JOURNAL OF GEOPHYSICAL RESEARCH, VOL. 108, NO. D23, 4742, doi:10.1029/2003JD003554, 2003

An automated cloud detection method for daily NOAA 16 advanced very high resolution radiometer data over Texas and Mexico

Pei-yu Chen
Blackland Research and Extension Center, Texas Agricultural Experiment Station, Temple, Texas, USA

Raghavan Srinivasan
Spatial Sciences Laboratory, Department of Forest Science, Texas A&M University, College Station, Texas, USA

Gunar Fedosejevs
Canada Center for Remote Sensing, Ottawa, Ontario, Canada

Received 3 March 2003; revised 7 August 2003; accepted 15 August 2003; published 13 December 2003.

[1] The advanced very high resolution radiometer (AVHRR) data acquired from the National Oceanic and Atmospheric Administration (NOAA) satellites have been widely applied to a variety of environmental research. A single AVHRR scene is seldom completely cloud-free. Maximum value compositing (MVC) to create a single image from multiple orbits and dates has become the most valuable method to minimize cloud contamination. Composite images are not absolutely cloud-free. Postcomposite cloud screening of the composite aggregates was developed to overcome the residual cloud contamination problem, but this is not possible for real-time delivery of composite data or not suitable for compositing based on AVHRR data from multiple NOAA satellites. Another approach is to detect and remove cloud-contaminated pixels from daily AVHRR scenes prior to applying the MVC method to provide real-time composite images. This study developed an automated cloud detection method for daily NOAA 16 AVHRR scenes over the state of Texas and Mexico. The accuracy of the cloud detection algorithm was greater than 93% based on a random test sample from 36 images. Unidentified cloud shadow pixels as well as misidentified barren land pixels and water pixels contributed to more than 5% of the accumulated errors. The error from misidentification of water pixels can be reduced by assigning different threshold values for channel 4 brightness temperature according to the geographical latitude of the data. The resulting daily cloud-free AVHRR data can be used to construct short-time period composite images valuable for detecting subtle but critical environment changes. In addition, compositing methods other than MVC, such as multidate averaging or minimum value selection, can be applied for various research purposes, once the daily AVHRR data are cloud-free. *INDEX TERMS:* 3359 Meteorology and Atmospheric Dynamics: Radiative processes; 3337 Meteorology and Atmospheric Dynamics: Numerical modeling and data assimilation; 3360 Meteorology and Atmospheric Dynamics: Remote sensing; *KEYWORDS:* NOAA 16, AVHRR, cloud detection

Citation: Chen, P.-Y., R. Srinivasan, and G. Fedosejevs, An automated cloud detection method for daily NOAA 16 advanced very high resolution radiometer data over Texas and Mexico, *J. Geophys. Res.*, *108*(D23), 4742, doi:10.1029/2003JD003554, 2003.

1. Introduction

[2] The normalized difference vegetation index (NDVI) values derived from the advanced very high resolution radiometer (AVHRR) data are commonly applied to research such as vegetation change detection [*Jakubauskas et al.*, 2002; *Remmel and Perera*, 2001], biomass estimation [*Labus et al.*, 2002; *Maselli et al.*, 2000] and land cover mapping [*Chen et al.*, 1999; *Hill et al.*, 1999] at various spatial scale (e.g., state to global scale). Frequent repeat coverage at no cost is a strong advantage of AVHRR data, but cloud contamination can be a severe shortcoming by obstructing the ocean and land surface, and thus limiting the potential to monitor sea surface and vegetation conditions, respectively [*Ohring and Clapp*, 1980]. In general, clouds have higher visible reflectances but lower NDVI values [*Chen et al.*, 2002b]. *Holben* [1986] proposed the maximum NDVI value compositing method to minimize the effects of cloud contamination on AVHRR data, because the presence of clouds, smoke, snow and ice in a pixel will reduce the NDVI value. Ten-day to biweekly NDVI composites are widely used for research [*Senay and Elliott*, 2002; *Emery and Baldwin*, 1998]. *Gutman and Ignatov* [1996] concluded that even 2-week NDVI composites were possibly cloud-contaminated, which is true for areas such as the state of

