Multi-feature combined cloud and cloud shadow detection in GF-1 WFV imagery

Zhiwei Li [a], Huanfeng Shen [a,b,c*], Huifang Li [a], Guisong Xia [d], Paolo Gamba [e], Liangpei Zhang [b,d]

[a] School of Resource and Environmental Sciences, Wuhan University, Wuhan, Hubei, 430079, China.

[b] The Collaborative Innovation Center for Geospatial Technology, Wuhan, Hubei, 430079, China.

[c] The Key Laboratory of Geographic Information System, Ministry of Education, Wuhan University, Wuhan, Hubei, 430079, China.

[d] The State Key Laboratory of Information Engineering in Surveying, Mapping and Remote Sensing, Wuhan University, Wuhan, Hubei, 430079, China.

[e] Department of Electrical, Computer and Biomedical Engineering, University of Pavia, Italy.

[*] Corresponding author. E-mail address: shenhf@whu.edu.cn

ABSTRACT

The wide field of view (WFV) imaging system onboard the Chinese GF-1 optical satellite has a 16-m resolution and four-day revisit cycle for large-scale Earth observation. The advantages of the high temporal-spatial resolution and the wide field of view make the GF-1 WFV imagery very popular. However, cloud cover is a common problem in GF-1 WFV imagery which influences its precise application. Cloud and cloud shadow detection in GF-1 WFV imagery is quite difficult due to the fact that there are only four visible and near-infrared bands. In this paper, an automatic multi-feature combined (MFC) method is proposed for cloud and cloud shadow detection in GF-1 WFV imagery. The MFC method first implements threshold segmentation based on the spectral features, and guided filtering to generate a preliminary cloud mask. The geometric features are then used in combination with texture

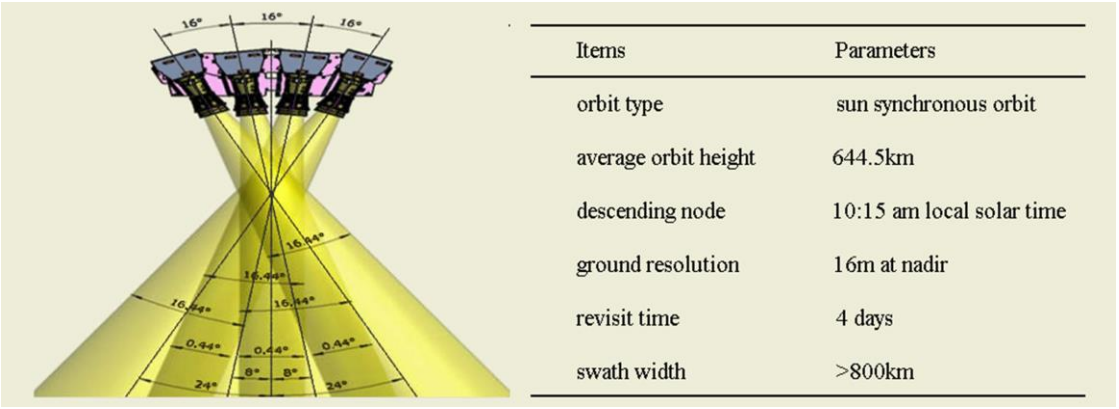
features to improve the cloud detection results and produce the final cloud mask. Finally, the cloud shadow mask can be acquired by means of the cloud and shadow matching and follow-up correction process. The method was validated on 16 scenes randomly selected from different areas of China. The results indicate that MFC performs well under different conditions, and the average cloud classification accuracy of MFC is as high as 98.3%. When relatively compared to Fmask, MFC achieved almost the same accuracy of the cloud and cloud shadow in GF-1 WFV imagery with less spectral bands.

Keywords: Cloud detection; Cloud shadow; GF-1; Multiple features; MFC;

1. Introduction

Clouds and the accompanying shadows are inevitable contaminants for optical imagery in the range of the visible and infrared spectra. The global annual mean cloud cover is approximately 66% according to the estimation of the International Satellite Cloud Climatology Project-Flux Data (ISCCP-FD) data set (Zhang et al., 2004). Cloud cover impedes optical satellites from obtaining clear views of the Earth's surface, and thus the existence of clouds influences the availability of optical satellite imagery. Cloud shadows cast by clouds are also a contaminant for imagery, and the dark effect of cloud shadows results in the spectral information of the imagery covered by cloud shadows being partly or entirely lost. The cloud and cloud shadows in the imagery affect the processing of the imagery, in applications such as classification, segmentation, feature extraction, etc. A number of cloud removal and image restoration methods (Zeng et al., 2013; Cheng et al., 2014; Li et al., 2014; Shen et al., 2014) can effectively repair cloud-contaminated imagery, but they do not provide a specific way to

extract the clouds. Accurately extracting clouds and cloud shadows from cloud-contaminated imagery can help to reduce the negative influences that cloud coverage brings to the application of the imagery. Furthermore, cloud cover estimation can be used for imagery availability evaluation. Therefore, cloud and cloud shadow detection in optical imagery is of great significance.



Items	Parameters
orbit type	sun synchronous orbit
average orbit height	644.5km
descending node	10:15 am local solar time
ground resolution	16m at nadir
revisit time	4 days
swath width	>800km

Fig. 1. The GF-1 WFV imaging system (sensors image credit: DFH Satellite Co. Ltd., China).

The GaoFen-1 (“gao fen” means high resolution in Chinese) satellite was launched by the China Aerospace Science and Technology Corporation (CASC) in April 2013. It was the first of a series of satellites in the civilian High-Definition Earth Observation Satellite (HDEOS) program to realize a high resolution and wide-swath optical remote sensing mission. The wide field of view (WFV) imaging system is one of the key instruments operating onboard the GF-1 satellite, as shown in Fig. 1, it includes four integrated cameras with a 16-m spatial resolution and four-day temporal resolution. Each WFV camera has four multispectral bands, and the spectral bands span the visible to the near-infrared spectral regions, which is almost the same as the Landsat ETM+ sensors in the first four bands (Table 1). The swath width of the GF-1 WFV imaging system increases to 800 km when the four cameras are combined, which significantly improves the capabilities for large-scale surface observation and monitoring. The

imagery of the GF-1 satellite has served a wide range of applications covering many topics. The typical applications include disaster prevention and relief, geographical mapping, environment and resource surveying, as well as precision agriculture support (Chen et al., 2015; Li et al., 2015; Li et al., 2015; Lu & Bai, 2015; Wang et al., 2015).

Table 1. Spectral range comparison of GF-1 WFV and Landsat ETM+ imagery.

Bandwidth (μm)	GF-1 WFV	Landsat ETM+
Band 1 (Blue)	0.45–0.52	0.45–0.52
Band 2 (Green)	0.52–0.59	0.52–0.60
Band 3 (Red)	0.63–0.69	0.63–0.69
Band 4 (NIR)	0.77–0.89	0.76–0.90
Band 5 (SWIR-1)	–	1.55–1.75
Band 6 (TIR)	–	10.4–12.5
Band 7 (SWIR-2)	–	2.08–2.35
Band 8 (Pan)	–	0.50–0.90

Cloud detection in GF-1 WFV imagery is a challenging task because of the unfixed radiometric calibration parameters and the insufficient spectral information. The GF-1 WFV imaging system lacks onboard calibration capabilities (Yang et al., 2015), which makes accurate calibration of GF-1 imagery difficult. In addition, this kind of imagery has no thermal infrared band or water vapor/CO_2 absorption band which are very important for cloud identification (Huang et al., 2010). Due to the lack of enough spectral information, it is not easy to separate clouds from some bright ground objects (such as snow, buildings, and coast lines) when only using the spectral features. Meanwhile, thin cloud is also difficult to capture in optical satellite imagery because of the different underlying surface. Moreover, in cloud shadow extraction, it is usually difficult to capture the complete cloud shadow location because of shadow screening and cloud shadow matching errors. In order to acquire better cloud and cloud shadow detection results based on limited spectral bands, more features such as geometric and

texture features should be taken into consideration.

2. Background

In recent years, scholars have undertaken a great deal of research into cloud and cloud shadow detection for different types of remote sensing data, such as AVHRR (Di Vittorio & Emery, 2002; Khlopenkov & Trishchenko, 2007), MODIS (Platnick et al., 2003; Luo et al., 2008), and Landsat series imagery (Irish et al., 2006; Zhu & Woodcock, 2012; Goodwin et al., 2013; Harb et al., 2016) The methods of cloud detection can be divided into two categories according to the single or multi-temporal scenes the algorithm uses.

Cloud detection methods based on a single scene are more popular than multi-temporal methods, due to the reduced requirement for input data. The automatic cloud cover assessment (ACCA) algorithm (Irish et al., 2006) was designed for the cloud cover assessment of Landsat-7 imagery. It is an official method and is included in the Landsat-7 Science Data Users Handbook (Irish, 2000). In order to further capture the thin clouds which cannot be detected by the ACCA algorithm in Landsat imagery, function of mask (Fmask) (Zhu & Woodcock, 2012; Zhu et al., 2015), which is a robust cloud detection method, was proposed for routine usage with Landsat images. Haze optimized transformation (HOT) (Zhang et al., 2002; Zhang et al., 2014) was also developed for the detection and characterization of haze/cloud in Landsat scenes, but it requires prior knowledge of the image to build a clear line in spectral space to separate haze/cloud from the clear surface. Le Hégarat-Mascle and André(2009) and Vivone et al. (2014) developed cloud detection algorithms based on Markov random fields. In addition, methods based on machine learning have also been applied in automatic cloud detection, including the spatial procedures for automated removal of cloud and shadow (SPARCS) algorithm (Hughes

& Hayes, 2014), which uses a neural network to identify cloud and cloud shadow in Landsat scenes, and a cloud image detection method based on support vector machine (Li et al., 2015).

Compared to the single-temporal cloud detection methods, multi-temporal cloud detection methods usually achieve a higher cloud detection accuracy. However, these methods require more scenes over a short time period to make ensure that the land cover in the same place does not change much. Therefore, multi-temporal cloud detection methods may be more suitable for relatively permanent land areas in high temporal resolution imagery. Examples of multi-temporal cloud detection methods include the multi-temporal cloud detection (MTCD) method (Hagolle et al., 2010), the multi-temporal cloud and snow detection algorithm (Bian et al., 2014) for the HJ-1A/1B CCD imagery of China, the multi-temporal mask (Tmask) for the automatic masking of cloud, cloud shadow, and snow for multi-temporal Landsat images (Zhu & Woodcock, 2014), and the optical satellite imagery cloud detection method using invariant pixels (Lin et al., 2015).

Cloud shadow detection is usually undertaken after the cloud detection (Luo et al., 2008; Hughes & Hayes, 2014; Braaten et al., 2015). Shadows in remote sensing imagery can be approximately divided into two categories, namely, terrain shadow and cloud shadow. Terrain shadow can be corrected or removed by topographic correction (Meyer et al., 1993), on the condition that the digital elevation model (DEM) and solar angle of incidence are provided, while the distribution of cloud shadow in imagery depends on the cloud location and the satellite viewing and solar angles. Cloud shadow location can be predicted by means of geometrical calculation if the location and height of clouds and the sun and satellite positions are known. Furthermore, DEM data can be used to refine cloud and cloud shadow detection results. Huang

et al. (2010) improved the projection of clouds onto the land surface with DEM data. Braaten et al. (2015) also incorporated DEM data and cloud projection to better separate cloud shadow from topographic shading and water.

