Remote Sensing of Environment 214 (2018) 135–153

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

An automatic method for screening clouds and cloud shadows in optical satellite image time series in cloudy regions

Xiaolin Zhu[a,*], Eileen H. Helmer[b]

[a] Department of Land Surveying and Geo-Informatics, The Hong Kong Polytechnic University, Hong Kong
[b] International Institute of Tropical Forestry, USDA Forest Service, Río Piedras, Puerto Rico, 00926, USA

ARTICLE INFO

Keywords:
Cloud detection
Cloud shadow
Mask
Optical satellite images
Time series

ABSTRACT

Clouds and cloud shadows block land surface information in optical satellite images. Accurate detection of clouds and cloud shadows can help exclude these contaminated pixels in further applications. Existing cloud screening methods are challenged by cloudy regions where most of satellite images are contaminated by clouds. To solve this problem for landscapes where the typical frequency of cloud-free observations of a pixel is too small to use existing methods to mask clouds and shadows, this study presents a new Automatic Time-Series Analysis (ATSA) method to screen clouds and cloud shadows in multi-temporal optical images. ATSA has five main steps: (1) calculate cloud and shadow indices to highlight cloud and cloud shadow information; (2) obtain initial cloud mask by unsupervised classifiers; (3) refine initial cloud mask by analyzing time series of a cloud index; (4) predict the potential shadow mask using geometric relationships; and (5) refine the potential shadow mask by analyzing time series of a shadow index. Compared with existing methods, ATSA needs fewer predefined parameters, does not require a thermal infrared band, and is more suitable for areas with persistent clouds. The performance of ATSA was tested with Landsat-8 OLI images, Landsat-4 MSS images, and Sentinel-2 images in three sites. The results were compared with a popular method, Function of Mask (Fmask), which has been adopted by USGS to produce Landsat cloud masks. These tests show that ATSA and Fmask can get comparable cloud and shadow masks in some of the tested images. However, ATSA can consistently obtain high accuracy in all images, while Fmask has large omission or commission errors in some images. The quantitative accuracy was assessed using manual cloud masks of 15 images. The average cloud producer's accuracy of these 15 images is as high as 0.959 and the average shadow producer's accuracy reaches 0.901. Given that it can be applied to old satellite sensors and it is capable for cloudy regions, ATSA is a valuable supplement to the existing cloud screening methods.

1. Introduction

Optical satellite images with bands ranging from visible to shortwave infrared are widely used for mapping land cover and land use, monitoring ecosystems, and estimating land surface parameters (Hansen and Loveland, 2012; Zhu and Liu, 2015, 2014). Unfortunately, optical satellite images are easily contaminated by clouds and cloud shadows. This contamination obscures land surface features and alters the reflectance of ground objects, reducing the availability of optical images for applications (Fisher, 2013; Zhu and Woodcock, 2014). Masking clouds and cloud shadows is often the first and a necessary step of image preprocessing in most optical remote sensing applications. Although manual digitization can obtain accurate cloud and shadow masks, it requires a lot of time and effort. Therefore, an automatic method for screening clouds and shadows is needed, especially when processing large numbers of images.

Automatic detection of clouds and cloud shadows is challenging (Zhu and Woodcock, 2014). First, different types of clouds have different spectral signatures and are easily confused with some cloud-free bright objects on the land surface, especially in images with limited spectral bands, such as Landsat Multispectral Scanner (MSS) images. The spectral signals of clouds are usually determined by cloud height, optical thickness, particle size, etc. (Platnick et al., 2003). As a result, cloud brightness ranges widely in visible and near infrared bands, and some clouds are easily confused with bright land surfaces, such as concrete surfaces, sand or snow. Second, blurry cloud edges and thin clouds partially obscure land surfaces, making their signal a mixture of cloud and land surface elements and making them difficult to separate from clear observations (Cahalan et al., 2001). Another challenge comes from cloud shadows. They are easily confused with dark land

* Corresponding author at: The Hong Kong Polytechnic University, Room ZS621, South Wing, Block Z, 181 Chatham Road South, Kowloon, Hong Kong.
E-mail address: xiaolin.zhu@polyu.edu.hk (X. Zhu).

https://doi.org/10.1016/j.rse.2018.05.024
Received 26 September 2017; Received in revised form 1 May 2018; Accepted 18 May 2018

surfaces, such as moist soil, water bodies and topographic shadow (Fisher, 2013).

