



Automated cloud and cloud shadow identification in Landsat MSS imagery for temperate ecosystems



Justin D. Braaten ^{a,*}, Warren B. Cohen ^b, Zhiqiang Yang ^a

^a Department of Forest Ecosystems and Society, Oregon State University, 321 Richardson Hall, Corvallis, OR 97331, United States

^b Pacific Northwest Research Station, USDA Forest Service, Corvallis, OR 97331, United States

ARTICLE INFO

Article history:

Received 14 October 2014

Received in revised form 27 July 2015

Accepted 6 August 2015

Available online xxxx

Keywords:

Landsat MSS

Automated cloud masking

Time series analysis

Change detection

Large area mapping

ABSTRACT

Automated cloud and cloud shadow identification algorithms designed for Landsat Thematic Mapper (TM) and Thematic Mapper Plus (ETM+) satellite images have greatly expanded the use of these Earth observation data by providing a means of including only clear-view pixels in image analysis and efficient cloud-free compositing. In an effort to extend these capabilities to Landsat Multispectral Scanner (MSS) imagery, we introduce MSS clear-view-mask (MSScm), an automated cloud and shadow identification algorithm for MSS imagery. The algorithm is specific to the unique spectral characteristics of MSS data, relying on a simple, rule-based approach. Clouds are identified based on green band brightness and the normalized difference between the green and red bands, while cloud shadows are identified by near infrared band darkness and cloud projection. A digital elevation model is incorporated to correct for topography-induced illumination variation and aid in identifying water. Based on an accuracy assessment of 1981 points stratified by land cover and algorithm mask class for 12 images throughout the United States, MSScm achieved an overall accuracy of 84.0%. Omission of thin clouds and bright cloud shadows constituted much of the error. Perennial ice and snow, misidentified as cloud, also contributed disproportionately to algorithm error. Comparison against a corresponding assessment of the Fmask algorithm, applied to coincident TM imagery, showed similar error patterns and a general reduction in accuracy commensurate with differences in the radiometric and spectral richness of the two sensors. MSScm provides a suitable automated method for creating cloud and cloud shadow masks for MSS imagery required for time series analyses in temperate ecosystems.

© 2015 Elsevier Inc. All rights reserved.

1. Introduction

The Landsat free and open data policy (Woodcock et al., 2008) provided the opportunity to realize the full potential of Landsat's unparalleled record of Earth observation data (Wulder, Masek, Cohen, Loveland, & Woodcock, 2012). This accomplishment, along with production and distribution of high quality standardized products by Earth Resources Observation and Science Center (EROS) (Loveland & Dwyer, 2012), has prompted development of powerful image processing and change detection algorithms that take advantage of Landsat's long record, high temporal dimensionality, and global coverage (Griffiths et al., 2014; Hansen & Loveland, 2012; Hilker et al., 2009; Huang et al., 2010; Kennedy, Yang, & Cohen, 2010; Kennedy et al., 2012; Roy et al., 2010; Zhu & Woodcock, 2014a; Zhu, Woodcock, & Olofsson, 2012). The benefit of these actions for landscape monitoring and mapping at previously impractical temporal and spatial scales cannot be overstated, and much credit is due to automated cloud masking, which plays an essential role in the implementation of these algorithms.

Clouds and their shadows block or reduce satellite sensors' view of Earth surface features, obscuring spectral information characteristic of clear-sky viewing. This spectral deviation from clear-sky view can cause false change in a change detection analysis and conceals true land cover, which can reduce the accuracy and information content of map products where cloud-free images are not available. As a result, cloud and cloud shadow identification and masking are important and often necessary pre-processing steps. Development of automated cloud and cloud shadow identification systems for Landsat Thematic Mapper (TM) and Enhanced Thematic Mapping Plus (ETM+) imagery (Goodwin, Collett, Denham, Flood, & Tindall, 2013; Huang et al., 2010; Hughes & Hayes, 2014; Irish, Barker, Goward, & Arvidson, 2006; Oreopoulos, Wilson, & Várnai, 2011; Zhu & Woodcock, 2012; Zhu & Woodcock, 2014b) have greatly relieved the costs of this traditionally time-consuming manual process, largely facilitating the proliferation and evolution of large volume Landsat data algorithms. However, no automated cloud and cloud shadow identification algorithm exists for Landsat Multispectral Scanner (MSS) imagery. The lack of such a system stands as a significant barrier to incorporating MSS into current large volume Landsat-based mapping and time series analysis efforts.

Including MSS as an integral data component in Landsat time series analysis is important for both the Landsat program and the

* Corresponding author.

E-mail address: justin.braaten@oregonstate.edu (J.D. Braaten).

science it supports. MSS provides rich temporal context for the current state of land use and land cover, which has been shown to increase the accuracy of predicting forest structure (Pflugmacher, Cohen, & Kennedy, 2012). Additionally, the extended temporal record it provides, better tracks long-term Earth surface changes, including forest dynamics, desertification, urbanization, glacier recession, and coastal inundation. It also has the benefit of increasing the observation frequency of cyclic and sporadic events, such as drought, insect outbreaks, wildfire, and floods. Furthermore, utilization of the full 42-plus years of Landsat imagery sets an example of effectual resource use and supports the need for continuity of Landsat missions to provide seamless spatial and temporal coverage into the future. Without comprehensive inclusion of MSS data the true power and benefit of the Landsat archive is not fully realized.

Recent and ongoing work promises to increase the capacity and suitability of MSS imagery for efficient and robust integration with TM, ETM+, and Operational Land Imager (OLI) (Gómez et al., 2011; Lobo, Costa, & Novo, 2015). Continual improvements of georegistration methods provide better spatial correspondence between coincident pixels of varying dates, and hold the potential to increase the proportion of analysis-ready images (Choate, Steinwand, & Rengarajan, 2012; Devaraj & Shah, 2014). Improved radiometric calibration coefficients developed by Helder et al., 2012 facilitate the future development of a standard surface reflectance model, and application of common cross-sensor spectral transformations (e.g., tasseled cap angle and NDVI) have proved successful in spectral harmonization methods for time series analysis (Pflugmacher et al., 2012).

In an effort to build on these developments and further enhance MSS usability, we present an automated approach to cloud and cloud shadow identification in MSS imagery. The algorithm, MSScvm (MSS clear-view-mask), is designed for the unique properties of MSS data, relying on a series of spectral tests on single band brightness and normalized band differences to identify cloud, cloud shadow, and clear-view pixels. It also incorporates a digital elevation model (DEM) and cloud projection to better separate cloud shadow from topographic shading and water.

2. Methods

2.1. MSScvm background

Successful automated identification of clouds and cloud shadows in imagery requires robust logic and fine tuning to minimize and balance commission and omission errors. Current TM/ETM+ methods generally achieve this through a series of multi-spectral tests, relying collectively on the full range of bands to identify a pixel's condition. Of particular importance is the Thermal Infrared (TIR) band, which leverages the characteristically cold temperature of clouds to separate them from similarly bright and white land cover such as barren sand/soil, rock, and impervious cover. TIR data are also used to identify cloud shadows by cloud projection and dark object matching using cloud temperature and adiabatic lapse rate to estimate cloud height, which is more accurate than spectral tests alone (Zhu & Woodcock, 2012).

Although MSS and TM/ETM+ are of the same Landsat affiliation, their spatial, spectral, and radiometric qualities differ, making application of TM/ETM+ systems difficult. By comparison, MSS is missing spectral representation from the blue, shortwave infrared (SWIR), and TIR windows, and has a reduced 6-bit radiometric resolution (scaled to 8-bit for distribution). These differences require that a new system be developed specifically for MSS data.

MSScvm is an automated rule-based algorithm for identifying clouds, cloud shadows, and clear-view pixels in MSS imagery (see Table 1 for a description of MSS spectral bands and the naming convention used throughout the text).

The input data are Landsat MSS Level 1 Product Generation System (LPGS) images converted to top-of-atmosphere (TOA) reflectance

Table 1

Landsat MSS band designations. Landsat MSS band alias is used as the band identifier throughout the text.

Landsat MSS band alias	Landsat MSS 1–3 band label	Landsat MSS 4 & 5 band label	Wavelength (μm)
B1	Band 4	Band 1	0.5–0.6
B2	Band 5	Band 2	0.6–0.7
B3	Band 6	Band 3	0.7–0.8
B4	Band 7	Band 4	0.8–1.1

(Chander, Markham, & Helder, 2009) and a corresponding DEM. The output is a binary raster mask with pixel values equal to one and zero, representing assignment as clear-view and obscured, respectively. This task is accomplished with five sequential steps: 1) cloud identification, 2) water identification, 3) candidate cloud shadow identification, 4) candidate cloud projection, 5) final mask class assignment (Fig. 1). In general, all algorithm decision rules were selected through a combination of trial and error and iterative optimization. We tested and selected the threshold values using 67 training images from 20 scenes in the western United States representing all five MSS sensors (Table 2). We conducted an accuracy assessment of the algorithm to demonstrate its utility and identify its strengths and weaknesses. An identical concurrent assessment of the widely used Fmask algorithm (Zhu & Woodcock, 2012), applied to coincident TM imagery, was conducted to provide a comparative standard for the performance of MSScvm.

