SPARC: New Cloud, Snow, and Cloud Shadow Detection Scheme for Historical 1-km AVHHR Data over Canada

KONSTANTIN V. KHLOPENKOV AND ALEXANDER P. TRISHCHENKO

Canada Centre for Remote Sensing, Earth Sciences Sector, Natural Resources Canada, Ottawa, Ontario, Canada

(Manuscript received 14 February 2006, in final form 10 July 2006)

ABSTRACT

The identification of clear-sky and cloudy pixels is a key step in the processing of satellite observations. This is equally important for surface and cloud–atmosphere applications. In this paper, the Separation of Pixels Using Aggregated Rating over Canada (SPARC) algorithm is presented, a new method of pixel identification for image data from the Advanced Very High Resolution Radiometer (AVHRR) on board the NOAA satellites. The SPARC algorithm separates image pixels into clear-sky and cloudy categories based on a specially designed rating scheme. A mask depicting snow/ice and cloud shadows is also generated. The SPARC algorithm has been designed to work year-round (day and night) over the temperate and polar regions of North America, for current and historical AVHRR/NOAA High-Resolution Picture Transmission (HRPT) and Local Area Coverage (LAC) data with original 1-km spatial resolution. The algorithm was tested and applied to data from the AVHRR sensors flown on board NOAA-6 to NOAA-18. The method was employed in generating historical clear-sky composites for the 1982–2005 period at daily, 10-day, and monthly time scales at 1-km resolution for an area of 5700 km x 4800 km centered over Canada. This region also covers the northern part of the United States, including Alaska, as well as Greenland and the surrounding oceans.

The SPARC algorithm is designed to produce an aggregated rating that accumulates the results of several tests. The magnitude of the rating serves as an indicator of the probability for a pixel to belong to the clear-sky, partly cloudy, or overcast categories. The individual tests employ the spectral properties of five AVHRR channels, as well as surface skin temperature maps from the North American Regional Reanalysis (NARR) dataset. These temperature fields are available at 32 km x 32 km spatial resolution and at 3-h time intervals. Combining all test results into one final rating for each pixel is beneficial for the generation of multisensor clear-sky composites. The selection of the best pixel to be used in the final clear-sky product is based on the magnitude of the rating. This provides much-improved results relative to other approaches or "yes/no" decision methods.

The SPARC method has been compared to the results of supervised classification for a number of AVHRR scenes representing various seasons (snow-free summer, winter with snow/ice coverage, and transition seasons). The results show an overall agreement between the automated (SPARC) and the supervised classification at the level of 80% to 91%.

1. Introduction

A primary component of any satellite data processing system is the pixel identification procedure. Knowledge of the pixel state, such as the presence or absence of clouds, snow, and ice within the pixel field of view, is critical for many applications. This information is used to separate pixels into clear-sky or cloudy categories. Once it is known that a pixel belongs to the clear-sky category, it can be used for retrievals of surface reflectance, surface temperature, emissivity, or more complex properties such as the leaf area index and the fraction of photosynthetically active radiation (Cihlar et al. 2004). The clear-sky pixels, as measurements of the top-of-the-atmosphere (TOA) radiance field, include information about atmospheric aerosols, vertical profiles, and column amounts of absorbers and, therefore, can be used in atmospheric sounding (Rogers 2000). The pixels containing clouds are employed for retrievals of cloud properties, such as cloud optical depth, water/ice

The Advanced Very High Resolution Radiometer (AVHRR) aboard the National Oceanic and Atmospheric Administration (NOAA) spacecraft is perhaps the most widely used satellite sensor for a broad range of applications (Cracknell 1997; Kidwell 1998; Goodrum et al. 2000). AVHRR data have been available since the early 1980s, thus providing a very long record of multispectral satellite observations with moderate spatial resolution. The need for the new cloud identification algorithm described in this paper stems from the development of the new AVHRR data processing system at the Canada Centre for Remote Sensing (CCRS) of the Department of Natural Resources Canada (NRCan) in the framework of the Climate Change program. To study the long-term changes of Canada’s landmass at high spatial and temporal resolution, it was decided to assemble and reprocess all available AVHRR observations at their original spatial resolution (~1.1 km × 1.1 km). The new AVHRR data processing system called the Earth Observation Data Manager (EODM) was designed and employed for this purpose (Latifovic et al. 2005). As one of its components, the Separation of Pixels Using Aggregated Rating over Canada (SPARC) algorithm has been developed for scene pixel identification as an advancement to the scheme employed originally at CCRS for generating clear-sky composite images (Cihlar et al. 2004). Previously, a straightforward approach was used based on the selection of the maximum Normalized Difference Vegetation Index (NDVI) and a simple threshold in the visible channel combined with postseasonal time series analysis. The SPARC algorithm is much more comprehensive. It utilizes all five AVHRR channels as well as some additional inputs. We attempted to make use of the best features of several known algorithms for cloud identification from AVHRR data and adapted them to our study area. The following schemes were considered: Clouds from AVHRR (CLAVR; Stowe et al. 1999), International Satellite Cloud Climatology Project (ISCCP; Rossow and Gardner 1993), the AVHRR Processing Scheme over Clouds, Land and Ocean (APOLLO; Kriebel et al. 2003), the European Organization for the Exploitation of Meteorological Satellites (EUMETSAT) scheme and the Swedish Meteorological and Hydrological Institute (SMHI) Cloud Analysis Model Using Digital AVHRR Data (SCANDIA) algorithm (Dybroe et al. 2005a; b; Karlsson 1996), the Clouds and the Earth’s Radiant Energy System (CERES) cloud analysis and determination scheme (CERES 1995), and the Support of Environmental Requirements for Cloud Analysis and Archive (SERCAA) scheme (Gustafson et al. 1994). Some ideas proposed in the Moderate Resolution Imaging Spectroradiometer (MODIS) cloud identification scheme (Ackerman et al. 2002; Platnick et al. 2003) were also considered. These include the need to detect shadows, which can produce substantial noise in surface and cloud property retrievals if not accounted for properly, and weighting the results in the groups of independent tests to get the final level of confidence in the clear-sky identification.

The paper is structured as follows. Section 2 contains an overview of the major cloud detection algorithms. Section 3 introduces and discusses the idea of aggregated rating as an alternative to “yes/no” branching approaches. Section 4 describes the basic features of the SPARC algorithm. Section 5 discusses the implementation of snow and ice identification, and section 6 lists a few additional tests of the algorithm. Section 7 explains the purpose and implementation of several correction factors required for better sensitivity to clear-sky conditions in some special situations such as sunglint areas. Section 8 describes the cloud shadow detection. Section 9 contains the results of validation against an independent approach based on the supervised classification. Section 10 concludes the paper with a summary and a discussion of the results.

2. Overview of major cloud detection algorithms

The basic idea in discriminating cloudy pixels from clear-sky pixels relies on the spectral information available from the AVHRR channels. The AVHRR instruments contain four, five, or six channels depending on the radiometer model. The AVHRR-1 radiometers flown onboard NOAA-6, -8, and -10 were four-channel instruments. They included channel 1 [visible (VIS): 0.58–0.68 μm], channel 2 [near-infrared (NIR): 0.725–1.10 μm]; channel 3 (shortwave infrared (SWIR): 3.55–3.93 μm]; and channel 4 [infrared (IR 4): 10.5–11.5 μm]. Channel 5 in the AVHRR-1 imagery was a repeat of the channel-4 data. The AVHRR-2 version of the instrument flown onboard NOAA-7, 9, 11, 12, 14 had one more channel than AVHRR-1: channel 5 (infrared (IR 5): 11.5–12.5 μm). The latest AVHRR-3 instruments onboard NOAA-15, -16, -17, and -18 have an additional sixth channel, 3A, that covers the wavelength range 1.58–1.64 μm in the shortwave infrared part of the spectrum. When channel 3A was introduced, the SWIR channel 3 (3.55–3.93 μm) was renamed channel 3B. Channels 3A and 3B on AVHRR-3 work interchangeably because only five channels can operate at any time. There are slight differences in the spectral response
functions of the same channels for different AVHRR radiometers that need to be taken into account when studying long-term trends (Trishchenko et al. 2002a; Trishchenko 2006). The calibration procedure and derivation of radiometric quantities, such as reflectance and brightness temperature for the corresponding AVHRR channels, are described by Latifovic et al. (2005).

