Object-based cloud and cloud shadow detection in Landsat imagery

Zhe Zhu *, Curtis E. Woodcock

Center for Remote Sensing, Department of Geography and Environment, Boston University, 675 Commonwealth Avenue, Boston, MA, 02215, USA

ABSTRACT

A new method called Fmask (Function of mask) for cloud and cloud shadow detection in Landsat imagery is provided. Landsat Top of Atmosphere (TOA) reflectance and Brightness Temperature (BT) are used as inputs. Fmask first uses rules based on cloud physical properties to separate Potential Cloud Pixels (PCPs) and clear-sky pixels. Next, a normalized temperature probability, spectral variability probability, and brightness probability are combined to produce a probability mask for clouds over land and water separately. Then, the PCPs and the cloud probability mask are used together to derive the potential cloud layer. The darkening effect of the cloud shadows in the Near Infrared (NIR) Band is used to generate a potential shadow layer by applying the flood-fill transformation. Subsequently, 3D cloud objects are determined via segmentation of the potential cloud layer and assumption of a constant temperature lapse rate within each cloud object. The view angle of the satellite sensor and the illuminating angle are used to predict possible cloud shadow locations and select the one that has the maximum similarity with the potential cloud shadow mask. If the scene has snow, a snow mask is also produced. For a globally distributed set of reference data, the average Fmask overall cloud accuracy is as high as 96.4%. The goal is development of a cloud and cloud shadow detection algorithm suitable for routine usage with Landsat images.

© 2011 Elsevier Inc. All rights reserved.

1. Introduction

The long history of Landsat data is one of the most valuable data-sets available for studying land cover change and human influences on the land surface (Cohen et al., 1998; Coiner, 1980; Coppin & Bauer, 1994; Seto et al., 2002), especially since the first Thematic Mapper (TM) sensor was launched in 1982, which provided higher spatial resolution and more spectral bands. However, many of the Landsat images are inevitably covered by cloud, especially in the tropics (Asner, 2001). The International Satellite Cloud Climatology Project-Flux Data (ISCCP-FD) data set estimates the global annual mean cloud cover is approximately 66% (Zhang et al., 2004). The presence of clouds and their shadows complicates the use of data in the optical domain from earth observation satellites. The brightening effect of the clouds and the darkening effect of cloud shadows influence many kinds of data analyses, causing problems for many remote sensing activities, including inaccurate atmospheric correction, biased estimation of Normalized Difference Vegetation Index (NDVI) values, mistakes in land cover classification, and false detection of land cover change. Therefore, clouds and cloud shadows are significant sources of noise in the Landsat data, and their detection is an initial step in most analyses (Arvidson et al., 2001; Irish, 2000; Simpson & Stitt, 1998). Generally, clouds can be divided into two categories:

thick opaque clouds and thin semitransparent clouds. The thick opaque clouds are relatively easier to identify because of their high reflectance in the visible bands. The identification of thin semitransparent clouds is difficult as their signal includes both clouds and the surface underneath (Gao & Kaufman, 1995; Gao et al., 1998, 2002).

Due to the high spectral variability of clouds, cloud shadows, and the Earth's surface, automated accurate separation of clouds and cloud shadows from normally illuminated surface conditions is difficult. Intuitively, it seems that clouds and cloud shadows are easily separable from clear-sky measurements, as clouds are generally white, bright, and cold compared to the Earth's surface, while cloud shadows are usually dark. Nevertheless, there are clouds that are not white, bright, or cold and cloud shadows even brighter than the average surface reflectance. Part of the difficulty arises from the wide range of reflectances and temperatures observed on the surface (Irish, 2000). One common approach is to screen clouds and cloud shadows manually. However, this approach is time consuming and will limit efforts to mine the Landsat archive to study the history of the Earth's surface.

Over the years, a number of methods were developed for cloud identification. However, most of them are designed for moderate spatial resolution sensors such as Advanced Very High Resolution Radiometer (AVHRR) and Moderate Resolution Imaging Spectroradiometer (MODIS). These sensors are usually equipped with more than one thermal band, or with water vapor/CO_2 absorption bands, both of which are useful for thin semitransparent cloud detection (Ackerman et al., 1998; Derrien et al., 1993; Saunders &

* Corresponding author. Tel.: +1 617 233 6031.
E-mail address: zhuzhe@bu.edu (Z. Zhu).

Kriebel, 1998). For high spatial resolution sensors like Landsat, with only one thermal band and 6 optical bands placed in atmospheric windows, accurate cloud identification is difficult. And, cloud shadow identification is even more difficult. Clouds cast shadows on any type of land cover. When cloud shadows fall on urban, snow, ice, or bright rocks, they can be very bright compared to the average surface reflectance. Moreover, when the cloud is semitransparent, the darkening effect of the cloud shadow can be subtle, making the cloud shadow hard to detect. Therefore, how to detect clouds, cloud shadows, and especially thin clouds and their shadows in Landsat images is still an important issue in the remote sensing community, particularly as we try to use increasingly automated methods to analyze large volumes of data.

2. Background

Historically, screening of clouds in Landsat data has been performed by the Automated Cloud Cover Assessment (ACCA) system (Irish, 2000; Irish et al., 2006). By applying a number of spectral filters, and depending heavily on the thermal infrared band, ACCA generally works well for estimating the overall percentage of clouds in each Landsat scene, which was its original purpose. However, it does not provide sufficiently precise locations and boundaries of clouds and their shadows to be useful for automated analyses of time series of Landsat images. Additionally, ACCA fails to identify warm cirrus clouds and falsely identifies snow/ice in high latitude areas as clouds (Irish, 2000; Irish et al., 2006). Wang et al. (1999) proposed the use of two multi-temporal Landsat TM images to find clouds and their shadows by image differencing. This method can successfully provide an accurate cloud and cloud shadow mask, but it is highly dependent on the input images. Since the Landsat sensors are not always turned on, it can be months between successive acquisitions. Also, it is possible that the next Landsat observation is still cloudy in the same location as the previous Landsat image, which would further limit the utility of the proposed algorithm. As cloud and snow/ice are very hard to distinguish from each other in high latitude areas, Choi and Bindschadler (2004) suggested a method for detecting clouds over ice sheets by using a shadow matching technique and an automatic Normalized Difference Snow Index (NDSI) threshold. This method matches the possible cloud and cloud shadow edges iteratively to find the optimal NDSI threshold for cloud detection. It works well over ice sheets but it is time consuming and only works for the surface of ice sheets. The Landsat Ecosystem Disturbance Adaptive Processing System (LEDAPS) atmosphere correction tool also generates an internal cloud mask (Vermote & Saleous, 2007). It uses two passes. There are four tests in the first pass and a thermal test in the second pass which is similar to ACCA, except that the second pass generates a cloud mask while the second pass of ACCA only provides the percentage of cloud cover. This algorithm needs other ancillary data like the surface temperature provided from National Centers for Environmental Prediction (NCEP) to help generate a coarse resolution surface temperature reference layer for cloud detection. This algorithm has already been used extensively for atmospheric correction of Landsat images and has shown a better method for cloud detection in low and middle latitudes compared to ACCA. However, it may not work well when the clouds cover a large percentage of the image (large amount of leakage were observed) or in sun glint and turbid water conditions (Vermote, 2010). Hégarat-Mascle and André (2009) developed an approach that uses only two bands, Green and Short Wave Infrared (SWIR), to generate a "clear-sky line" and use the distance from the tested points to this line to detect cloud pixels. This method was originally used by Zhang et al. (2002) to correct for haze in Landsat imagery. It has been shown to be accurate for retrieving clouds over vegetated areas, but it fails when the surface reflectance is bright, as is the case for rocks, snow, ice, sand, etc. (Zhang et al., 2002). By

implementing a cloud-mask algorithm originally developed for the MODIS Land bands on Landsat data, Oreopoulos et al. (2011) proposed an algorithm that performs on par with the ACCA algorithm without using the thermal band.

Detecting cloud shadow is more difficult than detecting cloud. Previously, cloud shadow identification was based on spectral tests. Though it works sometimes, most of the time it will inevitably include other dark surfaces that have similar spectral signatures (like topographic shadows or wetlands) and exclude cloud shadows that are not dark enough (Ackerman et al., 1998; Hutchison et al., 2009). Recently, geometry-based cloud shadow detection has been shown to be feasible and more accurate. Currently, there are three kinds of geometry-based cloud shadow detection methods in the literature: object matching, lapse rate, and scattering differencing. The object matching algorithm detects cloud shadow by matching cloud shadows with cloud objects (Berendes et al., 1992; Hégarat-Mascle & André, 2009; Simpson & Stitt, 1998; Simpson et al., 2000). The lapse rate method used a constant lapse rate to estimate cloud top height by brightness temperature and use the cloud pixels to cast shadows (Vermote & Saleous, 2007). This latter method works well for thick clouds but is not accurate when the clouds are semitransparent, as the brightness temperature will be a mixture of thin cloud and the surface, making cloud height estimation problematic. As cloud shadow scattering is stronger in the short wavelengths (especially Blue band), Luo et al. (2008) proposed to use this physical characteristic (scattering differences between the short wavelength and NIR or SWIR), combined with the geometry, to produce cloud shadow masks. This new method works well over vegetated area, but is less accurate when the cloud shadow falls on bright surfaces or the cloud shadow comes from a very thin cloud.