2.2. Cloud layer

Cloud identification is based on brightness in B1 and the normalized difference between B1 and B2 (Fig. 1a), according to the following spectral test (Eq. (1)).

$$\text{Cloud test} = (\text{B1} > 0.175 \text{ and } \text{NDGR} > 0.0) \text{ or } \text{B1} > 0.39, \quad (1)$$

where

$$\text{NDGR} = (\text{B1} - \text{B2}) / (\text{B1} + \text{B2}).$$

Pixels meeting these test criteria are classified as clouds. The first part of Eq. (1) identifies pixels that are relatively bright in B1 and have greater reflectance in B1 relative to B2. B1 pixels with TOA reflectance greater than 0.175 are moderately bright and represent both clouds and bright non-cloud features. Incorporating a positive NDGR value as a qualifier in this test typically separates clouds from non-cloud features. Occasionally, however, NDGR is not positive for very bright clouds, in which case the simple brightness test of B1 TOA reflectance values greater than 0.39 captures them. This threshold value is extremely bright and not often represented by non-cloud features except snow, which is not explicitly separated in the algorithm because bright cloud and snow are essentially spectrally indistinguishable in MSS data, based on our observations during algorithm development.

B1 plays an important role in these tests. The shorter wavelengths composing it are more sensitive to atmospheric scattering by cloud vapor and aerosols than the longer wavelengths recorded in the other bands (Chavez, 1988; Zhang, Guindon, & Cihlar, 2002). As a result, clouds, haze, and aerosol appear brighter in B1 than the others bands, making it a good simple brightness index, as well as a standard for relative band comparisons. B2 was selected as the contrasting band because initial evaluations showed that contrasts with B3 and B4 were less consistent in their relationship to B1 than B2, especially for vegetated pixels, which spike in these cover types because of high near-infrared scattering by green vegetation.

The high reflectivity of clouds in B1 and B2, however, can sometimes cause radiometric saturation, resulting in false differences between the bands, forcing NDGR values below the cloud-defining threshold.

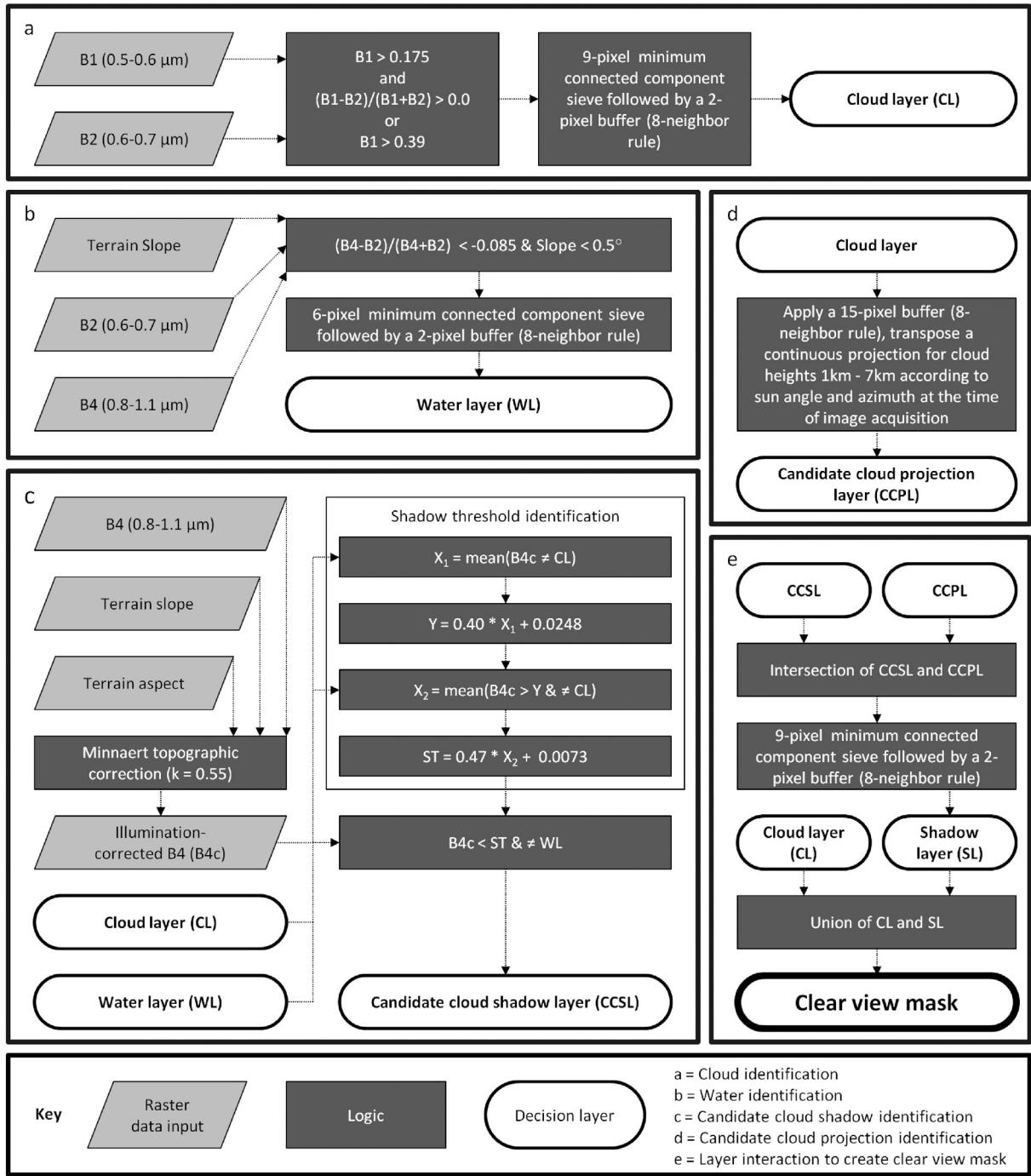


Fig. 1. MSScvm algorithm schematic showing the inputs, logic, and workflow for the multi-step process.

Karnieli, Ben-Dor, Bayarjargal, & Lugasi, 2004 suggest extrapolating the values of saturated pixels based on linear regression with other bands, but for our purpose, the simple B1 brightness test of B1 greater than 0.39 is sufficient to identify these affected pixels as cloud.

The cloud test produces a layer that contains a pattern of single and small group pixels that represent noise and small non-cloud features. A nine-pixel minimum connected component sieve is applied to eliminate these pixels from the cloud layer. Finally, a two-pixel buffer (eight-neighbor rule) is added to the remaining cloud features to capture the thin, semi-transparent cloud-edge pixels. The spatial filter and buffer parameters were determined by iterative trial and visual assessment with the goal of optimizing commission and omission error balance for the test scenes.

2.3. Water layer

A water layer dividing pixels between water and land is used to reduce the high rate of confusion between water and cloud shadows in the algorithm (Fig. 1b). The layer is produced by the following logic (Eq. (2)).

$$\text{Water test} = \text{NDVI} < -0.085 \text{ and slope} < 0.5, \quad (2)$$

where

$$\begin{aligned} \text{NDVI} &= (B4 - B2)/(B4 + B2), \\ \text{slope} &= \text{topographic slope (degrees) derived from a DEM.} \end{aligned}$$

Table 2

Images used to build the MSScvm algorithm.

