Automatic delineation of clouds and their shadows in Landsat and CBERS (HRCC) data

Mostapha Harb*, *Member, IEEE*, Paolo Gamba*, *Fellow,* and Fabio Dell'Acqua*, *Member, IEEE*

***Abstract*— The presence of clouds and their shadows is an obvious problem for maps obtained from multi-spectral images. As a matter of fact, clouds and their shadows create occluded and obscured areas, hence information gaps that need to be filled. The usual approach - pixel substitution - requires first to recognize the cloud/shadow pixels. This work presents a cloud/shadow delineation algorithm, the Cloud/Shadow Delineation Tool (CSDT) designed for Landsat and CBERS medium resolution multi-spectral data. The algorithm uses a set of literature indices, as well as a set of mathematical operations on the spectral bands, in order to enhance the visibility of the cloud/shadow objects.**

The performance of CSDT was tested on a set of scenes from the Landsat and CBERS catalogues. The obtained results showed more accurate and stable performance on Landsat data. In order to validate the proposed approach, this work presents also a comparison with the F-mask algorithm on Landsat scenes. Results show that the F-mask technique tends to overestimate the cloud cover, while CSDT slightly underestimates it. However, accuracy measures show a significantly better performance of the proposed method than the F-mask algorithm in our investigation.

***Index Terms*— Cloud extraction, Landsat imagery, CBERS HRCC imagery, land surface.**

I. OVERVIEW

The study of (historical) land surface dynamics requires monitoring tools that combine critical factors such as the availability of historical data, their easy accessibility, wide geographical coverage, and temporal/spectral/spatial resolution fine enough for the phenomena under investigation. For all these reasons, remote sensing techniques have attracted attention for their potentials, and the Landsat program occupies a unique position in tracking land changes. As a matter of fact, the Landsat archive spans over four decades and is freely accessible. Moreover, Landsat multi-spectral images show a convenient balance between spatial coverage and resolution [1], [2]. On a smaller temporal and geographical scale, but with similar aims, the China-Brazil Earth Resources Satellite (CBERS) program offers free data over Latin America since 2004 [3]. The sensors on board the CBERS satellites, specifically the High Resolution CCD Camera (HRCC), share the same operating bands of Landsat Thematic Mapper (TM) & Enhanced Thematic Mapper Plus (ETM+), with a spatial resolution of 20 m, finer than the 30 m by Landsat. Consequently, the CBERS program has attracted a growing attention for being a potentially complementing data source to Landsat in specific situations [4].

Many articles in technical literature report scientific applications of Landsat and CBERS products to monitor urban growth, crops, geological features, forestries, eco-regions, and global changes [5], [6], [7]. Although each application fields has its own challenges in using the satellite data sets, common sources of hindrance are quite noticeable. For instance, weather effects remain as the main impediment for information extraction from a large set of archived scenes in most of the applications [8], [9], [10].

Specifically, clouds and their shadows affect radiance values collected by multi-spectral sensors at various wavelengths [11], [10]. The problem reduces the accuracy of information extraction in change detection [12], land cover mapping [13], vegetation monitoring [14], and correction of atmospheric effects [15]. In general, the presence of clouds and their corresponding shadows is a cause for information gaps, and leads to confusion among land cover classes, making it more difficult to track their changes [16]. Additionally, even though the Landsat program is characterized by its long-timespan archive and a temporal resolution of 16 days, the frequency of the archived acquisitions varies dramatically with the geographical areas [17]. Even worse, the shorter operational period and the longer temporal resolution (26 days) of CBERS platforms prove to be similarly challenging [18]. For instance, while there are a few areas with acquisitions in close dates, other sites are seldom imaged, and show significant time gaps between acquisitions. Therefore, depending on the application and the test site, the user is occasionally forced to use all the available scenes, including partially cloudy scenes.

Therefore, the delineation of the cloud/shadow objects in a multi-spectral scene is a necessary pre-processing phase for many situations. To this aim, one may exploit physics and note that while cloud object generally are homogeneously bright, their shadows do not show the same level of homogeneity in darkness. The spatial and spectral diversity of the cloud shadows can be attributed to several causes. For instance, cloud density controls the sun light passage and leads to shadowed areas with various degrees of darkness, depending on the cloud opaqueness. Moreover, the reflection property of the shadowed area can also affect the shadow spectral signature creating different levels of shadow dimness

* Telecommunications and Remote Sensing Lab, Department of Electrical, Computer and Biomedical Engineering, University of Pavia, I-27100, Pavia, Italy.

(e.g, cloud shadows over urban and vegetation area show distinct properties). Additionally, there are other features that may help enhancing the visibility of the cloud/shadow objects such as cloud height, morphology, size, temperature, and density [19]. It should be pointed out at this stage that image shadows in multi-spectral images can be from diverse origins: buildings, topography, and clouds [20]. However, shadows from buildings are neglected in medium resolution imagery as the spatial resolution is coarse enough to average out the impact of shadows. Topographic shadow is critical and is an impediment for information extraction in mountainous areas, however it will not be handled in this work, as it is a different issue requiring different, ad-hoc remedies. Finally, cloud shadows is the problem treated in this study as it reduces the information quality in non-mountainous areas which can be critical for applications such as built up extraction or change detection.

II. State of the art

The problem of cloud opacity to multi-spectral radiation discussed in the previous section is tackled in technical literature in two typical ways:

- The first option is to use only cloud-free acquisitions, which limit the investigations only to dry season dates. Moreover, if the study is carried out on an area with scarce acquisitions, the user would end up using a small number of scenes, which in turn would reduce the statistical significance of the analysis on the temporal changes [9].
- The second option is to detect clouds and shadows, because cloudy scenes generally contain a significant amount of cloud-free - and thus useful - pixels [11]. An efficient method to solve the problem of cloud presence is 'pixel substitution' [11]. Recently, this approach has attracted larger attention, for instance being applied within Google Earth Engine, enabling the easy usage of image stacks to substitute a cloudy pixel by another cloud-free pixel taken from a scene in a different date [28].

The technical literature presents a large set of cloud/shadow extraction algorithms for the production of 'cloud-free' Landsat scenes [16, 18, 33, 43]. In most cases, these studies do not differentiate between TM and ETM+ sensors, due to the similarity of their spectral bands and spatial resolutions [43]. One approach that is used very often is the Automated Cloud Cover Assessment (ACCA), which gives the overall percentage of clouds per scene [18]. The cloud cover percentage is used for instance as a criterion in the USGS search engine as an optional filter to select the archived acquisitions. However, the cloud cover percentage is sometimes not sufficient, as the user needs more information on the cloud/shadow locations in order to benefit from the available acquisitions in tracking an object of interest (e.g. vegetation, built-up ...). To this aim, the Landsat program added the Operational Landsat Imager (OLI) sensor on Landsat 8 as an additional Short Wave Infra-Red band (1.36-1.38 μm) sensor to detect the cirrus cloud objects in the scene.

Methods designed for Landsat archived data rely on either automatic or empirical thresholds applied either directly to the spectral bands or indirectly to indices such as the Normalized Difference Snow Index (NDSI), the Haze Optimized Transformation (HOT), the Whiteness [30, 41, 6, 27]. Other approaches use multi-temporal scenes for cloud detection and benefit from the differences between two dates [20, 5, 11], while other techniques rely on classification trees [14]. Zhu et al. proposed the Function of mask (F-mask) routine for the automatic extraction of clouds, shadows and snow layers in [43], recently improved in [1] and [42]. The method is designed to work on Landsat data and relies on scene-based thresholds to establish a probabilistic cloud layer. Additionally, automatic thresholds, connected with physical properties of the clouds, were suggested and adopted after an extensive empirical study. The approach has been under development since 2012, and the latest algorithm version (V. 3.2.1) was issued recently (2015).