0148-0227/03/2003JD003554$09.00

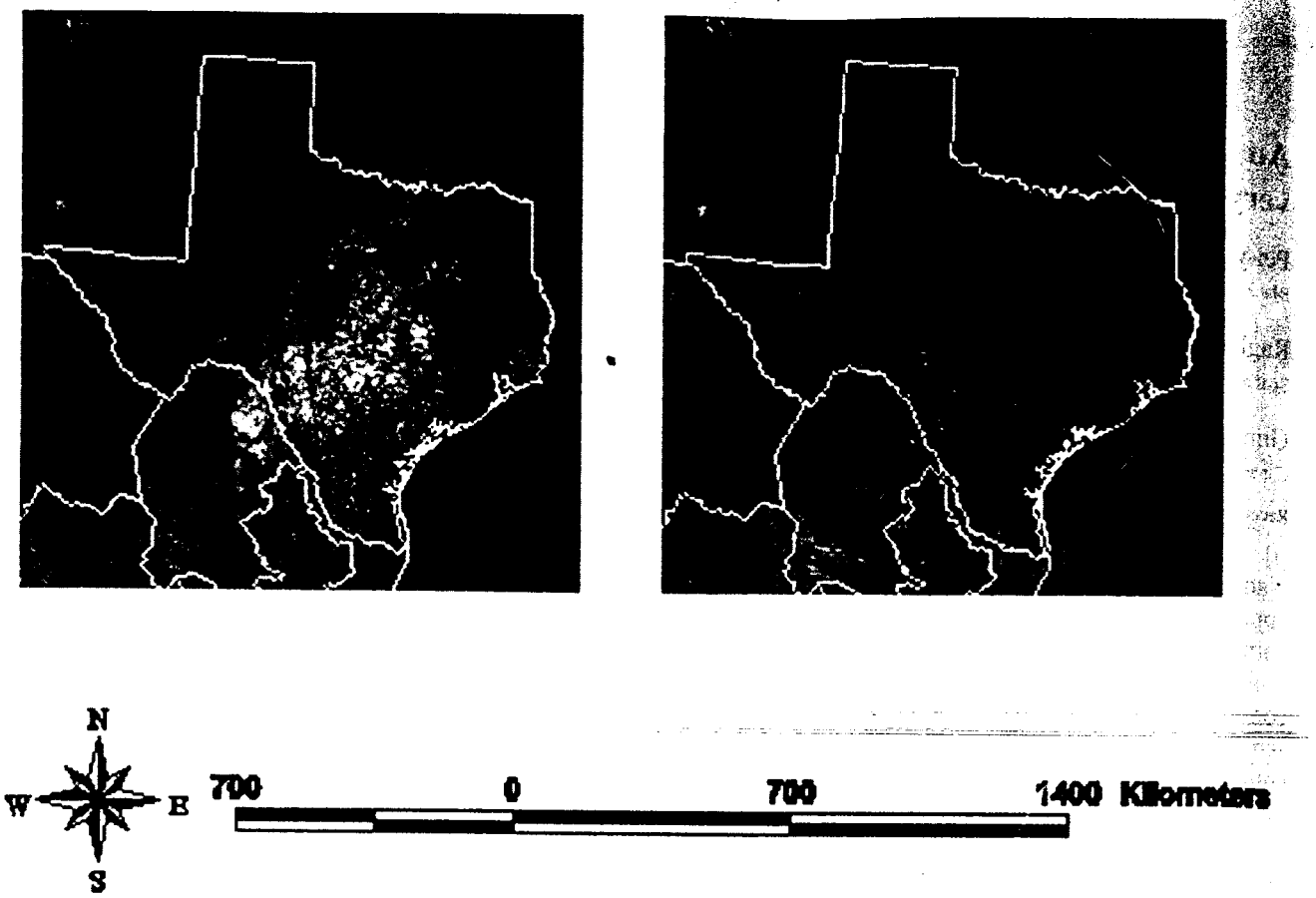

Figure 1. Composites of minimum visible reflectance: (a) high density of cloud coverage appeared in the weekly composite of Texas for 2–8 December 2001 and (b) a few clouds appeared in the biweekly composite for 2–15 December 2001.

Texas and Mexico next to the Gulf of Mexico (Figure 1). In addition, NDVI composite images for periods longer than 14 days are not suitable for monitoring agricultural crops with relatively short growing seasons of 3–4 months. Postseason cloud detection [*Cihlar*, 1996] and postcomposite cloud detection [*Gutman et al.*, 1994] were designed to provide cloud-free weekly or 10-day NDVI composite data. However, the former method based on temporal thresholding cannot deliver real-time composite data, and the latter method treats data from a single sensor and can not be used to composite data from multiple AVHRR sensors.

[3] Automated cloud detection for the daily National Oceanic and Atmospheric Administration (NOAA)-14 AVHRR data over Texas area had been developed to obtain cloud-free information for environmental research [*Chen et al.*, 2002b]. Cloud-contaminated pixels were identified and used to construct a cloud mask. Real-time NDVI composites based on daily clear-sky AVHRR data were successfully used to monitor crop growth in Texas [*Chen et al.*, 2003], and to estimate corn yield in Mexico [*Báez-González et al.*, 2002]. In addition, *Chen et al.* [2002a] merged the cloud-free AVHRR data from NOAA 14 and 15 to create NDVI composite images. The NOAA 14 AVHRR afternoon images were not appropriate for vegetation research since 2000 because of the negative impact of the large solar zenith angle on the reflectance retrieval accuracy. The NOAA 15 AVHRR morning data were mostly contaminated by cloud coverage over Texas and Mexico, and thus they were of little use to our study. The NOAA 16 satellite was launched on 20 September 2000 and became officially operational at the beginning of 2001. The ascending NOAA 16 afternoon overpasses were received between 1330 h and 1500 h Texas local time when the solar zenith angle was between 15° and 40° during March to October. The NOAA 16 AVHRR data with reflectance around solar noon in the summer were anticipated to provide better quality data for vegetation studies. Large-scale monitoring and estimation of crop yield is essential to ensure a stable and reliable food supply in developing countries. The crop growth modeling for Mexico is a long-term project using satellite and field data to predict grain yield. Previous work was done using NOAA 14 AVHRR data from 1999 and 2000. NOAA 16 AVHRR data were adopted for the large-scale real-time crop modeling starting in 2001.

[4] Each NOAA 14 AVHRR/2 image has five channels. In addition to the original five channels (0.58–0.68 μm, 0.72–1.00 μm, 3.55–3.93 μm, 10.3–11.3 μm and 11.5–12.5 μm), the NOAA 16 AVHRR/3 data has one extra channel with wavelengths of 1.58–1.64 μm (channel 3A) while the original channel 3 was renamed to channel 3B, but only five channels can be active at any one time. Channel 3A was active for NOAA 16 AVHRR after launch until the

end of April 2003 for daytime orbits. Starting in May 2003, channel 3B became active for fire studies. Published cloud detection algorithms for AVHRR data to date were developed for NOAA 14 and earlier satellites [*Saunders and Kriebel*, 1988; *França and Cracknell*, 1995; *Hutchison and Choe*, 1996; *Hutchison et al.*, 1997; *Dech et al.*, 1998; *Chen et al.*, 2002b]. In addition, NOAA/NESDIS also developed the algorithm CLAVR-1 (Clouds from AVHRR-Phase I) for deriving NOAA AVHRR cloud coverage and mapping total cloud amount [*Stowe et al.*, 1999]. These algorithms involve the use of solar reflectance, brightness temperatures, the ratio of solar reflectance, NDVI as well as the mathematical difference between reflectances or brightness temperatures in different channels. Several cloud detection algorithms were sensitive to surface inhomogeneity, because a series of universal thresholds and uniformity tests were applied to a 2×2 or a 3×3 pixel arrays. Channel 3B was employed in the CLAVR-1 algorithm to detect thinner but uniform clouds (>6%) as well as to restore misidentified clouds (<3%) [*Stowe et al.*, 1999]. Moreover, the brightness temperature difference between channels 3B and 4 is capable of identifying thin clouds and cirrus clouds of NOAA 14 AVHRR data [*Chen et al.*, 2002b]. However, because of NOAA programming of the NOAA 16 operational satellite, channels 3B data are not available for vegetation studies before May 2003.

[5] Cloud-free images are required for vegetation studies over Texas and Mexico. The existing NOAA 14 AVHRR daily cloud detection methods can not be transferred to NOAA 16 data without modification/validation. Part of this requirement hinges on the fact that the spectral response functions have changed for most of the channels on AVHRR/3. The objectives of this study were: (1) to develop an automated cloud detection method for NOAA 16 AVHRR data over the two study areas of Texas and Mexico, and (2) to assess the accuracy of this automated cloud detection method. Texas and Mexico are in the same geographical region, and both regions have similar weather conditions. Ideally, a single algorithm should be able to differentiate cloud-contaminated pixels from clear-sky pixels for both study areas.

2. Methodology

[6] NOAA 16 AVHRR data were downloaded from the High Resolution Picture Transmission (HRPT) receiving station located at Blackland Research and Extension Center (part of Texas Agricultural Experiment Station) in Temple, Texas. The NOAA 16 AVHRR data from ascending orbits covering Texas areas contain channel 3A for the daytime orbit segments from April 2001 to February 2003, when this study was conducted. While more than four orbits were received daily, data suitable for this study were only obtained from the afternoon orbits between 1330 h and 1500 h. The remaining orbits were received either at nighttime or at a time without sufficient daylight. It takes about 9 days for the NOAA 16 AVHRR to repeat an orbit. Thus one out of every 9 temporal images did not cover the entire state of Texas because of the satellite orbit location.