In this paper, an automatic multi-feature combined (MFC) method is proposed for cloud and cloud shadow detection in GF-1 WFV imagery. The MFC algorithm implements a local optimization strategy with guided filtering to refine the cloud and cloud shadow detection results. In addition, the geometric and texture features are used to decrease the commission error in cloud and cloud shadow detection. The experimental results suggest that MFC performs well in most of ground cover conditions, and it can also accurately detect thin clouds and cloud shadow just using the four optical bands.

3. The MFC algorithm

The input data for the MFC algorithm are the top of atmosphere (TOA) reflectance of all four bands in the GF-1 WFV imagery, because the TOA reflectance includes the surface reflectance of the Earth and atmospheric information, and a reduction in between-scene variability can be achieved by converting the digital number (DN) values to TOA reflectance values. The MFC algorithm first implements threshold segmentation by using spectral features, and a local optimization strategy with guided filtering to generate a refined cloud mask. The geometric features are then used in combination with texture features to improve the cloud detection results and produce the final cloud mask. Finally, the cloud shadow mask can be acquired by means of cloud and cloud shadow matching and correction. When a pixel is labeled as cloud as well as cloud shadow, a higher priority is set for cloud than cloud shadow in the integrated mask to generate the final cloud and cloud shadow mask. Fig. 2 shows the process flow of the MFC

algorithm.

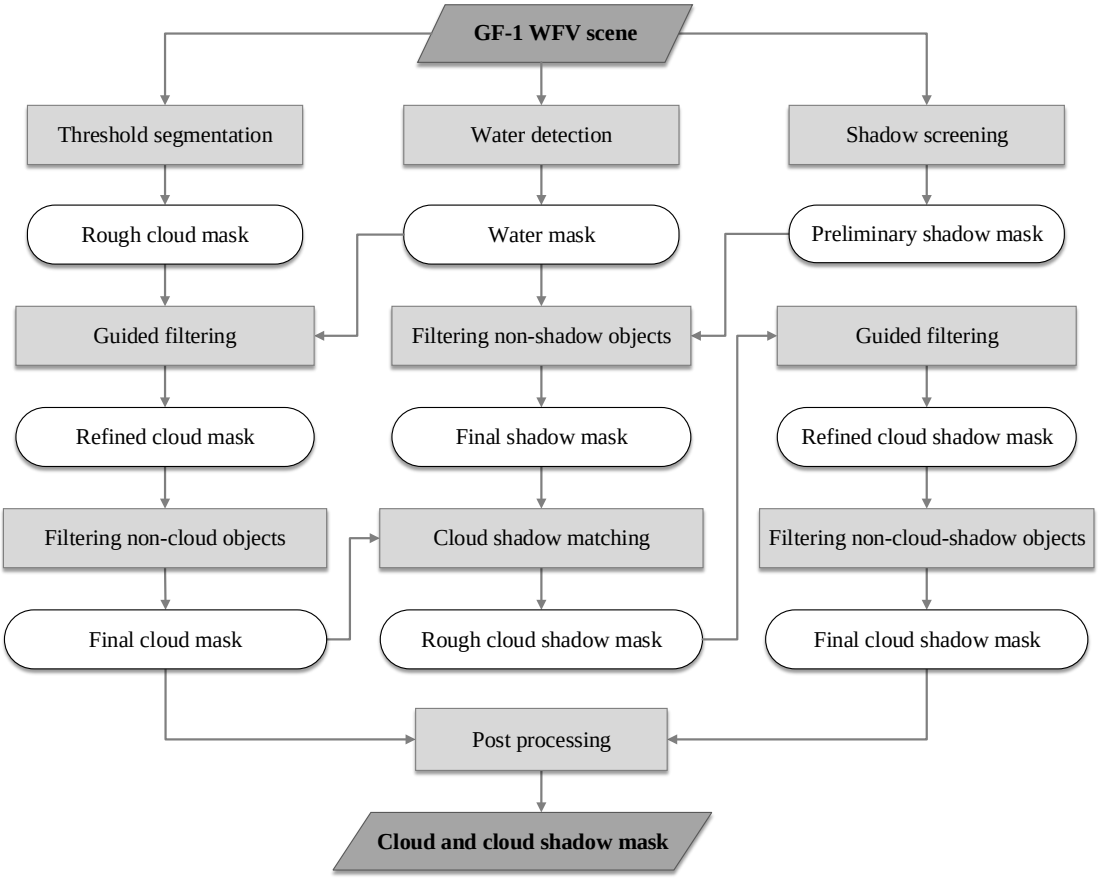


Fig. 2. Overall framework of the MFC algorithm.

3.1 Cloud detection

There are three steps to implementing cloud detection with the proposed method. MFC first produces a rough cloud detection result by applying threshold segmentation based on spectral features, and the core cloud regions are captured after this step. A finer results is then generated after guided filtering and binary image segmentation, and the thin clouds around cloud boundaries are included in the refined cloud mask. Finally, non-cloud bright objects are removed from the refined cloud mask by using the geometric and texture features. The reason why MFC does not directly acquire a refined cloud detection result is that, it cannot entirely exclude clear-sky pixels while ensuring that the thin clouds around cloud boundaries are not

missed at the same time.

3.1.1 Initializing a rough cloud mask using spectral features

The MFC algorithm first produces a rough cloud mask which includes most of the thick clouds. This is aimed at extracting the core cloud regions and attempting to make sure that the commission rate in the rough cloud mask is very low. The HOT index (Zhang et al., 2002) has been widely used for haze reduction and cloud detection (Vermote & Saleous, 2007; Zhang et al., 2014; Harb et al., 2016). It is used to separate cloud from clear-sky pixels considering the fact that the HOT values of cloud pixels are usually greater than clear-sky pixels. The HOT index used in MFC can be expressed as follows:

$$HOT = B1 - 0.5 \cdot B3 \tag{1}$$

where $B1$ and $B3$ denote the blue and red band reflectance.

HOT is an effective cloud and haze extraction method, and similar form has also been used in the LEDAPS internal cloud masking algorithm (Vermote & Saleous, 2007) and Fmask (Zhu & Woodcock, 2012). However, HOT cannot adequately suppress land surface information, and it often overestimates haze thickness over bright surfaces (Chen et al., 2015). As a result of the bands that the HOT index relies on, some ground objects with high reflectance in the visible bands or just the blue band, such as snow and blue buildings, cannot be excluded in the extracted result because of the high value of the HOT index. This results in some commission error in the cloud detection results.

Furthermore, the ratio of the minimal and maximal reflectance in the visible bands can be used to exclude ground objects with other blue, red, or green colored features. The visible band ratio (VBR) of a pixel

$$VBR = \frac{min(B1,B2,B3)}{max(B1,B2,B3)} \quad (2)$$

is close to one when the pixel is grey. Therefore, VBR can be used to exclude non-cloud pixels with salient color features from the extracted results. It is based on the idea that clouds in the optical imagery generally appear white or grey in RGB color space.

Meanwhile, cloud reflectance in the red band should be greater than 0.07, and a similar threshold is set in ACCA algorithm for Landsat imagery (Irish et al., 2006). In this case, a threshold is set for the red band reflectance to make sure that a pixel is more likely to be white than black. The formula used to set a first and rough cloud mask (CM_R) can be expressed as:

$$CM_R = (HOT > t_1)\ and\ (VBR > t_2)\ and\ (B3 > t_3) \quad (3)$$

where t_1, t_2, t_3 are the thresholds to initialize the rough cloud mask.

Thresholds used in MFC are all written as t in the following formulas, these parameters have been selected carefully by means of experiments, and it will be given and discussed later. By applying a binary segmentation to these spectral features, a pixel will be labeled as "cloud" when the above conditions are met, a rough cloud mask then is acquired in which the core cloud regions are included. However, there may still be some bright ground objects in the rough cloud mask, such as snow or buildings, which cannot be effectively excluded from the rough cloud mask merely by the visible and near-infrared spectral information alone.

3.1.2 Refining the cloud boundaries using a guided filtering

Although MFC can acquire an approximate cloud detection result through the above spectral tests, thin clouds around the cloud edges may be missed since the above cloud detection procedure mainly captures core cloud regions. In order to further capture the missed clouds, the statistical features, which combine the spectral information in the original image and the cloud

location information in rough cloud mask, are taken into consideration. In this paper, the guided filter proposed by He et al. (2013) is used to capture missing clouds around cloud boundaries by considering the combined statistical features, and improve the cloud detection results in GF-1 WFV imagery. It is based on the fact that thin clouds are usually distributed around the core cloud regions, and there is a transition from the core cloud regions to thin clouds in space around cloud boundaries.

The guided filter is a novel filter with both edge preserving and noise reducing properties. It has been applied to refine the cloud boundary detection for RGB color aerial photographs (Zhang & Xiao, 2014). The guided filter involves a guidance image I, an input image p, and an output image q. The key assumption of the guided filter is a local linear model between the guidance image I and the output image q, and q is supposed to be a linear transform of I in a square window w_k at pixel k:

$$q_i = a_k I_i + b_k \ (\forall i \in w_k) \tag{4}$$

where a_k and b_k are constant linear coefficients in w_k, and i denotes a pixel coordinate in the square window w_k.

The local linear model ensures that q has an edge only if I has an edge, since $\nabla q = a\nabla I$. To seek a solution that minimizes the difference between q and p while maintaining the linear model, the two coefficients a_k and b_k can be defined by Eqs. (5)–(6):

$$a_k = \frac{\frac{1}{|w|}\sum_{i \in w_k} I_i p_i - \mu_k \overline{p_k}}{\delta_k^2 + \varepsilon} \tag{5}$$

$$b_k = \overline{p_k} - a_k \mu_k \tag{6}$$

where μ_k and δ_k^2 are the mean and variance of I in w_k, ε is the regularization parameter, $|w|$ is the number of pixels in w_k, and $\overline{p_k}$ is the mean of p in w_k.

As a pixel i is involved in all the overlapping windows w_k that cover i, so the output value q_i should combine all of the overlapping windows, and the final output value of pixel i is defined as:

$$q_i = \bar{a} I_i + \bar{b} \tag{7}$$

where $\bar{a} = \frac{1}{|w|}\sum_{k \in w_i} a_k$ and $\bar{b} = \frac{1}{|w|}\sum_{k \in w_i} b_k$ are the average coefficients of all the windows overlapping pixel i.