Although several methodologies have been designed for Landsat data, only a limited number of them may operate on CBERS HRCC images. This discrepancy is basically due to the wider applicability of Landsat which has a particularly extensive historical archive, and thus permits multi-temporal exchange of information across images. Moreover, CBERS HRCC data sets are acquired only in four wavelength bands (three visible and one near-Infrared). Therefore, modifying a Landsat cloud extraction method to operate on CBERS data is anything but an easy task. Using a different approach, based on spatial, as opposed to spectral, properties of clouds in optical images, Pesaresi et al. presented a method designed for CBERS to delineate clouds from multi-spectral HR/VHR CBERS data. The method relies on spatial features of clouds regional maxima [36, 31]. In line with this idea, [31] also proposes another approach, exploiting morphological gradients to determine cloud boundaries from gradual intensity variations of CBERS panchromatic VHR Data.

According to the analysis of the previous paragraphs, there is a growing need for methodologies implementing cloudy pixel detection in datasets acquired by multiple sensors, capable of exploiting different numbers of bands, as well as both spectral and spatial information. To this aim, in the following section, a novel cloud/shadow extraction processing chain is introduced, labelled 'CSDT' and preliminarily presented in [15]. The method is designed to work on both Landsat (TM & ETM+) and CBERS (HRCC) data. It uses literature indicators to enhance the target-objects visibility followed by an unsupervised classification and an automatic selection of the target classes. The approach introduces a unified processing chain for clouds and shadows delineation.

III. Methodology

The assignment of empirical thresholds and the consequent assumption of their generic nature is common in different scientific fields including image processing. The thresholds are either scene-dependent or constant empirical values determined after analysis on a large set of reference data. However, particularly in image processing, the assumption of generality occasionally fails due to the frequent variations in the data acquisition environment.

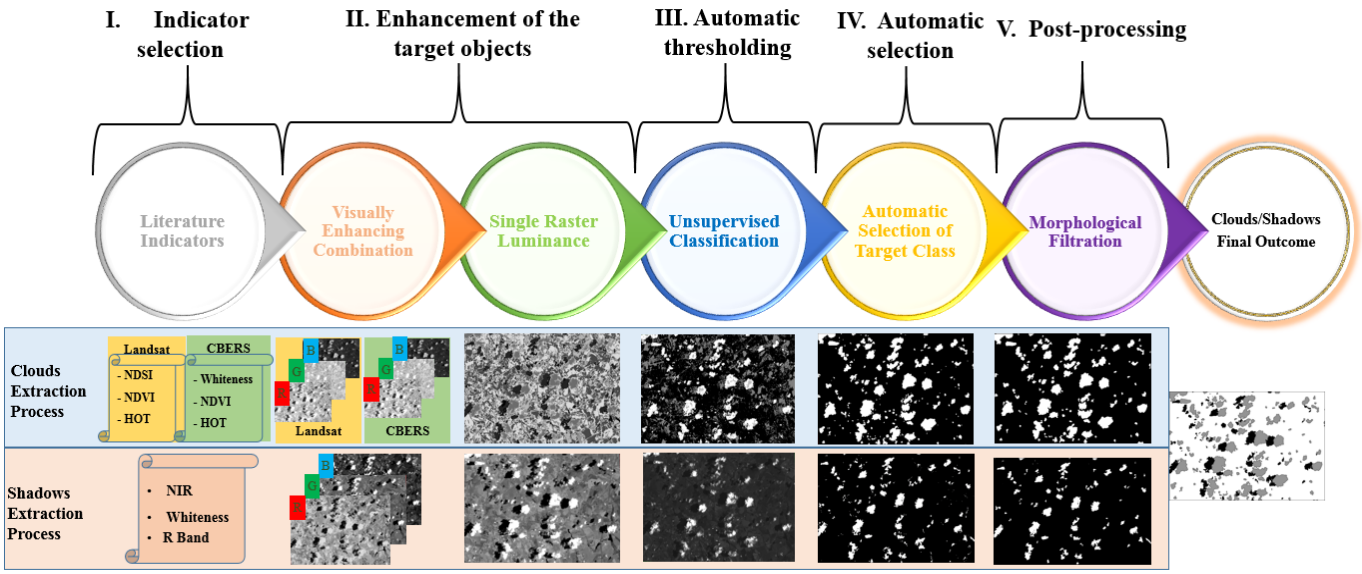
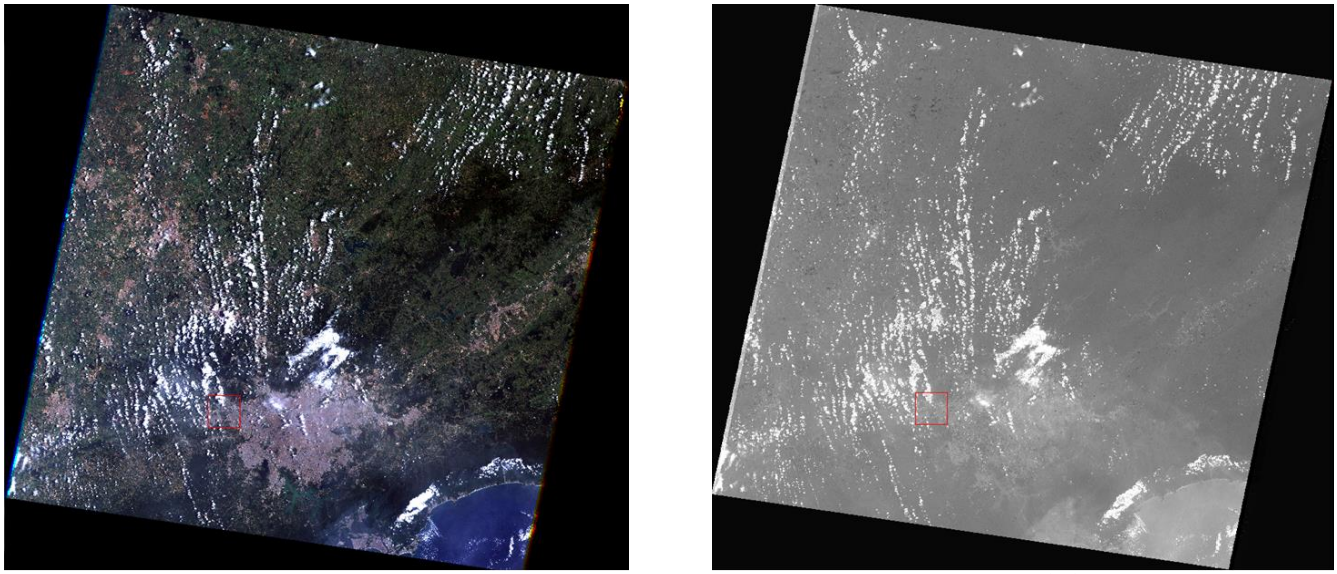


Figure 1. A graphical representation of the processing flow chart of the proposed approach for automatic cloud/shadow.



(a) Original CBERS scene (true colors)

(b) The luminance image useful for cloud extraction

Figure 2. Effect of the combination of the selected set of indicators on enhancing the visibility of cloud objects in a CBERS scene.

The approach proposed in this work implements a processing chain called the Cloud/Shadow Delineation Tool (CSDT), presented in Figure 1. The mechanism relies on several rules that reflect the physical properties of the monitored objects and reduce the environmental change effects. The adopted work-flow can be partitioned into five phases. These steps are described in the following paragraphs, while the selection of the parameters and their value are discussed in a dedicated subsection.