Although a limited number of radiometric tests can be applied to five AVHRR channels, their implementation differs among various algorithms. The major principle in detecting clouds is that clouds are bright objects in the solar-reflective bands but appear cold in the thermal bands. However, this simple idea may be invalid under certain conditions. The pixels corresponding to nonvegetated targets, especially targets representing arid areas, can also have high reflectance, and thus, be easily confused with clouds in the VIS and NIR bands. Sun glint areas over ocean are also quite bright and require special processing. The temperature contrast between low-altitude stratiform clouds and the surface can be too small for distinguishing between cloudy and clear-sky pixels based solely on their temperatures. Temperature inversions in the boundary layer can also complicate the analysis, because they cause lower-level clouds to appear warmer than the surface. The cases of snow and ice require special attention too. These targets are frequently brighter than clouds and are quite cold, and thus, a simple threshold based on the magnitude of reflectance in VIS or NIR channels and a temperature test can classify snow/ice pixels as clouds. Instead, a combination of channels is required to make a more reliable discrimination. For example, the difference in brightness temperature between AVHRR channels 4 and 5 (hereafter denoted as \( T_4 - T_5 \)), which corresponds to the difference in the spectral emissivity of clouds at these wavelengths, is employed for thin cloud (cirrus) detection (Stephens 1994). Channels 3A/3B need to be utilized together with other channels for snow/ice identification as they reveal unique spectral features of snow and ice in the SWIR bands.

Several schemes have been proposed over the years to deal with scene pixel identification in AVHRR observations. The earliest schemes were described by Coakley and Bretherton (1982) and Arking and Childs (1985). Below, we characterize several schemes that received broad attention in the AVHRR user community. The list is not comprehensive and serves mostly to assist our explanation of SPARC’s functionality.

a. **APOLLO scheme**

The APOLLO scheme was among the first to employ all five AVHRR channels for scene pixel identification (Saunders and Kriebel 1988). This scheme is designed to work for full spatial resolution High-Resolution Picture Transmission (HRPT) and Local Area Coverage (LAC) imagery as well as for reduced spatial resolution Global Area Coverage (GAC) data formats. Four categories of pixel scenes are considered: fully cloudy, partially cloudy, cloud free, and snow–ice. The APOLLO scheme employs a sequence of threshold tests to determine pixel status. Four general categories of the clear-sky spectral test are introduced: 1) ocean surfaces, 2) vegetated land, 3) arid land, and 4) snow and ice (Kriebel 1996; Kriebel et al. 2003). The pixel identification is carried out in three stages. During the first stage, up to five threshold tests are applied to each pixel. If all the tests point consistently to the clear-sky category, then the pixel is considered to be clear-sky. All other pixels with at least one clear-sky test failed are labeled as cloud contaminated. During the second stage, the threshold tests are reapplied to all pixels identified as cloud contaminated to separate the overcast pixels from the partially cloudy pixels. In the third stage, the identification of snow/ice pixels is carried out.

The five clear-sky tests that make up the first stage of APOLLO utilize AVHRR channels 1, 2, 4, and 5. These tests are 1) the gross temperature test, which thresholds the brightness temperature \( T_4 \); 2) the spatial coherence thermal test over sea surface, which thresholds the standard deviation of temperature; 3) the thin cirrus detection based on the difference \( T_4 - T_5 \); 4) the dynamic visible threshold test the reflectance in channel 1 \( r_1 \); and 5) the dynamic ratio test of reflectance in channels 1 and 2 \( r_1/r_2 \) over land and water. To set up the threshold values used in these tests, the APOLLO scheme calculates the statistics over a part of the image (either as large as one-third of the scene, or as small as 50 \( \times \) 50 pixels). In the second stage, in order to identify the fully cloudy pixels among the group of cloud-contaminated pixels, the dynamic ratio test (5) is repeated with a threshold generated from the statistics, and the spatial coherence test (2) is run again with slightly modified thresholds (Kriebel et al. 2003).

Snow–ice identification conducted during the third stage relies on the use of channel 3 (A or B, whichever is available). These tests modify the classification label of cloudy pixels if they are cold \( T_3 \), if they are bright in channels 1 and 2, and if the ratio of reflectances in channels 1 and 3 or channels 2 and 3 is above a threshold.

Despite being quite a sophisticated scheme, the APOLLO algorithm was developed using a threshold approach, and in the end provides only the pixel status. It is necessary to implement an additional analysis of the original data if some type of the image mosaic (i.e.,
clear-sky composite) needs to be generated. APOLLO lacks cloud shadow detection, which is required to produce good-quality clear-sky data products (Latifovic et al. 2005; Ackerman et al. 2002). Multiple passes through the data are required to build and apply the thresholds in the APOLLO algorithm, which makes the scheme demanding in terms of computational resources.

b. CLAVR scheme

CLAVR-1 (and its advanced version CLAVR-x described by Heidinger (2007)) is a major cloud identification algorithm employed by NOAA in global AVHRR data processing (Stowe et al. 1999). CLAVR is designed to work with AVHRR GAC data, which are available globally, unlike LAC and HRPT data collected at local scales. The GAC pixels are produced as an average of four AVHRR pixels along the scan line. As such, they are approximately 4.4 km × 1.1 km in the nadir direction. The GAC pixels are separated by one high-resolution pixel and produced for every third AVHRR scan line. Therefore, the GAC data do not actually represent contiguous imagery. The GAC image remapped onto a regular geographical projection would appear like a sieve with holes.

The CLAVR-1 algorithm provides a mask containing clear, mixed, and cloudy flags for an array of 2 × 2 GAC pixels. The CLAVR-x splits the mixed category into partly clear and partly cloudy categories. Like the APOLLO scheme, CLAVR uses a sequence of multispectral threshold tests to derive the cloud mask. CLAVR’s pixel identification puts major emphasis on the reliability of the clear-sky mask and ensures that no cloudy pixel is identified as clear sky, that is, the scheme is designed to be clear-sky conservative. Because of its operational near–real time application, CLAVR has strict requirements for speed and avoids multiple passes through the data. Ancillary datasets, such as surface type maps, digital elevation maps (DEM), and some monthly surface climatological fields, are used to set up the thresholds and to turn certain tests on and off. Daytime and nighttime scenes are processed. Several groups of tests in a decision tree are applied in CLAVR. They include the following tests [based on the CLAVR-x scheme (Heidinger 2007)]:

1) the gross contrast test for reflectance (NIR over water and VIS over land) and temperature (T_d) to identify bright and cold pixels corresponding to clouds;
2) the reflectance ratio (NIR/VIS) contrast test over water and land;
3) the channel 3A/3B albedo test over water and land;
4) the T_d–T_s or cirrus detection test to find large differences between brightness temperature in channel 3B and brightness temperature in channel T_s. The test is also used for the additional detection of optically thick clouds and clear-skies;
5) the uniform low stratus test to detect low (negative) values of temperature difference (T_d–T_s);
6) the T_d–T_s test to detect clear sky (small positive difference), thin cloud (large positive difference), or thick cloud (near zero or negative difference);
7) the channel 3B emissivity test: emissivity is below 1 for opaque water clouds and above 1 for semitransparent clouds;
8) the climatological sea surface temperature (SST) test;
9) the relative thermal gross contrast; and
10) the spatial uniformity tests (clear sky and cloudy, reflectance and temperature).

An updated scheme that uses the 8-day rotating clear-sky radiation dataset and dynamic thresholds derived from these data has also been proposed and is known as CLAVR-3 (Vemury et al. 2001).

c. EUMETSAT and SCANDIA schemes

The SCANDIA cloud identification scheme (Karls- son 1996) was a starting point in the development of a new cloud identification scheme in the framework of EUMETSAT Satellite Application Facility (SAF) projects (Dybbroe et al. 2005a,b). The SCANDIA scheme is similar to the schemes described above in many ways, except that it attempts to couple a series of tests together rather than to apply the threshold tests individually. New algorithms proposed (Dybbroe et al. 2005a,b) are designed to work for current and future AVHRRs onboard the NOAA satellites and EUMETSAT Meteorological and Operational Weather Satellites (METOPs). The EUMETSAT AVHRR cloud detection and analysis algorithm employs several dynamic input layers: water vapor column amount, surface skin temperature, and the temperatures at the 950, 850, 770, and 500 hPa and tropopause levels. The use of dynamic layers that define the current state of the atmosphere and surface thermal conditions increases the reliability of cloud/clear-sky detection and the determination of cloud types. The dynamic layers are derived from a short-range forecast of the High-Resolution Limited-Area Model (HIRLAM), which provides hourly forecasts at 44-km horizontal resolution. Substantial emphasis in the new EUMETSAT scheme is placed on cloud type determination. The scheme does not use AVHRR channel 2 (NIR) in the pixel identification, which may be considered as a weakness.
nals 1 and 2 are indeed very similar for cloudy scenes, channel 2 provides unique information for the clear-sky pixels over land. The EUMETSAT scheme does not include cloud shadow detection.

d. ISCCP scheme

The ISCCP scheme uses only two spectral channels: visible and thermal IR (11 μm). These channels are available either from AVHRR GAC data or from the data provided by geostationary satellites. The scheme is applied to pixels with the size of 4–7 km representing the regions of ~30 km × 30 km (Rossow and Gardner 1993; Rossow and Schiffer 1999). The ISCCP scheme employs a temporal approach to separate cloudy and clear-sky pixels instead of relying exclusively on spectral information. The ISCCP scheme includes five major steps: 1) the gross spatial thermal contrast test, 2) the gross temporal thermal contrast test, 3) the generation of spatiotemporal statistics for both thermal and visible channels, 4) the identification of clear-sky thresholds using the results of the previous step, and 5) the classification of pixels into three categories: clear, cloudy, and marginally cloudy using the derived thresholds.