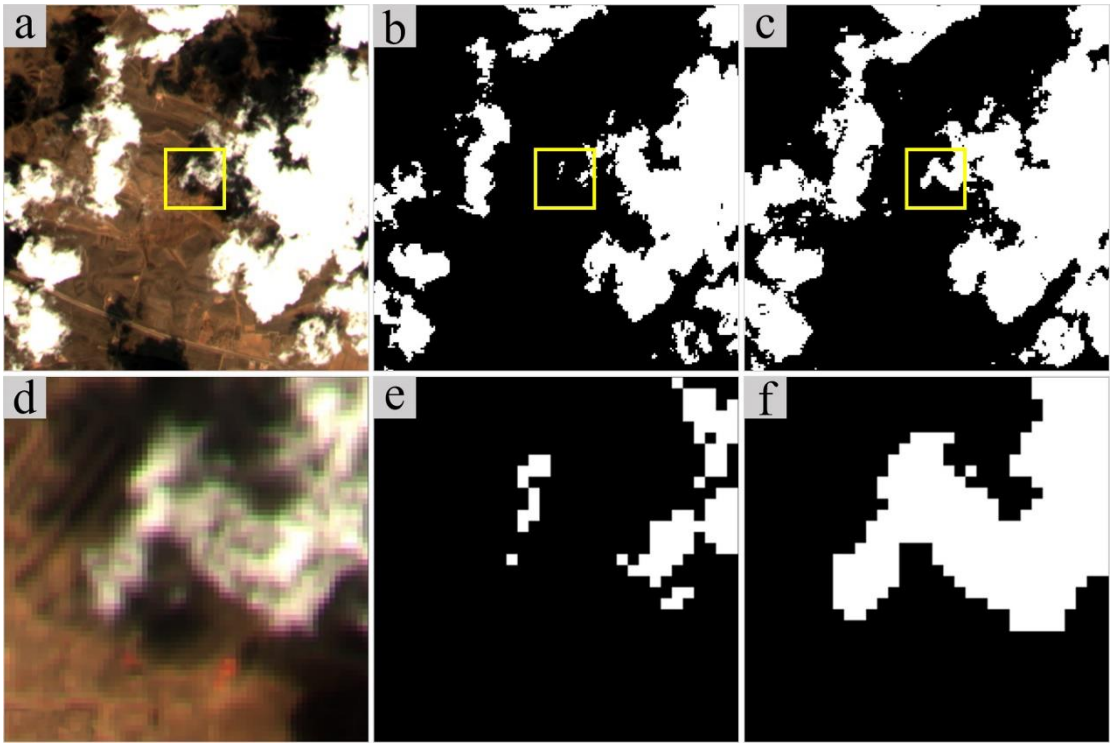


Fig. 3. Local optimization with guided filtering. (a) and (d) R-G-B composite guidance image. (b) and (e) Input binary cloud mask (i.e., the rough cloud mask). (c) and (f) Output binary cloud mask (i.e., the refined cloud mask). Upper row shows entire images, lower row are local enlargements of upper row images.

The MFC algorithm uses a guided filter as a guided feathering in which a binary cloud mask is refined to appear a gray alpha matte near the object boundaries. Here, we empirically set the window radius to 60, and regularization parameter ε equals 10^{-6} for the guided filter to better refine the cloud edges. An output image obtained by a linear transformation from the input image is produced. Here, the RGB composite TOA reflectance image is considered as the

guidance image, and the rough cloud mask CM_R acquired before is considered as the input image, the refined cloud mask CM_G can be generated by segmenting the output gray image acquired by guided filtering. Due to the spectral differences of the ground surface under the clouds, there are different rules for cloud refinement in land and water areas. Fig. 3 is an example of the results of the guided filter.

All pixels can be divided into water and land pixels via water identification. Water has a very low reflectance in the near-infrared band, but land shows a high reflectance in this band (Xie et al., 2016) Moreover, Normalized Difference Vegetation Index (NDVI) is a good indicator to separate water pixels form land pixels, because land NDVI values are generally above water NDVI values (Vermote & Saleous, 2007; Zhu & Woodcock, 2012). According to spectral features of water, near-infrared band reflectance and NDVI are applied to extract water. Here, considering that some turbid or eutrophic water pixels may have relatively large near-infrared band reflectance, thresholds for extracting clear and unclear water bodies are different. The water pixels in a scene are determined by the following test:

$$Water = (NDVI < t_4 \; and \; B4 < t_5) \; or \; (NDVI < t_6 \; and \; B4 < t_7) \tag{8}$$

where

$$NDVI = (B4 - B3)/(B4 + B3) \tag{9}$$

Thus, the process of producing the refined cloud mask CM_G can be expressed as follow:

$$CM_G = GuidedFiltering(RGB, CM_R) > t_8 \; and \; (HOT > t_9 \; or \; Water) \tag{10}$$

Considering that surface features in land areas are more complex than in water areas, the HOT index is used again in land areas to prevent involving non-cloud impurities around clouds into the refined cloud mask. As to the threshold selection for the output gray image after the

guided filtering, Otsu's threshold method (Otsu, 1979) is widely used to find the best threshold from a gray-level histogram to segment the gray image to a binary image. However, this kind of threshold selection method sometimes cannot fit complex conditions well (Zhang & Xiao, 2014; Liu et al., 2015). In this paper, MFC sets a fixed threshold for segmenting the output gray image to a binary mask. The refined cloud mask is then generated, which captures almost all the clouds, including thin clouds around cloud boundaries.

3.1.3 Filtering the non-cloud bright objects using geometric and texture features

Non-cloud bright objects like snow/ice, bright water bodies, and buildings which have similar spectral features to clouds are inevitably included in the refined cloud mask. This kind of impurities cannot be easily separated from the clouds because of the minor spectral differences in the visible and near-infrared bands. Instead, the geometric and textural features can be used to exclude the non-cloud bright objects from the refined cloud mask as much as possible. First, cloud pixels in the refined cloud mask which are connected in eight neighborhoods are merged to be an object. The geometric and texture features are then computed for every object. Next, a check procedure considering geometric and texture features of the object is implemented on the merged objects one by one to determine whether an object is a cloud object or not. Finally, we remove the objects from the refined cloud mask which are marked as non-cloud objects after this check.

A. Geometric features

There are various geometric metrics in FRAGSTATS (McGarigal & Marks, 1995) which is a spatial pattern analysis program for quantifying landscape structure, including area, contiguity index, perimeter-area ratio, etc. The perimeter-area ratio is a simple measure of shape

complexity. However, a problem with this metric as a shape index is that it varies with the size of the object. The fractal dimension index (FRAC) is a proxy to the complexity of an object's shape, which overcomes one of the major limitations of the straight perimeter-area ratio. Furthermore, the length to width ratio (LWR) reflects the relationship between width and length, which can be estimated by calculating minimum enclosing rectangle of the object. As a result, area, FRAC and LWR are considered as the three geometric metrics of the object in this paper. Here, area is the number of pixels involved in an object. The FRAC and LWR can be expressed as follows:

$$FRAC = \frac{2ln(perimeter/4)}{ln(area)} \tag{11}$$

$$LWR = \frac{max(length,width)}{min(length,width)} \tag{12}$$

where $perimeter$ and $area$ refer to each itself, $length$ and $width$ to the smallest rectangle enclosing it.

FRAC approaches 1 for shapes with very simple perimeters such as squares, and approaches 2 for shapes with highly convoluted perimeters. The value of LWR is greater than or equal to 1. The FRAC and LWR of cloud objects are relatively small for its low complexity in shape. The MFC algorithm excludes non-cloud bright objects such as coastlines, roads, and buildings by considering the above geometric features. It is based on the fact that these non-cloud bright objects usually have a higher LWR or FRAC value than cloud objects. Furthermore, the area of an object is considered to ensure that large-area cloud objects which may have high FRAC or LWR value won't be excluded from the cloud mask by mistake.

B. Texture features

Texture features have been successfully employed in object recognition and texture analysis,

and they have also been used for cloud classification and cloud detection (Tao et al., 2007; Xia et al., 2010; Hu et al., 2015; Cheng & Yu, 2015). Non-cloud bright objects, such as buildings or snow patches, do not have obvious geometric features to enable them to be effectively separated from cloud objects, and their shape can be similar to cloud objects. As a result, texture features are used in combination with geometric features to further distinguish cloud and non-cloud objects. In this paper, the local binary pattern (LBP) texture descriptor (Ojala et al. 1994) is implemented to extract the texture features of cloud and non-cloud objects due to its advantage of being illumination invariant and to its low computational cost. The LBP operator is a gray-scale texture operator that describes the spatial structure of the local image texture, and has since been extended to rotation invariant and uniform LBPs (Ojala et al., 2002). It labels each pixel in the image by computing the sign of the difference between the values of that pixel and its neighboring pixels. The result is a binary string or decimal number. The image can then be represented by the histogram of these binary strings or decimal numbers. The rotation invariant LBP for the central pixel is computed as:

$$LBP_{P,R}^{ri} = \min_{0 \le l \le P} \left\{ \sum_{p=0}^{P-1} s\left(g_p - g_c\right) \times 2^{[(p+l) mod P]} \right\} \quad (13)$$

where P is the total number of involved neighbors, R is the radius of the circle (which determines the distance between the neighbors and the central pixel), g_c is the gray value of the central pixel, and g_p is the gray value of the sample point which is angularly evenly distributed around the central pixel with radius R. The value of the step function $s(x)$ equals 1 when x is equal to or above zero, and is 0 if otherwise.

In our implementation, the texture extraction is based on a gray image which is converted from the mean TOA reflectance of the visible bands, and there are 36 levels in the $LBP_{8,3}^{ri}$

histogram. The LBP histogram templates of typical objects include two classes of cloud objects and two classes of non-cloud bright objects, and it were trained from 84 samples which were manually selected in 25 GF-1 WFV images used for training the MFC algorithm. The chi-square distance is an effective indicator to measure the histogram differences:

$$Chi\text{-}square = \sum_{i=1}^{level} \frac{(M_i - N_i)^2}{M_i + N_i} \quad (14)$$

where M_i and N_i are the two normalized histograms, and $level$ denotes the histogram levels.

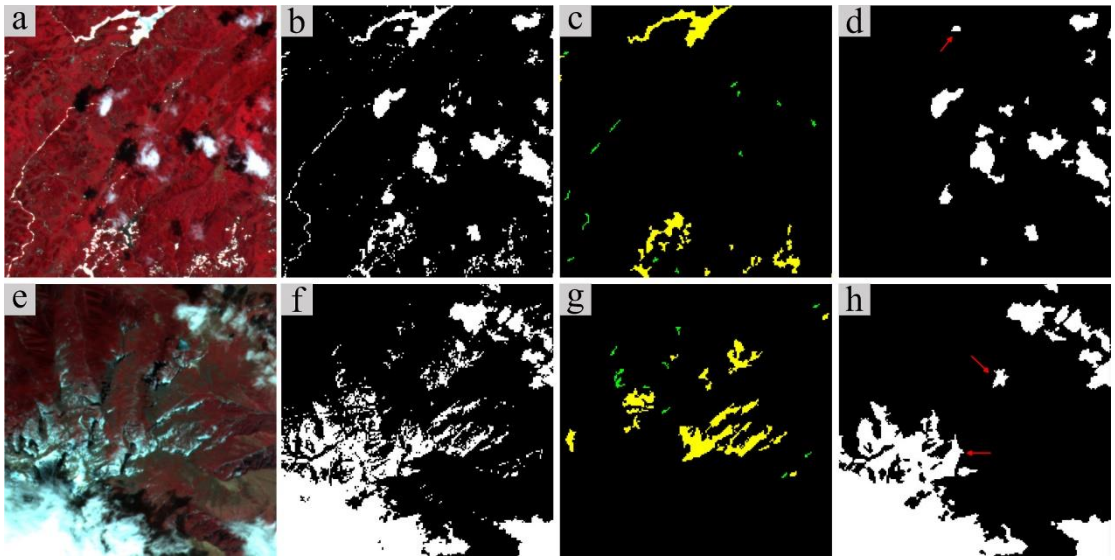


Fig. 4. Filtering non-cloud objects by using geometric and texture features. (a) False-color composite image. (b) Refined cloud mask. (c) Removed non-cloud bright objects (the green and yellow objects are excluded based on the geometric and texture features, respectively). (d) Cloud mask after filtering (objects marked with a red arrow denote non-cloud bright objects which are not excluded). (e)–(h) Another example of excluding snow from a cloud mask. Small-area objects less than five pixels are removed from the refined cloud mask in these two examples.