Platform	WRS type	Path/row	Year	Day	Platform	WRS type	Path/row	Year	Day
Landsat 1	WRS-1	035/033	1972	214	Landsat 2	WRS-1	049/029	1977	192
Landsat 1	WRS-1	035/034	1972	214	Landsat 2	WRS-1	049/029	1977	210
Landsat 1	WRS-1	035/034	1973	244	Landsat 2	WRS-1	049/029	1978	187
Landsat 1	WRS-1	036/031	1973	191	Landsat 2	WRS-1	049/029	1979	236
Landsat 1	WRS-1	036/031	1974	204	Landsat 2	WRS-1	049/029	1981	225
Landsat 1	WRS-1	036/032	1973	245	Landsat 2	WRS-1	049/029	1981	243
Landsat 1	WRS-1	036/032	1974	204	Landsat 2	WRS-1	049/030	1975	185
Landsat 1	WRS-1	036/033	1972	233	Landsat 2	WRS-1	049/030	1976	198
Landsat 1	WRS-1	036/034	1973	245	Landsat 2	WRS-1	049/030	1976	234
Landsat 1	WRS-1	037/030	1974	205	Landsat 2	WRS-1	049/030	1978	187
Landsat 1	WRS-1	037/030	1974	241	Landsat 2	WRS-1	049/030	1979	236
Landsat 1	WRS-1	037/031	1972	234	Landsat 2	WRS-1	049/030	1980	213
Landsat 1	WRS-1	037/034	1974	223	Landsat 2	WRS-1	049/030	1981	225
Landsat 1	WRS-1	037/034	1974	241	Landsat 3	WRS-1	038/030	1978	203
Landsat 1	WRS-1	038/031	1972	235	Landsat 3	WRS-1	049/029	1979	227
Landsat 1	WRS-1	038/031	1974	224	Landsat 3	WRS-1	049/029	1982	211
Landsat 1	WRS-1	038/032	1972	235	Landsat 3	WRS-1	049/030	1982	211
Landsat 1	WRS-1	039/030	1972	218	Landsat 4	WRS-2	033/033	1989	243
Landsat 1	WRS-1	049/029	1974	181	Landsat 4	WRS-2	045/030	1983	199
Landsat 1	WRS-1	049/029	1974	217	Landsat 4	WRS-2	045/030	1983	247
Landsat 1	WRS-1	049/029	1975	212	Landsat 4	WRS-2	045/030	1989	231
Landsat 1	WRS-1	049/030	1974	181	Landsat 4	WRS-2	045/030	1992	192
Landsat 1	WRS-1	049/030	1974	199	Landsat 4	WRS-2	045/030	1992	224
Landsat 1	WRS-1	049/030	1974	235	Landsat 5	WRS-2	033/033	1985	192
Landsat 1	WRS-1	049/030	1974	253	Landsat 5	WRS-2	033/034	1985	208
Landsat 1	WRS-1	049/030	1976	207	Landsat 5	WRS-2	034/030	1988	192
Landsat 2	WRS-1	036/031	1978	246	Landsat 5	WRS-2	045/030	1984	242
Landsat 2	WRS-1	037/032	1978	211	Landsat 5	WRS-2	045/030	1985	180
Landsat 2	WRS-1	048/030	1975	202	Landsat 5	WRS-2	045/030	1986	199
Landsat 2	WRS-1	048/030	1977	173	Landsat 5	WRS-2	045/030	1986	215
Landsat 2	WRS-1	048/030	1978	222	Landsat 5	WRS-2	045/030	1990	178
Landsat 2	WRS-1	048/030	1981	242	Landsat 5	WRS-2	045/030	1990	226
Landsat 2	WRS-1	049/029	1975	185	Landsat 5	WRS-2	045/030	1992	200
Landsat 2	WRS-1	049/029	1976	234					

Pixels meeting the test criteria are flagged as water in the water layer (WL). NDVI and slope are used as delineating indices because NDVI has been demonstrated to be especially useful for separating land and water (Zhu & Woodcock, 2012), and surface slope is a good qualifier for exceptions to this rule, as water bodies are generally flat at the 60 m resolution of MSS pixels. Zhu and Woodcock (2012) use two tests of NDVI with values of 0.01 and 0.1 as water-defining thresholds. Based on initial testing we found that these values were too liberal, even with the low-slope criteria, causing false positive water identification. We found that an NDVI value less than -0.085 with a topographic slope less than 0.5° to be a good comprise between commission and omission error for water identification.

The topographic slope layer was derived from Shuttle Radar Topography Mission 30 m DEMs. The DEMs were downloaded from the Global Land Cover Facility (<http://glcf.umd.edu/data/srtm/>) as their 1-arc second WRS-2 tile, Filled Finished-B product. For each MSS scene in this study, several DEMs were mosaicked together and resampled to 60 m resolution to match the extent and resolution of the MSS images. Topographic slope for each DEM was computed according to Horn (1981) using the R (R Core Team, 2014) raster package (Hijmans, 2015) terrain function.

A six-pixel minimum connected component sieve is applied to the water layer to eliminate single and small groups of pixels generally associated with non-water features. A two-pixel buffer (eight-neighbor rule) is then applied to capture shore and near-shore pixels that have both higher NDVI values and greater slope than central water body pixels. Like the cloud layer, the spatial filter and buffer values were determined by visual assessment with the goal of optimizing commission and omission error balance for the test images.

2.4. Candidate cloud shadow layer

Candidate cloud shadow identification is based on low spectral brightness in B4 (Fig. 1c). The near-infrared wavelengths composing

this band are well suited for taking advantage of the dark nature of cloud shadows to separate them from other image features. Shadows in B4 are particularly dark because diffuse illumination by atmospheric scatter is lower in B4 than the other bands. However, a simple spectral brightness test aimed at dark feature identification inevitably includes topographic shadows, water, and other dark features, which must be removed. Topographic shadows are eliminated by applying an illumination correction to B4, and water is removed using the previously identified water layer. Additionally, in following steps (Sections 2.5 and 2.6), pixels identified as candidate cloud shadow are matched against a liberal estimate of cloud projection to boost the probability that candidate shadow-identified pixels are associated with clouds.

2.4.1. Topographic correction

In the first step of candidate cloud shadow identification, B4 is radiometrically corrected to remove topographic shading using the Minnaert correction (Meyer, Itten, Kellenberger, Sandmeier, & Sandmeier, 1993; Teillet, Guindon, & Goodenough, 1982). The correction is expressed as follows (Eq. (3)).

$$L_H = L_T (\cos \theta_o / \cos i)^k, \quad (3)$$

where

- L_H B4 TOA reflectance observed for a horizontal surface,
- L_T B4 TOA reflectance observed over sloped terrain,
- θ_o Sun's zenith angle,
- i Sun's incidence angle in relation to the normal on a pixel, and
- k Minnaert constant.

The cosine of i , is calculated by the R raster package hillshade function, using inputs: slope, aspect, sun elevation, and sun azimuth. The slope and aspect variables are derived from DEMs described in

Section 2.3, while sun elevation, azimuth, and sun zenith angle are fetched from image metadata. We selected the value 0.55 for k by estimating the mean k across a range of slopes for near-infrared wavelengths represented by B4 (Ge et al., 2008).

2.4.2. B4 cloud shadow threshold calculation

After B4 is corrected for topographic shading (a layer now call B4c), a shadow threshold value is derived from a linear model based on mean B4c brightness of all non-cloud and provisionally-assigned cloud shadow pixels in a given image. The threshold value is modeled to account for inter-image differences in brightness as a result of varying atmospheric conditions and other non-stationary effects. Initial testing of a global threshold value showed inconsistency in commission and omission error for candidate cloud shadows between images. Further testing revealed that better constancy could be achieved through modeling the value based on image brightness.

The implemented model was developed from observations of a qualified image interpreter, who for the 67 test images (Table 2) identified an appropriate B4c threshold value that separated cloud shadow from other dark land cover features. Specifically, B4c was iteratively classified into shadow and non-shadow pixels based on a series of pre-defined threshold values. At each iteration a mask was created and overlaid on a false color representation of the image. The image analyst visually assessed each mask and selected the one that appeared to be the best compromise between cloud shadow commission and omission error. The threshold values corresponding to the best mask per image were then assessed for their linear relationship with a series of image summary statistics that describe the brightness of an image. Mean brightness of all clear-view (non-cloud and provisionally-assigned cloud shadow) pixels produced the best results achieving an r-squared equal to 0.56 (Fig. 2).