The gross spatial thermal contrast test (step 1) identifies pixels as cloudy if they are much colder than the others over a small spatial domain. The gross temporal thermal contrast test (step 2) is applied to a sequence of images over a 3-day interval and identifies a pixel as cloudy if it has sharply lower IR radiances compared to 1 day earlier or later. The generation of spatiotemporal statistics in step 3 is conducted over 5-day time intervals. During the final classification (step 5), the pixel is placed into the clear-sky (cloudy) category if visible and IR radiances pass the clear-sky (cloudy) thresholds. If the radiances fall in between, the pixel is assigned to the marginally cloudy category.

Although quite a robust scheme, the ISCCP approach has some limitations. One pixel of the size approximately 4–7 km represents quite a large area (~30 km × 30 km) that may reveal substantial variability in albedo and thermal state (e.g., the water–land mix). In such a case, it is possible that some clear-sky pixels are classified as marginal cloudy or cloudy, and vice versa. The scheme does not account for the temperature inversion when low-altitude clouds are warmer than the cold surface. This can be quite a frequent phenomenon in polar regions. The ISCCP scheme does not support cloud shadow identification. Cloud optical depth retrievals for such large pixels can lead to results that are biased toward smaller optical depths because a pixel identified as cloudy may actually represent a mixed scene (Wielicki and Parker 1992).

On the other hand, the scheme may miss some thin clouds, which will be detected as clear skies (CERES Science Team 1995).

e. CERES scheme

The CERES cloud identification scheme was developed to improve the determination of top-of-the-atmosphere radiative fluxes derived from coarse spatial resolution broadband CERES radiometer observations onboard the Tropical Rainfall Measuring Mission (TRMM), and the Terra and Aqua spacecrafts. The improvement is achieved through the use of high-resolution imagery available over the CERES pixels from the concurrent Visible Infrared Scanner (VIRS) on TRMM and/or MODIS on Terra/Aqua. Although it was designed to be coordinated with MODIS cloud detection, the CERES cloud identification scheme was developed and implemented independently of the MODIS scheme. A reason for this is that MODIS data are not available for CERES/TRMM data processing. The high-resolution imagery for CERES on TRMM is provided by the VIRS data processing.

The CERES cloud identification scheme has inherited many features of the CLAVR, ISCCP, MODIS, and SERCAA (Gustafson et al. 1994) cloud schemes. It was tested initially on AVHRR GAC data.

f. MODIS scheme

At the time of writing, the MODIS cloud detection scheme is probably the most comprehensive cloud detection scheme in terms of the amount of spectral information utilized (Ackerman et al. 2002). This scheme employs information from 19 out of 36 MODIS channels. It also requires several ancillary inputs, such as topography and geometry of observation for each 1-km pixel, land/water and ecosystem maps, and daily operational snow/ice data products from the NOAA and National Snow and Ice Data Center (NSIDC). The resulting product of the MODIS cloud identification is a 48-bit cloud mask that contains confidence flags (confident cloudy, uncertain, probably clear, and confident clear) and other flags indicating high cloud type, shadow, thin cirrus, snow/ice, sun glint, and results from the other tests, including the 16 values of the cloud flags for all 250 m × 250 m subpixels within the 1 km × 1 km field of view.

A unique and important feature of the MODIS cloud algorithm is the attempt to detect cloud shadows and the emphasis on their importance in producing uncontaminated clear-sky surface composite products. Cloud shadow detection is implemented in MODIS using the spectral (not geometrical) approach. The algorithm checks for cloud shadows once a confident clear-sky
pixel is found. Cloud shadow is detected if reflectance in the 0.94-μm channel is less than 0.07, the ratio of reflectances at 0.87 and 0.66 μm are greater than 0.3, and the reflectance in the 1.2-μm channel is less than 0.2 (Ackerman et al. 2002). This approach may confuse cloud shadows with shadows caused by uneven terrain and may also detect false shadows when the spectral signature of a clear-sky pixel is similar to a shadow pixel.

3. Aggregated rating versus branching approach

Most of the cloud detection schemes as described above are implemented as threshold/branching algorithms. At each stage of the scene identification process, a pixel is assigned to a certain category, for example, cloudy, clear-sky, or partly cloudy. The pixel’s status can remain uncertain if the pixel does not meet certain criteria. A series of sequential tests is applied to determine the pixel status. Although the pixel can change its status during the scene identification process, and can correspond to a variety of atmospheric and surface conditions within the same class, the resulting cloud mask may contain only the flags corresponding to the few selected categories.

This approach misses two important points. First, since each test makes a yes/no decision (i.e., cloud versus clear, snow versus snow-free, etc.), a pixel corresponding to an intermediate state may be classified incorrectly or its state may still remain uncertain. The second problem is related to the limited number of categories used for classification. A cloud mask corresponding to a small number of classes carries limited information about the degree of cloud or aerosol contamination (or the optical thickness) of the pixel. This information is not readily available after the cloud mask is produced and has to be retrieved by repeating the image processing. The degree of cloud or haze contamination may be important for certain applications, such as the generation of clear-sky composite images, or retrievals of surface properties or optical depth.

In the SPARC algorithm, the yes/no branching approach is replaced by the idea of a cumulative or aggregated rating, which is formed by summing the scores produced by the individual tests. With this convention, the score yielded by each test reflects the degree of a pixel’s cloud contamination. This idea is somewhat similar to the MODIS approach to computing the confidence level (from 0 to 1) for some tests (Ackerman et al. 2002). The MODIS scheme introduces three parameters (α, β, and γ) to determine the confidence level. The parameters α and γ define the cloudy conditions (confidence level 0) and the clear-sky conditions (confidence level 1), while β determines the location of the pass or fail threshold within the interval [α, γ]. The confidence level is assumed to be a linear function ranging from 0 to 1 on the interval [α, γ]. The SPARC algorithm extends the idea of the MODIS confidence levels by generating test scores and summing them to one final aggregated rating. Another approach for building a quantitative estimator is the method of the maximum likelihood estimator (MLE) as proposed by (Wielicki and Green 1989) for the Earth Radiation Budget Experiment (ERBE). The MLE estimator is also implemented for scene identification in several other schemes used in radiation budget research, such as the Scanner for Radiation Budget (ScaRaB; Kandel et al. 1998) and the ERBE-like data processing system of CERES (Wielicki et al. 1996). The new CERES data processing system uses the CERES cloud identification scheme mentioned above and a neural network method for radiance-to-flux conversion (Loukachine and Loeb 2004).

The aggregated rating method proposed in the SPARC algorithm has two major advantages. First, it provides the quantitative information (rating) for a straightforward selection of the best clear-sky pixel among several candidates. Also, by thresholding this rating at different levels, one can easily mask the pixels with the desired degree of confidence of cloud contamination, which can be useful for analysis of cloud fraction. Second, our cloud detection scheme incorporates the results of all major cloud tests to ensure the high reliability of the scene identification even if a single test fails. In the case of a yes/no decision and branching method, the failure of one particular test at any stage of the classification may carry the routine away from the correct branch, and to a wrong classification in the end. Although there are so-called “restore” tests in some schemes that can revert the pixel’s status under uncertain conditions, the yes/no branching methods are in general more vulnerable to uncertainties in the thresholds.

4. Description of the SPARC algorithm

The central part of the SPARC algorithm includes three major tests. These three major tests are

(a) the brightness temperature test in channel 4, which produces the $T$ score;
(b) the reflectance brightness test in channel 1 (VIS), which produces the $B$ score; and
(c) the reflectance test in channel 3 (A or B), which produces the $R$ score.
Several additional tests \((A_i)\) can also contribute to the aggregated rating depending on the spectral and auxiliary input information.

