Evaluating different NDVI composite techniques using NOAA-14 AVHRR data

P.-Y. CHEN, R. SRINIVASAN
Spatial Science Laboratory, Department of Forest Science, Texas A&M University, College Station, Texas 77843, USA

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

and J. R. KINIRY
USDA, Agricultural Research Service, 808 E. Blackland Road, Temple, Texas 76502, USA

(Received 8 August 2001; in final form 17 May 2002)

Abstract. Normalized Difference Vegetation Index (NDVI) data derived from Advanced Very High Resolution Radiometer (AVHRR) data are influenced by cloud contamination, which is common in individual AVHRR scenes. Maximum value compositing (MVC) of NDVI data has been employed to minimize cloud contamination. Two types of weekly NDVI composites were built for crop seasons in summer: one from all available AVHRR data (named the traditional NDVI composite) and the other from solely cloud-free AVHRR data (named the conditional NDVI composite). The MVC method was applied to both composites. The main objective of this study was to compare the two types of NDVI composites using Texas data. The NDVI seasonal profiles produced from the conditional NDVI composites agreed with the field measured leaf area index (LAI) data, reaching maximum values at similar times. However, the traditional NDVI composites showed irregular patterns, primarily due to cloud contamination. These study results suggest that cloud detection for individual AVHRR scenes should be strongly recommended before producing weekly NDVI composites. Appropriate AVHRR data pre-processing is important for composite products to be used for short-term vegetation condition and biomass studies, where the traditional NDVI composite data do not eliminate cloud-contaminated pixels. In addition, this study showed that atmosphere composition affected near-infrared reflectance more than visible reflectance. The near-infrared reflectance was increasingly adjusted through atmospheric correction.

1. Introduction
The Advanced Very High Resolution Radiometer (AVHRR) data acquired from the National Oceanic and Atmospheric Administration (NOAA)-14 afternoon orbits provide daily information for global vegetation monitoring. The Normalized Difference Vegetation Index (NDVI) is derived from the visible (VIS, 0.58–0.68 μm)
and near-infrared (NIR, 0.725–1.1 μm) reflectance data suggested by Rouse et al. (1973). It is defined by the following equation:

$$NDVI = \frac{NIR - VIS}{NIR + VIS}$$  \hspace{1cm} (1)

The NDVI values vary between –1.0 and +1.0. A high reflectance in the NIR channel and low reflectance in the VIS channel observed for dense green vegetation produce a positive NDVI value, and negative NDVI values indicate the presence of clouds, snow, water, or a bright non-vegetated surface (Yin and Williams 1997). A typical NDVI temporal profile for healthy green vegetation increases in the spring, reaches a peak or a plateau during the summer and declines with plant senescence in the fall (Viovy and Arino 1992). Several vegetation studies have been conducted using NDVI values based on the truth that the leaf area index (LAI) and NDVI were closely correlated to the fraction of PAR intercepted by canopies (Hatfield et al. 1984, Sellers et al. 1996, Dech et al. 1998, Shanahan et al. 2001). The basis for such a relationship comes from the fact that radiation in the visible wavelengths is absorbed by chlorophyll pigments in green leaves, and near-infrared radiation is scattered by the internal leaf structure (Nemani and Running 1989).