Hence, MFC calculates the chi-square distance between the current object's LBP histogram and the LBP histogram templates of cloud objects (the distance is denoted as D_c) and non-cloud objects (the distance is denoted as D_n). In the proposed method, the window size set for calculating the texture features is adaptive to the object size. Specifically, it expands according to the object size to make sure that there are enough pixels for the texture extraction. As a result of the object-based geometric feature extraction and the adaptive window size for the texture

feature extraction applied in the proposed method, there are no residues of pixels when a non-cloud object is determined to be excluded from the refined cloud mask. Fig. 4 shows examples in which snow and bright water bodies are excluded from the refined cloud mask. Fig. 5 describes the process flow of filtering non-cloud bright objects. Although there are still some non-cloud objects left in the cloud masks after the filtering, this step clearly improves the visual effect and decreases the commission error for cloud detection.

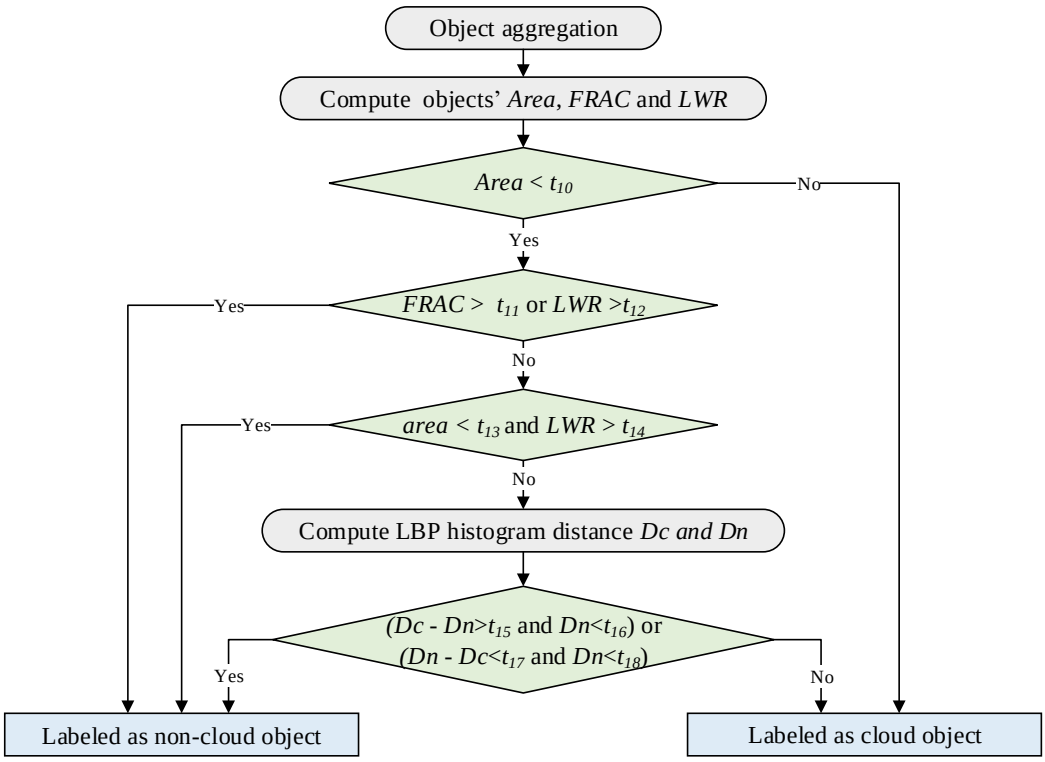


Fig. 5. Flow chart of filtering non-cloud objects using geometric and texture features.

Finally, in order to fill the cloud mask holes, each non-cloud pixel is examined and converted to cloud if at least five of its eight neighbors are cloud pixels. In addition, cloud objects of less than five pixels are removed from the cloud mask to avoid the influence of small-area bright impurities. Afterwards, the final cloud mask is generated.

3.2 Cloud shadow detection

Shadows in the land areas of GF-1 WFV scenes are extracted with the near-infrared band,

based on the fact that the dark effect of land shadow is more obvious in the near-infrared band than in the visible bands. In contrast, the dark effect is more obvious in the visible bands than in the near-infrared band for the shadows in water areas (Fig. 6), because water bodies show stronger absorption in the near-infrared band. A morphological transformation called flood-fill (Soille, 2013) is applied to extract the local potential shadow areas. The transformation is defined as the "reconstruction by erosion" of the input gray image, and it brings the intensity values of the dark areas that are surrounded by lighter areas up to the same intensity level as the surrounding pixels. In this case, the areas where the intensity difference is greater than zero after the transformation are likely to be shadow.

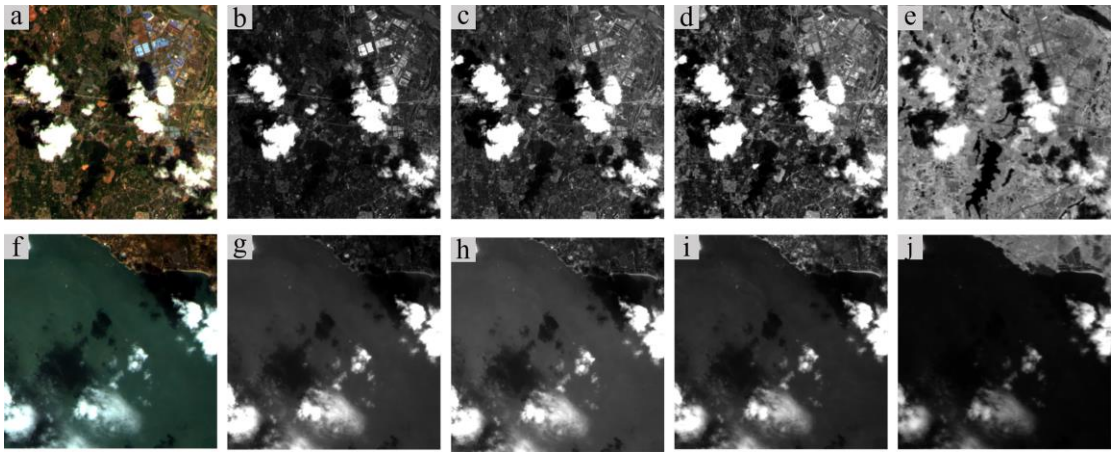


Fig. 6. Comparison of shadows in land and water areas. (a) and (f) The RGB composite images. (b)–(e) The blue, green, red, and near-infrared bands of the land area in a GF-1 WFV scene. (g)–(j) The blue, green, red, and near-infrared bands of the water area in a GF-1 WFV scene.

Therefore, the near-infrared band reflectance and mean visible band reflectance are used as the input of flood-fill to extract shadows in land and water areas, respectively. Since pixels are divided into land or water pixels, a rough shadow mask can be acquired by the following test:

$$Shadow = \begin{cases} flood_fill(B4) - B4 > t_{19} & (for\ land\ area) \\ flood_fill(MeanVis) - MeanVis > t_{20} & (for\ water\ area) \end{cases} \quad (15)$$

where

$$MeanVis = (B1 + B2 + B3)/3 \quad (16)$$

Water bodies can be easily detected as shadow, and there is almost no effective way to separate water from shadow based on their spectral characteristics (Li et al., 2015). In order to prevent water bodies from being wrongly matched as cloud shadow, the geometric features are used to exclude water objects from the rough shadow mask. For every object aggregated from the rough shadow mask, the water pixel percentage is computed to determine if the object is water or shadow. FRAC and LWR are also used considering the fact that some water bodies such as rivers have higher LWR values. Finally, a shadow mask is acquired in which most of the water bodies are excluded.

Object-based cloud and cloud shadow matching (Zhu & Woodcock, 2012) based on their geometric similarity can be implemented after the cloud mask and shadow mask are acquired. It is based on the idea that clouds and their shadows have similar geometric shapes, and the relative direction of cloud shadow can be estimated by the sun and satellite angles. Firstly, the cloud projection direction on the ground can be computed from the satellite viewing azimuth and zenith angle, and the cloud shadow projection direction is related to the solar azimuth and zenith angle. MFC computes the matching direction from the cloud to the cloud shadow according to the viewing and solar angle. The cloud height is then set dynamically based on the statistics, and is assumed to be from 200 m to 12 km according to the study of Luo et al. (2008). The cloud height iterates from the minimum to the maximum to match the cloud object to its shadow, and when the maximum similarity is greater than the similarity threshold, the matched shadow location will be labeled as cloud shadow. Finally, considering the fact that there may be some bias between the matching direction and the real cloud shadow projection direction, the matched shadow may not be integral and part of it may be missed, so MFC implements an

object-based cloud shadow correction process based on the shadow layer to generate the cloud shadow mask.

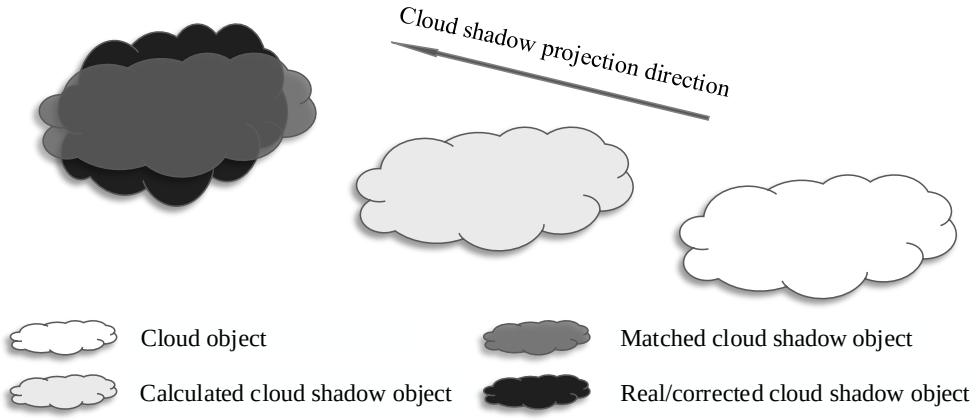


Fig. 7. Object-based cloud and cloud shadow matching and correction (improved from Fmask).

The cloud and cloud shadow matching used in the proposed method is simplified and improved from Fmask (Zhu & Woodcock, 2012), but the object-based cloud shadow correction is an extra step used for decreasing the omission error of the cloud shadow after the cloud and cloud shadow matching (Fig. 7). It first aggregates the shadow pixels which are connected in eight neighborhoods to be an object in the cloud shadow mask and shadow mask, respectively. The overlap area in the two masks for each cloud shadow object is then found. If the ratios of the overlap area to the corresponding shadow object area, and the current cloud shadow object area are both above the thresholds, then the correction condition can be met, and the current cloud shadow object location will be corrected to the corresponding shadow object location. The same correction step is repeated for every cloud shadow object. The original cloud shadow mask and the corrected part of the shadow layer are then merged to generate the rough cloud shadow mask.