In this paper, we provide a new algorithm for detecting both clouds and cloud shadows for Landsat TM and Enhanced Thematic Mapper Plus (ETM+) images that builds on the results of previous approaches. The cloud mask is computed from a probability mask and a scene-based threshold. Cloud shadows are calculated using a combination of previous methods (object matching and lapse rates) and a flood-fill transformation. This algorithm works well in high latitudes, separating clouds from shallow or turbid water accurately, and can also detect thin clouds and their shadows. If a Landsat scene has snow, Fmask also produces a snow mask.

3. The Fmask algorithm

The input data are Top of Atmosphere (TOA) reflectances for Bands 1, 2, 3, 4, 5, 7 and Band 6 Brightness Temperature (BT) (Table 1). For Landsat L1T images, Digital Number (DN) values are converted to TOA reflectances and BT (Celsius degree) with the LEDAPS atmosphere correction tool (Masek et al., 2006; Vermote & Saleous, 2007). Then, rules based on cloud and cloud shadow physical properties are used to extract a potential cloud layer and a potential cloud shadow layer. Finally, the segmented potential cloud layer and the geometric relationships are used to match the potential cloud shadow layer, leading to the production of the final cloud and cloud shadow mask. If the Landsat scene has snow, Fmask will also

Table 1
Landsat TM/ETM+ spectral bands.

TM bands (μm)	ETM+ bands (μm)
Band 1 (0.45–0.52)	Band 1 (0.45–0.515)
Band 2 (0.52–0.60)	Band 2 (0.525–0.605)
Band 3 (0.63–0.69)	Band 3 (0.63–0.69)
Band 4 (0.76–0.90)	Band 4 (0.75–0.90)
Band 5 (1.55–1.75)	Band 5 (1.55–1.75)
Band 6 (10.40–12.50)	Band 6 (10.40–12.50)
Band 7 (2.08–2.35)	Band 7 (2.09–2.35)

produce a snow mask in addition to the cloud and cloud shadow mask.

3.1. Layers of potential clouds, cloud shadows, and snow

3.1.1. Potential cloud layer — pass one

The Fmask algorithm first combines several spectral tests to identify the Potential Cloud Pixels (PCPs) — the pixels that may possibly be cloudy and may sometimes be clear pixels. Otherwise, the pixels are considered to be absolutely clear-sky pixels. This first pass includes a number of spectral tests as follows:

$$\text{Basic Test} = \text{Band } 7 > 0.03 \text{ and } \text{BT} < 27 \text{ and} \tag{1}$$
$$\text{NDSI} < 0.8 \text{ and } \text{NDVI} < 0.8$$

Where,

$$\text{NDSI} = (\text{Band } 2 - \text{Band } 5)/(\text{Band } 2 + \text{Band } 5)$$
$$\text{NDVI} = (\text{Band } 4 - \text{Band } 3)/(\text{Band } 4 + \text{Band } 3).$$

This "Basic Test" is one of the most fundamental tests for cloud identification. Due to the bright and cold nature of clouds, all kinds of clouds should have Band 7 TOA reflectance larger than 0.03 (heritage from LEDAPS internal cloud masking algorithm) and BT less than 27 °C (heritage from ACCA). The NDSI and NDVI values of clouds are usually around zero because of their "white" character in optical spectral bands. For certain cloud types, such as very thin clouds over highly vegetated area or icy clouds, the NDVI and NDSI values can be larger, but both of them cannot be higher than 0.8. ACCA also uses NDSI threshold of 0.8 to separate clouds from snow pixels in the first pass. Therefore, Fmask uses NDSI and NDVI thresholds of 0.8 to separate PCPs from some of the vegetated or snow covered areas.

$$\text{MeanVis} = (\text{Band } 1 + \text{Band } 2 + \text{Band } 3)/3$$

$$\text{Whiteness Test} = \sum_{i=1}^{3} \left| (\text{Band } i - \text{Mean Vis})/\text{Mean Vis} \right| < 0.7. \tag{2}$$

This "Whiteness" index was originally proposed by Gomez-Chova et al. (2007). As clouds always appear white due to their "flat" reflectance in the visible bands, they used the sum of the absolute difference between the visible bands and the overall brightness to capture this cloud property. This index works well in ENVIronmental SATellite (ENVISAT) Medium Resolution Imaging Spectrometer (MERIS) multispectral image as it has many narrow visible bands. However, it is not that useful for Landsat sensor which only has three visible bands. By dividing the difference by the average value of the visible bands, the new "Whiteness" index works well for Landsat imagery and 0.7 (sensitivity analysis of the global cloud reference dataset) appears to be an optimal threshold for excluding clear-sky pixels that exhibit high variability in the visible bands. All the sensitivity analyses in this paper are based on a set of 142 Landsat reference images (see Section 4.2 for details). To find the optimal threshold for "Whiteness", we let the "Whiteness" threshold vary from 0.5 to 0.9 (at 0.1 intervals) and chose the one with the highest average cloud overall accuracy (Fig. 1). The above "Whiteness" index is used to exclude pixels that are not "white" enough to be clouds. Note that this "Whiteness Test" may also include some pixels of bare soil, sand, and snow/ice as they may also have "flat" reflectance in the visible bands.

$$\text{HOT Test} = \text{Band } 1 - 0.5 \cdot \text{Band } 3 - 0.08 > 0. \tag{3}$$

This Haze Optimized Transformation (HOT) for Landsat data was firstly proposed by Zhang et al. (2002). It is based on the idea that
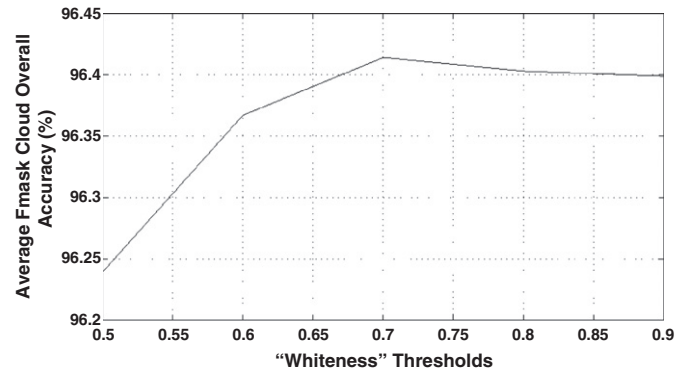

Fig. 1. An example of choosing an optimal threshold for "Whiteness" based on sensitivity analysis. Note that a threshold of 0.7 shows the highest average Fmask cloud overall accuracy.

the visible bands for most land surfaces under clear-sky conditions are highly correlated, but the spectral response to haze and thin cloud is different between the blue and red wavelengths. The "HOT" in Zhang et al. (2002) is built empirically from regression of DN values from clear-sky pixels. If we use TOA reflectance as inputs for regression, Eq. (3) is retrieved for most of the Landsat images. The results are especially helpful for separating haze and thin cloud from clear-sky pixels. Similar test has also been used in the LEDAPS internal cloud masking algorithm. All kinds of clouds (thin and thick) and possibly thick aerosols will be identified by this test. Note that this "HOT Test" may also include some bright pixels like rocks, turbid water, or snow/ice surface due to their large TOA reflectance in the visible bands.

$$\text{B4/B5 Test} = \text{Band } 4/\text{Band } 5 > 0.75. \tag{4}$$

This spectral test is similar to a test in ACCA (Irish, 2000) in which a Band 4 and Band 5 ratio larger than 1 is used to exclude bright rock and desert due to the fact that they tend to exhibit higher reflectance in Band 5 than in Band 4, whereas the reverse is true for clouds. However, this threshold may also exclude some thin clouds. Therefore, we reduced this threshold to 0.75 (sensitivity analysis of the global cloud reference dataset) in Fmask to include all possible cloud pixels. This test may also include other noncloud pixels, but the main focus of this test is separating most of bright rocks from clouds.

$$\text{Water Test} = (\text{NDVI} < 0.01 \text{ and } \text{Band } 4 < 0.11) \text{ or} \tag{5}$$
$$(\text{NDVI} < 0.1 \text{ and } \text{Band } 4 < 0.05).$$

This "Water Test" divides all pixels into two categories — water and land pixels. The thick clouds will be identified as land pixels whether they are over land or water (thick clouds block all information for land and water separation), while the thin clouds over water may still be able to be identified as water pixels. NIR band reflectance is a good indicator for water identification, as water is generally dark in this band while land is usually bright. Additionally, NDVI values are especially useful for separating water pixels from land pixels, as land NDVI values are generally above 0.10 and water NDVI values are usually less than 0.10 (Vermote & Saleous, 2007). Most of the water pixels are identified by NDVI less than 0.1 and Band 4 less than 0.05. Some of the water pixels may have relatively large Band 4 reflectance because of influence of thin clouds or turbid conditions, and they will be captured by using NDVI less than 0.01 and Band 4 less than 0.11. The "Water Test" is mainly used for separating cloud probability calculation over water and land in pass two.