1) The first phase builds on the literature experience in the form of some previously developed indices that are linear transformations of bands based on cloud/shadow physical properties (i.e. haze optimized transformation (HOT) [41], Normalized Difference Vegetation Index (NDVI) [21], normalized difference snow index (NDSI) [6], and Whiteness [6, 10], Near Infra-Red (NIR) [25]). The NDSI indicator was designed for snow detection [13], which shares the high reflectance property of cloud objects, and literature presents some approaches that apply a threshold on NDSI to separate cloudy and cloud-free pixels [6]. The NDVI was adopted, too, as cloudy pixels tend to take anomalous values [2]; NDVI was thus used as an additional clue on the position of cloud objects in the imagery. HOT was developed for detecting haze/cloud locations in Landsat data [41]. The whiteness indicator relies on the property of flat reflectance of some objects (e.g. cloud, dark cloud shadow, sand and snow) in the visible bands. Finally, the NIR is used to highlight cloud shadow objects which show low reflectivity in this band [25].
2) The second phase enhances the spectral properties of

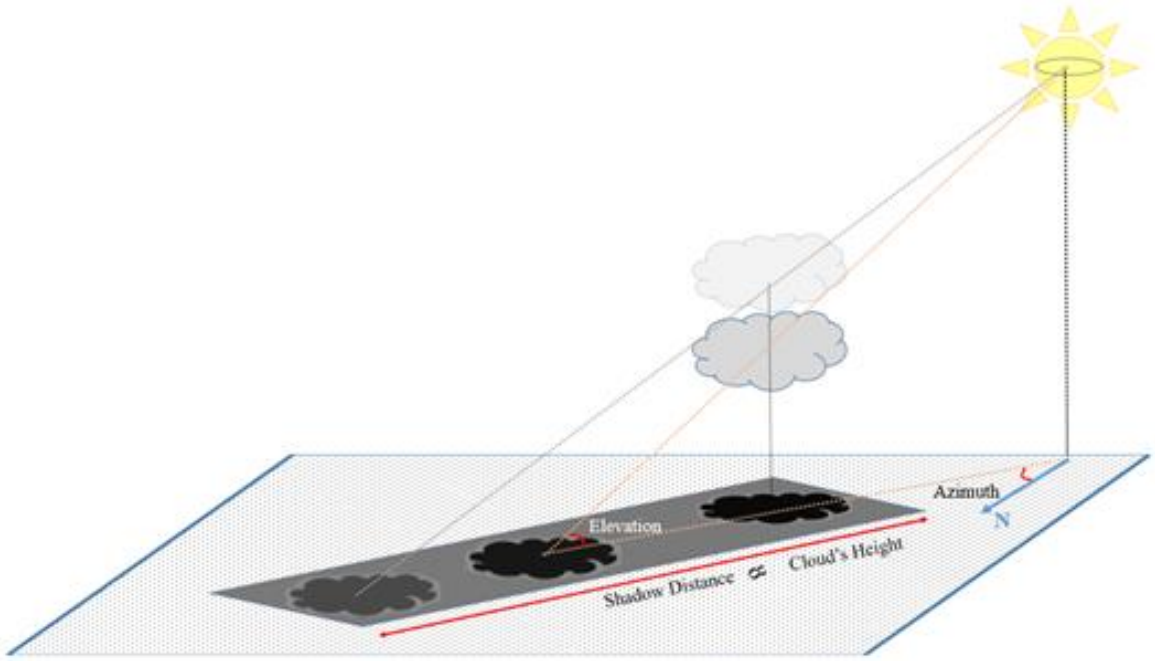


Figure 3. A graphical representation of the cloud/shadow geometry.

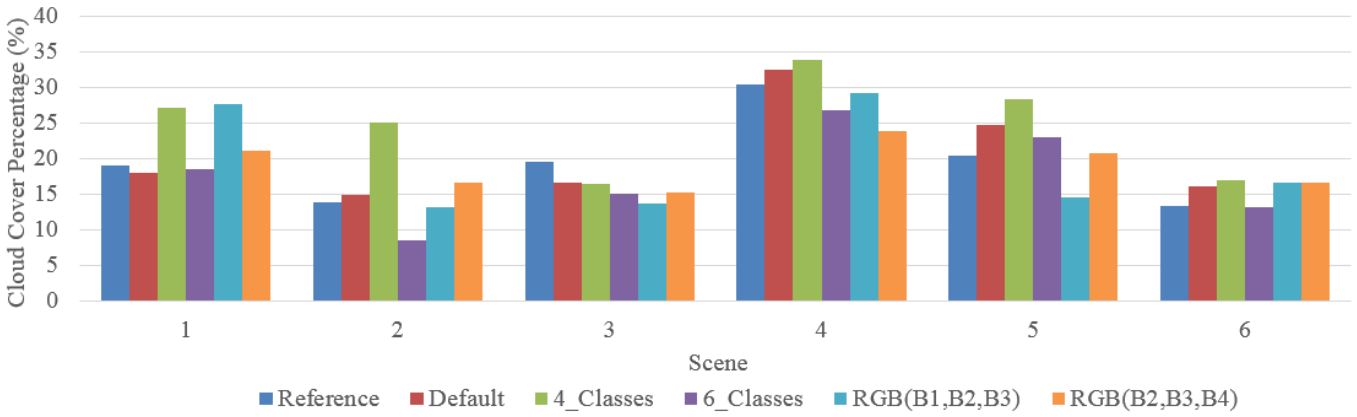


Figure 4. Chart representing quantitative results for performance of CSDT in detecting the cloud cover on 6 LS scenes. Results support selection of default parameters.

the target objects, where the pixels belonging to the two extracted classes will be denoted as C (cloud) and S (shadow) from here on. It starts from selecting three indicators to form a composite image that spotlights the clouds/shadows (I1, 2, 3(c), I1, 2, 3(s)).

	Landsat\CBERS		Landsat	CBERS
$I_1^{(s)}$	NIR	$I_1^{(c)}$	NDSI	Whiteness
$I_2^{(s)}$	Whiteness	$I_2^{(c)}$	NDVI	NDVI
$I_3^{(s)}$	Red	$I_3^{(c)}$	HOT	HOT

Then the bands are collapsed into a relative luminance raster calculated from the linear band components:

$$L = \sqrt{I_1{}^2 + I_2{}^2 + I_3{}^2}$$

The last step is justified by the need to produce a single raster band that simulates the human vision (see Figure 2). The set of used indicators varies according to the type of the traced object (cloud/shadow). The selection of optimal combinations between the indicators was based on two criteria: human vision and extensive empirical tests on the cloud/shadow extraction accuracy from different combinations (see Figure 4). It must be noted that while the group of indicators for cloud shadow extraction is shared between the two data sources, the sets used for cloud extraction are slightly different. The diversity is due to the lower number of available bands for CBERS HRCC data that prevents usage of the NDSI indicator, which was replaced by the whiteness indicator. Thus, indicators were selected considering the availability of

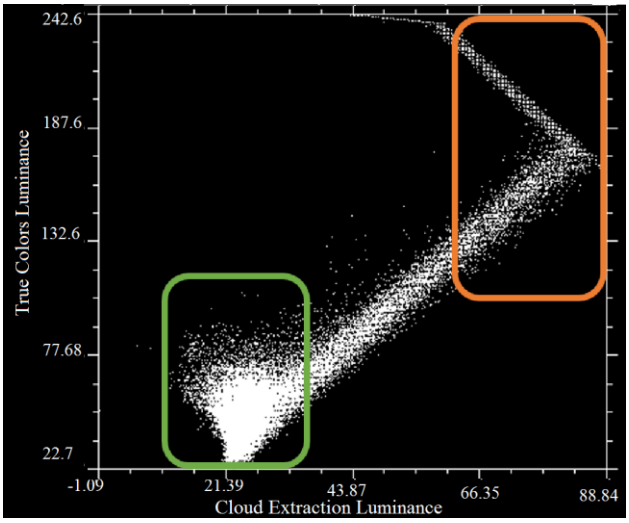


Figure 5. Scattergram showing two concentration centers in the combination used for cloud extraction on Landsat data

bands for Landsat and CBERS HRCC. Additionally, it is important to note that clouds and snow are partially mixed with this approach, although this issue is usually solved by exploiting the NDSI index [12].