The population of clear-view pixels is defined by exclusion of pixels identified as cloud in the previously described cloud layer (Section 2.2), and of provisionally-assigned cloud shadow pixels, which are identified by a spectral test of B4c against a threshold value determined from the mean of B4c excluding cloud pixels (Eq. (4)).

$$\text{PCSL test} = \text{B4c} < (0.4 * \text{MeanB4c}_1 + 0.0248), \quad (4)$$

where

$$\text{MeanB4c}_1 = \text{mean}(\text{B4c} \neq \text{cloud layer}).$$

Pixels meeting the test criteria are designated as the provisional cloud shadow layer (PCSL). The candidate cloud shadow layer (CCSL)

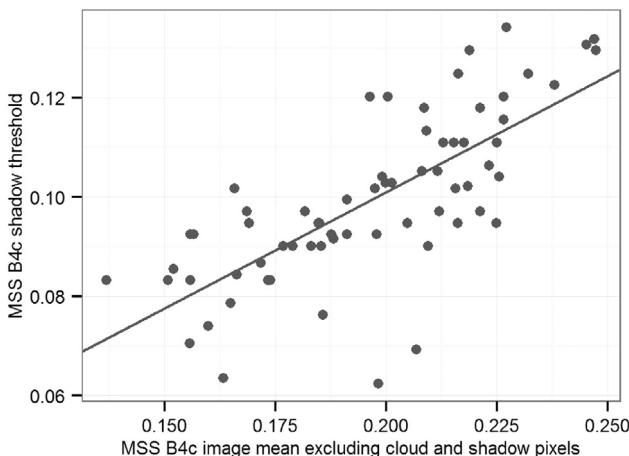


Fig. 2. Scatter plot and regression line ($r^2 = 0.56$ and $n = 67$) for the relationship between image mean brightness of MSS B4c (topographic-corrected TOA reflectance), excluding cloud and provisionally-assigned cloud shadow pixels (x-axis), and analyst-defined MSS B4c cloud shadow threshold values (y-axis).

is determined by a test of B4c mean against a threshold determined from the mean of B4c excluding both the cloud layer (CL) and PCSL followed by exclusion of water layer (WL) pixels (Eq. (5)).

$$\text{CCSL} = (\text{B4c} < (0.47 * \text{MeanB4c}_2 + 0.0073)) * \text{WL}, \quad (5)$$

where

$$\text{MeanB4c}_2 = \text{mean}(\text{B4c} \neq \text{CL and B4c} \neq \text{PCSL}).$$

Pixels meeting these test criteria are flagged as candidate cloud shadow pixels.

2.5. Candidate cloud projection layer

Within the candidate cloud shadow layer there is often commission error contributed by wetlands and dark urban features. To reduce this error, a candidate cloud projection layer is developed as a qualifier to ensure pixels identified as candidate cloud shadow are associated with clouds (Fig. 1d). Our method follows Luo, Trishchenko, and Khlopenkov (2008) and Hughes and Hayes (2014), where a continuous tract of cloud shadow pixels are cast from the cloud layer based on illumination geometry and a range of cloud heights (Fig. 3). The intersection of the candidate cloud shadow layer and this candidate cloud projection layer define cloud shadows. Similar approaches to incorporating spectral rules and cloud projection are used for TM/ETM+ (Huang, Thomas, et al., 2010; Zhu & Woodcock, 2012). However, these examples take advantage of the thermal band to estimate cloud height, which allows for a more precise estimate of cloud projection. In the absence of thermal data, an extended, continuous cloud projection field is a computationally efficient alternative for MSS imagery.

The candidate cloud projection layer is created by applying a 15-pixel buffer (900 m) to the previously described cloud layer (Section 2.2), which is then stretched out opposite the sun's azimuth for a distance equaling the projection of a range of cloud heights from 1 km to 7 km according to the sun's zenith angle. Sun azimuth and zenith angle are retrieved from image metadata. The cloud height range represents the typical elevations of low and medium height clouds. Higher clouds, such as cirrus, were excluded from this range because thin, semi-transparent cirrus clouds are often missed during cloud identification, and such an extended cloud projection layer increases the likelihood of cloud shadow commission error.

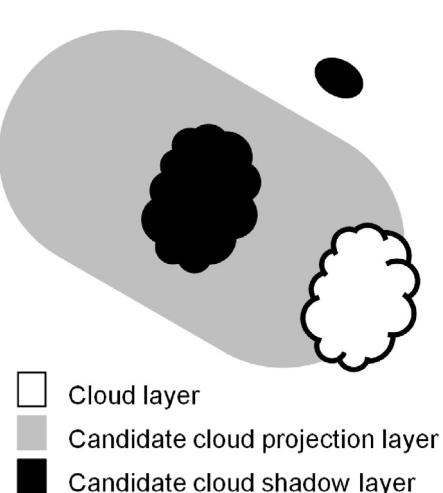


Fig. 3. Illustration of extended continuous tract cloud projection.

2.6. Final mask class assignment

The final step in the MSScvm algorithm is to aggregate the cloud layer, candidate cloud shadow layer, and candidate cloud projection layer into a single mask layer defining clear and obscured pixels (Fig. 1e). First, cloud shadow pixels are identified by the intersection of the candidate cloud shadow layer and the candidate cloud projection layer. A nine-pixel minimum connected component sieve is applied to these pixels to eliminate single and small groups of pixels generally associated with noise. A two-pixel buffer (8-neighbor rule) is added to the remaining pixels to capture shadow from thin cloud edges. This final cloud shadow layer is then merged with the cloud layer, where pixels representing cloud or cloud shadow are assigned a value of 0 and all other pixels a value of 1 to produce a clear-view mask.

2.7. Algorithm assessment

A point-based accuracy assessment of the MSScvm algorithm was performed on a sample of images to determine error rate and source. To provide context to the performance, the results were compared against an accompanying assessment of the Fmask algorithm (Zhu & Woodcock, 2012) applied to coincident TM images. In lieu of a comparable MSS masking algorithm, Fmask served as a surrogate standard from which to evaluate and discuss the accuracy of MSScvm. This is possible because Landsat 4 and 5 carried both the MSS and TM sensor. For a period of time, imagery was collected simultaneously by both sensors, which provides an excellent dataset for comparing these two algorithms, albeit operating on different data. The Fmask cloud and cloud shadow masks used for comparison were versions provided with the Landsat Surface Reflectance High Level Data Product available through USGS EarthExplorer.

Twelve images from 12 scenes across the United States representing a wide range of cover types and cloud types were used (Fig. 4). MSScvm was applied to each of the 12 MSS images and Fmask to each of the corresponding 12 TM images. One thousand points were randomly selected from within the extent of each image. The points were stratified by nine land cover classes and five cloud and cloud shadow classes (Table 3). Stratification by land cover provides information on error

Table 3

Accuracy assessment point sample frequency among land cover and mask stratification classes.

	Clear	Cloud core	Cloud edge	Shadow core	Shadow edge	Cover class total
Barren	49	49	47	39	48	232
Developed	50	49	45	45	49	238
Forest	50	50	45	38	50	233
Herbaceous upland	48	48	46	39	49	230
Open water	49	50	48	46	49	242
Perennial ice/snow	0	50	48	0	0	98
Planted/cultivated	50	50	44	45	49	238
Shrubland	50	50	45	41	50	236
Wetlands	48	49	44	43	50	234
Mask class total	394	445	412	336	394	1981

source, and stratification by cloud and shadow classes ensures equal sample representation of mask classes. Land cover classes were defined by the 1992 National Land Cover Database (NLCD) map (Vogelmann et al., 2001). Cloud and cloud shadow classes including clear, core cloud, cloud edge, core shadow, and shadow edge were based on post-processing of the MSScvm masks. Cloud and cloud shadow edge were respectively defined by an eight-pixel (480 m) region around clouds and cloud shadows, with six pixels (360 m) to the outside and two (120 m) inside. Core cloud and shadow were identified as cloud and shadow pixels not equal to edge pixels, and clear, as not equal to cloud/shadow edge or core pixels. These five mask classes were only used for sample point stratification; their parent classes (cloud, cloud shadow, and clear) were used for interpretation.

Sample pixels from each individual image were aggregated into a large database from which a subset of 50 points per combination of land cover and algorithm mask class were randomly selected. Points falling very near the edge of either the MSS or TM images were removed from the subset because of missing spectral data. Additionally, points that represented confusion between cloud and cloud shadow in either algorithm were also removed. This problem generally seemed to be the result of cloud and cloud shadow buffers and mask class priority in the assembly of the final mask classes. We did not want to penalize the algorithms for this mistake, since they ultimately identified an