[7] Piecewise linear (PWL) calibration coefficients for NOAA 16 derived by the Canada Center for Remote Sensing were applied to the Texas data. The PWL coefficients were based on a postlaunch calibration derived by NOAA for NOAA 16 AVHRR/3 data [*Heidinger et al.*, 2002]. The daily raw AVHRR data were preprocessed using an automated approach. Visible channels 1, 2 and 3A were calibrated from digital counts to top-of-atmosphere (TOA) reflectance in percent (%), and the thermal channels 4 and 5 were converted to TOA brightness temperature (BT) in degrees Kelvin (°K). The NDVI value was computed for every pixel in each image.

[8] A total of 36 NOAA 16 AVHRR images acquired between April 2001 and March 2002 were utilized to develop the cloud detection algorithm. Three images were selected on the 10th, 20th and 30th day of each month. Cloud-contaminated pixels were classified as thick cloud, thin cloud, cirrus cloud, cloud edge and cloud shadow based on the density, color, shape and reflectance of clouds. For example, thick clouds appear white (Figure 1) due to high solar reflectance, while the thin clouds appear in various shades of gray because of low densities of cloud particles. Moreover, most of the thin clouds were semi-transparent with visibly wispy shapes. Although the cloud shadow pixels were technically clear-sky pixels, the cloud shadows were treated as cloud contamination class in this study. Clear sky pixels for this study were separated into three classes: vegetated land, barren land and water according to solar reflectance, NDVI values and land cover maps developed in 1992 using LANDSAT Thematic Mapper (TM) images. Five pixels were chosen manually at random from each cloudy and noncloudy class. Therefore 40 pixels (25 cloudy pixels and 15 clear-sky pixels) were chosen for each sampled image. The individual TOA reflectance, BT and NDVI value were recorded for each sampled pixel. As well, the differences between TOA reflectance in channels 2 and 3A and between brightness temperature in channels 4 and 5 were computed for each sampled pixel. Each cloudy and clear-sky class had 180 randomly sampled pixels covering the entire year. A total of 1440 pixels were sampled from all eight classes for this cloud detection study.

3. Result

3.1. Solar Reflectance

[9] All of the cloudy and clear-sky pixels except for vegetated areas have similar reflectance values in the visible and near-infrared channels (Figure 2). Bright areas such as clouds and barren lands had relatively high solar reflectances. The thick cloud pixels had the highest reflectance, averaging at 0.79. The thin cloud, cirrus cloud and cloud edge pixels generally had similar reflectances greater than 0.2. The barren land pixels had reflectances between 0.1 and 0.4, which were similar to the reflectances of cloudy pixels. Cloud shadows and water bodies had relatively low solar reflectances less than 0.2. Green vegetation reflected strongly in the near-infrared wavelengths, which caused difficulties in differentiating vegetated pixels from most of the cloudy pixels (Figure 2). Since vegetated land pixels carried the most important information for continuous monitoring research projects, it was apparent that the visible reflectance (R_1) would perform more efficiently in terms of differentiating vegetated clear-sky pixels from cloud-contaminated pixels.

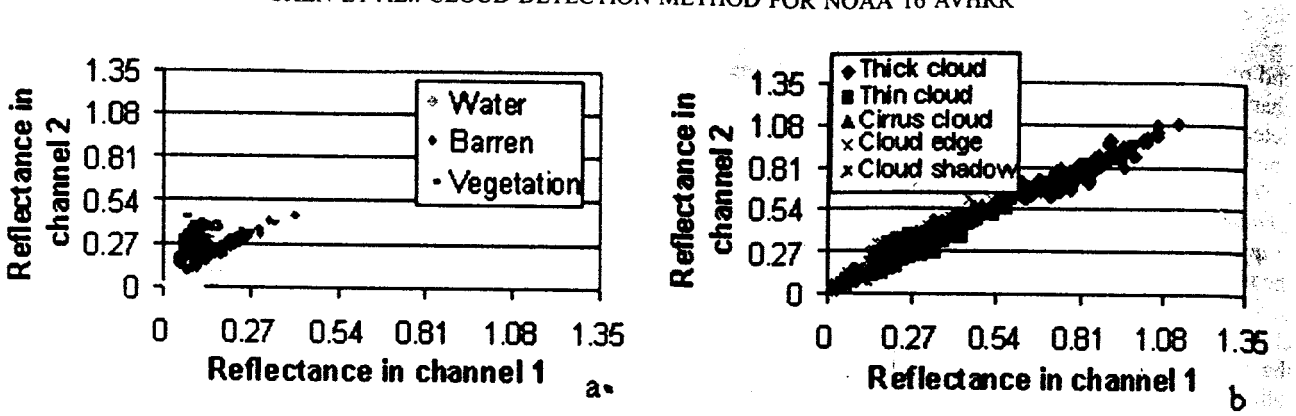

Figure 2. Top-of-atmosphere (TOA) reflectance in channel 1 (R_1) plotted against channel 2 (R_2) for (a) clear-sky pixels and (b) cloud-contaminated pixels. See color version of this figure at back of this issue.

[10] *Chen et al.* [2002b] chose a threshold of 0.27 for the R_1 for NOAA 14 AVHRR cloud detection to preserve clear-sky pixels while removing thick and most thin clouds as well as cirrus clouds. For this study of NOAA 16 data, both sampled vegetated and water pixels had R_1 less than 0.2. Eight out of 180 sampled barren land pixels had R_1 greater than 0.27 (Figure 2a). A lot of thin cloud, cirrus cloud and cloud edge pixels had R_1 greater than 0.27, while all of thick cloud pixels had R_1 greater than 0.47 (Figure 2b). Thus a threshold of 0.27 was adopted to eliminate thick clouds and some of the other cloudy pixels. According to the result, 505 out of 900 sampled cloudy pixels were correctly identified based on the threshold of 0.27, while only eight out of 180 barren pixels were accidentally removed or flagged as cloud-contaminated (Table 1). The R_1 was able to identify more than two thirds of the cirrus cloud pixels as well as half of the thin cloud pixels and cloud edge pixels. Unfortunately, none of the cloud shadow pixels were flagged by the R_1 test.