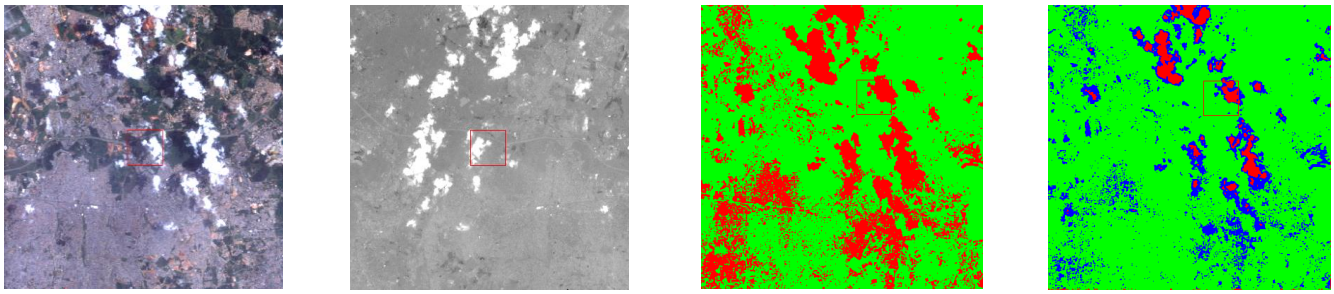
3) The third phase aims at clustering the obtained luminance image, using the unsupervised K-means approach, and automatically determine the thresholds of separation among an adjustable number of classes: the default selection is 5 and 8 classes for clouds and shadows extraction respectively. Therefore, the direct empirical thresholds on bands, usual in many approaches based on indicators, is substituted by an automatic threshold determined by the K-means clustering algorithm. The main assumption here is that the cloud/shadow objects are going to be concentrated in a single cluster. Hence, the adopted default numbers of classes for the cloud/shadow extraction process was selected after an experimental study on a large reference data set for the clustering efficiency of the target pixels. The selection of 5 classes for cloud detection was sufficient as clouds are often significantly brighter in the luminance composite image than other objects in the scene. Brightness in itself is a property shared by clouds with some mountainous and urban areas, but the usage of the indicators allows effective discrimination of clouds against these other objects. Instead, shadow extraction requires 8 classes. The main reason is cloud shadows share their main property of low reflection with several land covers (e.g. water, some vegetation areas, wet bare soil), but they are usually darker and can be discriminated using a large enough set of classes, as more accurately discussed later on.
4) The fourth step is automatic selection of the target class based on statistical comparison among the classes of the previous step. As previously discussed, the assumption is that target classes, clouds and shadows, occupy respectively the highest and lowest modal values in the scene histograms. As a result, the selection of the class with the highest modal value after the 5-class clustering and the lowest modal class in the 8-class clustering enables discriminating the target classes from the rest of the scene.
5) Finally, a post-processing step is applied to the extracted cloud and shadow masks, by means of mathematical morphology and cloud/shadow matching. Morphological filtering is applied to the processed image to overcome some limitations of the pixel-based approach and leads to a large improvement in the accuracy of the results. Two different sequences of filters are applied to the binary clouds/shadows layers obtained from the previous step:
   1a) a closing operation on the C/S layers with a square structuring element with width W_1
   1b) an opening operation on the C/S layers with a square structuring element with width W_2
   2a) a dilation operation on the C/S layers with a square structuring element with width W_3
   2b) a closing on the negative C/S layers with a square structuring element with width W_4
   2c) an exclusion operation on the C/S layers for objects area smaller than a threshold A_1

   The first sequence is meant to remove salt-and-pepper classification noise through two closing operations, with the aim to discard small unclassified parts within the extracted cloud/shadow objects. Similarly, an opening operation is used to remove extracted objects that are too small to be considered either as a cloud or as its associated shadow. Then, a dilation filter is used to expand the shapes of the extracted objects in order to get a smoother version of the originally extracted irregular boundaries.

   The second sequence starts with a so called “fill-holes operation”, which is basically an inside dilation to recover the remaining holes by invading them from the outer boundaries of each detected object. Finally, the last step is the exclusion of small objects in order to discard unwanted yet detected tiny objects (e.g., small urban areas), where "small" is here defined through a 4-pixel min size threshold, assumed as the lower limit for cloud/shadow size and adopted to reduce false alarms.

The second part of the post-processing procedure aims at retaining only shadows that are due to clouds. Since the low reflectance of cloud-shadowed areas varies with land surface characteristics, the previous clustering and cluster selection procedure is not always successful in isolating cloud shadows.

Therefore, the approach introduced in this work exploits cloud positions in the extracted cloud layer to select cloud shadows. The geometrical position of the sun, the scene location, the cloud height and size are the parameters determining the location and size of a cloud-shadow probability window. The main assumption made in this step is that the cloud shadow will cover equal or smaller area on the Earth surface than its associated cloud. After a labeling step



(a) (b) (c) (d)

Figure 6. The effect of applying the calculated luminance for cloud extraction (b) on a true colors scene (a). (c) and (d) Show the perturbation of the cloud extraction performance due to changes on the number of classes from the default set of parameter.

Figure 7. Global distribution of the Landsat (red) and CBERS (yellow) scenes used for the evaluation in Table I.

aimed at recognizing single clouds, the most probable location of the terrain can be extracted, as the above mentioned parameters are known. However, in addition to inherent inaccuracies in some of the parameters (e.g., cloud height), there are several factors which affect cloud shadow quality and its position on the Earth surface relative to the corresponding cloud. These factors may alter the shadow visibility which fades naturally on large distances due to the scattering effect of light. For instance, shadow dissipation is increased with characteristics such as small size, low density, and low sun elevation angles. Thus, due to the unstable properties of cloud shadows, the implemented approach depends on an adjustable distance parameter which determines the maximum distance for detectable cloud shadows in the direction determined by the sun azimuth (a graphical representation of the geometrical model is shown in Figure 3). The default value for this distance was determined to be 50 pixels, which, at the spatial resolution of the sensors considered in this work, approximately corresponds to lower clouds (nimbostratus, cumulus and low cumulonimbus, ~ 2 km height), as these clouds tends to be thick, large, more reflective, and create remarkably darkened shadows. Middle- and high-altitude clouds instead tend to create brighter shadows and cannot be detected by the developed shadow extraction method. However, these clouds tend to be partially transparent either because of their relatively low density (e.g, stratus, stratocumulus, altostratus, cirrostratus) or because they appear as a set of small distinct cloud objects (e.g, altocumulus, cirrus, cirrocumulus) [41].

A. Default parameters selection and robustness analysis

In this subsection, the sensitivity and robustness of the three main parameters of the proposed procedure, i.e.

- the set of indicators for the proposed work-flow;
- the cluster numbers;
- the morphological operations parameters: W_1; W_2; W_3; W_4; and A_1

About the first point, it must be noted that the default

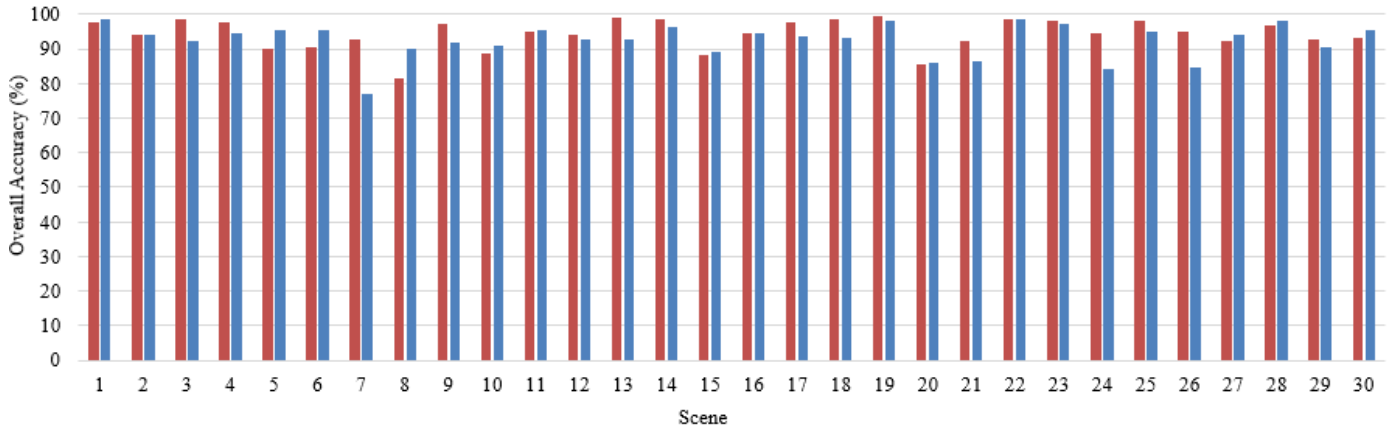


Figure 8. Quantitative comparison between overall accuracy values for CSDT (red) and F-mask (blue) on 30 Landsat test scenes.