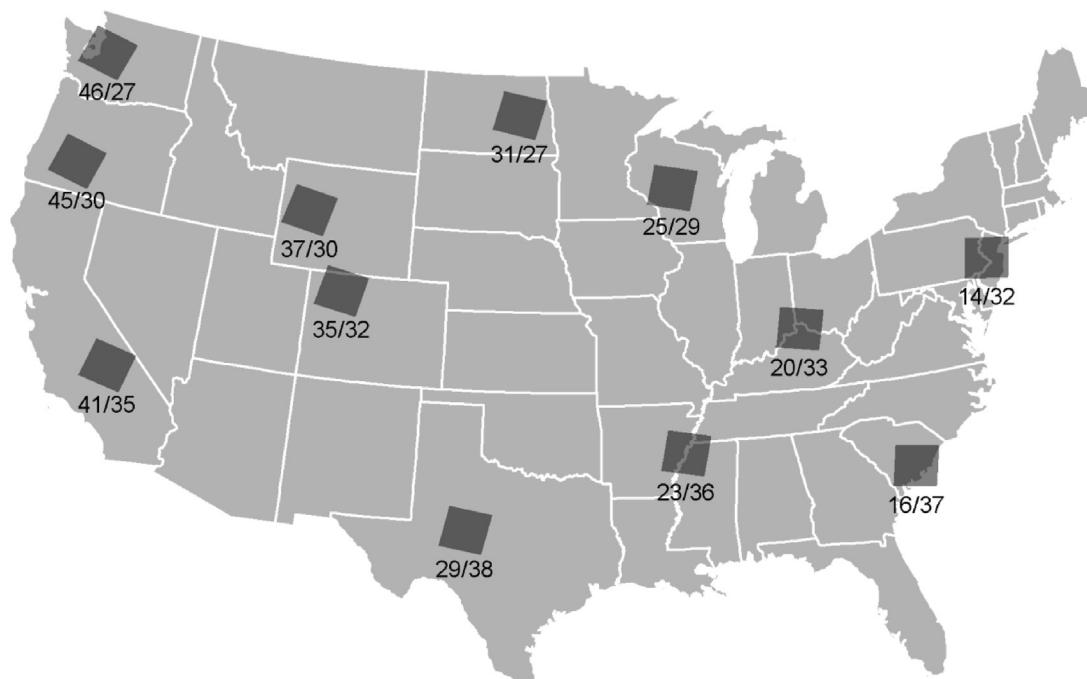


Fig. 4. Distribution of Landsat scenes within the United States used in the accuracy assessment of the MSScvm algorithm.

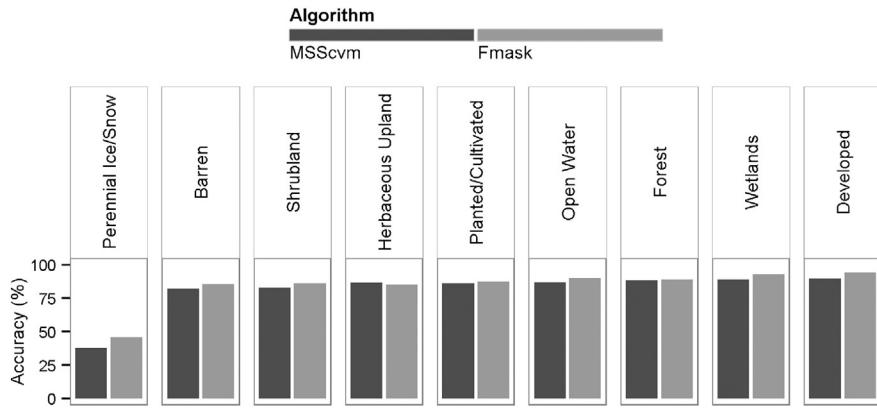


Fig. 5. Accuracy of MSScvm and Fmask cloud and cloud shadow identification algorithms for assessment points stratified by land cover class.

obscuring feature, which is the objective. The sample size for each sample stratification class is shown in Table 3. Note that the perennial snow/ice land cover class is missing representation from the clear and shadow mask classes. This is due to stratifying the point sample by the MSScvm masks, where these combinations did not exist.

For each subset point location, a qualified image analyst determined the condition of the intersecting pixel in the MSS image as cloud, cloud shadow, or clear-view through visual interpretation. These features were identified using elements of image interpretation including color, cloud/shadow association, pattern, size, and texture. These reference data were compared against the MSScvm and Fmask algorithm classification for the same points. To match the classes of MSScvm, Fmask classes clear land, clear water, and snow were aggregated as class clear, while cloud and cloud shadow remained unaltered. From these data, a series of error matrixes were created to describe accuracy and commission and omission error based on practices presented in Congalton and Green (2009).

3. Accuracy assessment results

Overall accuracy of the MSScvm algorithm was 84.0%, being 2.6% less accurate than Fmask (86.6%). Considering the differences in image information and algorithm complexity between MSS/MSScvm and TM/Fmask, the accuracy is commensurate. Results by land cover class (Fig. 5) shows that accuracy for perennial ice and snow was quite low, with MSScvm being 37.8% accurate, and Fmask 45.9%. The other eight classes, however, performed relatively well, ranging from 81.9% (barren) to 89.5% (developed) for MSScvm, and 85.2% (herbaceous/upland) to 94.1% (developed) for Fmask.

Overall error stratified by mask class (clear, cloud, and shadow) differed slightly between the algorithms, but their commission and omission error balance was very similar (Fig. 6). MSScvm clear-view pixel identification had the least error, followed by cloud shadow, and finally clouds. By contrast, Fmask had lower error for clouds than shadows. The greatest difference in overall error between the algorithms was attributed to cloud identification. MSScvm had about 10% greater cloud

omission error than Fmask, and about 5% greater commission error. Conversely, MSScvm had less commission and omission error for the cloud shadow class. Despite these differences, both algorithms showed a greater commission to omission error ratio for cloud identification, greater omission to commission error ratio for shadow identification, and nearly balanced error for clear-view pixels.

Error stratified by mask class and land cover class shows that overall error for each algorithm is driven by disproportionate confusion in just a few classes, and that for many cover type and mask class combinations, the algorithms perform almost equally (Fig. 7). As previously noted, perennial ice/snow was a major source of confusion for both algorithms, producing high cloud and cloud shadow commission error, and high clear omission error. For MSScvm, the error produced by perennial ice/snow changed the overall algorithm error ratio between commission and omission by shifting otherwise greater cloud omission error to greater commission error. Water was also responsible for an unusual amount of error, particularly for MSScvm, where it caused high cloud shadow commission error. Additionally, MSScvm had unusually high cloud shadow commission error in developed cover and cloud commission and omission error in barren cover. Disregarding assessment points representing snow/ice and water, MSScvm was generally biased toward cloud and cloud shadow omission error, whereas Fmask was just slightly biased toward greater cloud commission error and very near equal error for cloud shadow and clear.

4. Discussion and conclusion

The accuracy assessment comparing MSScvm operating on MSS imagery and Fmask operating on coincident TM imagery demonstrates that automated identification of clouds and cloud shadows in MSS imagery is feasible with accuracy at least proportionate to its data richness. Given the differences in image information and algorithm complexity between MSS/MSScvm and TM/Fmask, we consider their performance and biases to be closely aligned. The accuracy of MSScvm is a result of algorithm construction (training, logic, tuning, etc.) and the degree of disparity between image features in the input data. However, based

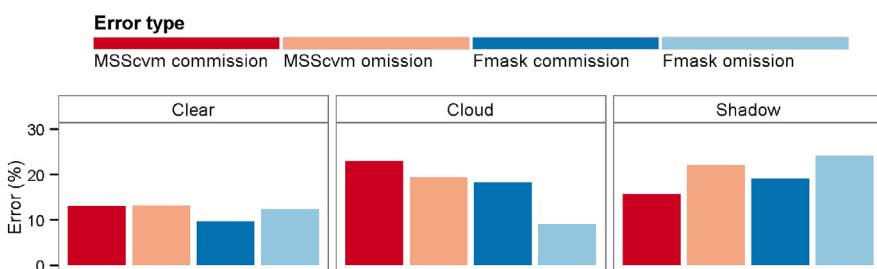


Fig. 6. Total percent commission and omission error for accuracy assessment of the MSScvm and Fmask cloud and cloud shadow identification algorithms stratified by mask class.