By applying the spectral tests above, Fmask will identify PCP as follows:

PCP = Basic Test (true) and Whiteness Test (true) and

HOT Test (true) and B4/B5 Test (true).   (6)

If the PCPs are more than 99.9% of the scene, they will be used for the final cloud mask directly, as there are not enough clear-sky pixels (approximately 50,000 pixels) for statistic analyses in the second pass. If the PCPs are less than 99.9% of the scene, the PCPs and the absolute clear-sky pixels will be sent to the second pass. As the algorithm tends to include all possible cloudy pixels (it overestimates cloud fraction) in the first pass, Fmask requires a relatively small percentage (0.1%) of the scene to be absolutely clear to allow the second pass to work.

3.1.2. Potential cloud layer — pass two

After identification of all PCPs, the rest of the pixels (absolute clear-sky pixels) can be used for computing cloud probability for all pixels in the image. As the temperature distributions and the range of reflectances for land and water can be quite variable in space and time, Fmask computes cloud probability separately for water and land. The water and land pixels are classified by the "Water Test" applied in pass one.

The cloud probability for water (*wCloud_Prob*) is a combination of temperature probability (*wTemperature_Prob*) and brightness probability (*Brightness_Prob*) computed as follows:

Temperature probability for water:

Clear−sky Water = Water Test (true) and Band 7 <0.03.   (7)

T_{water} = 82.5 percentile of Clear−sky Water pixels′ BT.   (8)

$wTemperature\_Prob = (T_{water}-\mathrm{BT})/4.$ (9)

The difference between the estimated clear-sky water temperature (T_{water}) and the pixels' BT are normalized by 4 °C to compute the temperature probability for water (Eq. 9). The clear-sky water pixels are identified with a "Water Test" and a low Band 7 reflectance threshold (Eq. 7). T_{water} is estimated with the upper level (82.5 percentile) of clear-sky water temperature, in purpose of exclude other atmospheric influences that are usually making water temperature colder. A constant of 4 °C is used for re-scale the temperature probability because a pixel would have a high probability of being a cloud pixel if its BT is 4 °C colder than the surface temperature (Vermote & Saleous, 2007). As temperature is one of the most important dimensions in cloud detection, the temperature probability can be higher than one if the BT of the pixel is more than 4 °C colder than the estimated clear-sky water temperature.

Brightness probability:

Water is generally dark, especially in Band 5 reflectance. The existence of clouds over water can increase Band 5 reflectance greatly. Fmask uses the normalized Band 5 reflectance to calculate the brightness probability for cloud detection over water. Usually Band 5 reflectance of water is less than 0.05, except for some turbid or shallow water pixels which may have higher reflectance. The brightest water may have Band 5 reflectance as high as 0.11. Fmask calculates the normalized brightness probability with Eq. (10).

$Brightness\_Prob = \min(\mathrm{Band}5, 0.11)/0.11.$ (10)

Cloud probability for water:

The cloud probability for water pixels is computed by combining both the temperature probability and the brightness probability (Eq. 11). The temperature probability may contribute more than the brightness probability for some very cold pixels because of its wider probability range.

$wCloud\_Prob = wTemperature\_Prob \cdot Brightness\_Prob.$ (11)

As BT and Band 5 reflectance for clear-sky water pixels are very homogenous, Fmask uses a fixed threshold to retrieve clouds over water. A water pixel is identified as a cloud pixel if *wCloud_Prob* is larger than 0.5. This fixed threshold works well for detecting clouds over water. By combining temperature and brightness probabilities, bright water pixels (like shallow or turbid water pixels) or cold water pixels (higher elevation water) will be easily excluded from cloud pixels because if one of the probabilities is close to zero, no matter how large the other probability is, the cloud probability for water will still be close to zero.

The cloud probability for land (*lCloud_Prob*) is a combination of temperature probability (*lTemperature_Prob*) and variability probability (*Variability_Prob*) computed as follows:

Temperature probability for land:

Clear−sky Land = PCP (false) and Water Test (false).   (12)

$\left(T_{low}, T_{high}\right) = (17.5, 82.5)$ percentile of Clear−sky Land pixels′ BT. (13)

If clear-sky land pixels cover less than 0.1% (minimum necessary pixels for statistic analysis) of the total observations in the scene, Fmask will use the clear-sky pixels (from both land and water) for calculating temperature probability instead of only using the clear-sky land pixels. T_{low} and T_{high} calculated from Eq. (13) provide the temperature interval for clear-sky land pixels. The 17.5% and 82.5% thresholds were derived from a sensitivity analysis of the global cloud reference masks. As land temperatures can differ greatly, Fmask uses the upper and lower level of the clear-sky land temperature to normalize the temperature probability for land (Eq. 14). Normally, if the pixel's BT is 4 °C colder than T_{low}, the pixel has a high probability of being a cloud. On the other hand, if the pixel's BT is 4° warmer than T_{high}, the pixel is most likely clear. Because temperature is one of the most important dimensions in cloud detection, the temperature probability for land can be larger than one if the BT of the pixel is more than 4 °C colder than T_{low}.

$lTemperature\_Prob = \left(T_{high} + 4-\mathrm{BT}\right)/\left(T_{high} + 4-(T_{low}-4)\right).$ (14)

Variability probability:

Due to the large variability of reflectance for land pixels, the brightness probability does not work well over land for cloud detection. However, as the cloud spectral reflectances in the optical bands are very consistent, Fmask uses the probability of the spectral variability to identify clouds over land. The NDVI, NDSI, and "Whiteness" values are used to capture the spectral variability in NIR/Visible, SWIR/Visible, and within the Visible. Fmask uses 1 minus the largest value among the three indices to represent the spectral variability. The NDVI and NDSI based spectral variability may not be accurate when dealing with saturated pixels. In this case, a modified NDVI and NDSI are used in Eq. (15). The NDSI and NDVI values are modified as follows: if a pixel is saturated in Band 2 and has Band 5 larger than

Band 2, Fmask gives a zero value for this pixel's NDSI; the same rule is applied for the modified NDVI, that is, if a pixel is saturated in Band 3 and has Band 4 larger than Band 3, Fmask gives a zero value for this pixel's NDVI. This is because compared to NIR and SWIR bands, the Landsat visible bands are easily saturated for bright pixels. Theoretically, for bright cloud pixels, all the optical bands TOA reflectance will be close to 1, making the NDSI and NDVI values close to 0. However, if visible bands become saturated at a small value, for example 0.5 (Dozier, 1989), while NIR and SWIR bands do not (close to 1), it would make the absolute values of NDSI and NDVI much larger than 0, making probability of spectral variability lower for cloud pixels.

$$Variability\_Prob = 1 - \max(\text{abs(modified NDVI)}, \quad (15)$$
$$\text{abs(modified NDSI)}, \text{and Whiteness}).$$

Cloud probability for land:

The cloud probability for land pixels is computed by combining both temperature probability and variability probability as follows. The temperature probability may contribute more than the variability probability for some very cold pixels because of its wider probability range.

$$lCloud\_Prob = lTemperature\_Prob \cdot Variability\_Prob. \quad (16)$$

The threshold for defining cloud over land is consisted by the upper level (82.5 percentile) of clear-sky land pixels' probability plus a constant of 0.2 (based on sensitivity analysis) shown in Eq. (17). Fmask identifies a pixel as cloud if the land pixel's $lCloud\_Prob$ is larger than this scene-based threshold.

$$Land\_threshold \quad (17)$$
$$= 82.5 \text{ percentile of } lCloud\_Prob \text{ (Clear}-\text{sky Land pixels)} + 0.2.$$

Therefore, by combining the cloud probability and the previously identified PCPs, Fmask generates the potential cloud layer in Eq. (18). Due to the possibility of omitting clouds in PCPs, Fmask finds missed cloud pixels if the $lCloud\_Prob$ is extremely large (more than 99%) over land or BT is extremely cold (35 °C colder than T_{low}).

Potential Cloud Layer is true if

(PCP (true) and Water Test (true) and $wCloud\_Prob > 0.5$) or
(PCP (true) and Water Test (false) and $lCloud\_Prob > Land\_threshold$) or
($lCloud\_Prob > 0.99$ and Water Test (false)) or (BT$<T_{low}-35$). (18)

Finally, Fmask will spatially improve the cloud mask by using the rule that sets a pixel to cloud if five or more pixels in its 3-by-3 neighborhood are cloud pixels; otherwise, the pixel stays clear.

3.1.3. Potential cloud shadow layer

Because beam solar radiation is blocked by clouds, the cloud shadows are mainly illuminated by scattered light. As the atmospheric scattering is stronger at shorter wavelengths (for example visible bands), the diffusive radiation in the shadows will be relatively smaller at longer wavelengths (for example NIR and SWIR bands), making the shadowed pixels darker than their surroundings (Luo et al., 2008). Moreover, as NIR reflectance is usually high (including vegetation, snow, ice, and rock), the darkening effect of cloud shadows is most obvious in this Band. Therefore, a morphological transformation called flood-fill is performed for Band 4 reflectance (NIR band) which brings the intensity values of dark areas that are surrounded by lighter areas up to the same intensity level as the surrounding pixels

(Soille, 1999). In field of morphology, the gray-scaled image is viewed as a "digital elevation model". Therefore, all cloud shadows are located at places with regional minima due to their relatively darker Band 4 reflectance compared to their surroundings. The flood-fill transformation is defined as the reconstruction by erosion of the input digital elevation model using a marker image which is set to the maximum height of the digital elevation model except along its borders and at the bottom of natural depressions where it inherits the values of the input digital elevation model (Soille et al., 2003). In this case, the difference between the filled Band 4 reflectance and the original Band 4 reflectance will include the darkening effect of the cloud shadows. If the cloud shadow is located at the edge of the scene, the flood-fill transformation will not be able to identify it. Therefore, Fmask fills the edge of the scene with the lower level (17.5 percentile) of the clear-sky land Band 4 reflectance to catch all potential cloud shadows.