The central idea of the SPARC algorithm is to produce an overall rating \(F\), which corresponds to the degree of cloud contamination and is a sum of the scores generated by the individual tests:

\[
F = T + B + R + \sum_i A_i. \tag{1}
\]

This summation is possible because of the way that the scores are calculated. Each score has an offset that makes its value negative for clear-sky scenes and positive for cloudy scenes. The values near zero correspond to an intermediate state. By construction, the higher the rating is, the greater the chance that a pixel contains clouds. The significance of each test in the total sum can be adjusted by scale factors in each term of Eq. (1). Details of the implementation for each test are explained further.

The temperature \((T)\) test uses the brightness temperature in channel 4 \((T_4)\) and compares it with a dynamic threshold. The dynamic threshold in the SPARC scheme is determined using the surface skin temperature data \(T_{\text{NARR}}\) from the North American Regional Reanalysis (NARR; Mesinger et al. 2006). The NARR data are available every 3 h at a spatial resolution of 32 km \(\times\) 32 km. Because the AVHRR data are available at higher spatial resolution and may be collected at different times than the NARR fields, the NARR temperature data are interpolated to the time of the AVHRR image acquisition and to the 1 km \(\times\) 1 km spatial grid using a trilinear interpolation routine (time, latitude, and longitude). Cubic spline interpolation was also tested instead of the trilinear method, but it was found to produce nonphysical distortions such as overshooting when interpolating the large temperature variations in coastal areas.

The \(T\) score is calculated as follows:

\[
T = (T_4 - T_{\text{NARR}} - T_{\text{offs}})T_{\text{scale}}, \tag{2}
\]

where \(T_{\text{offs}}\) and \(T_{\text{scale}}\) are the offset and the scale factor. Their values are provided in Table 1. The scale factor for the \(T\) test is negative, which results in a higher score value for a lower observed temperature, meaning a higher probability of cloud contamination for colder pixels (Stowe et al. 1999; Ackerman et al. 2002). The parameters listed in Table 1 were derived from extensive empirical analysis and then validated against supervised classification results (see section 9).

The brightness \(B\) test uses the reflectance \(r\) in AVHRR channel 1 (VIS) over land and channel 2 (NIR) over water because channel 2 is less affected by aerosols and Rayleigh scattering. The AVHRR NIR channel 2 cannot be used over land with the same efficiency as the visible channel because of its high values over vegetated areas. The \(B\) score is calculated as follows:

\[
B = (r - B_{\text{offs}})B_{\text{scale}}. \tag{3}
\]

The values of \(B_{\text{offs}}\) and \(B_{\text{scale}}\) are given in Table 1. Again, a negative value of the \(B\) score indicates a predominantly clear-sky condition (low reflectance) and a positive value corresponds to a more likely cloudy condition (higher reflectance).

The \(R\) score is calculated from the reflectance in channel 3 \((A\) or \(B\), whichever is available). In case of channel 3B, which includes both the thermal and solar flux components, the reflectance \(r_3\) is calculated from the radiance as

\[
r_3 = \frac{R_{3B} - B(\lambda_3, T_5)}{B_{\text{sun}}(\mu_0) - B(\lambda_3, T_5)}, \tag{4}
\]

where \(R_{3B}\) is the radiance in channel 3B expressed in \(W/(m^2\ \text{str} \ \mu m)^{-1}\), \(B(\lambda, T)\) is the blackbody radiance at wavelength \(\lambda\) and temperature \(T\) (brightness temperature in channel 5), \(B_{\text{sun}}\) is the solar radiance corrected to the sun–earth distance in the same units as \(R_{3B}\), and \(\mu_0\) is the cosine of the solar zenith angle. Various parameters, such as central wavelength \(\lambda\) and solar radiance \(B_{\text{sun}}\) for the AVHRR channel 3B on various NOAA platforms can be found in Trishchenko (2006).

For a low-sun geometry, the reflected radiation component in the total radiance observed in channel 3B is much smaller than the thermal emitted component. In such a case, the presence of noise in channel 3B, which is known to be substantial (Simpson and Yhann 1994; Trishchenko et al. 2002b; Ignatov et al. 2005), may lead to negative reflectances calculated by the simplified model (4). This is shown in Fig. 1. According to our assumption, these negative values should point to the

<table>
<thead>
<tr>
<th>Test</th>
<th>AVHRR channels involved</th>
<th>Offset</th>
<th>Scale factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>(T)</td>
<td>Channel 4</td>
<td>-6.0 K</td>
<td>-0.42 K(^{-1})</td>
</tr>
<tr>
<td>(B)</td>
<td>Channel 1 for land</td>
<td>0.30</td>
<td>67</td>
</tr>
<tr>
<td>(Cannel 2 for ocean)</td>
<td>0.24</td>
<td>75</td>
<td></td>
</tr>
<tr>
<td>(R_A)</td>
<td>Channel 3A</td>
<td>0.1</td>
<td>160</td>
</tr>
<tr>
<td>(R_B)</td>
<td>Channel 3B</td>
<td>0.06</td>
<td>33 if (r_3 &lt; 0.06)</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>110 if (r_3 \geq 0.06)</td>
</tr>
<tr>
<td>(N)</td>
<td>Channels 1 and 2</td>
<td>0.1</td>
<td>-30</td>
</tr>
<tr>
<td>(U_{ext})</td>
<td>Channel 2 for ocean</td>
<td>0.0034</td>
<td>1200</td>
</tr>
<tr>
<td>(U_{comp})</td>
<td>Channel 4</td>
<td>3.5 K(^2)</td>
<td>1.33 K(^{-1})</td>
</tr>
<tr>
<td>(C)</td>
<td>Channels 4 and 5</td>
<td>1.5 K</td>
<td>4.0 K(^{-1})</td>
</tr>
</tbody>
</table>
clear-sky scene. In the case of noise, that is, a high level
of uncertainty in the magnitude of reflectance, they
cannot, however, be used with confidence. To minimize
this effect, the range of the solar zenith angle is re-
stricted to the interval where the denominator in Eq.
(4) is greater than some small value. Solving the equa-
tion for the solar zenith angle (SZA) \( \theta_0 \) when the de-
nominator of Eq. (4) is equal to zero gives

\[
\theta_0 = \arccos \left( \frac{B(\lambda_3, T_5)}{B_{\text{sun}}} \right). \tag{5}
\]

To prevent the division by zero, a slightly smaller level
of SZA was selected as a threshold. The following ap-
proximation was derived for limiting SZA as function
of \( T_5 \) (in kelvins):

\[
\theta_0^{\text{max}} = 89.5^\circ - 0.000025(T_5 - 223)^3. \tag{6}
\]

Figure 2 displays the behavior of \( \theta_0 \) and \( \theta_0^{\text{max}} \) as func-
tions of \( T_5 \).

Channel 3B reflectance is calculated only for pixels
with SZA smaller than the threshold given by Eq. (6).
The histogram of the channel reflectance calculated us-
ing Eq. (4) and the SZA range restricted by Eq. (6)
show a significantly improved distribution, as demon-
strated by the results presented in Fig. 1.

The magnitude of the \( R_B \) score for channel 3B is
calculated as

\[
R_B = (r_{3B} - R_{\text{offs}})R_{\text{scale}}, \tag{7}
\]

Coefficients for the scale and offset are provided in
Table 1. It is also worth noting the dual scale approach
for the channel 3B score: \( R_{\text{scale}} = 110 \) if \( r_{3B} > R_{\text{offs}} \) and
\( R_{\text{scale}} = 33 \) otherwise (see Table 1). The scaling factor
for the region \( r_{3B} \leq R_{\text{offs}} \) (or the negative values of \( R_B \-
score \)) is reduced in order to decrease the influence of
the lower level of reflectance \( r_{3B} \), which does not nec-
essarily mean a clearer pixel. The reflectance \( r_{3B} \) is cer-
tainly high for water droplet clouds, but can be low for
ice clouds. Therefore, the contribution of negative \( R_{A} \-
score \) scores to the overall rating should be reduced, and this
is achieved by introducing a smaller scale factor. Also,
this reduction decreases the effect of noise in channel
3B, which is more noticeable at lower levels (Simpson
and Yhann 1994). Channel 3B also demonstrates cer-
tain sensitivity to the presence of snow and thus can be
exploited for snow detection, as explained in the next
section.

The reflectance score \( R_A \) for channel 3A is con-
structed differently. As an example, Fig. 3 shows the
scatterplot of reflectance in AVHRR channel 3A (\( r_{3A} \))
versus the reflectance in the visible channel for several
scenes acquired by AVHRR NOAA-16. The figure shows
that pixels corresponding to specific surface types or cloud conditions are grouped into relatively
tight clusters. Cloudy pixels occupy the middle part of
the plot (both \( r_{3A} \) and \( r_1 \) are similar to each other and
both are high) and snow pixels are distributed mainly in
the lower-right section (\( r_1 \) is high while \( r_{3A} \) is low),
while pixels for a mixed case corresponding to semi-
transparent optically thin clouds over snow are located
between these two groups. The clear-sky snow-free pix-
 els are distributed in the left and low sections of the plot
where \( r_1 \) is usually low and \( r_{3A} \) does not have high
values.