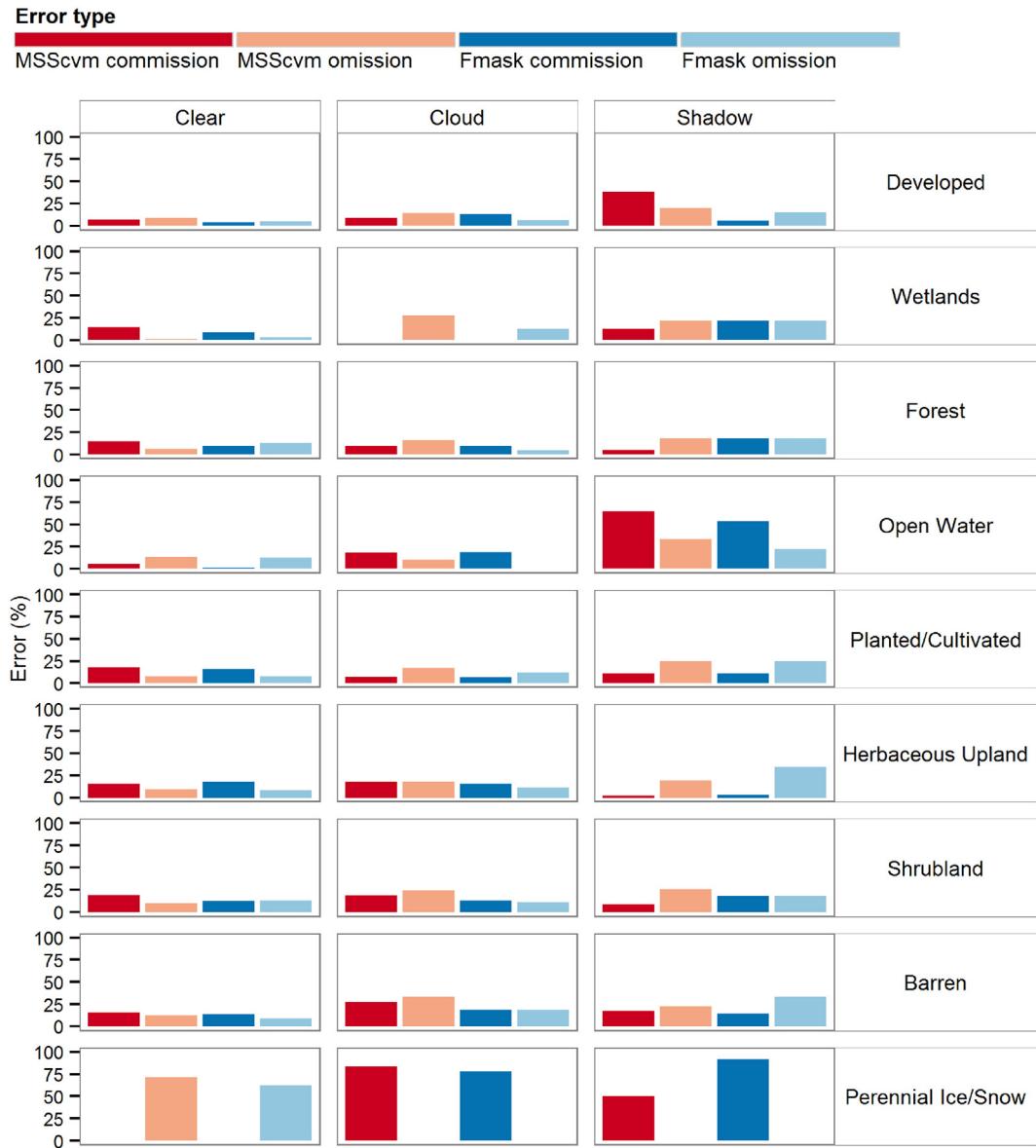


Fig. 7. Total percent commission and omission error for accuracy assessment of the MSScvm and Fmask cloud and cloud shadow identification algorithms stratified by mask class and land cover.

on many tests of various spectral and contextual analyses, we believe the limiting factor is the low radiometric and spectral depth of the imagery. The high spectral variability of land cover, clouds, and cloud shadows presents a classification challenge for the relatively low information content of MSS sensor data. The addition of a DEM to identify topographic shadows and water, and application of cloud projection to limit cloud shadow identification enhances the accuracy, but the fine spectral boundary between thin semi-transparent clouds, their shadows, and the land cover they obscure remains largely unresolved, causing the majority of error in the MSScvm algorithm.

As evident in the examples of cloud and cloud shadow masks displayed in Fig. 8, generally, thick clouds are well identified by MSScvm, regardless of land cover type, with the exception of very bright barren cover (Fig. 8b) and snow (Fig. 8c), where the algorithm falsely identifies these pixels as cloud. Fmask also misidentifies snow as cloud in Fig. 8c, but gets the bright barren cover correct in Fig. 8b, most likely with the aid of the thermal band. The thermal band is also probably responsible for the greater accuracy of Fmask with regard to identifying semi-transparent clouds, as represented in Fig. 8a and g. In these examples we see that MSScvm misses thin clouds, where Fmask captures

them. The increased cloud omission error and consequent clear-view commission error of MSScvm for these clouds are a compromise for lower overall clear-view omission error. If the cloud threshold rules were relaxed to include these thin clouds, there would be an increase in cloud commission error and clear-view omission error due to an increase in erroneously masked bright non-cloud features.

With regard to cloud shadow, both algorithms are biased toward omission error, where they incorrectly identify cloud shadows as clear-view pixels. Both MSScvm and Fmask rely on dark object identification coupled with cloud projection to assign pixels as cloud shadow. In the case of Fmask, the thermal band and lapse rate are used to estimate cloud base height, which combined with image metadata on sun elevation and azimuth provide the information for a good approximation of where cloud shadows fall on the landscape. After cloud projection, it performs a limited-area moving window routine to identify the best object match between projected clouds and dark pixels representing their shadows. Without a thermal band in MSS data, MSScvm creates an elongated candidate cloud projection region that incorporates a wide range of cloud heights and calculates its intersection with independently identified candidate cloud shadow pixels.

Both MSScvm and Fmask cloud shadow identification methods present classification problems. Fig. 8a shows that semi-transparent clouds produce relatively bright shadows. These bright shadows are often missed by MSScvm because they are brighter than the B4c cloud shadow threshold. However, relaxing the threshold to include these bright cloud shadow pixels would result in higher cloud shadow commission error and clear-view omission error. We error on the side of consistency in lower clear-view omission error, since bright cloud shadows are not always present in a given image. Fmask, on the other hand, generally identified bright cloud shadows very well, with the exception of cases where it appears cloud height was misinterpreted and non-cloud

shadow pixel were matched and incorrectly identified as cloud shadow, or actual cloud shadow pixels were missed.

In another example of cloud shadow error, Fig. 8c shows that both algorithms have misidentified snow as cloud. Fmask projects these misidentified clouds and finds corresponding relatively dark pixels on faintly illumination northwest facing slopes and incorrectly labels them as cloud shadow. Conversely, since MSScvm does not perform object matching and specifically eliminates topographic shading from the pool of candidate cloud shadows, it does not identify shadows for these falsely labeled clouds. The extended candidate cloud projection region and candidate cloud shadow intersection method used by MSScvm has