[11] The mid-infrared channel (3A) was new for the AVHRR/3 radiometer on board NOAA 15, -16 and -17. The performance of the mid-infrared reflectance (R_{3A}) was much different from both R_1 and the near-infrared reflectance (R_2) as seen in Figures 3 and 4. Cloud-filled and cloud edge pixels had a huge range from 0.08 to 0.6 for R_{3A}, which covered the reflectance range of barren lands and vegetated areas as well. Both cloud shadows and water bodies had very low R_{3A}, and many of the water bodies had a reflectance of 0. Overall, channel 3A reflectance alone was a poor discriminator of cloudy pixels.

3.2. Infrared Brightness Temperature

[12] All of the sampled pixels had a brightness temperature (BT) in channel 4 (BT_4) similar to that in channel 5 (BT_5). In general, cloudy pixels had lower BTs than clear-sky pixels (Figure 5). The BT_4 range for barren land pixels was very huge between 260°K and 328°K, which covered the BT_4 ranges of cloud shadow, cloud edge and cirrus cloud pixels. According to these results, the use of BT alone could only properly discriminate thick and thin clouds, which was the same as the use of R_1. Brightness temperature difference was widely applied in the cloud detection algorithm proposed by *Yamanouchi and Kawaguchi* [1992]. *Hutchison et al.* [1997] used the difference between brightness temperature in channel 3B (BT_{3B}) and BT_4. *Saunders and Kriebel* [1988] as well as *Dech et al.* [1998] used the difference between BT_4 and BT_5. Since BT_{3B} was not available from NOAA 16 ascending orbits, only the BT_4-BT_5 difference parameter was applicable to the current study. However, the results showed that both clear-sky and cloud-contaminated pixels had similar BT difference between −1°K and 7°K over Texas and Mexico areas, because the high water vapor absorption of BT_4 and BT_5 masked cloud signatures. The BT difference can detect cloudy pixels at higher latitude areas where water vapor is lower especially in winter. Overall, the BT difference was not capable of distinguishing cloud-contaminated pixels over the state of Texas and Mexico.

3.3. NDVI and Reflectance Difference

[13] The cloud shadow pixels had R_1 values and BT difference similar to the cloud-free pixels. The individual mid-infrared reflectance (R_{3A}) was not useful data for cloud detection, but the difference between R_2 and R_{3A} did provide valuable information to distinguish cloud shadow pixels from cloud-free pixels. The cloud shadows had higher R_2 than R_{3A}, while the barren lands had R_2 much less than R_{3A} (Figure 6). Barren lands had reflectance

Table 1. Stepwise Cloud Detection Algorithm and the Statistical Performance of Each Step

Processes of Identification	Classification of Sampled Pixels							
	Thick Cloud	Thin Cloud	Cirrus Cloud	Cloud Edge	Cloud Shadow	Water	Barren Land	Vegetation
$R_1 > 0.27$	180	89	150	86	0	0	8	0
R_2-$R_{3A} > -0.05$ & $0 \leq$ NDVI < 0.33	-	66	19	75	135	1	6	0
$R_{3A} \geq 0.09$ & NDVI < 0	-	18	7	7	7	2	0	0
$BT_4 > 300$°K/295°K & NDVI < 0	-	0	0	0	7	0	0	0
$R_1 < 0.1$ & $0 \leq$ NDVI < 0.33	-	0	0	0	8	0	4	0
Total Removed pixels	180	173	176	168	157	3	18	0
Accuracy	100%	96%	98%	93%	87%	98%	90%	100%

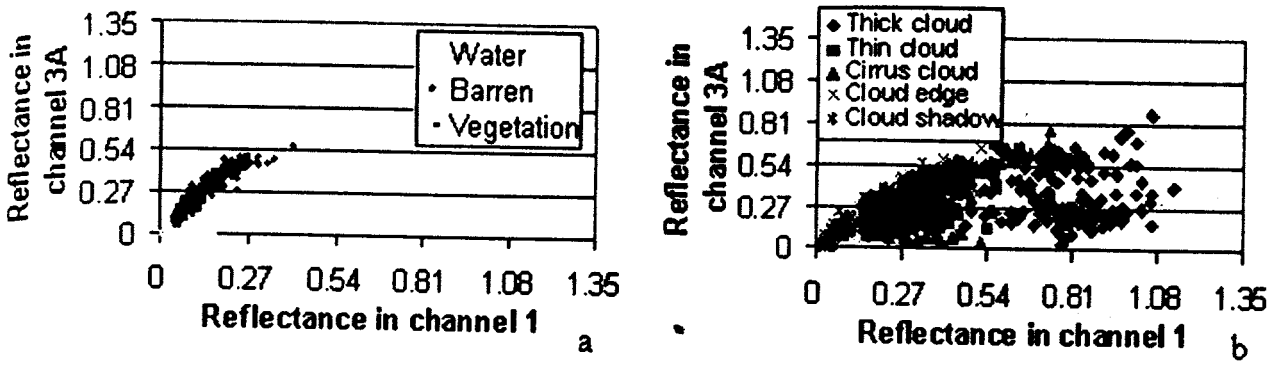

Figure 3. Top-of-atmosphere (TOA) reflectance in channel 1 (R_1) plotted against channel 3A (R_{3A}) for (a) clear-sky pixels and (b) cloud-contaminated pixels. See color version of this figure at back of this issue.

differences less than −0.05, while most cloud-contaminated pixels had reflectance differences greater than −0.05 similar to water bodies. Water bodies could be differentiated from cloud shadow pixels simply based on NDVI values. Results in Table 1 show that 295 out of 395 remaining cloud-contaminated pixels (900 − 505 = 395) were easily distinguished based on NDVI values being between 0 and 0.33 and the R_2-R_{3A} difference parameter being greater than −0.05. This process effectively identified 135 out of 180 cloud shadow pixels (Table 1). Six out of the 172 remaining barren land pixels were incorrectly removed through this process. This was an acceptable omission error for eliminating cloudy pixels.

3.4. Differentiation of Water From Cloud-Contaminated Pixels

[14] A total of 100 cloudy pixels remained because they had either NDVI values less than 0 (as did water bodies) or reflectance differences less than −0.05 (as did barren lands). Most water bodies had R_{3A} close to 0, and other cloudy pixels except cloud shadows had R_{3A} greater than 0.09. Moreover, several cloud shadows with negative NDVI values had BT_4 values greater than 300°K from April to October or greater than 295°K from November to March, while water bodies had lower BT_4 values. Thresholds for the two tests chosen were 0.09 for R_{3A} and 300°K (April–October) or 295°K (November–March) for BT_4 to identify cloudy pixels from all pixels with NDVI values less than 0. The R_{3A} test successfully removed 39 out of the remaining of 100 cloudy pixels. The R_{3A} threshold correctly identified 18 out of the remaining 25 thin cloud pixels, 7 out of the remaining 11 cirrus cloud pixels, 7 out of the remaining 19 cloud edge pixels and 7 out of the remaining 45 cloud shadow pixels (Table 1). The BT_4 threshold was used to identify cloud shadow pixels, and it effectively removed 7 out of the remaining of 38 cloud shadow pixels.