Potential Cloud Shadow Layer is true if (19)
Flood$-$fill Band 4 – Original Band 4 > 0.02.

3.1.4. Potential snow layer

Potential Snow Layer is true if (20)
NDSI > 0.15 and BT < 3.8 and B and $4 > 0.11$ and B and $2 > 0.1$.

Most of the spectral tests used here (BT less than 3.8, Band 4 more than 0.11, and Band 2 more than 0.1) are from the MODIS snow mapping algorithm (Hall et al., 2001). The only difference is the NDSI thresholds used. The MODIS snow algorithm uses NDSI larger than 0.4 as its threshold to identify pixels that are approximately 50% or greater covered by snow. We lower the NDSI threshold to 0.15 for Fmask to include pixels with snow coverage less than 50% and snow contaminated forest areas in which snow are partly blocked by the forests. At the same time, for all clear (snow and cloud free) land pixels in Landsat data, the NDSI values are always lower than 0.15. Therefore, with a NDSI threshold of 0.15, we can separate snow free and snow contaminated pixels accurately in Landsat data. This threshold has already been used for operational snow mapping in Meteosat Spinning Enhanced Visible Infra-Red Imager (SEVIRI) imagery (Wildt et al., 2007).

3.2. Object-based cloud and cloud shadow match

The basic idea of this cloud and cloud shadow matching approach is that by knowing the view angle of the satellite sensor, the solar zenith angle, the solar azimuth angle, and the relative height of the cloud, we can predict the cloud shadow location based on the geometric relationship between a cloud and its shadow. Because the first three factors are known, we can use them to calculate the projected direction of the cloud shadow. Along this direction, Fmask matches the cloud object with the potential shadow layer by using the idea that a cloud and its shadow have similar shape (Fig. 2). The original cloud object is excluded from the calculated shadow, as the pixels cannot be cloud and shadow at the same time. The match similarity for each cloud object is the ratio of the overlap area between the calculated shadow and the potential cloud or shadow layers to the calculated shadow area. To match the correct cloud shadow, iteration of the cloud height continues if similarity is increasing or not decreasing to 98% of the maximum similarity; otherwise, the iteration will stop. If similarity is larger than a given threshold, the cloud shadow is matched, otherwise, it is rejected. The similarity threshold can be any value from 0.2 to 0.5, which all provide similar cloud shadow results. A threshold of 0.3 is applied for Fmask as it keeps a balance between omission and commission errors of cloud shadows.
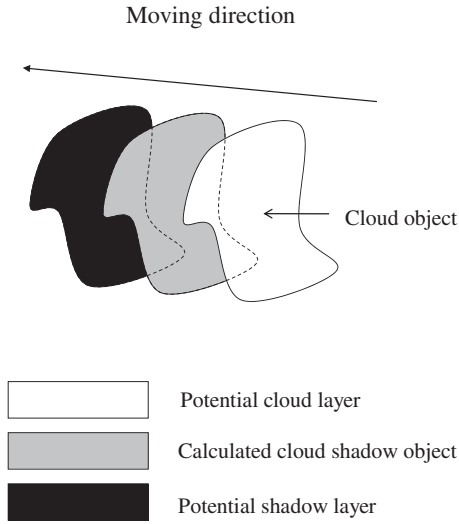
Moving direction

Cloud object

☐ Potential cloud layer

▨ Calculated cloud shadow object

■ Potential shadow layer

Fig. 2. Illustration of clouds and shadows matching based on similarity.

The cloud objects are derived by segmentation of the potential cloud layer, that is, the potential cloud pixels adjacent to other potential cloud pixels (using 8-way connectedness) are identified as one cloud object. The shapes of cloud objects are not always the same as their cloud shadows (Fig. 2 is an ideal case assuming cloud objects are flat and no other regional minima pixels are identified as potential cloud shadow), because some kinds of clouds having large vertical extents may cast cloud shadow that extend further than the flat cloud approximation allows. This can also occur with small vertical extent clouds at very low solar elevations angles. Therefore, Fmask treats each cloud as a 3D object with a base height retrieved by matching clouds and cloud shadows, and a top height estimated by a constant lapse rate and its corresponding base height.

As cloud base height can be any value from 200 m to 12,000 m, it would be time consuming and may cause false matches if we iterate cloud height across this entire range for each cloud object to find its cloud shadow. Fmask algorithm narrows the cloud base height range by using cloud object BT. For standard atmosphere conditions, the adiabatic lapse rate for dry air is $-9.8\,\text{K}\,\text{km}^{-1}$ and for wet air is $-6.5\,\text{K}\,\text{km}^{-1}$ (Hartmann, 1994). However, this is not always true for thin clouds, as their BT is influenced by the warmer ground surface underneath. In this case, Fmask uses a reduced wet adiabatic lapse rate of $-1\,\text{K}\,\text{km}^{-1}$ to capture thin cloud shadows. Therefore, we can predict the minimum and maximum cloud object base height range as:

$$H_{cloud\_base} = \left(\max\left(0.2, \left(T_{low}-4-T_{cloud\_base}\right)/9.8\right), \min\left(12, \left(T_{high}+4-T_{cloud\_base}\right)\right)\right)\text{km.} \tag{21}$$

For each cloud object, T_{cloud_base} should have the highest BT due to the fact that the cloud base pixels are the lowest cloud pixels that the sensor can detect. Nevertheless, for both thick and thin clouds, the warmest cloud pixels located at the edge do not represent the actual BT of the cloud base due to influences from the neighboring warm ground. Therefore, it is necessary to use pixels far enough from the edge of the cloud to represent the cloud base BT and adjust the edge pixels that are warmer than this value. For the purpose of simplify cloud base BT calculation, Fmask assumes each cloud object is round and 8 cloud edge pixels are influenced by the neighboring warm surface. If the calculated radius of cloud object is less than 8 pixels, Fmask uses the minimum BT of the cloud object as its

cloud base BT. Therefore, we can calculate the cloud base BT (Eq. 22) and adjust the influenced cloud BT with this value (Eq. 23) as follows:

$$\text{If } R \geq = 8 \quad T_{cloud\_base} = 100(R-8)^2/R^2 \text{ percentile of cloud object BT}$$

$$\text{Else} \quad T_{cloud\_base} = \min(\text{cloud object BT}) \tag{22}$$

Where,

$R = sqrt\,(\text{total pixels of a cloud object}/2\pi).$

$$\text{If } T_{cloud\_object} > T_{cloud\_base}$$

$$T_{cloud\_object} = T_{cloud\_base}. \tag{23}$$

Since within the cloud object the air is wet, Fmask assumes the lapse rate in the cloud is a constant of $-6.5\,\text{K}\,\text{km}^{-1}$. Therefore, the cloud top height can be estimated based on the cloud base height and relative BT difference between cloud base and cloud top:

$$H_{cloud\_top} = H_{cloud\_base} + 6.5\left(T_{cloud\_base}-T_{cloud\_top}\right)\text{km.} \tag{24}$$

Finally, as the matched cloud shadow may have holes, Fmask buffers by 3 pixels in 8-connected directions for each of the matched cloud shadow pixel to fill those small holes. Moreover, as the potential cloud shadow layer produced previously includes all shadow areas, Fmask further refines the cloud shadow mask by only choosing the overlap between the potential cloud shadow layer and the matched cloud shadow objects. For cloud objects less than 3 pixels, Fmask excludes them from cloud mask and does not match cloud shadows for them as most of them are misidentification of small bright cold noncloud pixels.

The details of the cloud and cloud shadow matching algorithm are shown in Fig. 3. Because snow pixels, cloud pixels, and cloud shadow pixels may overlap, Fmask sets cloud pixels to have the highest priority, cloud shadow pixels have the second highest priority, and snow pixels have the lowest priority. In this case, if the three classes overlap for a pixel, the class with the highest priority will be its label.