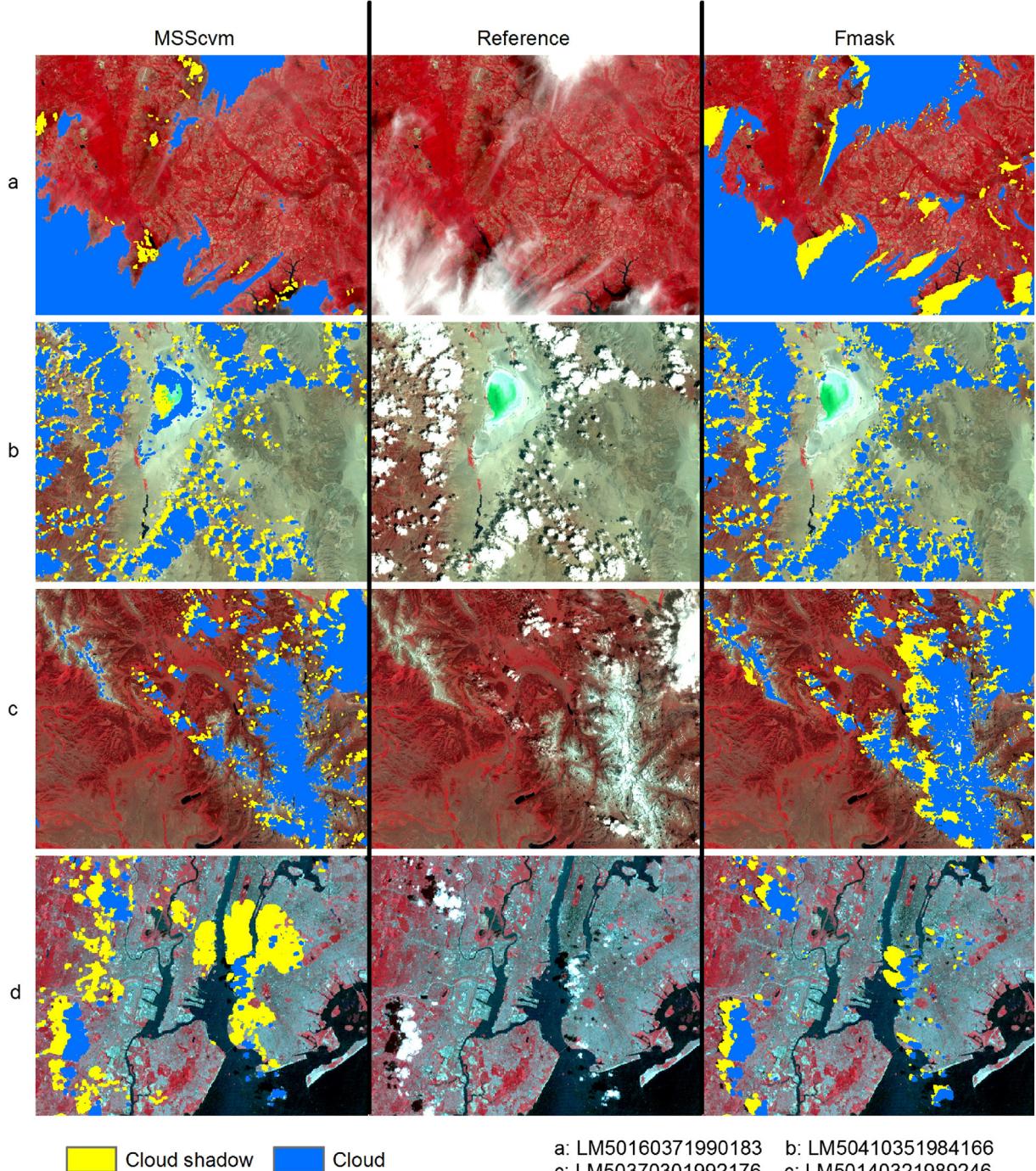


Fig. 8. Example masks produced by MSScvm algorithm operating on MSS imagery compared to Fmask applied to TM imagery for coincident dates by scene. a) Variable cloud transparency over a deciduous forest. b) Thick cumulus clouds over barren land. c) Cumulus clouds, mixed vegetation, and snow. d) Cumulus clouds, water, and developed impervious cover. e) Cumulus clouds, pothole lakes, and mixed agriculture. f) Cumulus clouds and rangeland/grassland. g) Mixed cloud types, Pacific Northwest conifer forest. h) Water, coastal wetland, and deciduous cover.

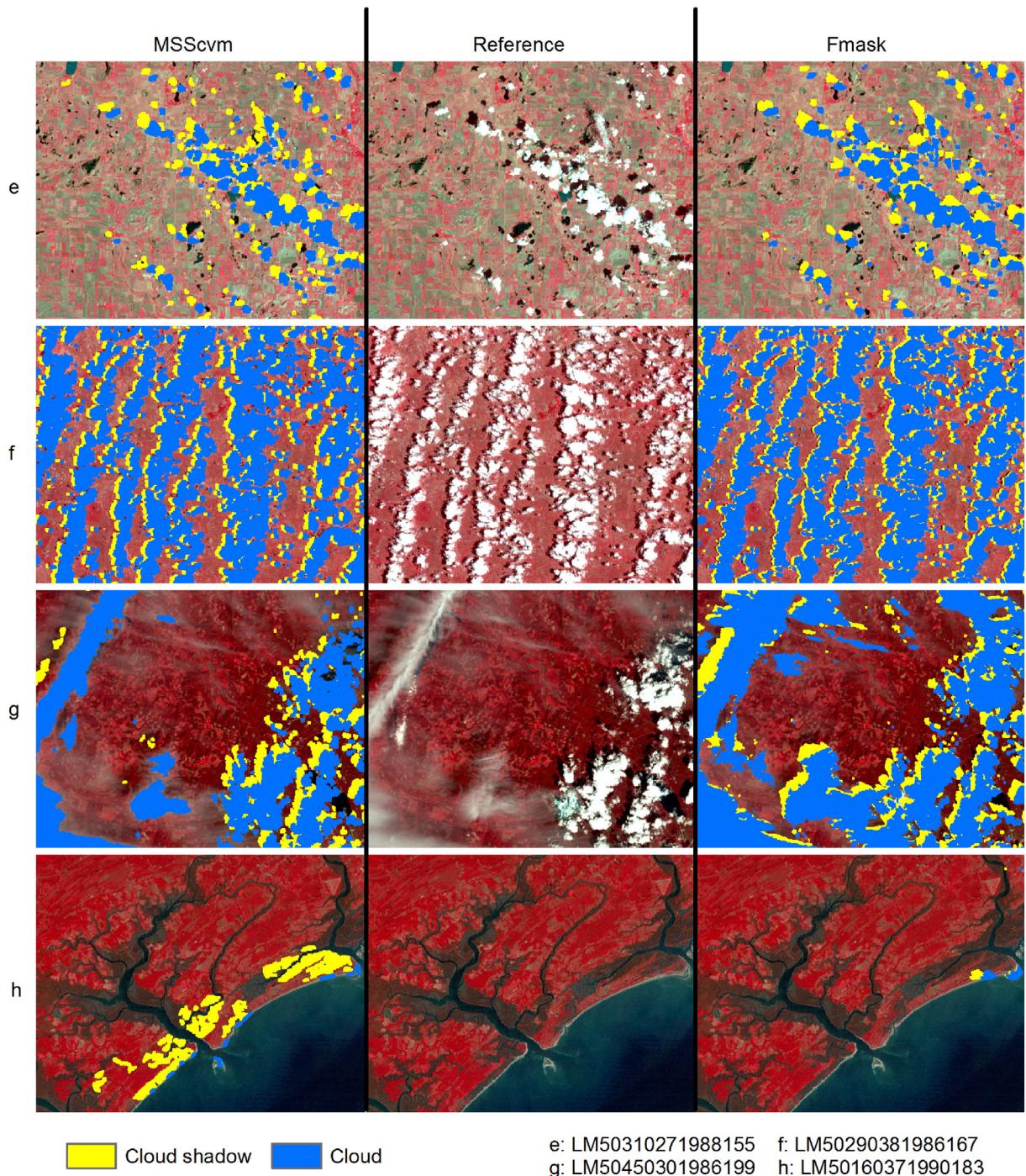


Fig. 8 (continued).

drawbacks though, as when dark pixels such as wetlands or some urban environments are near pixels identified as cloud. The problem is evident in Fig. 8d and h where candidate cloud projection regions from pixels identified as clouds intersect pixels falsely identified as candidate clouds shadow pixels, which produces false positive cloud shadows. In these cases, cloud projection based on estimated cloud height and dark object matching, implemented by Fmask, is more accurate.

Further testing of MSScvm is needed to fully understand the accuracy, but building a robust, global reference data source was beyond the scope of this project. Ideally, future efforts to test MSScvm or improve the methodology would use a spatially explicit, area-based reference data set similar to that used by Hughes & Hayes, 2014; Irish et al., 2006; Scaramuzza, Bouchard, & Dwyer, 2012; and Zhu & Woodcock,

2012, which consist of image masks developed through manual classification that represent both hemispheres, a range latitudes, and all variation of land cover and cloud types. These reference data would provide better accuracy assessment, as well as training data for machine learning algorithm construction, which could potentially improve semi-transparent cloud identification.

MSScvm is scripted as an R package and is completely automated, only requiring the input of prepared MSS images and corresponding LPGS metadata files and DEMs. The mask output can be simply multiplied by each band of a given image to set identified clouds and shadows to value zero or flagged as NA. This provides efficient use in time series analysis and mapping by eliminating cloud and shadows from imagery. Used in this context, the errors described in the results of the accuracy

assessment can propagate in two ways. First, cloud and cloud shadow omission error can result in false positive change in a change detection analysis and misclassification in predictive mapping. Second, cloud and shadow commission error can cause false negative change in a change detection analysis and eliminate affected pixels from inclusion in map prediction. However, with regard to change detection, pixels representing MSScvm omission errors are generally not significantly brighter or darker than the same pixel under clear-view conditions, and therefore, may not exceed a change detection threshold. Additionally, commission error can be relieved by merging multiple cloud-masked images from the same season to produce a near-cloud-free composite from which spatially comprehensive predictive mapping can be achieved.

The motivation for development of MSScvm was to provide a means of more easily incorporating MSS imagery in time series analysis with TM, ETM+, and OLI imagery by automating the time-consuming task of cloud and cloud shadow masking. The method presented is an initial effort to achieve this capability and offer a starting point to learn and expand from. Within the scope of North American temperate ecosystems, it performs well with the exception of thin semi-transparent clouds, their shadows, and snow/ice, as demonstrated in the accuracy assessment. MSS imagery is an important historical land surface data source, providing context for current conditions and offering rich temporal depth for studying trends and patterns in Earth surface changes. Automated cloud and cloud shadow masking overcomes a major hurdle to its effective use, however, we also identify the need for a robust MSS surface reflectance model similar to LEDAPS and L8SR, and development of spectral harmonization methods for cross-sensor time series analysis. Completion of these tasks will greatly improve the ability to efficiently include MSS in time series analysis with its successors to leverage the unprecedented 42-plus year Landsat archive for studying our dynamic Earth environment.

Acknowledgments

The development and accuracy assessment of MSScvm were made possible by the support of the Landscape Change Monitoring System (LCMS) project funded by the USDA Forest Service and by NASA's Carbon Monitoring System program (NNH13AW62I). This work was also highly dependent on the availability and easy access of free, high quality Landsat image data provided by USGS EROS. Their open data policy makes this work more relevant and widely useful. We would like to thank Joe Hughes for inspiring the implementation of a cloud projection technique, which greatly increased the algorithm's accuracy, and Dan Steinwand who provided valuable comments, citations, and perspective on the history of Landsat data processing. Additionally, we thank three anonymous reviewers for their helpful comments and suggestions.