3.5. Differentiation of Barren Lands From Cloud Shadow Pixels

[15] A total of 54 cloudy pixels were still not identified after applying the aforementioned threshold tests. Of these, 31 were cloud shadow pixels. Ten of the 31 cloud shadow pixels had NDVI values greater than 0, and were confused with barren land pixels having a similar BT difference and reflectance difference. Since cloud shadows tend to be dark, their R_1 values were slightly lower than for bright barren lands. A threshold of 0.1 was chosen for R_1 to distinguish cloud shadow pixels from barren land pixels. The R_1 threshold effectively identified 8 out of the 10 cloud shadow pixels with positive NDVI values, while 4 barren land pixels were also removed (Table 1).

3.6. Independent Validation

[16] Accuracy assessment has been required for research using satellite images to learn about the quality of information derived from remotely sensed data and to evaluate the performance of a research method [*Congalton and Green*, 1999]. A total of 36 daily images from June 2001 to May

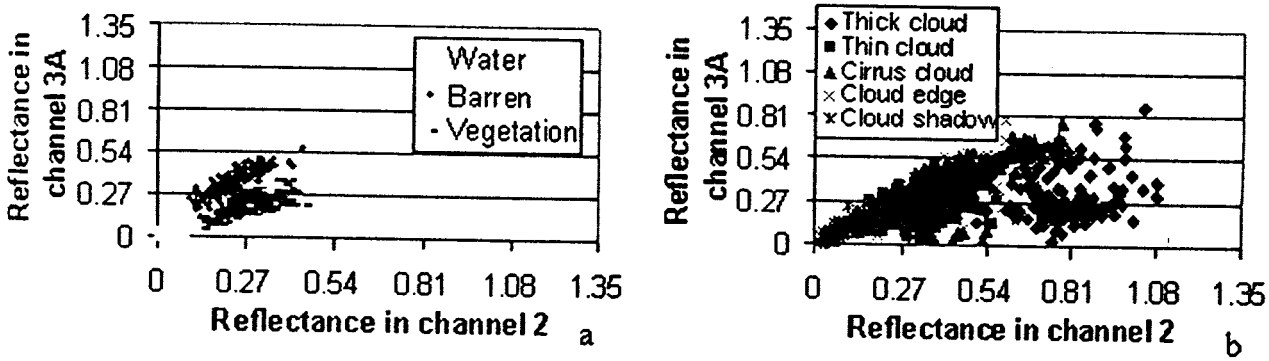

Figure 4. Top-of-atmosphere (TOA) reflectance in channel 2 (R_2) plotted against channel 3A (R_{3A}) for (a) clear-sky pixels and (b) cloud-contaminated pixels. See color version of this figure at back of this issue.

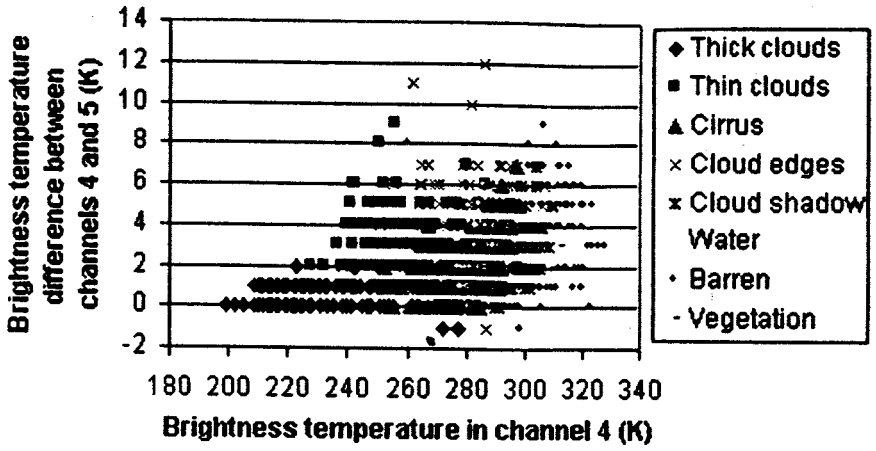

Figure 5. Brightness temperature in channel 4 (BT_4) plotted against brightness temperature difference. See color version of this figure at back of this issue.

2002 were used to validate the cloud detection algorithm. Three images were selected on the 5th, 15th and 25th day of each month. April and May 2002 were used for validation, because several daily AVHRR scenes from NOAA 16 in April and May 2001 were lost due to conflicts in data reception with NOAA 14. In addition, this cloud detection algorithm was developed for multitemporal applicability between years. The developed cloud detection method was applied to the 36 selected images. More than 1400 pixels were randomly sampled from the 36 images, including cloud-contaminated and cloud-free pixels. Each sampled pixel had been visually classified before being processed by our cloud detection algorithm. The results indicated that the sampled pixels of thick cloud, cirrus cloud and cloud edge were properly identified, while seven thin cloud pixels and 12 cloud shadow pixels were not correctly distinguished from clear-sky pixels (Table 2). On the other hand, 26 water pixels and 52 barren land pixels were misclassified as cloud-contaminated pixels. The overall accuracy was close to 93%.

4. Discussion

[17] It was unusual to have 26 water pixels misidentified as cloudy pixels. Eighteen of the 26 pixels were misclassified because BT_4 values were greater than the threshold value of 295°K. These pixels were either obtained from March or November images. In addition, they were distributed in water bodies in central or southern Mexico. According to the results of this study, the BT_4 for water bodies in Mexico was between 295 and 300°K from November to March, a couple of degrees higher than that in Texas. Moreover, the BT_4 was less influenced by seasonal changes, because the locations were closer to the equator. It was therefore necessary to increase the threshold value for BT_4 to 300°K for images from November to March for pixels with a latitude of less than 25° (Figure 7). There were only 4 water pixels misidentified when the threshold was increased to 300°K. Another option is to create a perennial water mask to eliminate water pixels before applying the cloud detection method.