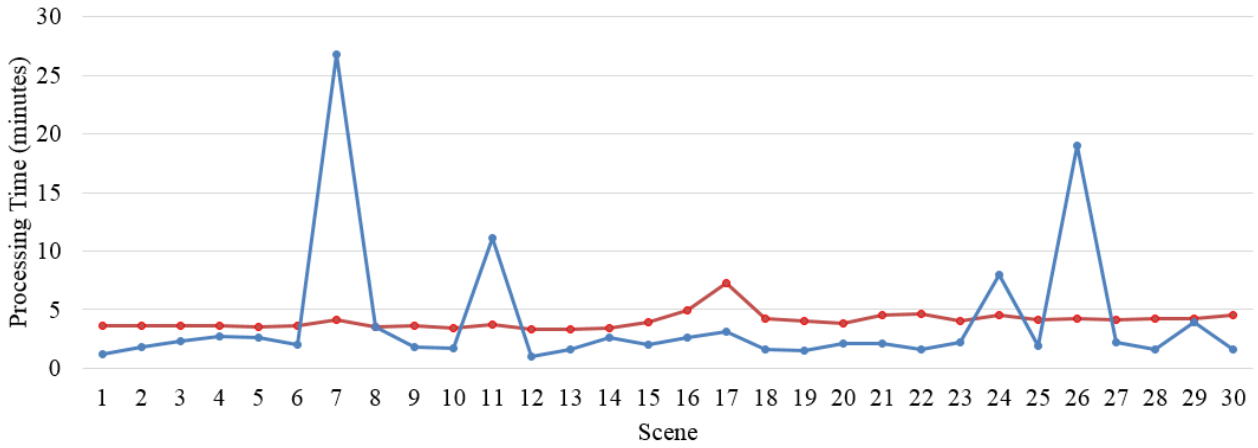


Figure 9. Quantitative comparison between processing time for CSDT (red) and F-mask (blue).

selection of the representative indicators in the extraction process of both clouds and their shadows considered two critical concepts: the spatial concentration of the target pixels in the reflection spectrum and the preservation of cloud/shadow ranking between the true color image (see Figure 6a), which suits the humans vision, and the cloud/shadow extraction luminance (see Figure 6b). While the former concept tends to concentrate the target classes in narrow ranges on the reflection spectrum (see Figure 5), the latter preserves the rational reflection hierarchy, where lowest to highest intensity values correspond to the darkest shadows, shadows with topological or opaque cloud origin, and cloud object pixels respectively. The retained hierarchy facilitates the automatic selection of the right target class through the following clustering step. The default selection of the set of bands (combinations) for the different work-flows were based on empirical tests following visual evaluation rules as just mentioned. The indicator sets with the highest accuracy score considering these criteria was eventually selected.

Moving to the second set of parameters, i.e. ~ the cluster number, it is important to note that the selection of the indicator set does not provide completely separable classes, but a gradual density variation spread on the reflection spectrum. This is why a clustering step is needed. Thanks to the efficiency of the chosen combinations in concentrating the target classes in low/high brightness parts of the spectrum, the clustering step is able to detect the target cover, although minor percentages of pixels may be missed. However, the excluded pixels are retrieved by applying the post-processing morphological operations.

For the clustering procedure, the adopted default number of classes equals 5 and 8 for clouds and shadows, respectively. Specifically, the selected bands for the cloud extraction workflow give far enough reflection centers. The example in Figure 5 shows that one of them refers to the background (vegetation, bare soil, urban areas, shadows, water), and the second center highlights the clouds. Accordingly, a selection of three classes, added to the target and the scene-margin class, is sufficient to include the reflectance variation from the other land covers. Instead, the set of bands for cloud shadows extraction is not able to create a separate center for the darkest cluster composed by cloud and topological shadows, even though the class always occupies the range at the low-value-tail of the reflection distribution. Therefore, to reduce the classifier confusion, it is necessary to increase the cluster number. In this case six classes, in addition to the target and

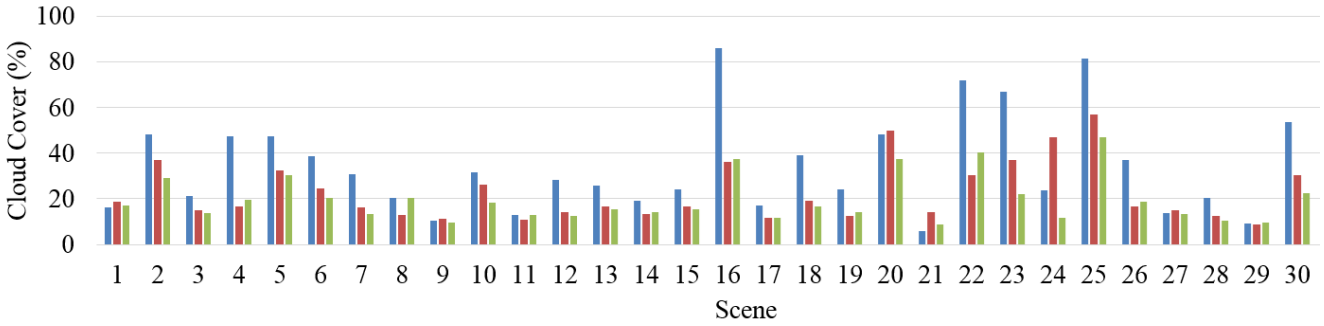


Figure 10. Quantitative comparison of cloud cover percentages in the scenes as detected by CSDT (red) and F-mask (blue), against the percentage in the reference layers (green).

Table I. Cloud extraction accuracy comparison between Landsat and CBERS data.

Landsat Data			**CBERS HRCC Data**		
Date	**Overall Accuracy**	**Kappa**	**Date**	**Overall Accuracy**	**Kappa**
29-11-1990	94.11	0.88	18-11-2003	84.11	0.68
29-08-1994	96.1	0.92	26-09-2004	98.05	0.96
27-1-1998	94.86	0.90	26-10-2004	78.98	0.58
26-2-2000	90.87	0.82	28-10-2004	82.15	0.64
14-3-2009	94.26	0.88	13-12-2004	80.64	0.61
2-10-2011	92.62	0.85	10-01-2006	99.25	0.98
29-7-2014	88.82	0.78	03-03-2006	80.11	0.60
6-10-2014	92.87	0.85	27-9-2006	97.23	0.94
17-9-2014	98.62	0.97	07-11-2006	99.1	0.98
7-10-2014	97.67	0.95	18-11-2006	86.71	0.73
27-10-2014	93.31	0.86	21-3-2007	86.24	0.72
22-2-2015	94.66	0.89	21-9-2008	78.6	0.56
Average	**94.06**	**0.88**	**Average**	**87.6**	**0.75**
Standard Deviation	**2.71**	**0.05**	**Standard Deviation**	**8.39**	**0.17**

the scene-margin classes, are required to discriminate the cloud and topological shadowed areas.