4. Fmask results and accuracy assessment

4.1. Fmask results

By comparing the results of Fmask with false color composites visually (Fig. 4a, b, c, d), it appears to work well in identifying cloud (yellow), cloud shadow (green), and snow/ice (cyan). Fig. 4a is one of the Sub-tropical South images with well-behaved clouds and cloud shadows over highly vegetated areas. Fmask showed its strong ability in identifying this kind of clouds (including some of the thin clouds) and their shadows. On the other hand, Fig. 4b is a Sub-tropical North image with large variability in surface reflectances. Fmask works well in terms of identifying clouds in areas of very bright rock and has no problem in labeling cloud shadows over this bright surface. Furthermore, Fig. 4c is one of the South Polar images with thick and thin clouds over very bright snow/ice. The snow/ice is accurately identified in cyan color, and the clouds (both thick and thin) are separated well from the bright snow/ice. Finally, Fig. 4d is a very difficult image, as it has extremely thin cirrus clouds (see red arrows) and bright turbid water (see yellow arrows). In the Fmask result, there is no commission errors of clouds from the bright turbid water and the extremely thin cirrus clouds are also identified with high accuracy. This sort of qualitative evaluation was an important
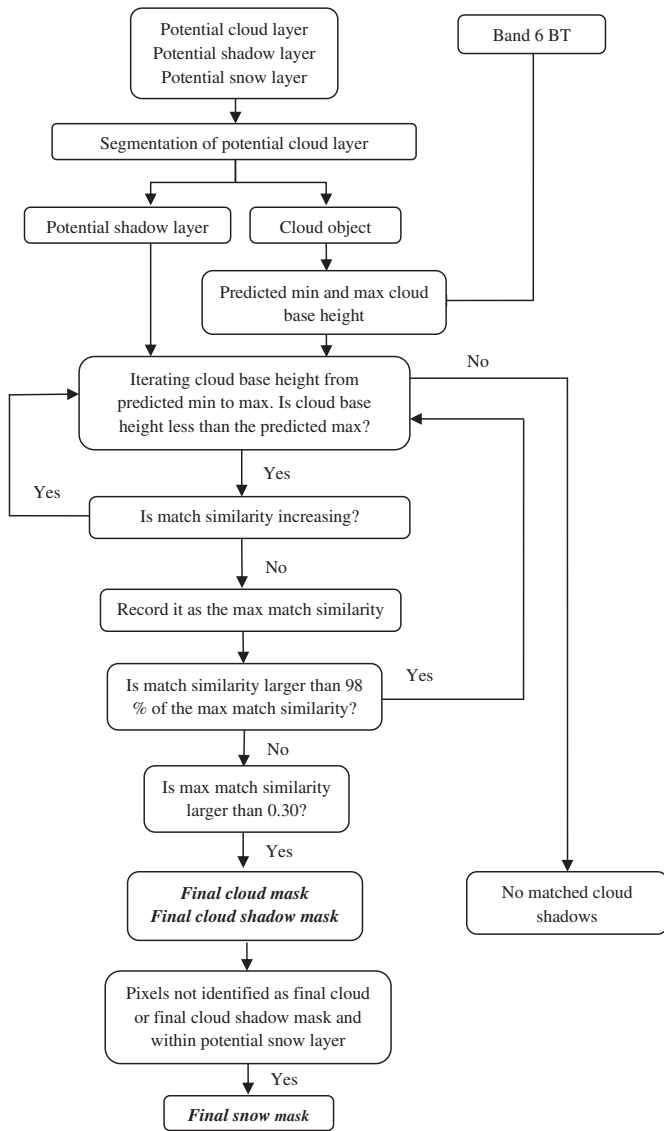
Fig. 3. Flow chart of object-based cloud and cloud shadow match algorithm.

part of the development of the algorithm. To more rigorously assess its accuracy, reference data are used.

4.2. Accuracy assessment of Fmask

The ACCA reference scenes are the only sample available at present designed to systematically cover the full range of global environments and cloud conditions (Irish et al., 2006). There are manual cloud masks for all reference scenes and a few of them have manual cloud shadow masks. A total of 188 random Landsat Worldwide Reference System (WRS) locations in nine latitudinal zones were chosen (Fig. 5).

There are 212 reference scenes evenly distributed among nine latitude zones. The manual mask was derived by visual assessment of full resolution scene in Adobe Photoshop using different combinations of bands (overlay the resampled thermal band if necessary) by three U. S. Geological Survey (USGS) image analysts. To obtain the approximate error of the manual masks, 11 scenes were examined by all three analysts, and the average difference was around 7% (Oreopoulos et al., 2011). Due to the difficulty in identifying cloud wood cloud shadow, not all the reference masks are accurate enough for accuracy

assessment of cloud mask at the pixel level. After carefully looking through the reference archive (by experts from Boston University and USGS), a total of 70 reference scenes were excluded, due to either low accuracy of the cloud manual mask or artifacts in the Landsat reference images. The remaining 142 reference scenes were used for accuracy assessment of Fmask results. The cloud shadow reference masks are not as accurate as the cloud reference masks as these reference scenes were originally interpreted to test estimates of percent cloud cover. In total, there are 26 scenes for accuracy assessment of cloud shadows. Five different accuracies were used to assess the accuracy of the algorithm results. Considering cloud and noncloud (including cloud shadow) as two classes, we have the following three accuracies for cloud accuracy assessment (Eqs. 25–27):

Cloud overall accuracy
$$= \frac{\text{agreement between manual mask and algorithm mask}}{\text{total pixels}}. \quad (25)$$

Cloud producers accuracy
$$= \frac{\text{agreement of cloud}}{\text{agreement of cloud} + \text{omission of cloud}}. \quad (26)$$

$$\text{Cloud users accuracy} = \frac{\text{agreement of cloud}}{\text{agreement of cloud} + \text{omission of cloud}}. \quad (27)$$

On the other hand, considering shadow and nonshadow (include cloud) as two classes, we have the following two accuracies for cloud shadow accuracy assessment (Eqs. 28–29):

Cloud shadow producers accuracy
$$= \frac{\text{agreement of cloud shadow}}{\text{agreement of cloud shadow} + \text{omission of cloud shadow}}. \quad (28)$$

Cloud shadow users accuracy
$$= \frac{\text{agreement of cloud shadow}}{\text{agreement of cloud shadow} + \text{comission of cloud shadow}}. \quad (29)$$

The cloud shadow overall accuracy is not used for accuracy assessment, because cloud shadows are usually much smaller in size compared to clouds, and this would make the cloud shadow overall accuracy always high even if cloud shadows identification is totally wrong. We suggest that producers accuracy is more important than users accuracy, because errors of omission of clouds or cloud shadows are more serious than errors of commission. If clouds or cloud shadows are missed they will greatly undermine future analyses like change detection or image classification. However if clear areas are masked as clouds or cloud shadows, the only consequence is a little lost data.

In addition to the per-pixel accuracies described above, Fmask results are compared with ACCA in terms of percent cloud cover. ACCA mainly consists of two passes in which the second pass of ACCA is only used to improve the scene-wide cloud cover percent and the first pass is the only phase that creates a per-pixel cloud mask (Scaramuzza, 2010). Therefore, we compared the Fmask cloud cover percent with results of the second pass of ACCA, and Fmask cloud accuracies at the pixel level with results from the first pass of ACCA.

Estimates of percent cloud cover from Fmask are very accurate (Fig. 6), with an R-square of more than 0.99. The slope of the regression line is 1.00, with a very small interception (0.83%), and relatively small Root Mean Square Error (RMSE) (3.25%). On the other hand, ACCA estimates of percent cloud cover are also accurate, with an R-square of 0.95 and the slope of the regression line is 0.95 with an interception of 0.39% and an RMSE of 6.56%. For the purpose of estimating percent cloud cover for a scene, Fmask appears to be an improvement over ACCA as except for the magnitude of interception, it has a higher R-square value, lower RMSE, and less bias in the slope of regression line.
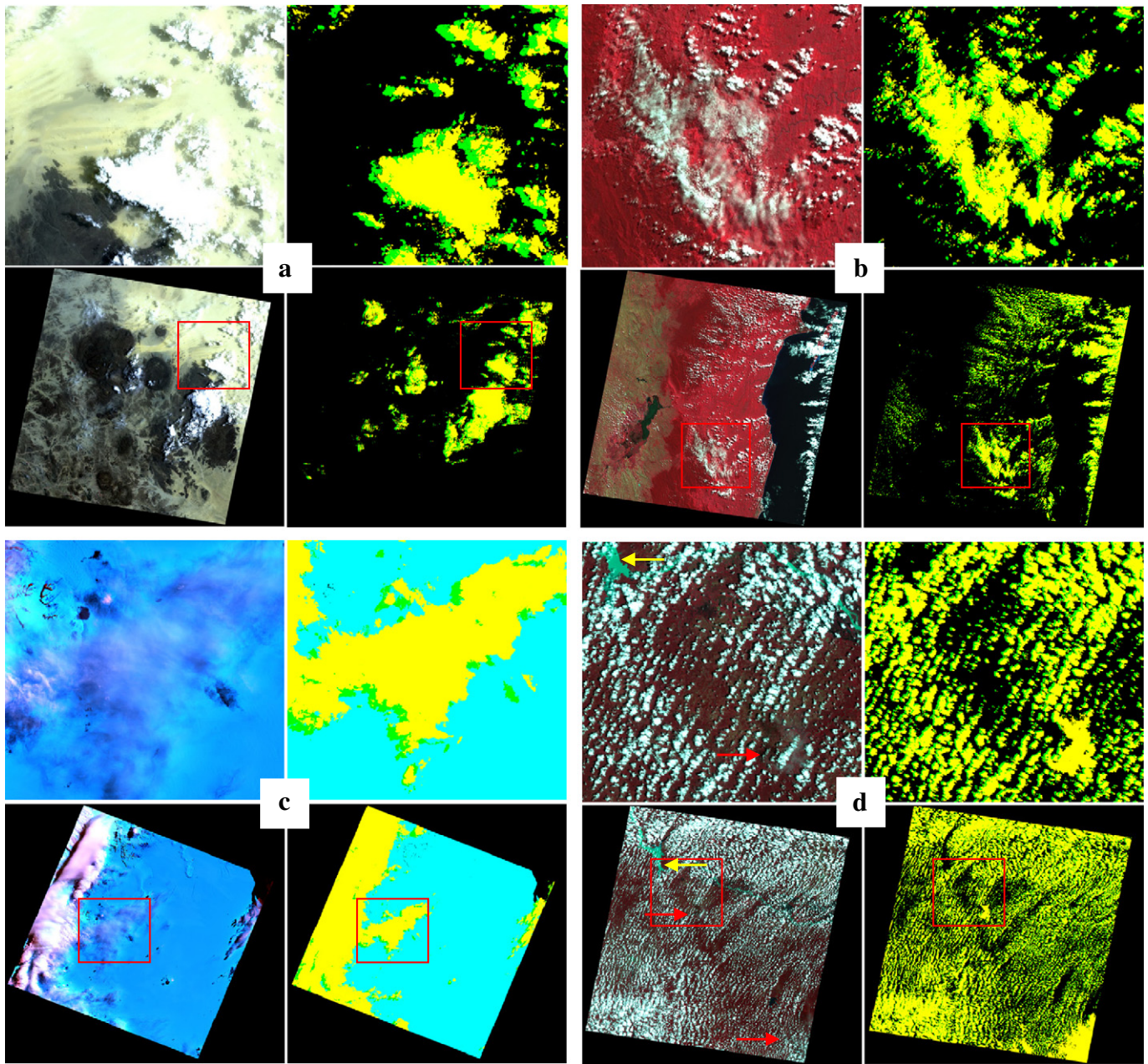
Fig. 4. Fmask results of four Landsat scenes. (a). Results of a Sub-tropical South Landsat scene (p31_r43 and 20010615). (b). Results of a Sub-tropical North Landsat scene (p189_r47 and 20010805). (c). Results of a Polar South Landsat scene (p217_r107 and 20011215). (d). Results of a Tropical Landsat scene (p190_r54 and 20010929). In each Landsat scene, (lower left) shows an entire false color composited Landsat image (Fig. 4a, b, and d are shown with Bands 4, 3, and 2 composited; Fig. 4c is shown with Bands 5, 4, and 4 composited). (Lower right) shows the corresponding Fmask cloud (yellow) and cloud shadow (green) mask for the whole scene. Black pixels are clear. (Upper left) and (upper right) images are enlargements of (lower left) and (lower right) images with a size of $60 \times 60 \text{ km}^2$.