A single AVHRR scene is seldom completely cloud-free. Cloud-free AVHRR observations are necessary for vegetation monitoring, primary production estimation, environmental transformation research and crop growth monitoring (Groten 1993, Penuelas et al. 1995, Walker et al. 1995, Baynes and Dunn 1997, Olson and Cochran 1998, Choubey and Choubey 1999). Cloud-contaminated pixels do not represent real information of land cover types. The maximum NDVI value compositing method is widely used to minimize the effects of cloud contamination on AVHRR data (Holben 1986). The maximum value compositing (MVC) method is based on the assumption that the presence of clouds, smoke, haze, snow, and ice in a pixel will reduce the NDVI value. In general, 7-day to bi-weekly NDVI composites are used for research and application development (National Oceanic and Atmospheric Administration 1990). However, a 7-day NDVI composite is seldom cloud-free. For study sites such as Texas where pixels are frequently contaminated by clouds or water vapours because of the Gulf of Mexico, bi-weekly NDVI composites could still include cloud-contaminated pixels. Gutman and Ignatov (1996) suggested that the maximum NDVI composite for obtaining cloud-free information could be effective in such cases if the composite period is long enough such as one month. For crops with growing seasons of three to four months, one-month cloud-free maximum NDVI composites may fail to indicate the precise crop growth stage or provide crop condition required on a weekly basis. Post-composite cloud detection proposed by Gutman and Ignatov (1996) or post-season cloud detection as proposed by Cihlar (1996) could provide cloud-free weekly NDVI composite data, but this is not possible for real-time delivery of composite data.

Currently most research is based on NDVI composites built from AVHRR scenes without cloud detection applied. In addition, most NDVI composites were created based on top-of-atmosphere (TOA) reflectance. However, atmospheric effects influence the signal received by sensors. The effects on the signal are mainly related to the wavelength of the signal, the composition of the atmosphere and the path from the sun to the surface and to the sensor. This study applied atmospheric correction to AVHRR data to adjust the effects of water vapour, ozone, carbon dioxide and
Rayleigh scattering. Each NDVI composite used in this study was produced using either TOA reflectance (before atmospheric correction) or surface reflectance (after atmospheric correction). The objectives of this study were: (1) to create weekly NDVI composites of Texas from all available AVHRR data using the MVC procedure (subsequently referred to as the ‘traditional NDVI composite’); (2) to build weekly NDVI composites of Texas from only cloud-free AVHRR data using the MVC procedure (subsequently referred to as the ‘conditional NDVI composite’); (3) to compare the NDVI data accuracy for the two types of composites; and (4) to observe the atmospheric effect on NDVI value distribution.

2. Method
2.1. AVHRR data processing

AVHRR High Resolution Picture Transmission (HRPT) data are downlinked daily from the NOAA-14 satellite to the receiving station located at the Blackland Research and Extension Center (part of Texas A&M University complex) in Temple, Texas. Four AVHRR scenes were received in Texas: two before sunrise, and another two between 1500 h and 1800 h. The early morning data were not appropriate for optical vegetation studies due to the absence of sunlight, and one of the afternoon overpasses was not suitable for this study because it covers only a small portion of East Texas. Thus, one scene per day between March and September 2000 was employed in this study.

Raw AVHRR data were first processed to level 1b data set. Metadata extracted from the AVHRR header file included orbit number, earth location in geographical coordinates, acquisition time of scenes and other relevant information. The metadata were further used for geocoding. The pre-processing of channel 1 (VIS) and 2 (NIR) data for this study included radiometric correction and calibration, reflectance conversion, atmospheric correction, cloud removal, satellite zenith, solar zenith, relative azimuth angle computation and NDVI calculation. The VIS and NIR data were calibrated to TOA reflectance factor. A simplified method for atmosphere correction (SMAC) developed by Rahman and Dedieu (1994) was applied to compute surface reflectance. The NDVI values were computed from TOA reflectance (before atmosphere correction) and surface reflectance (after atmosphere correction). A cloud detection method developed by Chen et al. (2002) for the state of Texas was applied for cloud removal.

Most AVHRR data for Texas have solar zenith angles between 40° and 70°. Pixels having a solar angle greater than 70° were not used in this study based on the suggestions by Holben (1986) and Singh (1988). AVHRR data have a spatial resolution of 1 km × 1 km at nadir and about 6 km × 6 km at the maximum scan angle of 55.4°. Pixels at extreme scan angles cover more than 30 km². Data for Texas with scan angles greater than 42° (≈ satellite zenith angle of 50°) were not used in NDVI composites (Lovell and Graetz 2000) to minimize random noise caused by large scan angles.