Eventually, as for the third group of morphological operations default parameters, a 3x3 structuring square element with connectivity equal to one was used as default in all cases except W_4 (i.e, $W_1 = W_2 = W_3 = 3$ pixels). The element size was chosen to smooth the outer boundary of the detected objects.

The default element on connectivity was used to ensure the closure of the objects' holes. Implementing the operation by means of a closing on the negative C/S layers, W_4 must be set to a large enough value to discard the biggest "holes" in the clouds. By default, $W_4 = 50$ was selected.

The A_1 threshold was set to 4 pixels, to correct uncertainties in detecting objects smaller than the structuring element of the morphological operations. The adoption of a larger values may lead to the omission of true detected objects, which may reduce the detection efficiency of small clouds (e.g, altocumulus, cirrus, cirrocumulus) and their associated shadows.

Finally, as a general statement, it is useful to understand that fine tuning of the default parameters, be it the set of indicators or the number of classes, is possible and would not always worsen the extraction results. However, according to our extensive tests on the reference data sets, the overall accuracy might be remarkably reduced. For instance, on the one hand increasing the number of classes would increase the size of the missed cloud pixels inside the cloud body, and would make it difficult to recover the created gaps using the post-processing morphological filters, as shown in Figure 6d. On the other hand, decreasing the number of classes would extend the cloud to include pixels from bright urban, sandy, or rocky areas, as shown in Figure 6c.

III. EXPERIMENTAL RESULTS

In order to evaluate cloud extraction performance, a reference layer of cloudy pixels for each test scene is required. A simple way to get it would be to extract the cloud cover manually. This procedure, however, is time- and effort-expensive and the products often include inherent subjectivity due to fluctuation of visual interpretation criteria within an image even for the same operator. Thus, in this work the manual step was replaced by a semi-automatic process. The procedure exploits machine learning techniques, in particular the supervised classification algorithm 'LIBSVM', to create the reference layers. To obtain more accurate reference layers, automatic results were also visually inspected and adjusted to reduce alarms represented by bright areas, water bodies, and topological shadows.

Once the reference layer is available, two different accuracy evaluation methods can be considered:

1) the first uses the full scene pixels (7000 * 8000 and 6000 * 6500 pixels for Landsat and CBERS scenes, respectively);
2) the second uses instead stratified sampling, and the reference is built by randomly selecting an equal set of sample pixels for each class, a set of 10000 pixels (5000 for each class) was used.

The main difference between the two approaches is the proportion of sample sets selected to compute the accuracy of each class. While accuracy in the full-area approach is calculated taking into account every single pixel in the scene, accuracy in the stratified sampling approach is measured according to an equal-size set, randomly selected from the scene. The literature highlights a crucial limitation of the full-area accuracy if applied to multiple scenes [42]. The intrinsic constraint is due to the use of all the pixels while computing the class accuracy in scenes with very different cloudy pixel percentages. This bias leads to a lack in consistency of the accuracy assessment, thus, the stratified sampling approach was adopted in the following analysis.

A. CSDT Performance on Landsat and CBERS data

In this subsection, the performance of the CSDT was compared on a set of 22 Landsat TM and CBERS HRCC data sets. The used scenes were as much as possible geographically distributed, on a global basis for Landsat, and mainly over Latin America for CBERS, because of its geographically limited accessibility. The scenes were selected considering diversity in type, size, and density of clouds (see Figure 7). A stratified random sampling was used to have a significant comparison across scenes and between Landsat and CBERS data. Results reflect the method efficiency in outlining cloud objects with a satisfying differentiation between clouds and built-up areas. In addition, results obtained on both data sets are excellent and comparable. Results using Landsat scenes were relatively better by an average of 6.5 % and 13.0 % for the overall accuracy and Kappa values respectively (see Table I).

In addition to the previous comment, Table I reveals a higher variability of the accuracy measures for CBERS data with respect to the values obtained from Landsat. This variability is quantified by means of the standard deviation of the numbers shown in Table I, whose value for CBERS is approximately 3 times larger than for Landsat. The obtained results show higher stability and better accuracy performance of CSDT applied to Landsat as opposed to CBERS data. The lower stability obtained on the CBERS data can be attributed to several factors. Primarily, Landsat data include a larger set of bands, allowing the computation of more representative indices (e.g. NDSI). Moreover, the difficulty to carry out radiometric calibration of CBERS scenes, using top of atmosphere reflectances instead of digital numbers, make the procedure less robust to atmospheric distortions. Lastly, CBERS data exhibit a larger sensitivity to variations in the data acquisition environment.

B. Comparison between the CSDT and F-mask approaches on a large set of Landsat scenes

Since F-mask [43] is one of the most valuable and used techniques for cloud layer extraction from Landsat data, it makes sense to compare the results of CSDT and F-mask on a relatively large data set of Landsat images. The chart in Figure 8 displays accuracy values for the two compared cloud extraction methods applied to 30 scenes from Landsat archives. Specifically, the results of the cloudy and cloud-free pixels reveal different patterns for the CSDT and the F-mask methods, respectively. These patterns are due to the tendency of each method to favor one of the two classes against the other.

According to the results, F-mask clearly adopts a more conservative criterion in extracting cloudy pixels. Even though the option is justified by the need to delineate semi-transparent clouds, however, this would also include parts of other bright areas (e.g., urban areas and rocky mountains). The CSDT method, instead, slightly underestimates the cloud cover. It therefore reduces errors in estimating cloud-free pixels such as the above mentioned bright urban areas, but it also excludes parts of the semi-transparent clouds. In conclusion, the tested techniques embrace opposite perspectives in the imperative trade-off to correctly extract either cloudy or cloud-free pixels in a scene.

In general, both techniques show very similar overall accuracy and Kappa values (see Table II). A slightly lower

Table II. CSDT and F-mask accuracy comparison.

	Kappa Value		Overall Accuracy	
	CSDT	**F-mask**	**CSDT**	**F-mask**
Min	0.63	0.54	81.7	77
Max	0.99	0.98	99.3	98.9
Average	0.89	0.85	94.4	92.6
St. Dev.	0.09	0.1	4	5

variance on the accuracy measures along the 30 images test set can be noticed for the CSDT method.

With respect to the computation time required by the two approaches, Figure 9 reports these numbers for the test scenes. Although the average processing times for the two techniques on the 30 scene set is similar (~ 4 minutes), F-mask is faster for most of the scenes, with the notable exception of sparsely cloudy scenes. The average processing time required by CSDT is slightly longer, but approximately alike among all scenes. It must be noted, however, that the time charts shown in Figure 9 corresponds to the processing times required by F-mask to extract clouds, shadows and snow covers, while the time for CSDT is just the time to delineate either the cloud or the shadow layer. Therefore, F-mask has a clear advantage with respect to the implementation time.

Finally, Figure 10 displays the percentage of the final cloud covers, as obtained by the two techniques and compared with the reference layer. The chart confirms the previously presented analysis. The cloud cover percentages by F-mask tend to be overestimated and the cloud layers by the CSDT method look closer to the reference covers, but with a notable tendency to underestimation.

IV. Conclusion

This paper presents a novel approach for an automatic cloud/shadow extraction from Landsat TM/ ETM+ and CBERS HRCC multi-spectral scenes. It defines a unified processing chain, called CSDT, to extract both clouds and cloud shadows layers. The conclusions of this research can be summarized into the following points:

- the flexibility of the CSDT processing chain enables its application on data from different sensors, with different spectral and spatial resolution and a minimum need for user intervention;
- the performance of the CSDT approach was tested on a set of Landsat and CBERS data scenes, and the results reveals a slightly more stable performance on Landsat data.
- as a validation, the study compares the method performance with the widely used F-mask algorithm, designed for Landsat data. The obtained results show different trends of overestimation and underestimation of the cloud cover by the F-mask and the CSDT methods, respectively. However, the final accuracy measures which consider the extraction ability of cloudy and cloud-free pixels shows a slightly better performance of CSDT.
- the differentiation of cloud objects in urban scenes is an achievement of the proposed method with respect to other comparable approaches such as F-mask.
- finally, the developed method revealed excellent performance, is open and has the potential for further refinements. Therefore, the margin of improvements to overcome the constraints would be an interesting investigation area for a future related work.