At the pixel scale, the average Fmask cloud overall accuracy is 96.41% with a small standard deviation of 3.2% (Fig. 7). It is a significant increase compared with ACCA whose average overall cloud accuracy is 84.8% with a standard deviation of 11.9%. Cloud producers and users accuracies for images with cloud cover less than 5% were not analyzed here, as producers and users accuracies computed for clouds with very small size may be biased greatly because of the definition of cloud boundaries. The average Fmask cloud producers accuracy is 92.1% (Fig. 8) with a standard deviation of 13.3% which is a significant improvement compared with ACCA whose average cloud producers accuracy is 72.1% with a standard deviation of 26.5%. Moreover, the

average Fmask cloud users accuracy is 89.4% (Fig. 9) with a standard deviation of 9.8% which is similar as ACCA whose average cloud users accuracy is 91.8% with a standard deviation of 12.3%. Considering that producers accuracy is more important (in our opinion) than users accuracy, the improvement of cloud identification in Fmask is significant compared with ACCA.

On the other hand, Fmask seems to overestimate cloud shadows, which is mainly caused by the 3 pixels buffered (in 8-connected neighborhood) for each cloud shadow pixel (Fig. 10). The average producers accuracy for cloud shadow is larger than 70%, and its average users accuracy is around 50% (images with cloud shadows
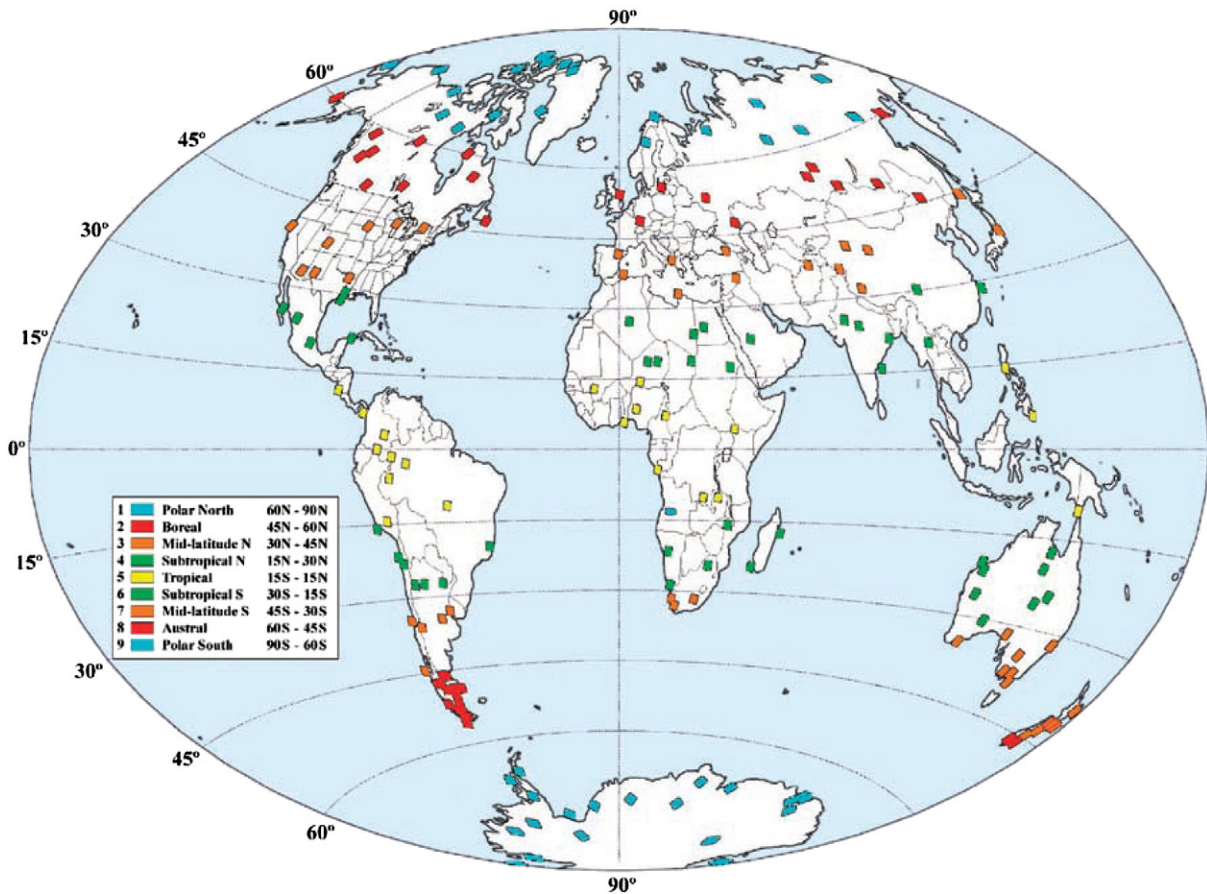
Fig. 5. Locations of the Landsat scenes used in the reference dataset (Plate 7. in Irish et al., 2006).

covering less than 1% of the image are not analyzed here). The lower accuracies are partly the result of errors in the manual cloud shadow masks and the relatively small size of cloud shadows compared to clouds in the scene. Even very small amounts of disagreement (differences in defining cloud shadow boundaries, mistakes in Fmask or reference shadow mask) reduce the cloud shadow users and producers accuracy greatly.

5. Discussion and conclusion

The Fmask algorithm effectively finds clouds and cloud shadows, which helps with a wide assortment of remote sensing activities. The goal is to provide an automated method for screening clouds and their shadows such that time series of Landsat images can be easily compiled. The need for effective cloud and shadow screening has
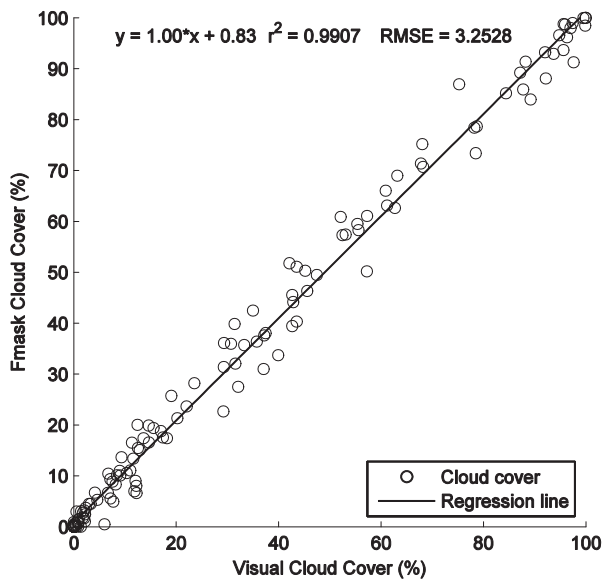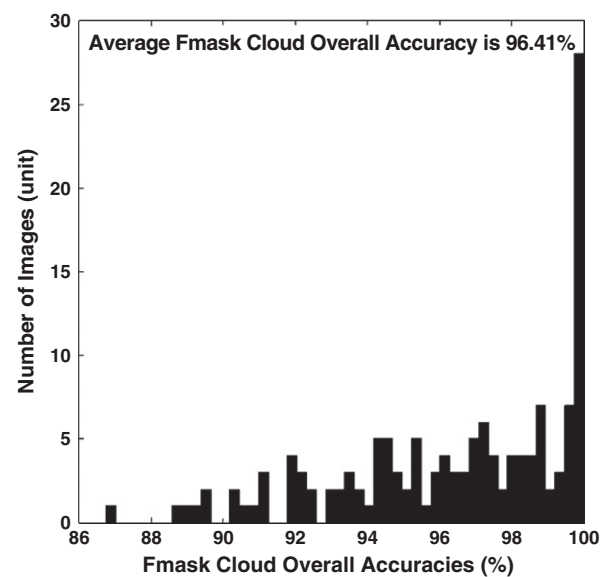

Fig. 6. Visual cloud cover vs. Fmask cloud cover.