Considering that not all cloud shadows can be matched with their corresponding clouds because of matching error, cloud shadow refinement with the guided filter is also implemented

in the cloud shadow detection, based on the idea that the missed cloud shadows are usually around the matched cloud shadows. In this case, we apply the guided filter again to capture the missed cloud shadows after cloud and cloud shadow matching and correction. In order to prevent that water bodies or other non-cloud-shadow objects are involved in the cloud shadow mask, geometry features are used to check every object in the refined cloud shadow mask, and filtering the non-cloud-shadow objects. More details are shown in Algorithm 1.

Algorithm 1. Cloud shadow refinement and filtering non-cloud-shadow objects.

Input image: $rough\ cloud\ shadow\ mask\ (CSM_R),\ NIR\text{-}R\text{-}G\ composite\ image\ (NRG)$
Output image: $final\ cloud\ shadow\ mask\ (CSM_F)$
Cloud shadow refinement:
$I_{shadow} = GuidedFilter(NRG, CSM_R)$
$CSM_G = (I_{shadow} > t_{21}\ and\ B4 < B4_treshold)\ or\ CSM_R$
$B4_{treshold}\ denotes\ the\ 17.5\ quantile\ of\ pixels'\ NIR\ reflectance\ in\ land\ area$
Filtering non-shadow objects:
$object\ aggregation\ in\ refined\ cloud\ shadow\ mask\ CSM_G$
$For\ first\ object\ to\ last\ object$
$Calculate\ Calculate\ \boldsymbol{FRAC}\ and\ \boldsymbol{LWR}\ according\ to\ Eq.\ (11)\text{-}(12)$
$If\ Area > t_{23}\ or\ FRAC > t_{22}$
$\quad remove\ the\ shadow\ object$
$else\ if\ LWR > t_{24}\ or\ (Area < t_{25}\ and\ LWR > t_{26})$
$\quad remove\ the\ shadow\ object$
$else\ continue$
End
$Generate\ the\ final\ cloud\ shadow\ mask\ CSM_F$

Finally, the same strategy is adopted to fill the holes in the cloud shadow mask. In addition, as previous study (Braaten et al., 2015; Harb et al., 2016) do in the postprocessing step, objects less than seven pixels generally associated with noise are removed from the cloud shadow mask, and one pixel dilation is necessary for the cloud shadow mask to capture cloud shadows form thin cloud edges. MFC sets a higher priority for cloud. Therefore, a pixel is labeled as cloud when it is also labeled as cloud shadow. The cloud and cloud shadow mask can be acquired

after the aggregation of the cloud mask and the cloud shadow mask.

3.3 Parameter selection analysis

The above parameters in MFC algorithm are fixed after a large number of experiments, and all the experimental results in this paper are produced with the same set of parameters. Here, recommended parameter settings in the MFC algorithm are provided in Table 2, and it can be directly applied without adjustment. Likewise, parameters in some cloud detection methods (Irish et al., 2006; Zhu & Woodcock, 2012; Braaten et al., 2015) which are trained by a great deal of data are certain. For instance, the ACCA algorithm consists of twenty-six specific decisions or filters which include 32 fixed thresholds and 3 dynamic thresholds. In the post-processing step of the proposed method, the morphological operation such as dilation is necessary to improve the cloud and cloud shadow detection results, and the parameter settings in this step need to ensure a finer visual effect and decrease the possible commission and omission errors.

Table 2. Recommended parameter settings in the MFC algorithm.

Recommended parameter settings for cloud detection							
t_1	0.13	t_2	0.7	t_3	0.07	t_4	0.15
t_5	0.2	t_6	0.2	t_7	0.15	t_8	0.12
t_9	0.08	t_{10}	4E4	t_{11}	1.56	t_{12}	6.3
t_{13}	4E3	t_{14}	5.4	t_{15}	0.02	t_{16}	0.10
t_{17}	0.02	t_{18}	0.03				
Recommended parameter settings for cloud shadow detection							
t_{19}	0.06	t_{20}	0.01	t_{21}	0.27	t_{22}	1.56
t_{23}	4E4	t_{24}	6.3	t_{25}	400	t_{26}	5.4

4. Experimental results

Due to the few radiation differences between the four cameras in the GF-1 WFV imaging system, scenes of all four cameras were used as experimental data to ensure the adaptability of the algorithm. To quantitatively evaluate the performance of the MFC algorithm, it was

validated on 16 GF-1 WFV full scenes from different seasons and areas across China from June 2013 to December 2014. In addition, the selected images were all L2A level products which were produced after systematic geometric correction. In order to validate the adaptability of the method for different land-cover types (sea/water, ice/snow, barren, forest, wetlands, rural/urban area, etc.), these different land-cover types were covered in the selected validation scenes. The list of scenes used for the accuracy evaluation is presented in Table 3, and the locations of these scenes across China are shown in Fig. 8.

Table 3. List of the GF-1 WFV scenes for method validation.

NO.	Date	Scene location		Camera	Central province
1	06/16/2013	E106.0	N26.9	WFV4	GuiZhou
2	07/09/2013	E87.1	N41.3	WFV1	XinJiang
3	07/14/2013	E89.3	N43.4	WFV4	XinJiang
4	08/13/2013	E119.2	N29.3	WFV2	ZheJiang
5	09/29/2013	E107.0	N30.2	WFV4	ChongQing
6	11/18/2013	E119.1	N36.3	WFV1	ShanDong
7	11/21/2013	E100.8	N23.9	WFV3	YunNan
8	03/02/2014	E102.0	N28.0	WFV1	SiChuan
9	04/23/2014	E115.5	N42.6	WFV2	Inner Mongolia
10	04/28/2014	E109.1	N39.3	WFV2	Inner Mongolia
11	05/10/2014	E109.6	N18.5	WFV4	HaiNan
12	07/02/2014	E89.3	N35.6	WFV3	Tibet
13	07/04/2014	E127.2	N45.9	WFV2	HeiLongJiang
14	07/16/2014	E99.4	N36.3	WFV1	QingHai
15	10/08/2014	E114.9	N23.5	WFV4	GuangDong
16	12/08/2014	E88.6	N28.0	WFV1	Tibet

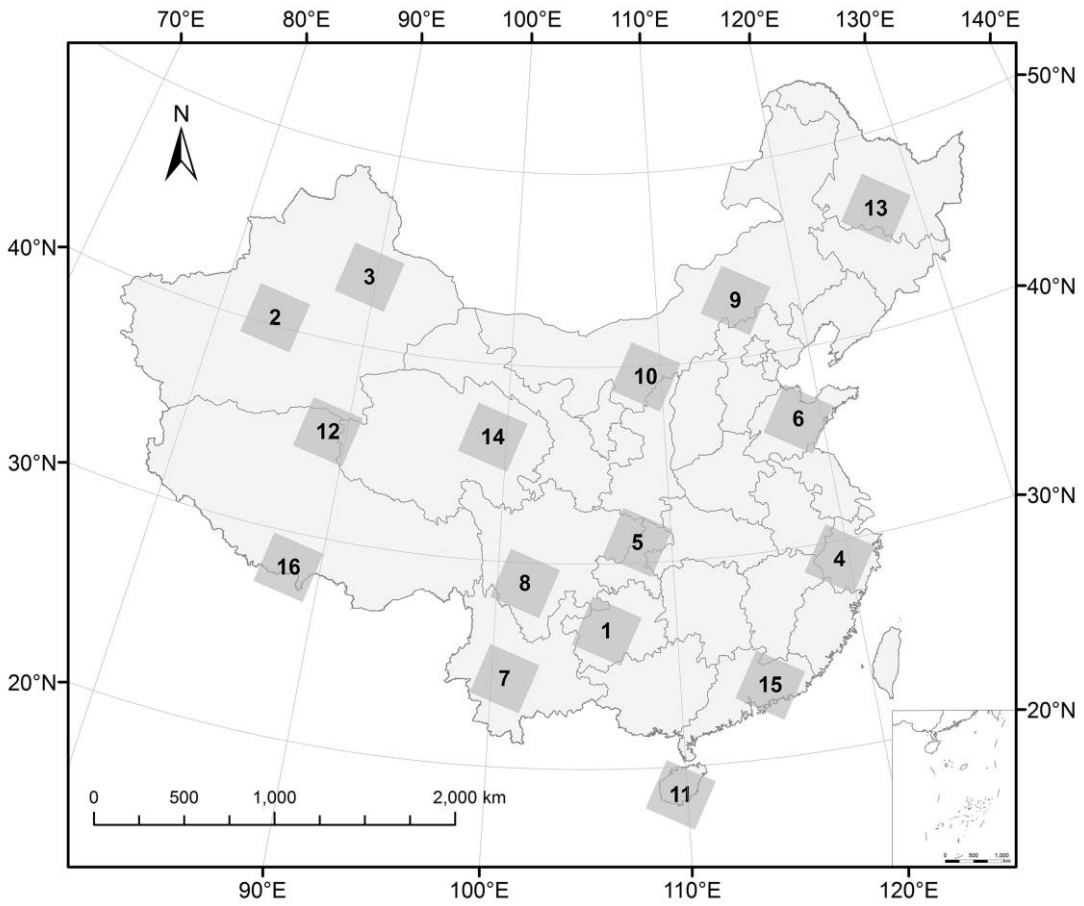


Fig. 8. Locations of GF-1 WFV scenes across China for method validation.

Since ground observations of cloud and cloud shadow are usually unavailable, it is difficult to perform an accurate quantitative validation of the proposed method results. Therefore, the reference masks for accuracy evaluation in this study were obtained by manually drawing cloud/cloud shadow borders after visual inspection by experienced users. Similar approaches were applied in previous studies of cloud detection (Irish et al., 2006; Scaramuzza et al., 2012) to acquire the reference masks for accuracy evaluation. As an indicator of image quality and availability, the cloud fraction of a single scene is also important in practical applications. Hence, in addition to the pixel scale evaluation method, the accuracy of the cloud fraction estimation can be used to evaluate the performance of the cloud detection algorithms.

4.1 Cloud fraction estimation

The cloud fraction denotes the cloud cover percentage in the imagery as a whole. In the header file of GF-1 WFV imagery, there is a parameter which indicates the cloud cover percentage. In this section, the cloud cover percentage in the header file is used to compare the accuracy with MFC in estimating the cloud fraction. The cloud fractions derived from the header file, the reference mask, and the MFC mask for 16 scenes are used for comparison. The mean absolute error (MAE) and the mean relative error (MSE) are used as indicators for error calculation:

$$MAE = \frac{1}{n}\sum_{i=1}^{n} |P_R(i) - P_M(i)| \tag{17}$$

$$MRE = \frac{1}{n}\sum_{i=1}^{n} \frac{|P_R(i) - P_M(i)|}{P_R(i)} \tag{18}$$

where $P_R(i)$ and $P_M(i)$ represent the cloud fraction from the reference mask and the MFC mask, respectively, and n is the amount of images used for the accuracy evaluation.

The method used to provide cloud cover percentage information in the header file of GF-1 WFV imagery by the data distributor is called official method, it helps search the data uses need in the data download website and filter out images which has a large-area cloud cover. Here, the errors of official method can be calculated by replacing the $P_M(i)$ with $P_O(i)$ which is the cloud fraction from header file.

Table 4. Error comparison of the cloud fraction estimation between the official method and MFC.