Despite the above challenges, several methods have been developed to automatically screen clouds and cloud shadows in optical images. These methods use one or more of the following rules based on cloud and cloud shadow properties: 1) clouds are generally brighter than ground objects, so they have high reflectance in visible, near and shortwave infrared bands; 2) clouds are generally colder than most ground objects, so they have lower brightness in thermal infrared bands; 3) shadows are generally darker than surrounding land surfaces, so they have lower reflectance in visible, near and shortwave infrared bands; 4) shadows are paired with clouds, so cloud location and solar angles can help locate cloud shadows; and 5) in a sequence of images, pixels affected by clouds and shadows have larger temporal variations than clear observations in the time series. In general, existing methods for masking clouds and cloud shadows can be divided into two categories: single-image methods (Choi and Bindschadler, 2004; Fisher, 2013; Helmer et al., 2012; Huang et al., 2010; Hughes and Hayes, 2014; Irish et al., 2006; Li et al., 2015, 2017; Luo et al., 2008; Martinuzzi et al., 2007; Roy et al., 2010; Scaramuzza et al., 2012; Wilson and Oreopoulos, 2013; Zhu and Woodcock, 2012) and multi-temporal or bi-temporal methods (Goodwin et al., 2013; Hagolle et al., 2010; Jin et al., 2013; Wang et al., 1999; Zhu and Woodcock, 2014).

Most existing single-image methods use either predefined thresholds or adaptive thresholds to screen clouds in individual images. For example, Luo et al. (2008) identify clouds in MODIS images if pixel reflectance satisfies these predefined thresholds: ($B1 > 0.18$ or $B3 > 0.20$) and $B6 > 0.16$ and Maximum $(B1, B3) > B6 \times 0.67$, where B1, B3, and B6 are reflectance of MODIS bands 1 (blue), 3 (red), and 6 (shortwave infrared), respectively. This MODIS cloud screening method was further adopted for Landsat-8 images (Wilson and Oreopoulos, 2013). Huang et al. (2010) use adaptive thresholds defined in the reflectance-temperature space to mask clouds in Landsat TM and ETM+ images. These adaptive thresholds are defined by the mean and standard deviation of pixel values of individual bands in the whole image. The Automatic cloud cover assessment (ACCA) algorithm consists of twenty-six filters and rules applied to Landsat bands to detect clouds (Irish et al., 2006). ACCA was used to produce web-enable Landsat data (WELD), a consistent, long-term, and large-area data record (Roy et al., 2010). The multi-feature combined (MFC) method uses thresholds in spectral, geometric and texture features to detect clouds in GaoFen-1 imagery (Li et al., 2017). Zhu and Woodcock (2012) proposed a method called function of mask (Fmask) for detecting clouds in Landsat TM and ETM+ images. Fmask uses all Landsat image bands and several band indices, such as the normalized difference vegetation index (NDVI) and the normalized difference snow index (NDSI). It employs > 20 predefined and adaptive thresholds to mask clouds. Besides the above methods using predefined or adaptive thresholds, machine-learning algorithms have been employed to model the complex relationships between image features and clouds using a training dataset. Then, the trained model is used to screen clouds in other images. These machine learning algorithms include decision trees (Scaramuzza et al., 2012), neural networks (Hughes and Hayes, 2014) and support vector machines (Li et al., 2015). Of several tested cloud and shadow masking algorithms that use only a single image, Fmask is globally the most accurate one that requires a thermal band (Foga et al., 2017). Of methods not requiring a thermal band, a version of ACCA (Irish et al., 2006) that uses a simulated thermal band is better overall, but it is not as accurate as Fmask with the thermal band (Foga et al., 2017). Recently, Fmask was further improved for mountainous areas through integrating Digital Elevation Models (DEMs) into the detecting process (Qiu et al., 2017).