2.2. NDVI composite processing

Daily AVHRR scenes between March and September 2000 were used in this study. Two types of NDVI composites were created: one was named the ‘traditional NDVI composites’ using all AVHRR data without applying cloud detection, and the other one was named the ‘conditional NDVI composite’ built from cloud-screened data. The MVC method was applied for both NDVI composite types. The
automated cloud detection algorithm for Texas was applied to daily AVHRR scenes for the conditional NDVI composites. Cloud-contaminated pixels were assigned a specific value to differentiate them from clear-sky pixels. Based on previous studies of automated cloud detection accuracy, more than 80% of the cloud-contaminated pixels were correctly masked out before creating the conditional NDVI composites (Chen et al., 2002). Many of the residual contaminated pixels were affected by cloud shadows.

Cloud contamination of AVHRR data appears to be an important problem in Texas due to the state’s proximity to the Gulf of Mexico. In general, less than seven scenes had clear-sky data for the weekly conditional NDVI composites. Continuous cloud coverage for an entire week resulted in missing NDVI values from the weekly conditional NDVI composites. The weekly traditional NDVI composites were created using all seven scenes in one week, thus no weekly maximum NDVI values were missing for the entire growing season.

Four agricultural sites in Texas were selected to compare differences between the two types of NDVI composites (figure 1). Three sites were planted with sorghum, and one site was planted with maize. Maize and sorghum are important summer crops in Texas. Two sites located in Northern Texas were facilitated with automatic irrigation systems, and the other two sites located in Central and Southeastern Texas were non-irrigated. The field data were collected in the four study sites between 1995 and 1999. Springlake data for 1995 and Gregory and Granger data for 1996 were not collected. Field data of each year included planting dates for each crop species and time of reaching silking for the maize or heading for the sorghum. Field data were not available after 1999. Crop species and planting times were observed to be similar for each of the four or five years prior to 2000 when field data were acquired.

The time of reaching silking or heading was similar in irrigated sites, and variable

![Map of Texas with study sites](image)

**Figure 1.** Locations of study sites in Texas. M: maize, and S: sorghum (lines are county boundaries).
in non-irrigated sites. The crop year 2000 was the same as previous years in Texas: same crop species and similar planting time (personal communication from field researchers). In addition, farming practices and crop management were very similar for these years. Therefore, the field data of 2000 were estimated by taking the average of the last four- or five-year field data of the study sites (table 1).

3. Results and discussion

3.1. Irrigated sites

The Springlake study site was planted with maize in the middle of April (about 20 April). Previous field data showed that the silking (tasseling) stage of maize occurred about 78 days after planting. The silking date is near the time when maize reaches maximum LAI. The LAI decreased slowly after silking, because leaves started withering and turning yellow. Since NDVI values are related to LAI data, the highest NDVI value was expected to be close to the time of maximum LAI, starting on 7 July 2000. The Lubbock study site was planted with sorghum in the middle of June (about 19 June). Field data showed that the sorghum starts heading about 57 days after planting. In addition, the sorghum reaches maximum LAI when heading starts. Observed NDVI values would be expected to have maximum values starting on 15 August 2000. The farming practices in irrigated sites are more predictable, and the estimation of the crop calendar for 2000 was more reliable than in non-irrigated sites.

The conditional composites of NDVI derived from surface reflectance using only clear-sky data were created for both irrigated sites (figure 2). The results showed that the maximum NDVI occurred between 8 July and 29 July 2000 in the maize site in Springlake and between 19 August and 2 September 2000 in the sorghum site in Lubbock. The maximum NDVI occurred very close to or at the time of maximum LAI. For the conditional NDVI composites in Springlake, a total of four NDVI values were missing on the weeks of 13, 20 May and 10, 24 June because of cloud contamination for an entire week. The cloud effect also caused two weekly NDVI values to be missing in the Lubbock site data. The conditional NDVI composites for the irrigated sites did not behave as an expected smooth curve after cloud detection and cloud removal. The slight NDVI fluctuation between weeks could be the vegetation response to environmental change.