	Official method	MFC
MAE	0.052	0.009
MRE	0.497	0.268

According to the accuracy evaluation results in Table 4, the MAE and MRE of the proposed method are 0.009 and 0.268, a significant decrease compared with the official method whose

MAE and MRE are 0.052 and 0.497. Moreover, the official method mostly underestimates the cloud cover percentage. The cloud fraction derived from the MFC cloud masks is more likely in agreement with the reference masks.

4.2 Cloud and cloud shadow distribution detection

The accuracy assessment for the cloud and cloud shadow distribution measures the agreements and differences between the cloud and cloud shadow in the MFC masks and the reference masks on a per-pixel basis. For the accuracy evaluation of cloud detection, cloud and non-cloud pixels are considered as two classes. However, it should also be noted that the overall accuracy of cloud shadow may be meaningless because of the low cloud shadow percentage in the imagery.

Table 5. Mean accuracy and standard deviation (in brackets) of MFC in cloud and cloud shadow detection.

	Overall acc.	Producer's acc.	User's acc.
Cloud	98.30% (0.013)	88.78% (0.194)	94.36% (0.042)
Cloud Shadow	— —	74.66% (0.131)	74.48% (0.142)

From the accuracy evaluation results in Table 5, the average overall accuracy of the cloud in MFC is 98.30%, the average producer's accuracy is 88.78% with a standard deviation of 0.194, and the average user's accuracy is 94.36% with a very small standard deviation (0.042). The reason why the cloud producer's accuracy is lower than the cloud user's accuracy, and the standard deviation of cloud producer's accuracy is greater than the standard deviation of cloud user's accuracy, is that there are large commission errors in the cloud mask of validation imagery covered with large-area non-cloud bright objects. MFC performs less accurately as for cloud shadow detection is concerned, as the average producer's accuracy and user's accuracy of cloud shadow are 74.66% and 74.48%, respectively.

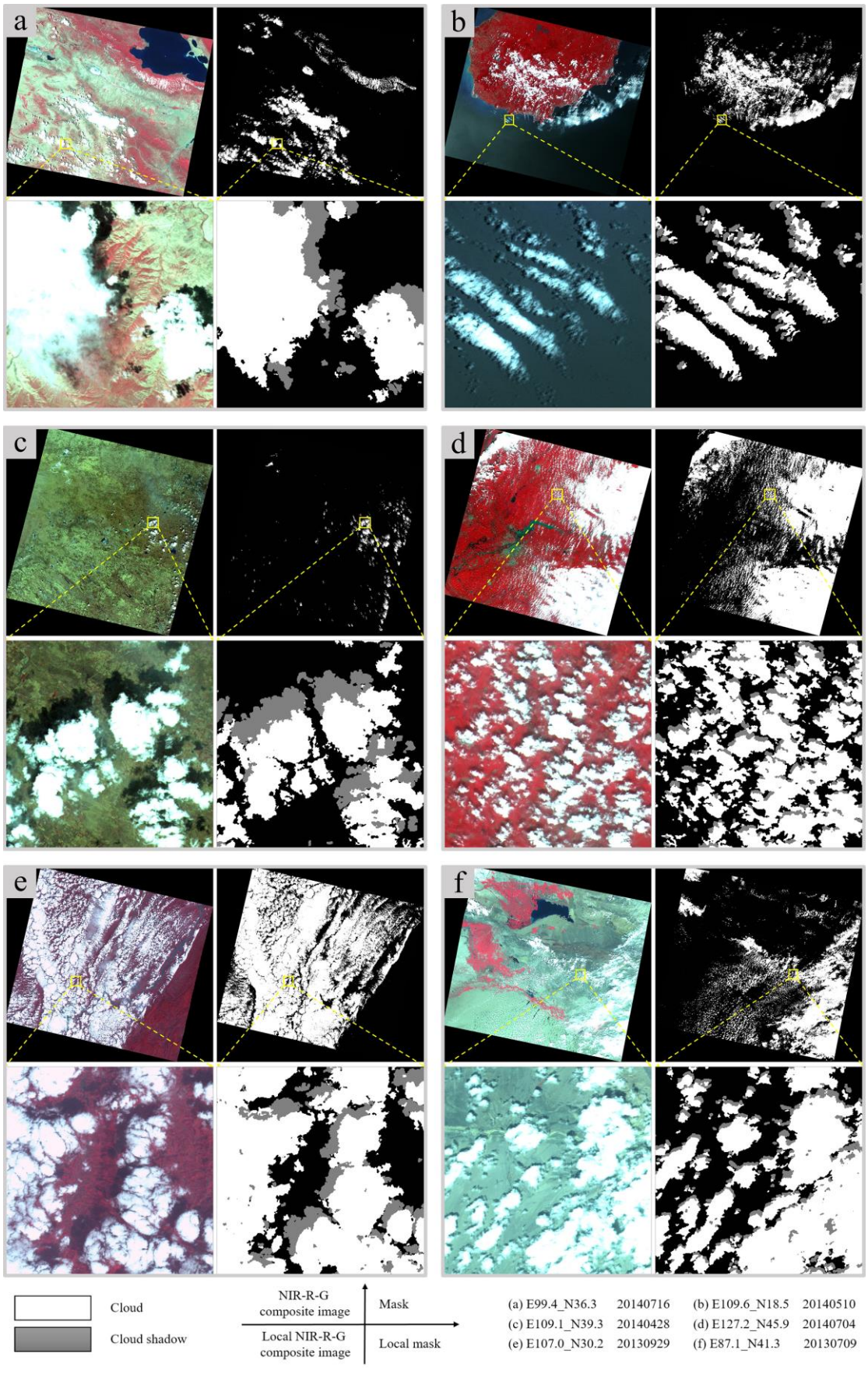


Fig. 9. Parts of GF-1 WFV scenes and MFC masks used for the accuracy evaluation.

The main errors in the cloud shadow mask by MFC come from terrain shadows, and cloud shadows cast by thin clouds which are not dark enough. By comparing the results of MFC with false color composite images (Fig. 9), it is possible to visually appreciate the strong ability of MFC to detect clouds, but also its deficiency in cloud shadow detection.

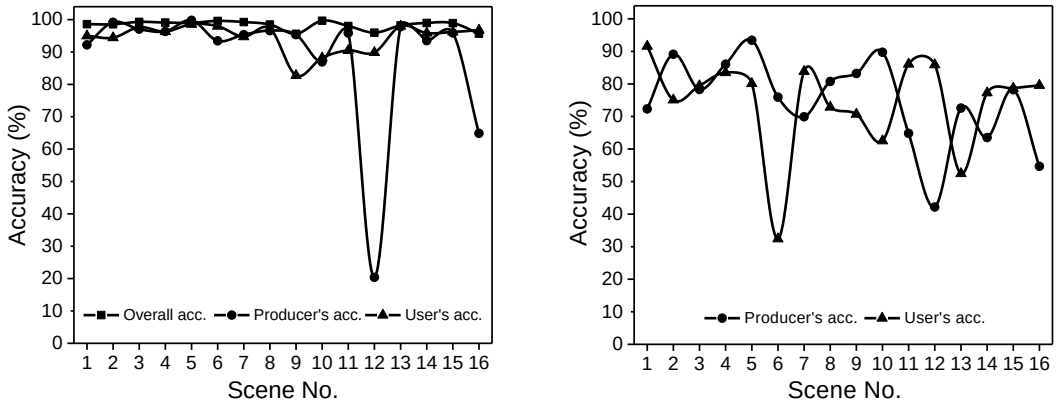


Fig. 10. Accuracy of MFC in validation imagery. (a) Cloud accuracy. (b) Cloud shadow accuracy.

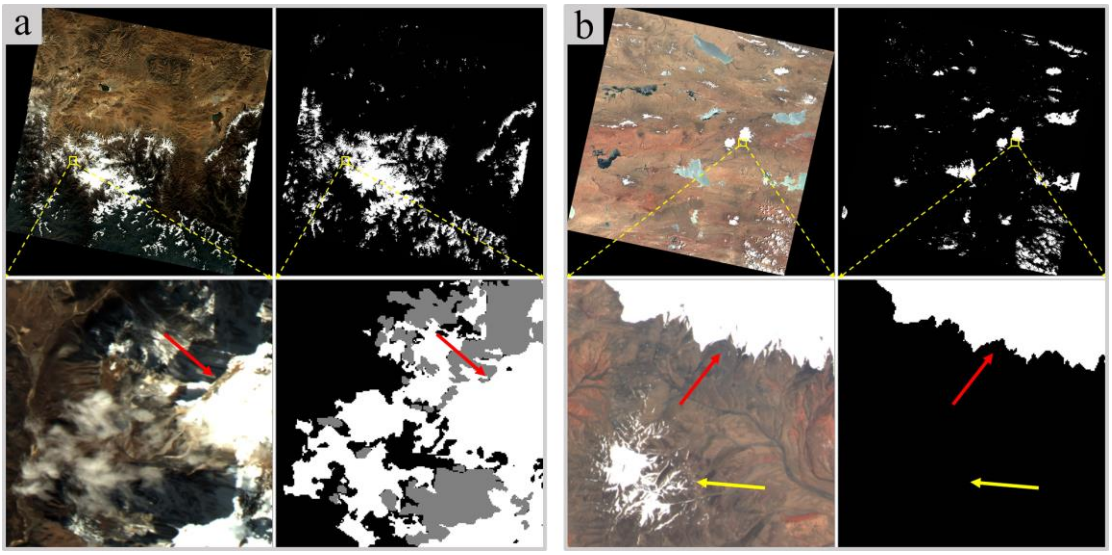


Fig. 11. Major cloud commission error of MFC in scenes covered with large-area snow and bright water bodies. (a) RGB composite image (E88.6 N28.0, December 2014) and MFC mask in which the snow object nearing clouds was misclassified as cloud (red arrow). (b) RGB composite image (E89.3 N35.6, July 2014) and MFC mask in which the large-area snow object was not excluded from the cloud mask (red arrow), while the smaller snow object was filtered out clearly (yellow arrow).

The accuracy values for the validation scenes are shown in Fig. 10. Due to the large area covered by snow, two scenes have low cloud producer's accuracy values. Fig. 11 shows these

two cases. MFC fails to exclude this kind of snow and bright water bodies objects from the cloud mask which have large area. Fortunately, MFC still achieves high cloud user's accuracy values even under this circumstances.

5. Discussion

5.1 Fast cloud fraction estimation

To meet the different application requirements, a so called "fast-mode" MFC has been implemented to rapidly estimate the cloud fraction in an image. This method can generate a rough cloud mask for a single GF-1 WFV scene (about 17000×16000 pixels) within 20 s, while precise-mode MFC results are usually obtained in 3–5 min. The main difference between the two modes is the different ratios of downsampling for the original scene. Fast mode downsamples the original scene to a smaller size than precise mode. Furthermore, the cloud shadow detection procedure is discarded in fast mode to save processing time.