[18] It was unfortunate that a total of 52 barren land pixels were misidentified as cloudy pixels. The main reason was that one quarter of the barren land pixels were collected in the bright deserts of northern Mexico or West Texas where the R_1 was greater than 0.27. The surface of the deserts was as bright as the clouds, though the deserts had a much higher BT_4 than the clouds. The results indicated that all of the barren lands with bright reflectances had a BT_4 of

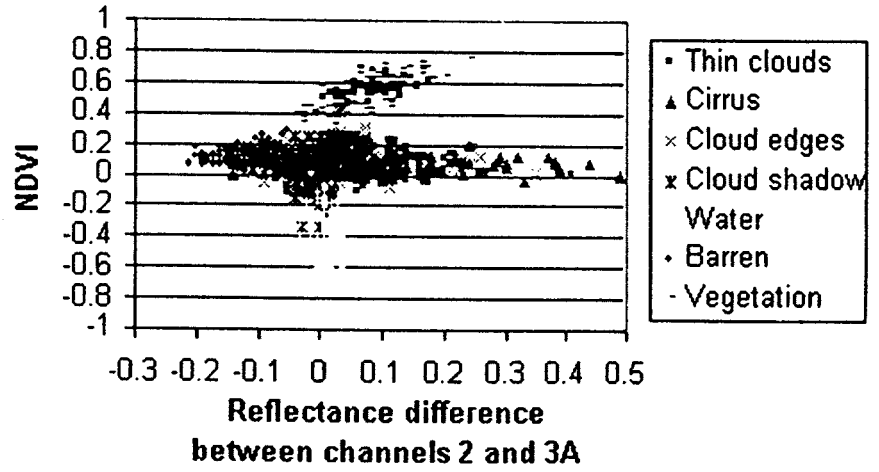

Figure 6. Reflectance difference plotted against NDVI values. See color version of this figure at back of this issue.

Table 2. Results of the Independent Validation of the Noaa 16 AVHRR Cloud Detection Algorithm[a]

	Reference Data							
	Cloud Types					Land Cover Types		
Cloud Detection Results	Thick Clouds	Thin Clouds	Cirrus Clouds	Cloud Edge	Cloud Shadow	Water	Barren Land	Vegetation
Cloud-contaminated pixels	180	173	180	180	168	26	52	0
Clear-sky pixels	0	7	0	0	12	154	128	180
Accuracy	100%	96.1%	100%	100%	93.3%	85.6%	71.1%	100%

[a]Overall accuracy = (180 + 173 + 180 + 180 + 168 +154 + 128 + 180)/1440 = 93.26%.

greater than 300°K, while all types of cloud-contaminated pixels had a maximum BT_4 of less than 299°K. Thus the R_1 threshold combined with the BT_4 threshold would positively improve the accuracy of the algorithm (Figure 7). In addition, many of barren land pixels were misidentified as cloud shadows, because the difference between R_2 and R_{3A} was greater than -0.05 or the R_1 was less than 0.1. Most of the misidentified pixels had a dark surface and were collected from Mexico data between October and February. Changing threshold values can increase the accuracy of identifying barren land pixels, but may fail to detect some of the cloud-contaminated pixels, especially cloud shadows. It was not easy to find a suitable threshold value to accurately classify the cloud-contaminated and clear-sky pixels. Moreover, this cloud detection algorithm was mainly developed to provide cloud-free data for vegetation monitoring and green biomass estimation.

[19] Pixels contaminated by cloud shadows were classified as cloudy pixels in this study. A couple of cloud shadow pixels with negative NDVI values were not correctly identified by our automated cloud detection algorithm. Two thirds of the unidentified cloud shadow pixels were collected from images acquired between November and March. These unidentified cloud shadow pixels appeared as a black surface like water bodies, and their R_2 was lower than R_1. In addition, few cloud shadow pixels had high NDVI values greater than 0.33 similar to the surrounding vegetated pixels, but their dark surfaces were readily identified as cloud shadow pixels by visual inspection. In general, those few remaining cloud shadow pixels would not cause a strong bias in the yield forecast for the composites.

[20] The CLAVR-1 algorithm was compared with the cloud detection algorithm developed in this study using the same AVHRR data for Texas and Mexico areas. Two tests in the CLAVR-1 algorithm involving channel 3B were neglected. Results showed the CLAVR-1 algorithm correctly identified most cloudy pixels except cloud shadow pixels. Many of the barren land pixels were classified as cloudy or mixed pixels when applying CLAVR-1 algorithm, and few water and vegetated pixels were misidentified as mixed pixels as well. Most of the misidentifications were caused by the R_1 and BT_4 uniformity tests for each 2 × 2 array because of surface inhomogeneity. More than 100 of the cloud shadow pixels were classified as clear-sky pixels, and these misidentified cloud shadow pixels occurred randomly in each month. Many of the cloud shadow pixels in Southern US and Mexico were missed because they had a BT_4 of greater than 293°K. A lot of barren land pixels were classified as mixed or cloudy pixels, particularly the deserts within the border of Texas and Mexico where the ratio of R_2 to R_1 fell into the cloudy range between 0.9 and 1.1. *Stowe et al.* [1999] mentioned that the CLAVR-1 tests employing thermal channels (3B, 4 and 5) became unreliable and meaningless when the BT_4 exceeded 315°K, because the thermal channels approached saturation values. The desert areas frequently had a BT_4 of greater than 315°K between June and September, and therefore the CLAVR-1 test using the difference between BT_4 and BT_5 was not appropriate for Texas and Mexico.

[21] Although the cloud detection algorithm developed in this study included only 5 tests, each of threshold values was chosen based on data for an entire year, which would accommodate seasonal and climatic variations expected in Texas and Mexico. This algorithm utilized R_{3A} as essential procedures to effectively identify cloud-contaminated pixels. Two out of five cloud detection tests required the R_{3A}, and 142 out of 180 cloud shadow pixels were identified because of the inclusion of R_{3A} data. Less than half of cloud shadow pixels were correctly identified for NOAA 14 AVHRR data, because the R_{3A} was not available for NOAA 14 and earlier satellites [*Chen et al.*, 2002b]. In addition, 285 out of 900 cloud contaminated pixels from the data for independent validation were identified by the second and third tests utilizing R_{3A}. The use of R_{3A} successfully identified 30% of thin cloud pixels, 30% of cloud edge pixels and 85% of cloud shadow pixels according to the results of the independent validation, while 17% of barren land pixels were incorrectly removed.

[22] The major function of R_{3A} in this study was to identify remaining cloudy pixels (including cloud shadow

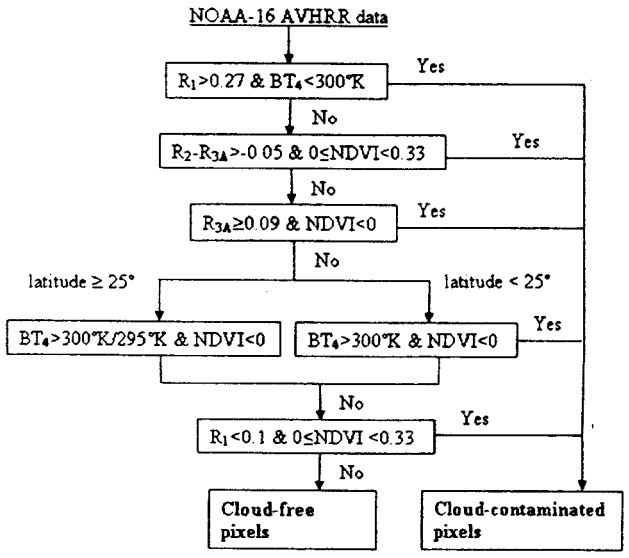

Figure 7. Cloud detection algorithm of NOAA 16 AVHRR for the state of Texas and Mexico.