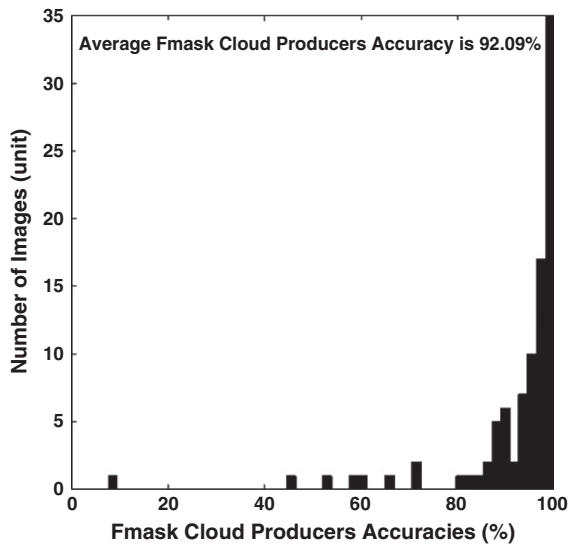**Fig. 7.** Histogram of Fmask cloud overall accuracies.

Fig. 8. Histogram of Fmask cloud producers accuracies.
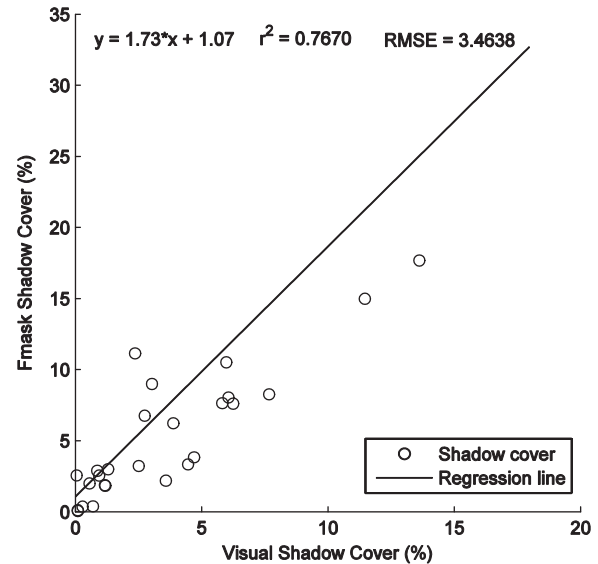

Fig. 10. Visual cloud shadow cover vs. Fmask cloud shadow cover.

grown tremendously for two primary reasons. First, the Landsat L1T format now provides accurate enough registration of images that they can be compiled into a time series without significant attention to registration issues. And second, free access to the archive is changing what we consider a useful Landsat image. Now that images are free, it can be worth processing images even if substantial portions of the images are cloudy to extract the cloud free observations. As a result, more images with more clouds are being used and the need for automated cloud and cloud shadow screening is growing.

The estimates of percent cloud cover from Fmask are a slight improvement compared with ACCA estimates. The cloud masks generated from Fmask are significantly better than from the first pass of ACCA, with cloud overall accuracy of 96.41% (84.8% in ACCA), cloud producers accuracy of 92.1% (72.1% in ACCA), and cloud users accuracy of 89.4% (91.8% in ACCA). The cloud probability mask generated from Fmask will be beneficial for customizing cloud masking results, as instead of a binary mask, it can provide the probability of a pixel being cloudy. Users can make their own decisions in choosing the confidence level (e.g. 50%) for defining a cloudy pixel for their specific
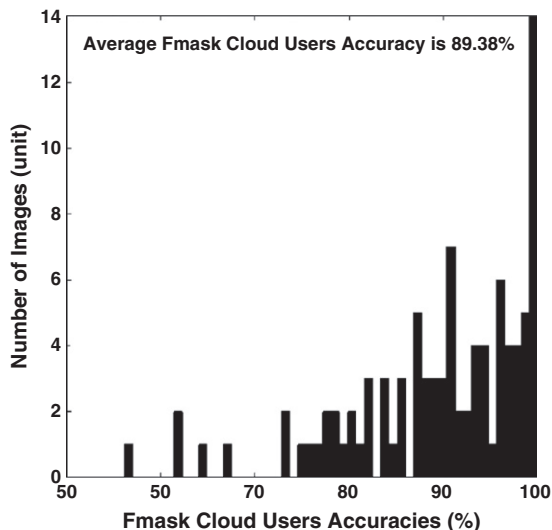
applications. Fmask has achieved producers accuracy for cloud shadow of more than 70% and users accuracy higher than 50%. The reliability of these estimates is questionable due to the frequency of errors in the reference datasets. By examining each cloud shadow mask carefully, we find Fmask identified many cloud shadows that are not included in the cloud shadow reference masks. Therefore, more accurate reference data for cloud shadows are necessary for better assessment of cloud shadow detection algorithms.

There are some limitations in Fmask cloud detection. First, Fmask may fail to identify a cloud if it is both thin and warm (Fig. 11 upper left and upper right images). These errors of omission may not be that important, as usually thin and warm "clouds" are actually haze or aerosols and they can be further removed by atmospheric correction (Vermote & Saleous, 2007). Second, Fmask may also identify other very bright and cold land features (salt pans, cold snow etc.) as clouds (Fig. 11 lower left and lower right images). We think commission is better than omission in cloud detection, as this kind of error will only remove a few clear-sky pixels from subsequent remote sensing applications. Finally, as Fmask uses a scene-based threshold and applies this same threshold to all pixels in the image, it may not work well for some images with very complex surface reflectances.

The cloud accuracy of Fmask would increase if we included other ancillary data such as a land/water mask, DEM data, and multi-temporal data. The time required for Fmask to process one scene is determined by the percentage cloud cover. Generally, it takes 0.5 to 6 min for Fmask (1.6.0 Matlab version) to process one scene on an 8 core Linux server. Thanks to the LEDAPS atmosphere correction tool, Fmask can be completely automated (without any manual inputs or tuning) based on the TOA reflectance files generated by LEDAPS.

One issue needing further attention is the establishment of a standard definition for clouds and cloud shadows. For example, what category should we put pixels that are shadows from a high cloud falling on a lower cloud? Should we include smoke, thick aerosols, and haze in the cloud mask? Shall we still validate cloud shadows over water? Answers to these questions would facilitate future accuracy assessments and comparison of alternative methods.

In conclusion, a new cloud and cloud shadow masking algorithm (Fmask) is the result of combining ideas from many past approaches and integrating an innovative new object-based approach. It exhibits high accuracy for cloud detection and is an improvement over the traditional ACCA cloud algorithm. At the same time it can also provide an accurate cloud shadow mask which has limited remote sensing activities for a long time.
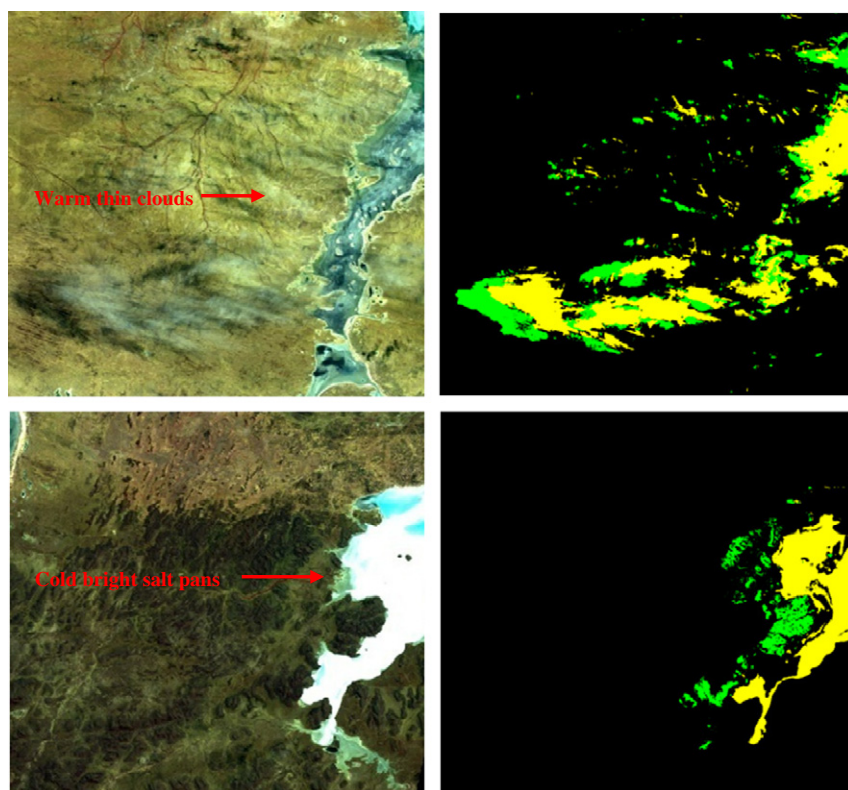

Fig. 9. Histogram of Fmask cloud users accuracies.