3.2. Non-irrigated sites

The study site at Gregory was planted with sorghum between March and July every year. According to previous field data, LAI was estimated to reach a maximum value starting on 19 May 2000, about 76 average days after planting. The sorghum at the Granger site had a similar planting calendar to the Gregory area. The highest LAI was assumed to occur on 6 June 2000, about 92 average days after planting.
Both sites were located in dry-land areas without irrigation. The time to reach heading was very different from year-to-year in non-irrigated sites depending on the soil-water holding capacity and the amount of local rainfall.

The conditional NDVI composites based on surface reflectance indicated that maximum NDVI values for both non-irrigated sites occurred in the week of 10 June (figure 3). For the Gregory site, the maximum NDVI occurred about 2 weeks later than the estimated greatest LAI occurred. The maximum NDVI and the maximum LAI occurred at a similar time at the Granger site. Three out of four study sites showed that the conditional NDVI profiles provided information parallel to the field data. The two-week time difference that occurred in the non-irrigated site was most likely caused by incorrect estimation of the heading date. The non-irrigated Gregory site in 2000 very possibly experienced moisture stress during the growing season. Weekly NDVI values fluctuated more frequently in non-irrigated sites than in irrigated sites because of the limited water supply. The sorghum had maximum NDVI values about 0.7 in both irrigated and non-irrigated sites, and the maize had a lower maximum NDVI value of about 0.65. Although some of the weekly NDVI values were missing because of cloud contamination effects, the conditional NDVI profiles still provided useful information for crop monitoring.
3.3. Atmospheric influence on NDVI profiles

The traditional NDVI composite products derived from top-of-atmosphere (TOA) reflectance and surface (SUR) reflectance are presented in figure 4. Although both NDVI profiles had the same patterns throughout the crop seasons, the surface reflectance produced higher NDVI values. The main reason was that water vapour absorption and Rayleigh scattering strongly reduce measurements of the near-infrared channel (0.75–1.1 µm) observed by the satellite sensors. The visible channel (0.58–0.68 µm) was weakly affected by ozone absorption. As a result, the near-infrared reflectance observations are more affected by atmospheric effects than the visible reflectance observations. The near-infrared reflectance increasingly improved after atmospheric correction, and NDVI values were therefore increased.

Chilar and Huang (1994) used TOA reflectance data to generate composites based on maximum NDVI values, because atmospheric correction may introduce a systematic bias towards the selection of off-nadir pixels for the composites. However, this present study showed that the NDVI profiles based on TOA and surface reflectance had the same patterns over the growing seasons, which may be a result of the constrained sun and satellite zenith angles. As suggested by Lovell and Graetz (2000), the study was restricted to pixels with a maximum scan angle of 42° for composites. In addition, the scan angle effect on solar reflectance in low- to mid-latitude areas was not as strong as in high-latitude areas, such as Canada. Further studies are necessary to identify a suitable scan angle threshold for maximum NDVI compositing for the state of Texas.

Figure 4. Traditional NDVI seasonal profiles derived from top-of-atmosphere (TOA) and surface (SUR) reflectance at (a) Springlake and Lubbock sites and at (b) Gregory and Granger sites.
3.4. Conditional vs traditional NDVI composites

The traditional weekly NDVI composite product based on surface reflectance was strongly influenced by clouds, for example the data of 24 June at the Springlake site and 13 May at the Gregory site. Some maximum weekly NDVI values were lower compared to the weeks before and after because of cloud influence as well; for example, the data of 10 June at the Springlake site and 1 April at the Gregory site. Although the MVC method was developed primarily to eliminate cloudy pixels, the results of traditional NDVI composites suggested that a composite time of seven days was too short to obtain cloud-free data. The choice of a proper composite time period has been important in terms of excluding cloudy pixels and observing meaningful change (Eidenshink and Faundeen 1994). The MVC method is applicable to traditional composites only when a minimum of one clear-sky pixel was available in a week.