The main limitations of CSDT can be summarized into two main points:

- first of all, the process lacks the ability to extract all semi-transparent clouds, and in turn this leads to a slight underestimation of the cloud cover.
- Additionally, another drawback is the confusion between snow and cloud objects, as the adopted criteria of clouds distinction rely on the property of high brightness, shared between cloud and snow layers.

ACKNOWLEDGMENT

Funding of this work by the EU through the FP7 project 'TOLOMEO' (call: FP7-PEOPLE-2010-IRSES, project ID: 269115) is gratefully acknowledged. The author would also like to thank Prof. Feitosa at PUC-RIO for his support during the visits by M. Harb and P. Gamba to Brazil In the Autumn 2014.

References

[1] A. Lewis and T. Malthus, "Data continuity and new opportunities for land monitoring," in *Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International.* IEEE, 2013, pp. 3285–3288.

[2] H. Bagan, W. Takeuchi, T. Kinoshita, Y. Bao, and Y. Yamagata, "Land cover classification and change analysis in the Horqin Sandy Land from 1975 to 2007," in *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 3, no. 2, pp. 168–177, 2010.

[3] CBERS Data Policy. National Institute For Space Research (INPE).Online. Available at: http://www.inpe.br. [Accessed: 10- Feb- 2016].

[4] J. Choi, S. Khanal, and S. Ambinakudige, "A comparative analysis of CBERS-2 CCD and LANDSAT-TM satellite images in vegetation mapping," Revista Brasileira de Cartografia, no. 63/1, pp. 115–122, 2011.

[5] E. H. Wilson and S. A. Sader, "Detection of forest harvest type using multiple dates of Landsat TM imagery," in *Remote Sensing of Environment*, vol. 80, no. 3, pp. 385–396, 2002.

[6] X. Jinyong, Z. Zengxiang, W. Changyou, Z. Xiaoli, L. Bing, and Y. Ling, "Urban expansion monitoring and driving forces analysis: a case study of Jiangsu Province, China," in *Urban Remote Sensing Event, 2009 Joint.* IEEE, 2009, pp. 1–6.

[7] D. G. Hadjimitsis, A. Agapiou, K. Themistocleous, D. D. Alexakis, and A. Sarris, "Remote sensing for archaeological applications: Management, documentation and monitoring," Remote Sensing of Environment-Integrated Approaches, Hadjimitsis, DG (Ed.). InTech Publisher, pp. 57–95, 2013.

[8] F. Li, D. L. Jupp, and M. Thankappan, "Issues in the application of digital surface model data to correct the terrain illumination effects in Landsat images," International Journal of Digital Earth, no. ahead-ofprint, pp. 1–23, 2013.

[9] G. Mancino, A. Nolè, F. Ripullone, and A. Ferrara, "Landsat TM imagery and NDVI differencing to detect vegetation change: assessing natural forest expansion in Basilicata, southern Italy," *iForest-Biogeosciences and Forestry*, vol. 7, no. 2, p. 75, 2014.

[10] Y. Zhang and N. Kerle, "Satellite remote sensing for near-real time data collection," Geospatial information technology for emergency response, vol. 6, pp. 75–94. London: Taylor & Francis, 2008.

[11] Z. Zhu and C. E. Woodcock, "Object-based cloud and cloud shadow detection in Landsat imagery," *Remote Sensing of Environment*, vol. 118, pp. 83–94, 2012.

[12] Z. Zhu and C. E. Woodcock, "Automated cloud, cloud shadow, and snow detection in multitemporal landsat data: An algorithm designed specifically for monitoring land cover change," *Remote Sensing of Environment*, vol. 152, pp. 217–234, 2014.

[13] Y. Zhang, B. Guindon, and J. Cihlar, "An image transform to characterize and compensate for spatial variations in thin cloud contamination of Landsat images," *Remote Sensing of Environment*, vol. 82, no. 2, pp. 173–187, 2002.

[14] A. Huete, K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira, "Overview of the radiometric and biophysical performance of the MODIS vegetation indices," *Remote sensing of environment*, vol. 83, no. 1, pp. 195–213, 2002.

[15] E. F. Vermote, N. Z. El Saleous, and C. O. Justice, "Atmospheric correction of MODIS data in the visible to middle infrared: first results," *Remote Sensing of Environment*, vol. 83, no. 1, pp. 97–111, 2002.

[16] N. R. Goodwin, L. J. Collett, R. J. Denham, N. Flood, and D. Tindall, "Cloud and cloud shadow screening across Queensland, Australia: An automated method for Landsat TM/ETM+ time series," *Remote Sensing of Environment*, vol. 134, pp. 50–65, 2013.

[17] L. G. Olmanson, M. E. Bauer, and P. L. Brezonik, "A 20-year Landsat water clarity census of Minnesota's 10,000 lakes," *Remote Sensing of Environment*, vol. 112, no. 11, pp. 4086–4097, 2008.

[18] Q. Liu and G. Liu, "Using tasseled cap transformation of CBERS-02 images to detect dieback or dead Robinia pseudoacacia plantation," in *Image and Signal Processing*, 2009. CISP'09. 2nd International Congress on. IEEE, 2009, pp. 1–5.

[19] S. Platnick, M. D. King, S. A. Ackerman, W. P. Menzel, B. A. Baum, J. C. Riédi, and R. A. Frey, "The MODIS cloud products: Algorithms and examples from terra," *Geoscience and Remote Sensing*, IEEE Transactions on, vol. 41, no. 2, pp. 459–473, 2003.

[20] A. Shahtahmassebi, N. Yang, K. Wang, N. Moore, and Z. Shen, "Review of shadow detection and de-shadowing methods in remote sensing," *Chinese geographical science*, vol. 23, no. 4, pp. 403–420, 2013.

[21] T. Danaher, P. Scarth, J. Armston, L. Collett, J. Kitchen, and S. Gillingham, "Remote sensing of tree-grass systems: The eastern Australian woodlands," pp. 175–194, 2010.

[22] D. McInerney and P. Kempeneers, "Conclusions and future outlook," in *Open Source Geospatial Tools*. Springer, 2015, pp. 331–335.

[23] C. Huang, N. Thomas, S. N. Goward, J. G. Masek, Z. Zhu, J. R. Townshend, and J. E. Vogelmann, "Automated masking of cloud and cloud shadow for forest change analysis using Landsat images," *International Journal of Remote Sensing*, vol. 31, no. 20, pp. 5449–5464, 2010.

[24] R. R. Irish, J. L. Barker, S. N. Goward, and T. Arvidson, "Characterization of the Landsat-7 ETM+ automated cloud-cover assessment (ACCA) algorithm," *Photogrammetric Engineering & Remote Sensing*, vol. 72, no. 10, pp. 1179–1188, 2006.

[25] D. P. Roy, J. Ju, K. Kline, P. L. Scaramuzza, V. Kovalskyy, M. Hansen, T. R. Loveland, E. Vermote, and C. Zhang, "Web-enabled Landsat data (WELD): Landsat ETM+ composited mosaics of the conterminous United States," *Remote Sensing of Environment*, vol. 114, no. 1, pp. 35–49, 2010.

[26] L. Oreopoulos, M. J. Wilson, and T. Várnai, "Implementation on Landsat data of a simple cloud-mask algorithm developed for MODIS land bands," *Geoscience and Remote Sensing Letters*, IEEE, vol. 8, no. 4, pp. 597–601, 2011.