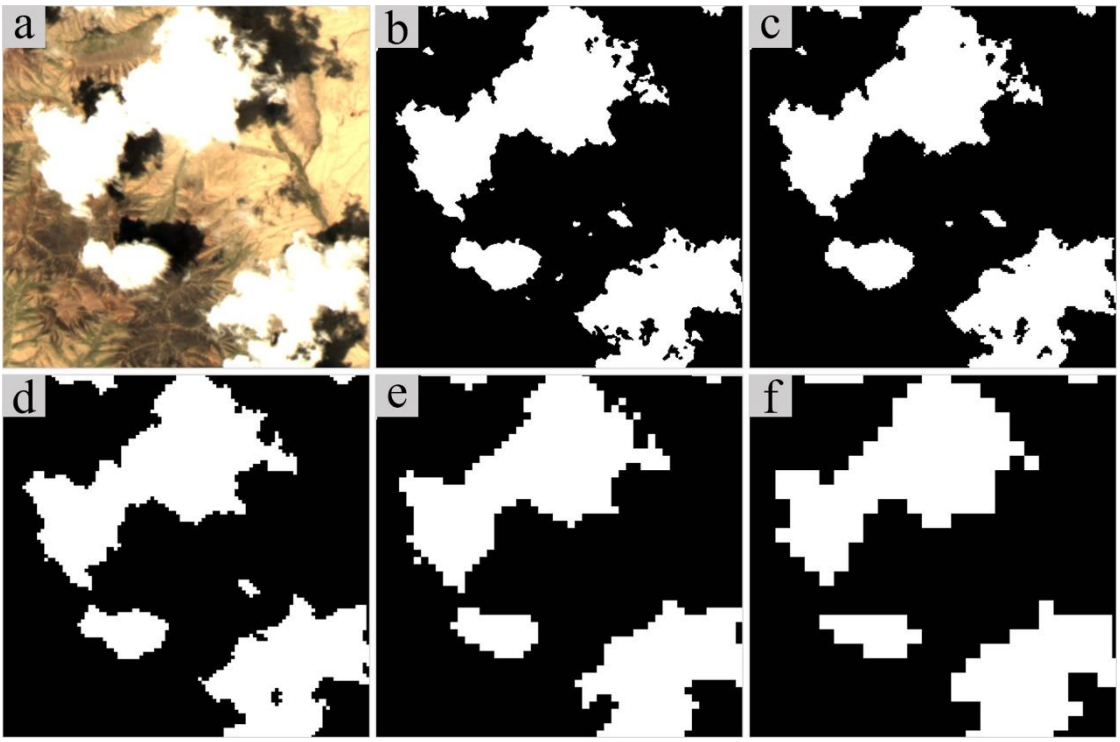


Fig. 12. Comparison of MFC masks acquired with different downsampling scales for the input scene. (a) RGB composite image. (b) The mask acquired without downsampling, i.e. scale=1. (c) Scale=2. (d) Scale=4. (e) Scale=8.

(f) Scale=16.

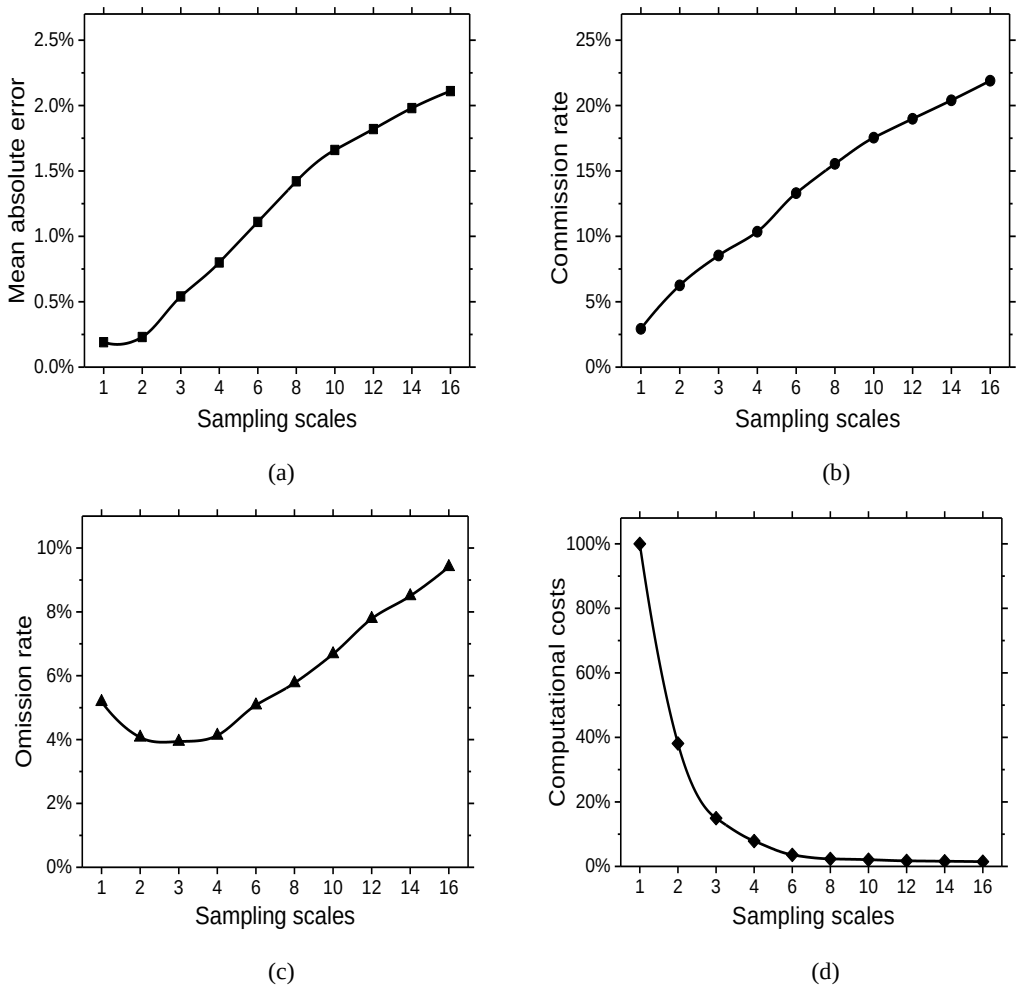


Fig. 13. Accuracy loss and computational costs for different downsampling scales. (a) The relationship between the MAE of fast cloud fraction estimation and different downsampling scales. (b) The commission rate in cloud detection with different downsampling scales. (c) The omission rate in cloud detection with different downsampling scales. (d) The relationship between the computational costs and different downsampling scales.

An acceleration strategy which downsamples the original image to save processing time is very common in image processing, and the scale of the downsampling is the key factor in this strategy. Different downsampling scales were tested in MFC. A comparison of the MFC masks obtained with different downsampling scales for the input scene is shown in Fig. 12. The MAE for cloud fraction estimation, commission and omission rate at a pixel scale, and the costs are the indicators which are chosen to measure the effect of the different scales (Fig. 13). Note that, in order to better compare the results for different downsampling scales, the images which are

covered by large-area snow are not included in the test imagery in this section. Moreover, an appropriate scale can be decided only if it results in an apparent reduction in computational costs with a slight accuracy sacrifice. Thus, considering the comprehensive influence of the acceleration strategy, the default subsampling ratio is set to 2 for the precise mode instead of the original size. This leads to an obvious improvement in the processing speed and only a slight reduction in the accuracy. Finally, a subsampling ratio of 6 is recommended for the fast mode, because it maintains a good balance between the processing time and accuracy for the cloud fraction estimation.

The fast-mode MFC has apparent advantages because of fast cloud fraction estimate, and it can provide more accurate results than the cloud fraction results provided by official method in the header file. Precise-mode MFC, instead, aims at providing a pixel-scale precise cloud and cloud shadow mask which can be used for cloud and cloud shadow removal, land-cover change detection, and so on. The computational cost of MFC is low: our experimental program was coded in C++ language and run in parallel on a laptop with an Intel Core i5-4210M CPU. It takes less than 20 s to estimate the cloud fraction, and 3–5 min. to generate a precise cloud and cloud shadow mask for one GF-1 WFV scene. As a result of this good performance and efficiency, it is expected that MFC will be used for automatic cloud and cloud shadow detection for a large number of GF-1 WFV images in land cover monitoring project.

5.2 Comparative analysis on detection accuracy

Note that there is no specific method appearing in the published papers for cloud detection in GF-1 WFV imagery, which leads to difficulties in comparison. The official method which generates the cloud cover percentage information in the header file of GF-1 WFV imagery is

not public, as a result, our method only can be quantitatively compared to the official method in cloud cover percentage estimation on the whole. The validation results indicate that the MFC algorithm performs better than the official method.

Additionally, since the GF-1 WFV sensor has the similar spectral setting as Landsat ETM+ sensor in the first four bands, the proposed method can be relatively compared to the cloud detection methods for Landsat imagery. The Landsat imagery cloud detection methods including ACCA (Irish et al., 2006), Fmask (Zhu & Woodcock, 2012) and MSScvm (Braaten et al., 2015) are relatively compared in accuracy and number of input bands (Table 6). The ACCA and Fmask algorithms respectively use five and seven of the eight spectral bands in Landsat ETM+ imagery, and MSScvm uses all four spectral bands in Landsat MSS imagery and corresponding DEM data, while MFC uses only all four visible and near-infrared bands in GF-1 WFV imagery, and achieves almost the same accuracy as Fmask which performances best in cloud detection between the three methods in Table 6. By the relative comparison to Landsat imagery cloud detection methods in accuracy and the number of input bands, the proposed method achieves a high cloud accuracy with less spectral bands.

Table 6. Accuracy and input bands comparison of different methods for cloud and cloud shadow detection.

Method	Sensors	Input bands	Item	Overall accuracy	Producer's accuracy	User's accuracy
ACCA (Irish, 2006)	Landsat ETM+	5	Cloud	84.8%	72.1%	91.8%
			Cloud shadow	— —	— —	— —
Fmask (Zhu, 2012)	Landsat ETM+	7	Cloud	96.41%	92.10%	89.40%
			Cloud shadow	— —	>70%	>50%
MSScvm (Braaten, 2015)	Landsat MSS	4	Cloud	84.0%	— —	— —
			Cloud shadow		— —	— —

Since both the ACCA and Fmask algorithms exploit shortwave infrared and thermal infrared bands in Landsat ETM+ imagery, the thermal infrared band plays an important role in cloud

detection, and the shortwave infrared bands is critical for snow identification. Note that the spectral setting of GF-1 WFV sensor doesn't include shortwave infrared and thermal infrared regions, the proposed method combines multiple features to improve the ability of identifying cloud pixels.

5.3 Limitations

There are still some errors in the cloud and cloud shadow masks generated by MFC, due to the limitations of the algorithm. For cloud detection, because of the minor differences of both geometric and texture features between cloud and some large-area non-cloud bright objects, there might be non-cloud bright objects in cloud masks which still were not removed. This is especially true were wide snowy areas and bright water bodies. As for cloud shadow detection, terrain shadow and water bodies around clouds are easily misclassified as cloud shadows when they have low reflectance in the near-infrared band. Although a more accurate cloud shadow mask can be acquired after the cloud shadow correction process most of the time, the correction process may increase the cloud shadow commission error when the cloud shadows are connected with terrain shadow or water bodies which are not excluded from the potential shadow layer.

A further source of uncertainty is that the radiometric calibration coefficients for GF-1 WFV imagery are not stable for a whole year (Yang et al., 2015), because the cameras on the WFV imaging system are not state-of-the-art instruments. Likewise, the radiometric calibration parameters for GF-1 imagery used in the MFC algorithm are not absolutely accurate over time. As a result of the radiometric calibration error, the cloud and water reflectance in different scenes are not always the same, which may lead to errors in water and cloud detection.

