-BT$_{11}$ test, and orange represents the 3.85-3.88 $\mu$m slope test. All other colors in the image represent various combinations of the four tests used in the AIRS cloud mask algorithm. As the red color would indicate, the BT$_{11}$-BT$_{3.9}$ test is either overdetecting clouds on the left-hand side of the image or it is detecting lower, warmer clouds that were not detected by the other tests. The gray color in the image represents the pixels that were

![Figure 4. AIRS spectrum (3.7-4.1 $\mu$m region) for 23 April, 2003 at 19:30 UTC.](image)

![Figure 5. Color cloud mask product for 1 August, 2003. BT$_{11}$ test is white, BT$_{11}$-BT$_{3.9}$ is red, BT$_{7.3}$-BT$_{11}$ is blue, and 3.85-3.88 $\mu$m slope is orange. All other colors represent different combined tests.](image)
detected as cloudy by all the tests in the algorithm, and one could therefore conclude that these clouds are probably thicker, colder clouds that have a signature detectable by all tests. Figure 6 shows an example of an intercomparison between the AIRS cloud mask and the MODIS-derived AIRS mask used as ‘ground truth’. There were considerably more clouds detected by the MODIS cloud mask (‘ground truth’ in the figure) than the AIRS cloud mask. Again, this could be due to a number of different factors such as MODIS overdetecting clouds, the AIRS cloud detection algorithm underdetecting clouds, or a mixture of both. In this case, the particular test used in the AIRS mask tended to underdetect cloud edges, but overdetect very little according to the ‘ground truth’. Table 1 provides a list of the preliminary thresholds that will be used in the AIRS cloud mask algorithm. To further understand the behavior of the AIRS cloud mask algorithm and its results, the algorithm was run for several different cases for each season of the year. The results are presented below.

Table 1. Table of threshold values of each test for day and night passes.

<table>
<thead>
<tr>
<th>Test</th>
<th>Daytime Pass</th>
<th>Nighttime Pass</th>
</tr>
</thead>
<tbody>
<tr>
<td>BT11</td>
<td>&lt; 289 K</td>
<td>&lt; 268 K</td>
</tr>
<tr>
<td>BT11-BT3.9</td>
<td>&lt; -9 K</td>
<td>&lt; -6 K</td>
</tr>
<tr>
<td>BT7.3-BT11</td>
<td>&lt; -27 K</td>
<td>&lt; -11 K</td>
</tr>
<tr>
<td>3.85-3.88 slope</td>
<td>&lt; 0 K µm⁻¹</td>
<td>&lt; 36 K µm⁻¹</td>
</tr>
</tbody>
</table>

Table 6. AIRS cloud mask (top), MODIS-derived AIRS mask (middle), and intercomparison of the two (bottom) for 23 April, 2003. White represents cloud, black represents clear. In bottom image, red represents underdetection and green represents overdetection.

4. VALIDATION AND RESULTS

Initial validation of the AIRS cloud mask was performed qualitatively using a MODIS-derived AIRS cloud mask as “truth”. The MODIS-derived AIRS mask was created using the EOS MODIS cloud mask product (MOD35) (Ackerman et. al. 2002). MODIS was selected for validation because it is located on the same platform as AIRS, therefore having similar geometries and scan times. The domain of the MODIS cloud mask product was then converted into the AIRS domain (90 columns, 135 rows; approximately 12150 pixels). The AIRS L1B latitude and longitude information was used, as well as a ‘clear test’ and threshold. The ‘clear test’ was essentially a value that determined whether to use MODIS cloudy, cloudy and uncertain clear, or cloudy, uncertain clear, and probably clear as a cloudy pixel in the respective AIRS pixels. The threshold for the program represented the percentage (0-100) of MODIS pixels required to be cloudy for the corresponding AIRS pixel to be labeled as cloudy. For the validation of the AIRS mask, a ‘clear test’ value representing ‘cloudy and uncertain clear’ was used along with a threshold of 33%. Therefore, there had to be at least 33% of MODIS pixels that were cloudy or uncertain clear in a particular AIRS pixel for that AIRS pixel to be labeled as cloudy. These values seem reasonable based on the spatial resolution of the data sets.
Although MODIS was selected to be used as ‘ground truth’, one has to keep in mind that MODIS does not perform perfectly, which may possibly skew results of the AIRS mask. The performance of the EOS MODIS cloud algorithm varies with about an 80-95 percent efficiency for the continental US and surrounding ocean regions (Haines et. al. 2004). The EOS MODIS cloud algorithm tends to overdetect for areas of snow cover, in which it would classify pixels as ‘cloud’ instead of ‘clear’. Its accuracy is also reduced at night, over the oceans, and in areas of sunglint.

For the preliminary validation of the AIRS cloud mask, fifteen to twenty days were selected from each season of the year (spring, summer, fall, and winter). Within the case study days, data sets were selected that corresponded to daytime passes within a selected region over the southwest United States shown in Figure 7. For each of the days, a MODIS-derived AIRS mask (for validation) and the AIRS mask were created, and then compared to determine the overall performance of the AIRS mask.

Figure 7. Approximate case study domain for AIRS cloud mask.