[27] H. Choi and R. Bindschadler, "Cloud detection in Landsat imagery of ice sheets using shadow matching technique and automatic normalized difference snow index threshold value decision," *Remote Sensing of Environment*, vol. 91, no. 2, pp. 237–242, 2004.

[28] S. Martinuzzi, W. A. Gould, and O. M. R. González, "Creating cloudfree Landsat ETM+ data sets in tropical landscapes: cloud and cloudshadow removal," US Department of Agriculture, Forest Service, International Institute of Tropical Forestry Río Piedras, Puerto Rico, Tech. Rep. General Technical Report IITF-GTR-32, 2007.

[29] R. E. Kennedy, W. B. Cohen, and T. A. Schroeder, "Trajectory-based change detection for automated characterization of forest disturbance dynamics," *Remote Sensing of Environment*, vol. 110, no. 3, pp. 370– 386, 2007.

[30] W. Bin, K. Muramatsu, and N. Fujiwara, "Automated detection and removal of clouds and their shadows from Landsat TM images," *IEICE Transactions on Information and Systems*, vol. 82, no. 2, pp. 453– 460, 1999.

[31] M. C. Hansen, D. P. Roy, E. Lindquist, B. Adusei, C. O. Justice, and A. Altstatt, "A method for integrating MODIS and Landsat data for systematic monitoring of forest cover and change in the Congo Basin," *Remote Sensing of Environment*, vol. 112, no. 5, pp. 2495–2513, 2008.

[32] Z. Zhu, S. Wang, and C. E. Woodcock, "Improvement and expansion of the F-mask algorithm: cloud, cloud shadow, and snow detection for landsats 4–7, 8, and sentinel 2 images," Remote Sensing of Environment, vol. 159, pp. 269–277, 2015.

[33] P. Soille, Morphological image analysis: principles and applications. Springer-Verlag New York, Inc., 2003.

[34] M. Pesaresi, G. Huadong, X. Blaes, D. Ehrlich, S. Ferri, L. Gueguen, M. Halkia, M. Kauffmann, T. Kemper, L. Lu et al., "A global human settlement layer from optical HR/VHR RS data: concept and first results," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 6, no. 5, pp. 2102–2131, 2013.

[35] M. Harb, D. De Vecchi, P. Gamba, F. DellAcqua, and F. Raul, "Automatic clouds shadows extraction method from CBERS 2 CCD and Landsat data," in International Geoscience and Remote Sensing Symposium (IGARSS), July 2015, pp. 4594–4597.

[36] F. Kriegler, W. Malila, R. Nalepka, and W. Richardson, "Preprocessing transformations and their effects on multispectral recognition," in *Remote Sensing of Environment*, VI, vol. 1, 1969, p. 97.

[37] L. Gomez-Chova, G. Camps-Valls, J. Calpe-Maravilla, L. Guanter, and J. Moreno, "Cloud-screening algorithm for ENVISAT/MERIS multispectral images," *IEEE Transactions on Geoscience and Remote Sensing*, , vol. 45, no. 12, pp. 4105–4118, 2007.

[38] Y. Luo, A. P. Trishchenko, and K. V. Khlopenkov, "Developing clearsky, cloud and cloud shadow mask for producing clear-sky composites at 50-meter spatial resolution for the seven modis land bands over Canada and north america," *Remote Sensing of Environment*, vol. 112, no. 12, pp. 4167–4185, 2008.

[39] D. K. Hall, G. A. Riggs, and V. V. Salomonson, "Development of methods for mapping global snow cover using moderate resolution imaging spectroradiometer data," *Remote sensing of Environment*, vol. 54, no. 2, pp. 127–140, 1995.

[40] Landsat Calibration Records. Landsat Missions. Online. Available:http://landsat.usgs.gov/science-calibration.php. [Accessed: 10- Feb- 2016].

[41] R. Scorer and H. Wexler, "A colour guide to clouds," *Elsevier*, 2013, vol. 1.

[42] R. G. Congalton and K. Green, "Assessing the accuracy of remotely sensed data: principles and practices," CRC press, 2008.

[43] L. Yu, S. B. Ball, C. E. Blinn, K. Moeltner, S. Peery, V. A. Thomas, and R. H. Wynne, "Cloud-sourcing: Using an online labor force to detect clouds and cloud shadows in landsat images," *Remote Sensing*, vol. 7, no. 3, pp. 2334–2351, 2015.

Mostapha Harb (M'12). He received the B.S. in physics from the Lebanese University in 2004, and a M.S. degrees in Radiation Physics from the University of Damascus 2009. From 2007 to 2011, he was a Researcher Assistant at the Lebanese Atomic Energy Commission (LAEC).
In 2012, Mr. Harb received a Master degree in Risk and Emergency Management from the Istituto Universitario di Studi Superiori (IUSS) in Pavia. Since 2012 he is a Ph.D. Candidate in the Department of Electrical, Computer and Biomedical Engineering, University of Pavia. His recent work is mainly concentrated on developing automatic methodologies for extracting risk-related information at different scales from Space borne data, mainly multispectral data.

Paolo Gamba (SM'00, F'13) is Associate Professor of Telecommunications at the University of Pavia, Italy, where he also leads the Telecommunications and Remote Sensing Laboratory. He received the Laurea degree in Electronic Engineering "cum laude" from the University of Pavia, Italy, in 1989, and the Ph.D. in Electronic Engineering from the same University in 1993.

He is a Fellow of IEEE. He served as Editor-in-Chief of the IEEE Geoscience and Remote Sensing Letters from 2009 to 2013, and as Chair of the Data Fusion Committee of the IEEE Geoscience and Remote Sensing Society from October 2005 to May 2009. Currently, he is the Chair of the Chapters' Committee of the same Society.

He has been the organizer and Technical Chair of the biennial GRSS/ISPRS Joint Workshops on "Remote Sensing and Data Fusion over Urban Areas" since 2001. He also served as Technical Co-Chair of the 2010 IEEE Geoscience and Remote Sensing Symposium, Honolulu, Hawaii, July 2010, and Technical Co-Chair of the 2015 IEEE Geoscience and Remote Sensing Symposium, in Milan, Italy.

He has been the Guest Editor of special issues of IEEE Transactions on Geoscience and Remote Sensing, IEEE Journal of Selected Topics in Remote Sensing Applications, ISPRS Journal of Photogrammetry and Remote Sensing, International Journal of Information Fusion and Pattern Recognition Letters on the topics of Urban Remote Sensing, Remote Sensing for Disaster Management, Pattern Recognition in Remote Sensing Applications.
He has been invited to give keynote lectures and tutorials in several occasions about urban remote sensing, data fusion, EO data and risk management. He published more than 110 papers in international peer-review journals and presented more than 250 research works in workshops and conferences.

Fabio Dell'Acqua (M'00; SM'08) obtained his 5-y degree cum laude in Electronics Engineering at the University of Pavia, Italy, in 1996. In 1999 he completed his PhD in Remote Sensing at the same University. He is currently associate professor of remote sensing at the University of Pavia, Italy. He teaches courses in remote sensing at the University of Pavia and at the International Risk and Emergency Management MSc at the UME School in Pavia, Italy. His research interests focus on radar data processing and radar/optical data fusion for risk-related applications. In this area he is/has been participating to several research projects, both at national and international level (FP6, FP7, INTERREG, Marie-Curie, Global Earthquake Model). In February 2014 he co-founded a University spin-off company, named Ticinum Aerospace, to commercially exploit research results in the use of EO data for risk management. He is Life Member of the Technical and Scientific Board of the Lombardy Aerospace Cluster; co-founder and Earth Observation coordinator of the International Center for Astronomical and Remote-sensing Observations at the Istituto Universitario di Studi Superiori (IUSS) in Pavia; Principal for the Global Earthquake Model foundation at the Intergovernmental Organization "Group on Earth Observations – GEO". His publication records currently include over 50 journal papers, over 150 conference papers, and 15 contributions to books.