References

- Chander, G., Markham, B.L., & Helder, D.L. (2009). Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sensing of Environment*, 113(5), 893–903.
- Chavez, P.S., Jr. (1988). An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data. *Remote Sensing of Environment*, 24, 459–479.
- Choate, M., Steinwand, D., & Rengarajan, R. (2012). *Multispectral Scanner (MSS) Geometric Algorithm Description Document: USGS Landsat Project Documentation, LS-IAS-06*.
- Congalton, R.G., & Green, K. (2009). *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*. CRC Press.
- Devaraj, C., & Shah, C.A. (2014). Automated Geometric Correction of Landsat MSS L1G Imagery. *IEEE Geoscience and Remote Sensing Letters*, 11, 347–351.
- Ge, H., Lu, D., He, S., Xu, A., Zhou, G., & Du, H. (2008). Pixel-based Minnaert correction method for reducing topographic effects on a Landsat 7 ETM+ image. *Photogrammetric Engineering & Remote Sensing*, 74, 1343–1350.
- Gómez, C., White, J.C., & Wulder, M.A. (2011). Characterizing the state and processes of change in a dynamic forest environment using hierarchical spatio-temporal segmentation. *Remote Sensing of Environment*, 115(7), 1665–1679.
- Goodwin, N.R., Collett, L.J., Denham, R.J., Flood, N., & Tindall, D. (2013). Cloud and cloud shadow screening across Queensland, Australia: an automated method for Landsat TM/ETM+ time series. *Remote Sensing of Environment*, 134, 50–65.
- Griffiths, P., Kuemmerle, T., Baumann, M., Radloff, V.C., Abrudan, I.V., Lieskovsky, J., et al. (2014). Forest disturbances, forest recovery, and changes in forest types across the Carpathian ecoregion from 1985 to 2010 based on Landsat image composites. *Remote Sensing of Environment*, 151, 72–88.
- Hansen, M.C., & Loveland, T.R. (2012). A review of large area monitoring of land cover change using Landsat data. *Remote Sensing of Environment*, 122, 66–74.
- Helder, D.L., Karki, S., Bhatt, R., Micijevic, E., Aaron, D., & Jasinski, B. (2012). Radiometric calibration of the Landsat MSS sensor series. *IEEE Transactions on Geoscience and Remote Sensing*, 50, 2380–2399.
- Hijmans, R.J. (2015). raster: Geographic data analysis and modeling. Retrieved from <http://CRAN.R-project.org/package=raster>
- Hilker, T., Wulder, M.A., Coops, N.C., Linke, J., McDermid, G., Masek, J.G., et al. (2009). A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MODIS. *Remote Sensing of Environment*, 113, 1613–1627.
- Horn, B.K. (1981). Hill shading and the reflectance map. *Proceedings of the IEEE*, 69, 14–47.
- Huang, C., Goward, S.N., Masek, J.G., Thomas, N., Zhu, Z., & Vogelmann, J.E. (2010). An automated approach for reconstructing recent forest disturbance history using dense Landsat time series stacks. *Remote Sensing of Environment*, 114, 183–198.
- Huang, C., Thomas, N., Goward, S.N., Masek, J.G., Zhu, Z., Townshend, J.R.G., et al. (2010). Automated masking of cloud and cloud shadow for forest change analysis using Landsat images. *International Journal of Remote Sensing*, 31, 5449–5464.
- Hughes, M., & Hayes, D. (2014). Automated detection of cloud and cloud shadow in single-date Landsat imagery using neural networks and spatial post-processing. *Remote Sensing*, 6, 4907–4926.
- Irish, R.R., Barker, J.L., Goward, S.N., & Arvidson, T. (2006). Characterization of the Landsat-7 ETM+ automated cloud-cover assessment (ACCA) algorithm. *Photogrammetric Engineering & Remote Sensing*, 72, 1179–1188.
- Karnieli, A., Ben-Dor, E., Bayarjargal, Y., & Lugasi, R. (2004). Radiometric saturation of Landsat-7 ETM+ data over the Negev Desert (Israel): problems and solutions. *International Journal of Applied Earth Observation and Geoinformation*, 5, 219–237.
- Kennedy, R.E., Yang, Z., & Cohen, W.B. (2010). Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr – temporal segmentation algorithms. *Remote Sensing of Environment*, 114, 2897–2910.
- Kennedy, R.E., Yang, Z., Cohen, W.B., Pfaff, E., Braaten, J., & Nelson, P. (2012). Spatial and temporal patterns of forest disturbance and regrowth within the area of the Northwest Forest Plan. *Remote Sensing of Environment*, 122, 117–133.
- Lobo, F.L., Costa, M.P.F., & Novo, E.M.L.M. (2015). Time-series analysis of Landsat-MSS/TM/OLI images over Amazonian waters impacted by gold mining activities. *Remote Sensing of Environment*, 157, 170–184.
- Loveland, T.R., & Dwyer, J.L. (2012). Landsat: Building a strong future. *Remote Sensing of Environment*, 122, 22–29.
- Luo, Y., Trishchenko, A., & Khlopenkov, K. (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(12), 4167–4185.
- Meyer, P., Itten, K.I., Kellenberger, T., Sandmeier, S., & Sandmeier, R. (1993). Radiometric corrections of topographically induced effects on Landsat TM data in an alpine environment. *ISPRS Journal of Photogrammetry and Remote Sensing*, 48, 17–28.
- Oreopoulos, L., Wilson, M.J., & Várnai, T. (2011). Implementation on Landsat data of a simple cloud-mask algorithm developed for MODIS Land bands. *IEEE Geoscience and Remote Sensing Letters*, 8, 597–601.
- Pflugmacher, D., Cohen, W.B., & Kennedy, R.E. (2012). Using Landsat-derived disturbance history (1972–2010) to predict current forest structure. *Remote Sensing of Environment*, 122, 146–165.
- R Core Team (2014). *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing (Retrieved from <http://www.R-project.org/>).
- Roy, D.P., Ju, J., Kline, K., Scaramuzza, P.L., Kovalsky, V., Hansen, M., et al. (2010). Web-enabled Landsat Data (WELD): Landsat ETM+ composited mosaics of the conterminous United States. *Remote Sensing of Environment*, 114, 35–49.
- Scaramuzza, P.L., Bouchard, M.A., & Dwyer, J.L. (2012). Development of the Landsat data continuity mission cloud-cover assessment algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, 50(4), 1140–1154.
- Telleit, P., Guindon, B., & Goodenough, D. (1982). On the slope-aspect correction of multispectral scanner data. *Canadian Journal of Remote Sensing*, 8, 84–106.
- Vogelmann, J.E., Howard, S.M., Yang, L., Larson, C.R., Wylie, B.K., & Van Driel, N. (2001). Completion of the 1990s National Land Cover Data Set for the conterminous United States from Landsat Thematic Mapper data and ancillary data sources. *Photogrammetric Engineering and Remote Sensing*, 67(6).
- Woodcock, C.E., Allen, R., Anderson, M., Belward, A., Bindschadler, R., Cohen, W., et al. (2008). Free access to Landsat imagery. *Science (New York, NY)*, 320(5879), 1011.
- Wulder, M.A., Masek, J.G., Cohen, W.B., Loveland, T.R., & Woodcock, C.E. (2012). Opening the archive: how free data has enabled the science and monitoring promise of Landsat. *Remote Sensing of Environment*, 122, 2–10.
- Zhang, Y., Guindon, B., & Cihlar, J. (2002). An image transform to characterize and compensate for spatial variations in thin cloud contamination of Landsat images. *Remote Sensing of Environment*, 82, 173–187.
- Zhu, Z., & Woodcock, C.E. (2012). Object-based cloud and cloud shadow detection in Landsat imagery. *Remote Sensing of Environment*, 118, 83–94.
- Zhu, Z., & Woodcock, C.E. (2014a). Continuous change detection and classification of land cover using all available Landsat data. *Remote Sensing of Environment*, 144, 152–171.
- Zhu, Z., & Woodcock, C.E. (2014b). Automated cloud, cloud shadow, and snow detection in multitemporal Landsat data: an algorithm designed specifically for monitoring land cover change. *Remote Sensing of Environment*, 152, 217–234.
- Zhu, Z., Woodcock, C.E., & Olofsson, P. (2012). Continuous monitoring of forest disturbance using all available Landsat imagery. *Remote Sensing of Environment*, 122, 75–91.