Based on the \( r_{3A} \) versus \( r_1 \) scatterplot, shown in Fig.
3, the following function was constructed for computing
the \( R_A \) score:

\[
R_A = \left[ \frac{r_1^2 + r_{3A}^2}{2r_1^2 + 2r_{3A}^2} - R_{\text{offs}} \right]R_{\text{scale}}. \tag{8}
\]
The contour plot of the \( R_A \) score derived using Eq. (8) is superimposed on the scatterplot of Fig. 3. The shape of the isolines provides a good separation of the cloudy pixels from the others. The clear-sky pixels are assigned negative values of the \( R_A \) score and the cloudy pixels receive the positive score. The contour line \( R_A \) = 0 follows the approximate boundary between cloudy and clear-sky pixels. The highest score is achieved when a pixel’s spectral reflectance approaches the middle part of the plot around the diagonal. The direction of the gradient in the middle lower part of the plot represents the direction from the snow pixels through the thin cloud pixels toward pixels with more optically thick clouds. In the area where \( R_A > 0 \), the \( R_A \) score increases almost linearly toward the region of cloudy pixels, but for clear-sky pixels, the gradient is much lower, which is achieved by the use of the squared terms in Eq. (8). The reason for this difference is the same as for the two different scales in the \( R_B \) score: a lower reflectance in channel 3A should have a smaller effect on the overall rating.

5. Snow identification

A widely used approach in the identification of snow from satellite observations employs the Normalized Difference Snow Index (NDSI; Dozier 1989; Hall et al. 2002). Because it is not possible to implement the NDSI method in a systematic way for AVHRR data, it is necessary to construct an alternative approach that can be used interchangeably for channel 3A or 3B. We use the \( R_A \) and \( R_B \) scores combined with additional tests for this purpose.

Figure 4 shows the flowchart of the part of the SPARC algorithm responsible for snow detection. If the data from channel 3A are not available, the procedure begins with the testing of the solar zenith angle, which must be smaller than the threshold \( \theta_{\text{max}} \) given by Eq. (6), and then computes the reflectance \( r_{3B} \) by Eq. (4) and the \( R \) score from Eq. (7). If channel 3A is available, then the \( R \) score is calculated from Eq. (8). If the \( R \) score of a pixel is greater than the gross cloud limit (which is set to 12) then the pixel is considered to be cloudy and further tests are not applied in order to save processing time.

Once the \( R \) score is known to be smaller than the gross cloud limit, a two-stage snow test is applied. In the first stage, for a pixel to pass the snow test it has to be sufficiently bright (\( B_{\text{score}} \) score is more than 3.0 for pixels over land or more than \(-2.0 \) for water pixels), its \( R \) score has to be smaller than 3.0 (unlikely cloud contamination), and its brightness temperature has to be close to the surface level predicted by the temperature map (\( T_{\text{score}} \) score is less than 3.0), but lower than the freezing point \( T_{\text{froze}} \). The freezing point is calculated as a sine function oscillating between a maximum of \(+2^\circ\)C in the spring season and a minimum of \(-2^\circ\)C in the fall. This dependence approximately accounts for the thermodynamics of the freezing and melting phases. If snow is detected, then the pixel is marked with a snow flag and has to pass the second stage that analyzes the presence of thin clouds/haze over snow. During the second stage, two parameters (the \( R \) score and the \( T \) score) are analyzed. If both are smaller than zero, then the pixel is considered to be clear-sky snow, and the processing jumps to the thin cirrus test (described below). Otherwise, the pixel is considered to be snow covered and cloud contaminated, and undergoes additional cloud tests. If no snow is detected in the first stage, then the pixel is considered to be potentially cloud contaminated and undergoes additional tests, beginning with the simple ratio test.

6. Additional tests

The entire flowchart of the SPARC algorithm is shown in Fig. 5. The additional tests identified as \( A_i \) in Eq. (1) include the simple ratio test (see Fig. 4) that produces the \( N \) score, the uniformity test that produces the \( U \) score, and the thin cirrus (Ci) test that produces the \( C \) score. The \( N \), \( U \), and \( C \) scores are defined below. Not all of these tests may appear in the processing chain. Thus, they are considered as secondary and are used to provide more information about the pixel status.
in complex or uncertain situations. If these tests are applied, their score values are subsequently added to the aggregated rating $F$ calculated in Eq. (1). These tests are applied to the pixels that did not get definitive scene identification after the three major tests described in section 3 (approximately 10% of all pixels). For the majority of pixels, the additional tests are not used for computational efficiency.

a. Simple ratio test

The simple ratio test uses the ratio of the reflectance in channels 2 and 1, which is used to construct the following $N$ score:

$$N = \frac{|r_2/r_1 - 1| - N_{offs}N_{scale}}{N_{offs}}. \quad (9)$$

Because the simple ratio $\rho = r_2/r_1$ is uniquely related to the NDVI ($\text{NDVI} = (\rho - 1/\rho + 1)$), we use the notation $N$ for this score. The offset and scale factors are given in Table 1. Because $N_{scale}$ is chosen to be negative ($-30$), the $N$ score is larger when $\rho$ is closer to 1. The $N$ score decreases when the simple ratio deviates from unity in either direction.

The physical meaning of this test is to identify cloudy scenes, which are usually spectrally neutral, with a positive $N$ score. Its maximum value $N = 3$ is achieved when $r_1 = r_2$. A negative $N$ score is assigned to potentially clear-sky pixels when $\rho$ exceeds 1.1 (clear-sky land pixel: vegetated or barren land) or falls below 0.9 (clear-sky water pixels). The smallest allowed $N$ score is $-3$, which is assigned to all scenes where the ratio $\rho$ is smaller than 0.8 or larger than 1.2. This threshold ensures equal contribution to the aggregated rating for pixels with different magnitudes of NDVI.

b. Uniformity test

The uniformity test includes reflectance (texture) uniformity and the thermal uniformity examination. The texture uniformity test uses the reflectance in channel 2 and the thermal uniformity test employs the chan-
nel-4 brightness temperature. The uniformity test calculates the variability of pixel values (reflectance or brightness temperature) for the central point and its four nearest neighbors on each side (not diagonal). This test produces two scores, one for the texture \( U_{\text{tex}} \) and one for the thermal uniformity \( U_{\text{temp}} \):

\[
U_{\text{tex}} = \left[ \frac{1}{5} \sum_{i=1}^{5} (r_i - 7)^2 - U_{\text{offs}}^{\text{tex}} \right] U_{\text{scale}}^{\text{tex}},
\]

\[
U_{\text{temp}} = \left[ \frac{1}{5} \sum_{i=1}^{5} (T_{4,i} - T_{5,i})^2 - U_{\text{offs}}^{\text{temp}} \right] U_{\text{scale}}^{\text{temp}},
\]

For computational efficiency, the calculation of standard deviation, which usually includes the square root operation, is replaced with calculation of the variance. The offset and scale factors are adjusted appropriately to account for this substitution. The magnitude of the \( U \) scores is limited to the range from -4 to +4. From a performance point of view, although the uniformity test is the most time consuming it is only executed when the major tests indicate marginal cloudiness. Our analysis showed that the uniformity test gives adequate results for water-covered scenes but was found to be less effective for land pixels. For this reason, this test is applied only to water pixels.

c. Thin cirrus test

The final test of the SPARC algorithm is the thin cirrus test, which employs the difference in brightness temperature between channels 4 and 5 \( (T_4 - T_5) \). This difference arises because the emissivity of ice crystal clouds is slightly different for AVHRR thermal IR channels 4 and 5 centered around 11 and 12 \( \mu \)m (Stephens 1994). The satellite-observed radiance for thin Ci clouds contains a mixture of the radiation emitted by the underlying surface and the cloud. The brightness temperature \( T_r \) is close to \( T_z \) for a clear-sky atmosphere and for optically thick clouds, whereas a large difference arises for thin crystal clouds (Stephens 1994). Several cloud detection schemes such as CLAVR (Stowe et al. 1999; Heidinger 2007) and MODIS (Ackerman et al. 2002) attempt to parameterize \( T_4 - T_5 \) difference as a function of the brightness temperature \( T_r \). However, the brightness temperature \( T_r \) is normally higher for thin clouds and lower for optically thick clouds. The above dependence can be modeled if one knows the temperature of the underlying surface, the atmospheric state, the cloud optical depth, and the particle size distribution. Because of the climate conditions and geographical location of Canada, a wide range of