Besides cloud contamination, insufficient water supply, insect damage, nutrient shortage and other environment-related factors could also affect vegetation growth. The unexpectedly low NDVI values could have occurred because of cloud contamination or environmental change. Observing the seasonal NDVI profile is one approach to monitoring vegetation growth and studying the interaction between vegetation and related environmental factors. Hence, the process of cloud removal would certainly eliminate cloud influence on NDVI values. An ideal crop NDVI profile should be a relatively smooth curve during the growing season. The conditional NDVI composites (figures 2 and 3) illustrated temporal profiles close to real field conditions. The results from conditional NDVI composites indicated that the automated cloud detection algorithm removed most cloud-contaminated pixels accurately. In addition, the restrictions on scan and sun angles in this study also helped to refine data accuracy. For crops with growing seasons of a few months, crop information on a weekly basis can help to determine precise growth conditions. These results indicate that conditional weekly NDVI composites were appropriate for monitoring crop growth, while the traditional composite data resulted in NDVI values that fluctuated frequently over a few weeks because of cloud effects. The traditional NDVI compositing method applied over longer composite periods has a higher possibility of excluding cloud-contaminated pixels, but is not suitable for critical, real-time crop monitoring. The long period composites can be applied to forest, soil or water shed studies, because these land features seldom exhibit significant change within short time periods.

3.5. Conditional composites vs field data

The field data showed that sorghum needed four to five more weeks to reach heading in non-irrigated sites than in irrigated sites. Furthermore, the time needed to reach heading was more consistent in irrigated sites than in non-irrigated sites. It was apparent that water supply was an important factor for sorghum growth, although farming practices, soil properties and other climate factors might cause time differences. The conditional NDVI composites revealed that it took sorghum nine weeks to start reaching maximum NDVI in the irrigated site and 13–14 weeks in the non-irrigated sites. Although missing NDVI values caused information loss in the current study, it could be corrected by observing the values of one week before and after. Overall, the conditional NDVI composites accurately depicted in-field crop conditions.
4. Conclusions

The conditional weekly NDVI composites produced in this study appear to be a promising tool for crop monitoring in Texas. The NDVI composites agreed with the field-measured LAI data, reaching maximum values at similar times. The weekly traditional NDVI composites with considerable noises caused by clouds were not suitable for short-term crop monitoring. The irregular up and down behaviour of the NDVI seasonal profile data could cause confusion in data interpretation. The results indicated that accurate crop monitoring was achievable using AVHRR data with proper requisite pre-processing including cloud detection and angle constraints. Automated cloud detection for single date AVHRR scenes is strongly recommended before producing NDVI composites. Many NOAA-14 AVHRR overpasses for Texas were received in the late afternoon when the solar zenith angle was greater than 45°, and the reflectance was not as high as for data received around noon when the solar zenith angle was close to 0°. This data behaviour is typical for the long time series acquired from a single satellite that drifts into a later-and-later orbit throughout its lifetime. The NOAA-16 satellite was launched in September 2000, and became officially operational in early 2001. NOAA-16 ascending afternoon orbits occur between 1330 h and 1430 h Texas time, which is anticipated to provide better quality data for vegetation studies in the future.

Acknowledgments

The authors would like to thank the Texas Agricultural Experiment Station (TAES) for supporting a portion of this research, Dr H. Rahman and Dr G. Dédieu of Centre National d'Etudes Spatiales in France for providing a simplified method for atmospheric correction (SMAC), Dr Robert Baker of the Department of Forest Science at Texas A&M University in the USA for providing many helpful comments and PCI Geomatics Group in Canada and Quorum Communications Inc. in USA for providing technical assistance in using their software.

References