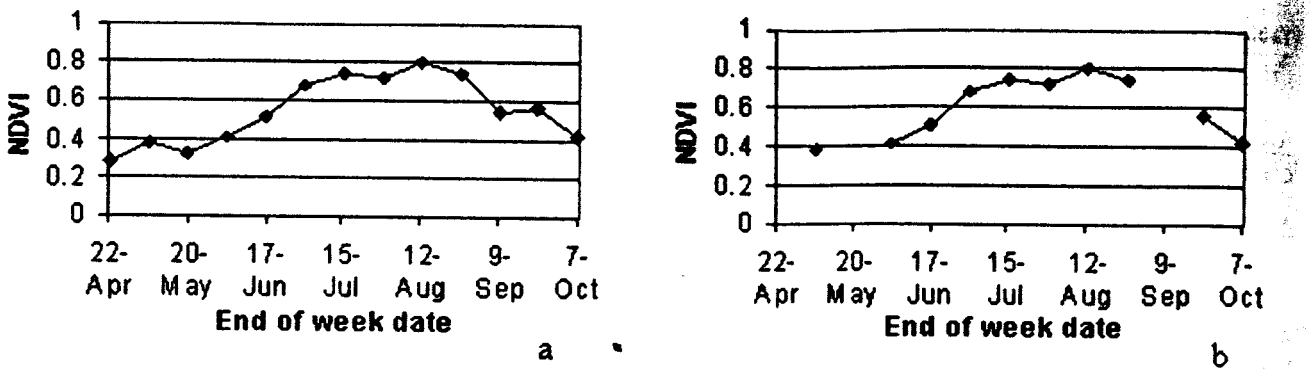

Figure 8. NDVI temporal profiles for corn growth monitoring in northern Texas, (a) traditional NDVI composites and (b) conditional NDVI composites.

pixels) that were undetected by the first R_1 test (Table 1), though R_{3A} was originally included to distinguish snow/ice from clouds. For the NOAA 14 data, the brightness temperature difference between channel 3 (BT_{3B}) and BT_4 could identify clouds but not cloud shadows, because the cloud shadows had a BT_4 much higher than cold clouds [*Chen et al.*, 2002b]. Both NOAA 14 and NOAA 16 are afternoon satellites, and the NOAA 15 and NOAA 17 are morning satellites. BT_{3B} data from NOAA 14 AVHRR provided valuable information for the research of global fire monitoring and detection. Therefore channel 3B has been selected for NOAA 16 operation at all times starting May 2003 by NOAA for more efficient fire monitoring, while channel 3A has been selected for NOAA 17 daylight orbits. In the future, a new approach is considered necessary for cloud shadow detection with NOAA 16 AVHRR data when the R_{3A} is not available.

[23] Both snow and ice/frost are not mentioned in this study, because they rarely occur in our study areas. In general, bright snow pixels will produce very strong visible reflectances similar to those for thick cloud pixels. The snow pixels unlike the cloudy pixels do not have associated shadow pixels. Terrain shadows and cloud shadows can be confused for pixels with large solar zenith angles. Ice/frost pixels could not be simply identified by visual inspection, but they typically have low NDVI values and low BTs as do cloudy pixels. Therefore both snow and ice/frost covered pixels would be classified as various cloud pixels according to the cloud detection algorithm. This algorithm was developed based on the characteristics of each cloud contamination type for Texas and Mexico, and it could be certainly applied elsewhere. Each threshold value was chosen to identify/remove cloud-contaminated pixels while retaining clear-sky pixels based on observed land and water conditions. Texas and Mexico have short mild winters. It would be necessary to modify the threshold values before applying this algorithm to other areas in the United States, particularly during the winter.

[24] Corn, a major crop species in Texas, is mainly planted in April and harvested in September or October. The biweekly maximum value NDVI composites derived from NOAA 16 AVHRR data have been used to monitor corn growth stage in Texas and to predict corn yield in Mexico. The NDVI composites based on the MVC method (traditional NDVI composites) for corn monitoring in Northern Texas had an unusually low NDVI value for the date of 9 September when corn reached maturity (Figure 8a). The NDVI composites based on the MVC method using daily cloud-free AVHRR scenes (conditional NDVI composites) showed that the NDVI value for the date of 9 September was missing because of continuous cloud contamination for 14 days in that location (Figure 8b). In addition, two NDVI values based on conditional NDVI composites were missing because of cloud contamination at the beginning of the corn season between the end of April and May, while the traditional NDVI composites suggested that the NDVI values were influenced because of environmental changes. The traditional NDVI temporal profiles could not resolve whether the unusually low NDVI values were caused by environmental stress such as drought and insect damage or by continuous cloud contamination. The conditional NDVI profile showed that the low NDVI value between 15 July and 12 August was not contaminated by clouds, which suggested a possible drought situation and/or other environmental causes.

5. Conclusion

[25] This newly developed cloud detection algorithm for NOAA 16 AVHRR data successfully classified all of the thick cloud, cirrus cloud as well as cloud edge pixels. The accuracy of this algorithm was greater than 93%. Less than 2% of the cloudy pixels were unidentified, while none of the vegetated pixels were misclassified. NDVI composite images based on daily cloud-free data will provide real-time information beneficial to agricultural monitoring and short-term environmental studies. Moreover, they provide more accurate and reliable data for short time period composites. In general, 14-day composites are long enough to obtain cloud-free information. However, our study sites in Texas and Mexico were frequently covered by clouds for periods longer than two weeks at a time in winter. Thus the daily cloud detection algorithm is required before creating composite images, though some biweekly composites may lack any pixel values because of continuous cloud coverage. Moreover, daily cloud-free AVHRR data can be used to create composite images using compositing methods other than MVC such as multidate averaging, minimum value selection or other mathematical operators, since the cloud contaminated pixels have been removed. *Hansen et al.*

[2002] successfully used minimum visible reflectance compositing and maximum NDVI compositing to study tree cover. Multidate averaging compositing can be useful for temperature studies where the MVC method is likely to produce biased surface temperature information not representative of the composite period. The switch from channel 3A to channel 3B starting in May 2003 limits the use of this algorithm to NOAA 16 AVHRR data acquired between September 2000 and April 2003. Research is planned to modify our cloud detection methodology to accommodate data collection from channel 3B instead of channel 3A since May 2003.