![Figure 6](https://example.com/fig6.png)

**Fig. 6.** Scatterplot of \( T_4 - T_5 \) difference versus \( T_4 \) temperature is shown in Fig. 6. Scenes of various parts of our study region were used, including snow-covered and snow-free land, ocean surface, and the Greenland ice sheet. These scenes included predominantly clear-sky and thin Ci fields. The scene selection and identification were conducted by visual inspection of the imagery. Figure 6 demonstrates the problems with the variable CLAVR thresholds parameterized against the \( T_r \) temperature. Figure 6 shows that a moderate \( (T_4 - T_5) \) difference \((<3\, \text{K})\) is observed for both clear-sky and thin cloud pixels, and this threshold varies little over the wide range of \( T_4 \) temperatures. To make the thin cirrus test more robust, a constant threshold independent of \( T_4 \) temperature was proposed. The effective location of this threshold is shown in Fig. 6 as the hatched area. The corresponding score \( C \) for this test is computed as

\[
C = (T_4 - T_5 - C_{\text{offs}})C_{\text{scale}}.
\]

The \( C \) score is limited to positive values only because its major purpose is to detect optically thin clouds. If the test fails (negative \( C \) score), this does not mean a clear-sky scene; it may also occur when optically thick clouds are present. Therefore, such results should not decrease the aggregated rating accumulated already for a cloudy
scene from the previous tests. This is achieved by constraining $C \geq 0$. Figure 5 shows that the thin cirrus test is applied to pixels with a marginal cloudiness level and to those detected as clear sky in order to check for the possible presence of thin clouds.

7. Correction factors

The SPARC algorithm combines the outcome of the several tests by a summation of the test scores. The relative significance of a particular test is controlled by the scaling factors (Table 1). Although this approach provides an appropriate general weighting for each of the tests, some additional adjustments are still required. These adjustments include (a) the correction of the brightness test $B$ score for snow-covered pixels; (b) the processing of observations in the specular reflection region (sun glint over water); and (c) the processing of observations in the day–night transition zone.

a. Correction factor for snow conditions

For pixels with snow, the brightness test $B$ is not efficient since clear-sky snow and clouds can be equally bright in AVHRR channels 1 and 2. Therefore, its score should be reduced for pixels with snow. The SPARC scheme achieves this by introducing a correction factor $s$ for snow conditions. Initially, the snow factor varies from 0.1 to 1.0 and is estimated from the model surface skin temperature predicted by NARR as

$$s = 1 - (T_{\text{freeze}} - T_{\text{NARR}})0.9/7,$$

if $T_{\text{freeze}} - 7 < T_{\text{NARR}} < T_{\text{freeze}}$;

$$s = 0.1, \quad \text{if } T_{\text{NARR}} \leq T_{\text{freeze}} - 7;$$

$$s = 1.0, \quad \text{if } T_{\text{NARR}} \geq T_{\text{freeze}}.$$

Once the pixel passes the first snow test as shown in Fig. 4 and the snow condition is determined, the $s$ factor is changed to 0.1 for a cloud-contaminated snow scene or 0 for clear-sky snow scene depending on the outcome of the second snow test. The use of the $s$ factor reduces the contribution of the $B$ test to the final rating and can turn it off fully if the presence of snow is established with confidence.

b. Sun glint correction factor

Over ocean areas where sun glint conditions occur, the brightness $B$ test [Eq. (3)] may produce large scores due to the high reflectance of the water surface under specular reflection geometry. This high score can be attributed mistakenly to the presence of clouds although the scene may actually correspond to the clear-sky conditions. The CLAVR and MODIS cloud detection schemes process observations in the sun glint area in a special way (Ackerman et al. 2002; Stowe et al. 1999) and generate special flags that mark these conditions. For sun glint areas, the SPARC algorithm introduces the correction factor $g$. Its function is to reduce the contribution from the reflectance tests ($B$ test, $R$ test, and $N$ test) and to assign more weight to the thermal tests ($T$ test and $U_{\text{temp}}$ test). For land pixels, this correction is not applied and the $g$ factor is zero. For water pixels, it is calculated as follows:

$$g = (0.95 - b)^2,$$

$$b = (300 - 180\theta_0)(1 - \cos\delta)^2, \quad 0 \leq b \leq 0.95,$$

$$\cos\delta = \cos\theta \cos\theta_0 - \sin\theta \sin\theta_0 \cos\phi.$$

The shape of the $g$ factor versus scattering angle $\delta$ for various solar zenith angles is shown in Fig. 7. Angle $\theta$ in Eq. (14) denotes viewing zenith angle, and angle $\phi$ denotes relative azimuth. The $g$ factor reaches a maximum at $\delta = 0$ and reduces to values near zero for $\cos\theta > 0.90 - 0.96$ depending on the solar zenith angle. The correction is applied as a $(1 - g)$ multiplication factor. Therefore, the larger $g$ is, the stronger the correction. This correction factor is applied to all the scores derived from visible and NIR channels. For correction of the $R_b$ score obtained from channel 3B reflectance, the $g$ factor is scaled by a factor of 0.6 in order to compensate for the lower reflectance of water in the 3.7-μm spectral region. For ice-covered water pixels, the sun glint has a much smaller effect and is therefore reduced by the $s$ factor, indicating the presence of snow.

c. Day–night transition zone

At large solar zenith angles, the calculated reflectance in channels 1, 2, and 3A/3B can be large because
of the properties of the surface bidirectional reflectance distribution function (Luo et al. 2005). Reflectance becomes also sensitive to surface roughness (3D) effects and atmospheric spherical properties, which are often treated as plane parallel. Under such conditions, the results of the reflectance tests \((B \text{ and } R_{A,B})\) become less reliable and more weight should be given to the tests involving thermal channels. For that reason, many cloud detection schemes limit the application of their daytime algorithms to the range \(\theta_0 < 85^\circ\). To extend this range, and also to ensure the smooth transition from daytime to nighttime conditions, the SPARC scheme introduces the nighttime correction factor \(n\):

\[
n = (88^\circ - \theta_0)/3. \tag{15}
\]

The \(n\) factor is calculated from Eq. (15) only when \(85^\circ < \theta_0 < 88^\circ\). The \(n\) factor is applied as a multiplicative factor to the scores of reflectance tests and as coefficient \((2 - n)\) for the thermal \(T\) test. Hence, this factor increases the weight of the thermal test compared to the reflectance tests in the day–night transition zone. For angles \(\theta_0 \geq 88^\circ\), the \(n\) factor is set to zero, which disables all the reflectance tests and switches the algorithm into the nighttime mode.

d. Aggregated rating including all correction factors

The final expression for the aggregated rating \(F\) that replaces Eq. (1) and includes all correction factors and additional tests as described above is:

\[
F = B(1 - g)sn + T(1 + g)(2 - s)(2 - n) + R(1 - 0.6g)(2 - s)n + N(1 - g)sn + U_{\text{temp}}(1 - g)sn + U_{\text{temp}}(1 - g)sn + C. \tag{16}
\]

Note that, in Eq. (16), the scaling factor \((1 - g)\) is also applied to the \(N\) and the \(U_{\text{temp}}\) scores.

Equation (16) shows that for situations when the significance of the reflectance tests is reduced, the thermal tests receive more weight. This is designed to maintain the overall scale of the aggregated rating for various scenarios, such as no-glint versus sun-glint pixels and snow-free versus snow-covered areas, as well as the daytime versus the day–night transition. For nighttime processing, the SPARC scheme uses essentially the same routine, excluding the tests for channels 1, 2, and 3A.

The output of the SPARC algorithm contains the aggregated rating stored as 1 byte per pixel. The rating ranges from 1 to 255, and zero is reserved for missing data. The output rating is calculated as

\[
F_{\text{output}} = 4F + 128 \tag{17}
\]

and is rounded to the nearest integer within the range 1–255.

8. Shadow detection

Cloud shadows are an important scene type that should be identified and excluded from the set of clear-sky and cloudy pixels. The application of atmospheric, land, or cloud retrieval algorithms for pixels containing cloud shadows leads to erroneous results. Cloud shadow contamination can become a source of substantial defects in the clear-sky land products (Latifovic et al. 2005). It may also introduce systematic biases and degrade the quality of the climate records of cloud properties. The MODIS cloud detection system includes cloud shadows as a special scene type that is detected using a spectral approach (Ackerman et al. 2002). So far, little attention has been paid to this issue in generating climate datasets from AVHRR observations because of technical difficulties with shadow identification and the demands for computational resources required for implementation.