In addition, drawing the reference masks for the validation imagery is a time-consuming task. Unavoidable manual drawing errors and the limited amount of scenes used for the accuracy assessment lead to minor bias in the accuracy assessment. Considering that there are enough pixels involved in accuracy evaluation, this minor bias can be reduced.

6. Conclusions

In general, it is hard to obtain satisfactory results for cloud and cloud shadow detection when using images which only include visible and near-infrared spectral bands. As a result of the insufficient spectral information of GF-1 WFV imagery for precise cloud and cloud shadow detection, thin clouds are not easy to capture, and non-cloud bright objects are frequently labeled as “cloud” in the cloud mask. In the proposed method, a local optimization strategy with guided filtering is implemented to capture the thin clouds around cloud boundaries and decrease the omission error of the cloud mask. In addition, the geometric features are used in combination with texture features to reduce the commission errors by excluding non-cloud bright objects from the cloud mask, non-shadow objects from the shadow mask, and non-cloud-shadow objects from the cloud shadow mask. To some degree, the use of multiple features, such as geometric and texture features, makes up for the deficiency of the spectral information for cloud and cloud shadow detection in GF-1 WFV imagery.

In conclusion, the proposed multi-feature combined method (MFC) is a promising algorithm for cloud and cloud shadow detection. Indeed, it achieves good results with limited spectral bands, and will be applied to land cover monitoring project. In general, a specific cloud detection method is not universally applicable for different kinds of satellite imagery because of the differences in spectral setting and spectral response characteristic. As a result, it may not

work well if directly apply our method to other kind of imagery without changing parameters. In our future study, the general framework of cloud and cloud shadow detection proposed in this paper will be extended to other optical satellite imagery which has a similar spectral setting.

Acknowledgements

This work was supported by the Cross-disciplinary Collaborative Teams Program for Science, Technology and Innovation of the Chinese Academy of Sciences. The GF-1 WFV imagery used in this paper were provided by the China Centre for Resources Satellite Data and Application (CRESDA) and the China Land Surveying and Planning Institute (CLSPI). We would like to gratefully thank the authors of Fmask for the source code of the Fmask algorithm, which is a good reference for MFC. Thanks also to Qi Wu, who is a software engineer at Wuhan University, for helping us to implement MFC more effectively.

References

Bian, J., Li, A., Jin, H., Zhao, W., Lei, G., & Huang, C. (2014). Multi-temporal cloud and snow detection algorithm for the HJ-1A/B CCD imagery of China. In, *2014 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 501-504).

Braaten, J.D., Cohen, W.B., & Yang, Z. (2015). Automated cloud and cloud shadow identification in Landsat MSS imagery for temperate ecosystems. *Remote Sensing of Environment, 169*, 128-138.

Chen, N., Li, J., & Zhang, X. (2015). Quantitative evaluation of observation capability of GF-1 wide field of view sensors for soil moisture inversion. *Journal of Applied Remote Sensing, 9*, 97097.

Chen, S., Chen, X., Chen, J., Jia, P., Cao, X., & Liu, C. (2016). An iterative haze optimized transformation for automatic cloud/haze detection of Landsat imagery. *IEEE Transactions on Geoscience and Remote Sensing, 54*, 2682-2694.

Cheng, H., & Yu, C. (2015). Multi-model solar irradiance prediction based on automatic cloud classification. *Energy, 91*, 579-587.

Cheng, Q., Shen, H., Zhang, L., Yuan, Q., & Zeng, C. (2014). Cloud removal for remotely sensed images by similar pixel replacement guided with a spatio-temporal MRF model. *ISPRS Journal of Photogrammetry and Remote Sensing, 92*, 54-68.

Di Vittorio, A.V., & Emery, W.J. (2002). An automated, dynamic threshold cloud-masking algorithm for daytime AVHRR images over land. *IEEE Transactions on Geoscience and Remote Sensing, 40*, 1682-1694.

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.

Hagolle, O., Huc, M., Pascual, D.V., & Dedieu, G. (2010). A multi-temporal method for cloud detection, applied to FORMOSAT-2, VENμS, LANDSAT and SENTINEL-2 images. *Remote Sensing of Environment, 114*, 1747-1755.

Harb, M., Gamba, P., & Dell Acqua, F. (2016). Automatic delineation of clouds and their shadows in Landsat and CBERS (HRCC) data. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 9*, 1532-1542.

He, K., Sun, J., & Tang, X. (2013). Guided image filtering. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 35*, 1397-1409.

Hu, X., Wang, Y., & Shan, J. (2015). Automatic Recognition of Cloud Images by Using Visual Saliency Features. *IEEE Geoscience and Remote Sensing Letters, 12*, 1760-1764.

Huang, C., Thomas, N., Goward, S.N., Masek, J.G., Zhu, Z., Townshend, J.R., & Vogelmann, J.E. (2010). Automated masking of cloud and cloud shadow for forest change analysis using Landsat images. *International Journal of Remote Sensing, 31*, 5449-5464.

Huang, C., Thomas, N., Goward, S.N., Masek, J.G., Zhu, Z., & Townshend, J.R.G. ô. (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. (2000). Landsat 7 science data users handbook. *National Aeronautics and Space Administration, Report, 2000*, 415-430.

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 and Remote Sensing, 72*, 1179-1188.

Khlopenkov, K.V., & Trishchenko, A.P. (2007). SPARC: New cloud, snow, and cloud shadow detection scheme for historical 1-km AVHHR data over Canada. *Journal of Atmospheric and Oceanic Technology, 24*, 322-343.

Le Hégarat-Mascle, S., & Andr é, 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.

Li, J., Chen, X., Tian, L., Huang, J., & Feng, L. (2015). Improved capabilities of the Chinese high-resolution remote sensing satellite GF-1 for monitoring suspended particulate matter (SPM) in inland waters: Radiometric and spatial considerations. *ISPRS Journal of Photogrammetry and Remote Sensing, 106*, 145-156.

Li, P., Dong, L., Xiao, H., & Xu, M. (2015). A cloud image detection method based on SVM vector machine. *Neurocomputing, 169*, 34-42.

Li, Q., Wang, H., Zhang, H., Du, X., Zhao, L., Li, D., & Yu, H. (2015). Agricultural production investigation using GF-1 CCD data in Beijing-Tianjin-Hebei region. In, *2015 Fourth International Conference on Agro-Geoinformatics (Agro-geoinformatics)* (pp. 188-191).

Li, S., Sun, D., Goldberg, M.E., & Sjoberg, B. (2015). Object-based automatic terrain shadow removal from SNPP/VIIRS flood maps. *International Journal of Remote Sensing, 36*, 5504-5522.

Li, X., Shen, H., Zhang, L., Zhang, H., Yuan, Q., & Yang, G. (2014). Recovering quantitative remote sensing products contaminated by thick clouds and shadows using multitemporal dictionary learning.

IEEE Transactions on Geoscience and Remote Sensing, 52, 7086-7098.

Lin, C.H., Lin, B.Y., Lee, K.Y., & Chen, Y.C. (2015). Radiometric normalization and cloud detection of optical satellite images using invariant pixels. *ISPRS Journal of Photogrammetry and Remote Sensing, 106*, 107-117.

Liu, S., Zhang, L., Zhang, Z., Wang, C., & Xiao, B. (2015). Automatic cloud detection for all-sky images using superpixel segmentation. *IEEE Geoscience and Remote Sensing Letters, 12*, 354-358.

Lu, C., & Bai, Z. (2015). Characteristics and typical applications of GF-1 satellite. In, *2015 IEEE International Geoscience and Remote Sensing Symposium (IGARSS)* (pp. 1246-1249).

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*, 4167-4185.

McGarigal, K., & Marks, B.J. (1995). FRAGSTATS: spatial pattern analysis program for quantifying landscape structure.

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.

Ojala, T., Pietikäinen, M., & Harwood, D. (1994). Performance evaluation of texture measures with classification based on Kullback discrimination of distributions. In, *Proceedings of the 12th IAPR International Conference on Pattern Recognition* (pp. 582-585).

Ojala, T., Pietikäinen, M., & Mäenpää, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 24*, 971-987.

Otsu, N. (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man and Cybernetics, 9*, 62-66.

Platnick, S., King, M.D., Ackerman, S.A., Menzel, W.P., Baum, B.A., Riedi, J.C., & Frey, R.A. (2003). The MODIS cloud products: algorithms and examples from terra. *IEEE Transactions on Geoscience and Remote Sensing, 41*, 459-473.

Scaramuzza, P.L., Bouchard, M., & Dwyer, J.L. (2012). Development of the Landsat data continuity mission cloud-cover assessment algorithms. *IEEE Transactions on Geoscience and Remote Sensing, 50*, 1140-1154.

Shen, H., Li, H., Qian, Y., Zhang, L., & Yuan, Q. (2014). An effective thin cloud removal procedure for visible remote sensing images. *ISPRS Journal of Photogrammetry and Remote Sensing, 96*, 224-235.

Soille, P. (2013). *Morphological image analysis: principles and applications*: Springer Science & Business Media.

Tao, D., Li, X., Wu, X., & Maybank, S.J. (2007). General tensor discriminant analysis and gabor features for gait recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 29*, 1700-1715.

Vermote, E., & Saleous, N. (2007). LEDAPS surface reflectance product description. *College Park: University of Maryland Department of Geography*.

Vivone, G., Addesso, P., Conte, R., Longo, M., & Restaino, R. (2014). A class of cloud detection algorithms based on a map-mrf approach in space and time. *IEEE Transactions on Geoscience and Remote Sensing, 52*, 5100-5115.

Wang, L., Yang, R., Tian, Q., Yang, Y., Zhou, Y., Sun, Y., & Mi, X. (2015). Comparative analysis of GF-1 WFV, ZY-3 MUX, and HJ-1 CCD sensor data for grassland monitoring applications. *Remote Sensing, 7*, 2089-2108.

Xia, G., Delon, J., & Gousseau, Y. (2010). Shape-based invariant texture indexing. *International Journal of Computer Vision, 88*, 382-403.

Xie, H., Luo, X., Xu, X., Pan, H., & Tong, X. (2016). Evaluation of Landsat 8 OLI imagery for unsupervised inland water extraction. *International Journal of Remote Sensing, 37*, 1826-1844.

Yang, A., Zhong, B., Lv, W., Wu, S., & Liu, Q. (2015). Cross-calibration of GF-1/WFV over a desert site using Landsat-8/OLI imagery and ZY-3/TLC data. *Remote Sensing, 7*, 10763-10787.

Zeng, C., Shen, H., & Zhang, L. (2013). Recovering missing pixels for Landsat ETM + SLC-off imagery using multi-temporal regression analysis and a regularization method. *Remote Sensing of Environment, 131*, 182-194.

Zhang, Q., & Xiao, C. (2014). Cloud detection of RGB color aerial photographs by progressive refinement scheme. *IEEE Transactions on Geoscience and Remote Sensing, 11*, 7264-7275.

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.

Zhang, Y., Guindon, B., & Li, X. (2014). A robust approach for object-based detection and radiometric characterization of cloud shadow using haze optimized transformation. *IEEE Transactions on Geoscience and Remote Sensing, 52*, 5540-5547.

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: Atmospheres (1984–2012), 109*.

Zhu, Z., Wang, S., & Woodcock, C.E. (2015). Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. *Remote Sensing of Environment, 159*, 269-277.

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. (2014). 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.