[26] **Acknowledgments.** The authors would like to thank the USDA-Agricultural Research Service (ARS) in Temple, Texas, for supporting this research through the Specific Cooperate agreement 58-6206-1-005.

References

Báez-González, A. D., P. Y. Chen, M. Tiscareño-López, and R. Srinivasan, Using satellite and field data with crop growth modeling to monitor and estimate corn yield in Mexico, *Crop Sci.*, *42*(6), 1943–1949, 2002.

Chen, P. Y., R. Srinivasan, G. Fedosejevs, A. D. Báez-González, and P. Gong, Assessment of NDVI composites using merged NOAA-14 and NOAA-15 AVHRR data, *Geogr. Inf. Sci.*, *8*(1), 31–38, 2002a.

Chen, P. Y., R. Srinivasan, G. Fedosejevs, and B. Narasimhan, An automated cloud detection method for daily NOAA-14 AVHRR data for Texas, USA, *Int. J. Remote Sens.*, *23*(15), 2939–2950, 2002b.

Chen, P. Y., R. Srinivasan, G. Fedosejevs, and J. R. Kiniry, Evaluating different NDVI composite techniques using NOAA-14 AVHRR data, *Int. J. Remote Sens.*, *24*(17), 3403–3412, 2003.

Chen, X. W., R. Tateishi, and C. Y. Wang, Development of a 1-km land cover dataset of China using AVHRR data, *J. Photogramm. Remote Sens.*, *54*(5–6), 305–316, 1999.

Cihlar, J., Identification of contaminated pixels in AVHRR composite images for studies of land biosphere, *Remote Sens. Environ.*, *56*, 149–163, 1996.

Congalton, R., and K. Green, *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*, CRC Press, Boca Raton, Fla., 1999.

Dech, S. W., P. Tungalagsaikhan, C. Preusser, and R. E. Meisner, Operational value-adding to AVHRR data over Europe methods, results, and prospects, *Aerospace Sci. Technol.*, *2*(5), 335–346, 1998.

Emery, W., and D. Baldwin, Online access to weather satellite imagery through the World Wide Web, *IEEE Trans. Geosci. Remote Sens.*, *36*(5), 1367–1375, 1998.

França, G. B., and A. P. Cracknell, A simple cloud masking approach using NOAA AVHRR daytime data for tropical areas, *Int. J. Remote Sens.*, *16*(9), 1697–1705, 1995.

Gutman, G., and A. Ignatov, The relative merit of cloud/clear identification in the NOAA/NASA pathfinder AVHRR land 10-day composites, *Int. J. Remote Sens.*, *17*(16), 3295–3304, 1996.

Gutman, G., A. Ignatov, and S. Olson, Towards better quality of AVHRR composite mages over land: Reduction of cloud contamination, *Remote Sens. Environ.*, *50*, 134–148, 1994.

Hansen, M. C., R. S. DeFries, J. R. G. Townshend, R. Sohlberg, C. Dimiceli, and M. Carroll, Towards an operational MODIS continuous field of percent tree cover algorithm: Examples using AVHRR and MODIS data, *Remote Sens. Environ.*, *83*, 303–319, 2002.

Heidinger, A. K., C. Cao, and J. T. Sullivan, Using Moderate Resolution Imaging Spectrometer (MODIS) to calibrate advanced very high resolution radiometer reflectance channels, *J. Geophys. Res.*, *107*(D23), 4702, doi:10.1029/2001JD002035, 2002.

Hill, M. J., P. J. Vickery, E. P. Furnival, and G. E. Donald, Pasture land cover in eastern Australia from NOAA-AVHRR NDVI and classified Landsat TM, *Remote Sens. Environ.*, *67*, 32–50, 1999.

Holben, B. N., Characteristics of maximum-values composite images from temporal AVHRR data, *Int. J. Remote Sens.*, *7*(15), 1417–1434, 1986.

Hutchison, K. D., and N. Choe, Application of 1.38 μm imagery for thin cirrus detection in daytime imagery collected over land surface, *Int. J. Remote Sens.*, *17*(17), 3325–3342, 1996.

Hutchison, K. D., B. J. Etherton, and P. C. Topping, Cloud top phase determination from the fusion of signatures in daytime AVHRR imagery and HIRS data, *Int. J. Remote Sens.*, *18*(15), 3245–3262, 1997.

Jakubauskas, M. E., D. L. Peterson, J. H. Kastens, and D. R. Legates, Time series remote sensing of landscape-vegetation interactions in the southern Great Plains, *Photogramm. Eng. Remote Sensing*, *68*(10), 1021–1030, 2002.

Labus, M. P., G. A. Nielsen, R. L. Lawrence, R. Engel, and D. S. Long, Wheat yield estimates using multi-temporal NDVI satellite imagery, *Int. J. Remote Sens.*, *23*(20), 4169–4180, 2002.

Maselli, F., S. Romanelli, L. Bottai, and G. Maracchi, Processing of GAC NDVI data for yield forecasting in the Sahelian region, *Int. J. Remote Sens.*, *21*(18), 3509–3523, 2000.

Ohring, G., and P. F. Clapp, The effects of changes in cloud amount on the net radiation at the top of the atmosphere, *J. Atmos. Sci.*, *37*(2), 447–454, 1980.

Remmel, T. K., and A. H. Perera, Fire mapping in a northern boreal forest: Assessing AVHRR/NDVI methods of change detection, *For. Ecol. Manage.*, *152*(1–3), 119–129, 2001.

Saunders, R. W., and K. T. Kriebel, An improved method for detecting clear sky and cloudy radiance from AVHRR data, *Int. J. Remote Sens.*, *9*(1), 123–150, 1988.

Senay, G. B., and R. L. Elliott, Capability of AVHRR data in discriminating rangeland cover mixtures, *Int. J. Remote Sens.*, *23*(2), 299–312, 2002.

Stowe, L. L., P. A. Davis, and E. P. McClain, Scientific basis and initial evaluation of the CLAVR-1 global clear/cloud classification algorithm for the advanced very high resolution radiometer, *J. Atmos. Oceanic Technol.*, *16*(6), 656–681, 1999.

Yamanouchi, T., and S. Kawaguchi, Cloud distribution in the Antarctic from AVHRR data and radiation measurements at the surface, *Int. J. Remote Sens.*, *13*(1), 111–127, 1992.

P.-Y. Chen, Blackland Research and Extension Center, Texas Agricultural Experiment Station, 720 East Blackland Road, Temple, TX 76502, USA. (pchen@brc.tamus.edu)

G. Feodsejevs, Canada Center for Remote Sensing, 588 Booth Street, Ottawa, Ontario, Canada K1A 0Y7.

R. Srinivasan, Spatial Sciences Laboratory, Department of Forest Science, Texas A&M University, College Station, TX 77843, USA.