The cloud shadow detection technique in the SPARC algorithm relies on the cloud detection routine described above and uses geometrical computation of the cloud shadow projection on the surface. The SPARC cloud shadow routine determines cloud shadow presence only over potentially clear-sky parts of the imagery; the cloud shadows from high cloud tops over lowlying cloud decks are not analyzed at this time. The cloud shadow information is stored as part of an 8 bit/pixel mask of the status flags. In addition to the shadow flag, the mask contains the land/water flag, the day/night flag, and the snow flag. As long as the aggregated rating reaches a certain cloudiness level, the pixel is considered to be cloudy and the geometrical routine can compute the location of the shadow by projecting the cloud top onto the surface level.

An important step in determining the location of cloud shadow is the estimation of cloud-top height \(H\). For optically thick clouds, \(H\) is derived using surface skin temperature \(T_{\text{sk}}\), cloud-top temperature \(T_{\text{ct}}\), and the atmospheric temperature gradient \(G\).

The cloud-top temperature of thick cloud is assumed to be equal to AVHRR channel-5 brightness temperature \(T_5\). The temperature gradient \(G\) is set
equal to 4.5 K km$^{-1}$. This value is slightly smaller than an average saturated adiabatic lapse rate (about 5 K km$^{-1}$; Gill 1982). The smaller value for the temperature gradient $G$ is chosen intentionally because it leads to some overestimation of the cloud-top height, and hence predicts longer shadows. This provides a more reliable identification of the shadow pixels, thus reducing the risk of identifying a shadow contaminated pixel as a pure clear sky.

The cloud-top height $H$ is computed as

$$H = \frac{T_{\text{NARR}} - T_{\text{CT}}}{G}. \quad (18)$$

For semitransparent clouds however, this height needs to be corrected because the observed brightness temperature $T$ represents a mixture of the cloud-top temperature and the surface temperature. This mixing can be described with the following expression:

$$T_s = T_{\text{CT}} + (T_{\text{NARR}} - T_{\text{CT}})M,$$  \quad (19)

where $M$ denotes the mixing between $T_{\text{NARR}}$ and $T_{\text{CT}}$ and ranges from 0 to 1. This mixing depends on cloud optical thickness and can be estimated from the difference $T_4 - T_5$ (Stephens 1994). Empirical analysis shows that adequate correction of the cloud-top height can be achieved by assuming that the mixing factor $M$ changes linearly with the difference $T_4 - T_5$. The latter is replaced with the $C$ index obtained by Eq. (12), which not only depends linearly on this difference, but also turns the cloud-top height correction off when the difference $T_4 - T_5$ is smaller than $C_{\text{off}}$. With this substitution, the following dependence is derived:

$$T_s = T_{\text{CT}} + 0.039C(T_{\text{NARR}} - T_{\text{CT}}), \quad (20)$$

where the $C$ index is limited to the range of $0 < C < 18$.

Substituting Eq. (20) into Eq. (18), one obtains an expression for cloud-top height:

$$H = \left(\frac{T_{\text{NARR}} - T_s}{G}\right) \frac{1}{1 - 0.039C}. \quad (21)$$

A comparison of the cloud-top height calculation from Eq. (21) with the results available from cloud data product MOD06 containing data collected from MODIS on the Terra platform (Platnick et al. 2003) showed a good overall level of consistency between the two algorithms. About 681,000 pixels were compared for 10 scenes selected over the period of 2002–06. The average correlation coefficient for all pixels in 10 scenes was 0.76 and the average bias between two algorithms was around 100 m, although large differences up to several kilometers were also sometimes present (the average absolute difference between cloud-top heights from two algorithms was around 2 km). On the other hand, the correlation coefficients for continuous cloud areas varied typically between 0.88 and 0.93, and the bias on average was within 500 m. Some larger differences and reduced correlation were detected for broken cloud fields and near cloud edges, where the MOD06 5 km $\times$ 5 km product frequently showed significantly lower cloud-top pressure, and therefore elevated cloud-top height, relative to the adjacent cloud system. In such situations, the SPARC cloud-top estimation was considered to be more consistent.

Once the cloud-top height $H$ is determined, the problem of finding the corresponding shadow pixels can be solved geometrically knowing the solar zenith, viewing zenith, and the relative azimuth angles (Simpson et al. 2000). An example of the cloud scene with shadows is shown in Fig. 8, where the left panel displays the channel-1 image, and the right panel shows the cloud mask (red depicts the pixels flagged as shadows).

In the status flag mask, the shadow flag is set not only for the identified shadow pixel, but also for its three neighbors that are closest in the horizontal, vertical, and diagonal directions toward the original cloud pixel. Such a procedure compensates for the cloud layer vertical extent below the estimated cloud-top level and ensures more contiguous shadow areas, especially for broken clouds. For larger solar zenith angles, the effect of vertical extension on the length of shadows is stronger. In this case, the procedure flags proportionally more shadow pixels on the line toward the cloud pixel.

The cloud shadow detection algorithm is applied to cloudy pixels only, since the clear-sky pixels do not produce shadows. This also saves processing time. However, a decision needs to be taken about the pixel status (clear-sky or cloudy) at this point. By construction, the scaled aggregated rating ranges from 1 to 255 and the level of 128 serves as the threshold between cloudy and clear-sky conditions. It was later found that this threshold often misses shadows from thin clouds. For higher flexibility of choosing shadow-free pixels during multiscene compositing (Latifovic et al. 2005), two thresholds of the aggregated rating are established that separate thin and thick cloud status. Accordingly, three levels of the shadow status are introduced: shadow-free, thin cloud shadow, and thick cloud shadow. These levels are stored as two bits in the status flag mask: the “thin cloud shadow” bit is set when the aggregated rating level is higher than 112, and both bits are set when the aggregated rating exceeds 144, indicating thick cloud shadow.

9. Validation

The validation of the results obtained by the SPARC algorithm in automated mode was conducted against
the results obtained with a supervised classification. For this purpose, 12 AVHRR scenes were selected from different seasons. These scenes represented a variety of weather and surface conditions, including summer snow-free and winter snow–ice-covered land and water regions, as well as a number of scenes for spring and fall transitional seasons. The validation was limited to daytime NOAA-16 orbits in 2002.

The supervised image classification was carried out using the maximum likelihood classifier (MLC) routine from the multispectral analysis package of PCI Geomatica 2003 (http://www.pcigeomatics.com/). The MLC routine was applied to each scene as a sequence of supervised iterations. This process was repeated until a reliable separation of classes was achieved, which represented visually the best classification that could be obtained for a particular scene.

The best MLC classification results were obtained with a maximum likelihood classifier that uses Gaussian probability distribution and the Mahalanobis minimum distance criterion to cluster pixels in spectral space (Duda et al. 2000). The MLC method was applied to the dataset of five AVHRR channels, where the AVHRR channel 5 was replaced with \( T_4 - T_5 \) difference. It was found that this substitution improves the stability of the MLC routine, because a straightforward use of channel 5 was less efficient due to the high correlation between channels 4 and 5. To increase the sensitivity of the MLC method to lower reflectance levels, which is important for correct scene recognition in the majority of cases, some preprocessing is needed for reflectance in channels 1, 2, and 3. This can be achieved by means of a gamma correction of the dynamic range (Russ 1992). For the optimal operation of the MLC routine, the correction was implemented as

![Table 2. Comparison of cloud detection results derived by the SPARC algorithm and the supervised classification with MLC. Numbers show the percentage of pixels in each class determined by the SPARC algorithm (as columns) and the supervised MLC routine (as rows). The total number of pixels in the statistics is \( 1.4 \times 10^8 \). Numbers in boldface show diagonal elements.](image)

![Fig. 8. Example of the shadow detection: (a) AVHRR channel-1 image; (b) cloud mask with the shadow pixels marked in red.](image)
where \( r \) is reflectance in AVHRR channel 1, 2, or 3 and \( \gamma \) is the correction factor (\( \gamma = 0.05 \) for channels 1 and 2 and \( \gamma = 0.01 \) for channel 3). After the correction of the dynamic range, the distribution of clusters in multispectral space becomes more uniform and a more reliable separation between classes is achieved.

For the validation of the SPARC algorithm, the results from the supervised MLC classification were reduced to six basic classes: 1) opaque cloud, 2) thin cloud over water or land, 3) thin cloud over snow–ice, 4) snow–ice, 5) water, and 6) land. The land–water mask was taken from a geographical database (Latifovic and Pouliot 2005). For comparison with the MLC classes, the aggregated ratings generated by the SPARC algorithm were converted into three classes: clear sky, thin cloud, and opaque cloud. The threshold separating clear-sky and thin cloud classes was set equal to 112 and the value separating thin cloud and opaque cloud classes was set equal to 144. With the help of the snow/ice flag taken from the SPARC status mask, two more classes were produced: snow–ice clear sky and thin cloud over snow–ice. For opaque clouds the snow flag was considered to be undefined.