Figure 8 shows the performance of the AIRS cloud detection algorithm for the case study days for each season. There is underdetection and over-detection for each season, with an average over-detection of approximately 10 percent and an average under-detection of approximately 10 percent, which could be the result of several different reasons. The AIRS mask thresholds could be a little too conservative, the MODIS-derived cloud mask is not as accurate as it should be (or is believed to be), or particular tests are performing in ways for which thresholds do not compensate.

An evaluation of the individual days for each season indicated that in some instances, there is great variation in the accuracy of the AIRS mask, with some days averaging above 80 percent correct, and the following day having an accuracy between 20 and 50 percent. This phenomenon still needs to be investigated, but several factors could be causing the drastic change in accuracy. The view of the AIRS instrument might be different from one time to the next (i.e. further south for an earlier pass and more northerly for the next), a particular AIRS file may have been corrupted, or thresholds for particular regions do not perform as well as in others. The season with the largest variation of accuracy was fall, with winter coming in next, although its changes in accuracy were not as drastic. Although fall and winter had noticeable changes in accuracy, the spring and summer seasons were fairly constant in their accuracies, with very little variability from day to day, with an average accuracy of approximately 83 percent. Both spring and summer had some over- and underdetection, but the spring season had the least of both with less than 10 percent. There was a slight decrease in accuracy for the fall and winter seasons, and one feature of interest was the tendency for the winter season to have a much larger over-detection than under-detection. Although it is still not precisely clear as to why there is much under- and over-detection, several characteristics about each test in the AIRS cloud mask algorithm have been discovered, of which could possibly contribute to the overall over- and under-detection. The BT_{11} test performs well when detecting cloud edges, but tends to over-detect for some areas of water and under-detects smaller scale and warmer clouds. The under-detection of warmer clouds could be a result of the threshold used. If the threshold is too cool, then warm clouds may be determined as clear. As for the BT_{11}-BT_{3.9} test, scattered and smaller scale clouds are detected better than with the BT_{11} test, although it slightly over-detects cloud edges in most cases. A drawback to the BT_{11}-BT_{3.9} test is the effect of sunglint, which is also sometimes present in the 3.85-3.88 µm slope test. The BT_{7.3}-BT_{11} test does well in the detection of high, cold clouds, as does the BT_{11} test, although the BT_{7.3}-BT_{11} also tends to under-detect clouds in cases where the BT_{11} test detects the same clouds. The 3.85-3.88 µm slope test seems to over-detect around nadir, but still does fairly well at detecting smaller scale scattered clouds. Although each test contributed to the AIRS cloud mask algorithm, it seems as if the BT_{11} and BT_{11}-BT_{3.9} tests dominated. As a result, and keeping in mind that these two tests over- and under-detect pixels in various ways, may be

Figure 8. Accuracy of the AIRS cloud mask.
5. CONCLUSIONS

The AIRS cloud mask algorithm currently being developed has been compared to the EOS MODIS cloud mask used as ‘ground-truth’. The AIRS cloud detection algorithm has the potential to be much more robust in that it makes use of the hyperspectral capabilities of the AIRS instrument as compared to MODIS, which uses only 19 of its 36 spectral bands in its cloud mask algorithm. The AIRS cloud mask algorithm expands upon a few tests already used by the MODIS cloud mask algorithm and also takes advantage of its hyperspectral capabilities by using slopes present in the AIRS spectrum to develop slope tests to overcome its coarse spatial resolution.

The results presented are from initial validation of daytime AIRS passes. This validation is subject to error as a result of the varying accuracy of the MODIS cloud mask algorithm depending on the region of interest. The results obtained from case study days for each season are promising, but further investigation has to be done to understand varying accuracies in some seasons and the percentage of over- and underdetection present for each season.

Although there are still some issues with the overdetection and underdetection of the AIRS cloud mask algorithm, the percentage of pixels correct (approximately 81% overall) is promising. When compared to the accuracy of the MODIS cloud mask algorithm, this accuracy is even more promising.

6. ONGOING AND FUTURE WORK

Continuing effort will be put towards the validation of the daytime passes with possible changes in the ‘clear test’ and threshold used to develop the MODIS-derived AIRS mask used for validation. A more ‘manual’ validation approach will be implemented for nighttime passes due to the greatly reduced accuracy of the MODIS cloud mask at night (Haines et al. 2004). The approach will be applied to nighttime AIRS data with adjustments to the test thresholds. This manual validation may also be used as a supplemental validation of the daytime passes. A perceived problem with the daytime passes is the presence of sunglint. This still seems to be a problem in other already developed cloud masks, as well, but continued effort will be put into filtering out the sunglint from the AIRS cloud mask.

7. ACKNOWLEDGEMENTS

This research was partially funded by NASA’s Earth Science Enterprise through the Short-term Prediction and Research Transition (SPoRT) Center at Marshall Space Flight Center and by the Office of Naval Research (ONR) as part of the Multidisciplinary University Research Initiative (MURI) program through a sub-contract (G066835) to the University of Alabama in Huntsville from the University of Wisconsin – Madison.

8. REFERENCES