In these single-image methods, shadow detection is often subsequent to cloud detection. In general, the possible shadow locations can be calculated from the geometric relationship between sun, sensor, and clouds. The calculation requires cloud heights, which can be estimated with brightness temperature derived from thermal infrared bands, because temperature declines with elevation (Qiu et al., 2017; Zhu and Woodcock, 2012). Some methods also use the fact that cloud shadows are dark to confirm whether the possible shadow location estimated from geometry is real cloud shadow, including Fmask (Zhu and Woodcock, 2012) and MFC (Li et al., 2017). In Fmask, predefined thresholds in the near infrared (NIR) band are used to produce a potential shadow mask, which is further compared to the possible shadow locations. If there is a high similarity between potential shadow masks and possible shadow locations, the shadow pixels are finally confirmed (Zhu and Woodcock, 2012).

For multi-temporal methods, temporal information in the images acquired at different times is used to detect clouds and shadows. Wang et al. (1999) used the brightness difference between a target image and a reference cloud-free image to detect clouds. Lyapustin et al. (2008) developed an algorithm, abbreviated as MAIAC CM, to detect clouds in time series of MODIS images. The general idea of MAIAC CM is to use the low covariance between reference cloud-free image blocks and cloudy image blocks as a criterion to identify clouds in the time series. Hagolle et al. (2010) computes differences in the blue band between a target image and a cloud-free reference image. It then flags cloud pixels if variations are larger than a threshold. Goodwin et al. (2013) uses filters to smooth the time series and then identify clouds and shadows based on reflectance differences between each point in the time series and the smoothed time series. Zhu and Woodcock (2014) propose a new algorithm called multiTemporal mask (Tmask) to improve Fmask. Tmask fits a time series model of each pixel using remaining clear pixels based on an initial cloud mask from Fmask. Then, it compares model estimates with observations in the time series to detect cloud and shadow pixels which are omitted in the initial screening by Fmask. In general, these multi-temporal methods are better at detecting clouds and cloud shadows than single-image methods. The temporal information is a valuable complement to the spectral information for differentiating cloud, cloud shadow and clear observations over land surfaces (Goodwin et al., 2013; Zhu and Woodcock, 2014).

However, these multi-temporal methods still face challenges in areas with persistent cloud cover, such as tropical and subtropical regions (Ju and Roy, 2008). First, in these areas cloud-free observations may be the exception rather than the rule, making it difficult to know whether the fit of a time series represents clear or cloudy conditions, which limits the application of existing time-series methods (Foga et al., 2017). Example limitations include the requirement by the MAIAC CM method of a cloud free image as a reference image (Lyapustin et al., 2008), and the recommendation for Tmask of 15 cloud-free observations for estimating the time series model (Zhu and Woodcock, 2014). Second, most existing methods were designed for images of a specific sensor, so they lack flexibility. For example, Fmask and Tmask were designed for Landsat TM, ETM+, and OLI images, so they cannot be directly applied to the old Landsat MSS data with limited bands. Third, most of the current methods use predefined fixed thresholds to detect clouds and shadows in an entire scene. For instance, in Tmask, a pixel with observed green band reflectance of 0.04 higher than the time series model estimation will be identified as cloud. Considering the complex situation of clouds and shadows and the diversity of objects on land surfaces and in coastal areas, these fixed thresholds may not always obtain satisfactory results.

To overcome the above limitations of existing methods in cloudy regions, the objective of this study is to develop a new automatic method for accurately screening clouds and cloud shadows in multi-temporal optical images in places with persistent clouds. Our scope of inference is landscapes where are so cloudy that the typical frequency of cloud-free observations of a pixel is too small to use existing methods to mask clouds and shadows with image time series. The new method should have the following strengths: 1) it needs fewer predefined parameters; 2) it is suitable for areas with persistent clouds; and 3) it needs a minimal number of bands. Automatic Time-Series Analyses