A summary of the comparisons between the SPARC and supervised MLC methods as an average statistics for all scenes is presented in Table 2. It contains the percentage of pixels that were identified as the corresponding class by each method. The numbers on the diagonal show the percentage of overlap between the two algorithms. The rightmost column provides the sums of each row. These numbers show the relative abundance of the corresponding class as identified by the MLC algorithm. Table 2 shows that the overall agreement is very good: the sum of the diagonal elements (hereafter referred to as the quality index) is 84.1%. The misclassification represented by the off-diagonal elements is relatively small (15.9%) and observed mostly between the clear sky and thin cloud classes (2.6% over land, 0.4% over water, and 3.3% over snow–ice), and between the thin cloud and the opaque cloud classes (3.2% over land, 1.0% over water, and 1.2% over snow–ice). This is explained by the uncertainty in the subjective definition of the transition between the adjacent classes and by the use of fixed thresholds for all seasons. The total percentage of pixels misclassified by the two schemes for the two most distinct classes (clear sky versus opaque cloud and vice versa) is very low (only 0.3% for all classes). This demonstrates the overall reliability of the SPARC algorithm, which is run in the automated mode, relative to the supervised classification MLC scheme.

As for the validation of the snow–ice detection, the percentage of incorrectly identified snow/ice pixels is 1.9% for the clear-sky and 2.3% for the thin cloud condition. The latter result is explained by the difficulties in the detection of snow through the clouds. The former
result occurs most likely due to tradeoffs in setting the threshold for snow detection in complex situations such as the snowmelt, ice breakup, and snow–ice onset during transition seasons (Simic et al. 2004).

Table 3 shows the overall snow–ice detection efficiency that includes all clear-sky and thin-cloud pixels over land and water for warmer and colder seasons. Table 3a shows an average of seven winter scenes (from November to May) and Table 3b gives an average of five summer scenes (from June to October). Again, each cell contains the percentage in each class and the last column shows the total abundance of each class as identified by the MLC. The quality index is 88.2% for winter scenes and 91.9% for the summer, although the abundance of the snow classes is much lower for summer scenes.

The overall cloud detection efficiency over all surface types (land, water, and snow/ice) is shown in Table 4. Again, the results are provided separately for winter (Table 4a) and summer (Table 4b) cases. The comparison quality index is 86.5% for winter and 89.4% for summer cases. The percentage of misclassification between clear sky and opaque cloud classes is only 0.2% for winter and 0.4% for summer.

The overall performance of the SPARC algorithm is shown in Fig. 9. It shows the percentage of agreement between the two schemes characterized by the quality index for different times of the year. The overlap between the two schemes is in the range of 86% to 90% for the warmer season and around 80% to 86% for the colder season. The quality of snow classification varies from 86% in winter to 94% in summer. To evaluate the potential effect of possible uncertainties in the sensor calibration we conducted multiple tests with AVHRR scenes in different seasons by changing the calibration for all optical channels by ±10%, which represents the range of uncertainty in the absolute calibration (Rossov and Schiffer 1999). We found that the relative changes to the numbers presented in Tables 2–4 are on average within ±5%, which means that the response of the SPARC system is significantly weaker than the input variation of calibration coefficients, even if all changes in the calibration were accumulated with the same sign. This demonstrates that the SPARC scheme is robust enough to uncertainties in the calibration coefficients.

Only one winter scene acquired on 22 January 2002 was found particularly difficult for snow detection (81.3% overlap), producing an overall quality index of only 74.9%. Further detailed analysis showed that one
large region with snow on the ground was the reason for the reduced level of agreement. The region was covered by thin low-level clouds, and the analysis of the temperature distribution from NARR surface temperature fields indicated a possibility of temperature inversion in the boundary layer.

An example of the winter scene and classification results for the NOAA-16 AVHRR image taken on 2 February 2002 is shown in Figs. 10a–d. Figure 10a shows reflectance in AVHRR channel 1; Fig. 10b is the T rating computed with Eq. (2); and Figs. 10c,d display the MLC and SPARC classification results for six basic classes. The overall agreement between the supervised classification and the automated SPARC algorithm is visually very good. Some discrepancy can be noticed for the areas classified as thin cloud over snow–ice. Also, for the northern region close to the terminator, some opaque clouds are identified as thin clouds for pixels at large solar zenith angle.

An example of a summer scene is the NOAA-16 AVHRR image taken on 2 August 2002 as presented in Figs. 11a–d. The panels are defined in the same way as those defined above for Fig. 10. Figure 11 shows that the efficiency of the SPARC algorithm for a summer scene is very good and looks better than the winter case presented in Fig. 10. Some minor discrepancies are observed in the northern part for the thin cloud class. This is similar to the winter case and points to the difficulties in scene identification for low sun elevations.
10. Conclusions

The novel scene identification scheme SPARC (Separation of Pixels Using Aggregated Rating over Canada) is proposed here for clear-sky and cloudy pixel identification. This scheme also produces flags for snow, ice, and cloud shadows. The SPARC scheme was designed for application in the processing of historical and current 1-km AVHRR imagery in HRPT and LAC formats. However, SPARC can also be used for any other sensor that provides spectral information similar to AVHRR (e.g., MODIS), although some correction may be needed for the difference in the spectral response function between AVHRR and this sensor (Trishchenko et al. 2002a; Trishchenko 2006). Additional input data used in the SPARC algorithm include the surface skin temperature from North America Regional Reanalysis and the land–water mask. The algorithm was tuned to work most efficiently over temperate and polar regions that correspond to Canadian climate and surface conditions. The scheme was successfully applied to the generation of historical climate data records over Canada from AVHRR since 1981 (Latifovic et al. 2005).

A distinctive feature of the algorithm is the computation of an aggregated rating that combines the results of several individual tests into one final score. It was demonstrated that this rating carries all of the essential information necessary to discriminate between clear-sky and cloudy scenes. Major tests implemented in the SPARC scheme include a brightness test in the AVHRR channel 1, a brightness temperature test in channel 4, a reflectance test in channel 3 (A or B), a simple ratio test, and a thin cirrus test and a spatial uniformity test. The principal difference from other schemes, which are based on the “yes/no” branching approach, is the aggregation of results from all of the individual tests (weighted appropriately). This not only improves the overall reliability of the algorithm, but also provides a quantitative parameter for the comparison and decision about the degree of a pixel’s cloud contamination. This methodology was used in generating the historical clear-sky data datasets of Canada’s landmass suitable for climate change applications (Latifovic et al. 2005).

Another important component of the SPARC scheme is the cloud shadow detection technique. A simple yet robust and adequate model is proposed for estimation of the cloud-top height, which compares well to MODIS results. Cloud shadows are determined for thin and opaque clouds using the calculated cloud-top height and observation geometry.

Special correction factors were proposed and the parameterization provided to enhance data processing over sun-glint areas and at the large solar zenith angles near the terminator line. Special tuning of the test scores was applied to snow–ice pixels.
Fig. 11. (Continued)
The SPARC algorithm produces two output products: the 8-bit status flag mask and the 8-bit aggregated rating ranging from 1 to 255. The status flags include the snow–ice and shadow flags, among others. The aggregated rating carries information about the degree of cloud contamination.

The results produced by the SPARC algorithm, which runs in automated mode, were validated by comparison with respect to supervised MLC classification. The comparison was conducted for 12 scenes that spanned all seasons and thus represented a variety of weather and surface conditions. We have showed that the overall level of agreement between the automated and supervised classifications is very good and varies from about 80%–90% (84.1% on average for all the 12 scenes analyzed). In general, the agreement is slightly better during the warmer season than the colder season. The level of major classification errors (clear-sky versus opaque cloud and vice versa) was demonstrated to be low (around 0.3% for all classes). Most of the difficulties and the largest uncertainties (up to ~4.5%) are associated with snow detection errors. These errors may be especially high for some regions during transition seasons and under conditions of thin cloudiness, although on average the reliability of snow/ice detection remains quite high (between 86% and 94%).

Acknowledgments. This work was conducted at the Canada Centre for Remote Sensing (CCRS), Earth Sciences Sector of the Department of Natural Resources Canada as part of the Project J28 of the program “Reducing Canada’s Vulnerability to Climate Change.” This work was partially supported by the Canadian Space Agency under the Government Related Initiative Program (GRIP) grant to CCRS and the International Astronautical Federation/European Space Agency under the GMES/BEAR initiative. The authors thank Gunar Fedosejevs, Philippe M. Teillet, and Andrew Davidson for their help with editing and the critical review of the manuscript at CCRS.

REFERENCES