Fig. 11. Two subset Landsat images from Mid-latitude South (p100_r82 and 20011212).(Lower left) shows a subset of Landsat image where commission of clouds was observed (Bands 4, 3, and 2 composited). (Lower right) shows the corresponding Fmask cloud (yellow) and cloud shadow (green) mask. Black pixels are clear. (Upper left) shows a subset of Landsat images where omission of clouds was observed (Bands 4, 3, and 2 composited). (Upper right) shows the corresponding Fmask cloud and cloud shadow mask. The red arrows point to the omission and commission errors of clouds in Fmask results as examples.

Acknowledgments

References

Ackerman, S. A., Strabala, K. I., Menzel, W. P., Frey, R. A., Moeller, C. C., & Gumley, L. E. (1998). Discriminating clear sky from clouds with MODIS. *Journal of Geophysical Research*, *103*(24), 32141–32157.

Arvidson, T., Gasch, J., & Goward, S. N. (2001). Landsat-7's long-term acquisition plan — An innovative approach to building a global imagery archive. *Remote Sensing of Environment*, *78*(1–2), 13–26.

Asner, G. P. (2001). Cloud cover in Landsat observation of the Brazilian Amazon. *International Journal of Remote Sensing*, *22*(18), 3855–3862.

Berendes, T., Sengupta, S. K., Welch, R. M., Wielicki, G. A., & Navar, M. (1992). Cumulus cloud base height estimation from high spatial resolution Landsat data: A Hough transform approach. *IEEE Transactions on Geoscience and Remote Sensing*, *30*, 430–443.

Choi, H., & Bindschadler, R. (2004). Cloud detection in Landsat imagery of ice sheets using shadow matching technique and automatic normalized difference snow index threshold value decision. *Remote Sensing of Environment*, *91*, 237–242.

Cohen, W. B., Fiorella, M., Gray, J., Helmer, E., & Anderson, K. (1998). An efficient and accurate method for mapping forest clear cuts in the Pacific Northwest using Landsat imagery. *Photogrammetric Engineering and Remote Sensing*, *64*(4), 293–300.

Coiner, J. C. (1980). Using Landsat to monitor changes in vegetation cover induced by desertification processes. *Symposium on Remote Sensing of the Environment*, *3*(14), 1341–1347.

Coppin, P. R., & Bauer, M. E. (1994). Processing of multitemporal Landsat TM imagery to optimize extraction of forest cover change features. *IEEE Transactions on Geoscience and Remote Sensing*, *32*, 918–927.

Derrien, M., Farki, B., Harang, L., LeGleau, H., Noyalet, A., Pochic, D., & Sairouni, A. (1993). Automatic cloud detection applied to NOAA-11/AVHRR imagery. *Remote Sensing of Environment*, *46*(3), 246–267.

Dozier, J. (1989). Spectral signature of alpine snow cover from the Landsat Thematic Mapper. *Remote Sensing of Environment*, *28*, 9–22.

Gao, B. C., & Kaufman, Y. J. (1995). Selection of the 1.375-μm MODIS channel for remote sensing of cirrus clouds and stratospheric aerosols from space. *American Meteorological Society*, *52*, 4231–4237.

Gao, B. C., Kaufman, Y. J., Han, W., & Wiscombe, W. J. (1998). Correction of thin cirrus path radiance in the 0.4–1.0 μm spectral region using the sensitive 1.375 μmc cirrus detecting channel. *Journal of Geophysical Research*, *103*(24), 32169–32176.

Gao, B. C., Yang, P., Han, W., Li, R. R., & Wiscombe, W. J. (2002). An algorithm using visible and 1.38-μm channels to retrieve cirrus cloud reflectances from aircraft and satellite data. *IEEE Transactions on Geoscience and Remote Sensing*, *40*(8), 1659–1668.

Gomez-Chova, L., Camps-Valls, G., Galpe-Maravilla, J., Guanter, L., & Moreno, J. (2007). Cloud-screening algorithm for ENVISAT/MERIS multispectral images. *Geoscience and Remote Sensing*, *45*(12), 2 4105–4118.

Hall, D. K., Riggs, G. A., & Salomonson (2001). *Algorithm theoretical basis document (ATBD) for the MODIS snow and sea ice-mapping algorithms.*

Hartmann, D. L. (1994). Global physical climatology (1st ed.). San Diego: Academic Press (Chapter 3).

Hégarat-Mascle, S. L., & André, C. (2009). Use of Markov Random Fields for automatic cloud/shadow detection on high resolution optical images. *ISPRS Journal of Photogrammetry and Remote Sensing*, *64*, 351–366.

Hutchison, K. D., Mahoney, R. L., Vermote, E. F., Kopp, T. J., Jackson, J. M., Sei, A., & Iisager, B.D. (2009). A geometry-based approach to identifying cloud shadows in the VIIRS cloud mask algorithm for NPOESS. *Journal of Atmospheric and Oceanic Technology*, *26*(7), 1388–1397.

Irish, R. (2000). Landsat-7 automatic cloud cover assessment algorithms for multispectral, hyperspectral, and ultraspectral imagery. *The International Society for Optical Engineering*, *4049*, 348–355.

Irish, R., Barker, J. L., Goward, S. N., & Arvidson, T. (2006). Characterization of the Landsat-7 ETM+ Automated Cloud-Cover Assessment (ACCA) algorithm. *Photogrammetric Engineering and Remote Sensing*, *72*(10), 1179–1188.

Luo, Y., Trishchenko, A. P., & Khlopenkov, K. V. (2008). Developing clear-sky, cloud and cloud shadow mask for producing clear-sky composites at 250-meter spatial

resolution for the seven MODIS land bands over Canada and North America. *Remote Sensing of Environment, 112*, 4167–4185.

Masek, J. G., Vermote, E. F., Saleous, N., Wolfe, R., Hall, E. F., Huemmrich, F., et al. (2006). A Landsat surface reflectance data set for North America, 1990–2000. *Geoscience and Remote Sensing Letters, 3*, 68–72.

Oreopoulos, L., Wilson, M., & Várnai, T. (2011). Implementation on Landsat data of a simple cloud mask algorithm developed for MODIS land bands. *IEEE Transactions on Geoscience and Remote Sensing, 8*(4), 597–601.

Saunders, R. W., & Kriebel, K. T. (1998). An improved method for detecting clear sky and cloudy radiances from AVHRR data. *International Journal of Remote Sensing, 9*(1), 123–150.

Scaramuzza, P. (2010, Nov). *United States Geological Survey Earth Resources Observation Systems (EROS) Data Center (EDC), Sioux Falls, SD, private communication.* .

Seto, K. C., Woodcock, C. E., Song, C., Huang, X., Lu, J., & Kaufmann, R. K. (2002). Monitoring land-use change in the Pearl River Delta using Landsat TM. *International Journal of Remote Sensing, 23*(10), 1985–2004.

Simpson, J. J., & Stitt, J. R. (1998). A procedure for the detection and removal of cloud shadow from AVHRR data over land. *Geoscience and Remote Sensing, 36*(3), 880–890.

Simpson, J. J., Stitt, J. R., & Jin, Z. (2000). Cloud shadow detection under arbitrary viewing and illumination conditions. *IEEE Transactions on Geoscience and Remote Sensing, 36*, 880–897.

Soille, P. (1999). *Morphological image analysis: Principles and applications* (pp. 173–174). : Springer-Verlag.

Soille, Pierre, Vogt, Jugen, & Columbo, Roberto (2003). Carving and adaptive drainage enforcement of grid digital elevation models. *Water Resources Research, 12*, 1–13.

Vermote, E. (2010, Jan.). *Landsat Science Team Meeting* Mountain View, California, USA.

Vermote, E., & Saleous, N. (2007). *LEDAPS surface reflectance product description, Version 2.0.*

Wang, B., Ono, A., Muramatsu, K., & Fujiwara, N. (1999). Automated detection and removal of clouds and their shadows from Landsat TM images. *IEICE Transactions on Information and Systems, E82-D*, 2.

Wildt, M. D. R. D., Seiz, G., & Gruen, A. (2007). Operational snow mapping using multitemporal Meteosat SEVIRI imagery. *Remote Sensing of Environment, 109*, 29–41.

Zhang, Y., Guindon, B., & Cihlar, J. (2002). An image transform to change characterize and compensate for spatial variability in thin cloud contamination of Landsat images. *Remote Sensing of Environment, 82*(2–3), 173–187.

Zhang, Y., Rossow, W. B., Lacis, A. A., Oinas, V., & Mishchenko, M. I. (2004). Calculation of radiative fluxes from the surface to top of atmosphere based on ISCCP and other global data sets: Refinements of the radiative transfer model and the input data. *Journal of Geophysical Research, 109*(19), 19105.