



Boletim de Ciências Geodésicas

ISSN: 1413-4853

bcg_editor@ufpr.br

Universidade Federal do Paraná

Brasil

Xiao, Yang; Ouyang, Zhiyun; Zhang, Zhiming; Xian, Chaofan
A COMPARISON OF HAZE REMOVAL ALGORITHMS AND THEIR IMPACTS ON
CLASSIFICATION ACCURACY FOR LANDSAT IMAGERY

Boletim de Ciências Geodésicas, vol. 23, núm. 1, enero-marzo, 2017, pp. 55-71

Universidade Federal do Paraná

Curitiba, Brasil

Available in: <http://www.redalyc.org/articulo.oa?id=393950135004>

- How to cite
- Complete issue
- More information about this article
- Journal's homepage in redalyc.org

redalyc.org

Scientific Information System

Network of Scientific Journals from Latin America, the Caribbean, Spain and Portugal

Non-profit academic project, developed under the open access initiative

A COMPARISON OF HAZE REMOVAL ALGORITHMS AND THEIR IMPACTS ON CLASSIFICATION ACCURACY FOR LANDSAT IMAGERY

Comparação entre algoritmos de remoção de neblina e seu impacto na acurácia da classificação de imagens Landsat

Yang Xiao¹

Zhiyun Ouyang^{1,*}

Zhiming Zhang²

Chaofan Xian¹

¹ State Key Laboratory of Urban and Regional Ecology, Research Center for Eco-Environmental Sciences, Chinese Academy of Sciences, Beijing 100085, China. Email: xyrcees@163.com; zyouyang@rcees.ac.cn; chaofan136@163.com

² Institute of Ecology and Geobotany, Yunnan University, Yunnan 650091, China. Email: zhiming_zhang76@hotmail.com.

Abstract:

The quality of Landsat images in humid areas is considerably degraded by haze in terms of their spectral response pattern, which limits the possibility of their application in using visible and near-infrared bands. A variety of haze removal algorithms have been proposed to correct these unsatisfactory illumination effects caused by the haze contamination. The purpose of this study was to illustrate the difference of two major algorithms (the improved homomorphic filtering (HF) and the virtual cloud point (VCP)) for their effectiveness in solving spatially varying haze contamination, and to evaluate the impacts of haze removal on land cover classification. A case study with exploiting large quantities of Landsat TM images and climates (clear and haze) in the most humid areas in China proved that these haze removal algorithms both perform well in processing Landsat images contaminated by haze. The outcome of the application of VCP appears to be more similar to the reference images compared to HF. Moreover, the Landsat image with VCP haze removal can improve the classification accuracy effectively in comparison to that without haze removal, especially in the cloudy contaminated area.

Keywords: haze contamination; HF; VCP; classification accuracy

Resumo:

A qualidade de imagens Landsat em áreas úmidas é consideravelmente degradada pela neblina em termos do padrão de resposta espectral, o que limita a possibilidade da sua aplicação usando as bandas do visível e infravermelho próximo. Diversos algoritmos de remoção de neblina têm sido propostos para corrigir os efeitos de iluminação indesejados causados pela contaminação por neblina. O propósito deste estudo é ilustrar a diferença entre os dois principais algoritmos (a filtragem homomórfica (HF) e a nuvem de pontos virtual (VCP)) considerando sua efetividade na solução da contaminação da neblina com variação espacial, e para avaliar os impactos da remoção da neblina na classificação da cobertura da terra. Um estudo de caso com utilização de imagens Landsat TM e climáticas (claras e com neblina) em várias áreas úmidas da China

comprovou que estes algoritmos de remoção têm bom desempenho no processamento das imagens contaminadas por neblina. O resultado da aplicação do VCP parece ser mais similar às imagens de referência quando comparado ao HF. Além disso, a imagem Landsat com a remoção de neblina por VCP pode efetivamente melhorar a acurácia da classificação em comparação àquelas sem a remoção da neblina, especialmente em áreas contaminadas por nuvens.

Palavras-chave: Contaminação por neblina; HF; VCP; Acurácia da classificação.

1. Introduction

Cloud and haze can change the brightness values of pixels at different levels of saturation by masking the land surface, which causes to many problems of remote sensing, including exotic errors in land cover classification and false detection of land cover change, etc., especially where large areas are covered by haze (thin cloud) (Zhu and Woodcock, 2012). Due to the variations of climate and weather, it is difficult to obtain cloud and haze free satellite image data completely, which significantly inhibits their application, in humid areas (Asner, 2001).

In order to remove or alleviate the impact of haze on images, several haze removing approaches have been developed during the past 20 years (Chavez, 1988; Liang et al., 2001; Liang et al., 2002; Zhang et al., 2002; He et al., 2010; Liu et al., 2011). Generally, the image-based haze removal technique can be grouped into two algorithms : 1) spatial filtering (i.e. wavelet analysis (Du et al., 2002), the homomorphic filtering (HF) (Fan and Zhang, 2011; Adelman, 1998), and the improved HF (Cai et al., 2011)); 2) spectral transformation (i.e. Tasseled Cap (TC) transformation (Richter, 1996), haze optimized transformation (HOT) (Zhang et al., 2002), and advanced thin cloud optimized transformation (AHOT) (He et al., 2010; Liu et al., 2011).

The algorithm of spatial filtering can remove the lower frequency layer where haze is distributed in the frequency domain. Under the first approach, the HF has been widely used to remove haze from large areas by combining frequency filtering and grayscale change to discriminate haze from background, and subsequently removing the influence of haze based on the illumination–reflectance model. However, the limitations of the selected filter structure and cut-off frequency result in the discard of some useful information. In order to overcome these limitations, many improved haze removal algorithms based on the HF have been presented (Cai et al., 2011), which can effectively deal with the spatial variation of haze contamination with the auto-adjusting filter cut-off frequency.

Unlike spatial filtering, spectral transformation can remove haze contamination with spectral information. TC Transformation is the algorithm of an orthogonal linear transformation to remove haze by discarding the fourth component of the TC transformation which is considered to be thin cloud and other noise (Richter, 1996). HOT is an advanced form of TC transformation with only using blue and red bands for haze detection, and dark object subtraction (DOS) is implemented in each band of each slice after density slicing the HOT image (Zhang, 2002). Recently, in order to retrieve the digital number (DN) contaminated by the spatial variation of ever-present haze, an improved technique that can remove haze by the virtual cloud point (VCP) algorithm was presented by He and Liu (He et al., 2010; Liu et al., 2011).

The improved HF and VCP are the two of most frequently-used haze removal approaches. Majority of peer studies mainly focused on the development of the haze removal algorithm.

However, less research was dedicated to studying the effects of these two haze removal algorithms on land cover classification. Meanwhile, the relationships between haze removal and image classification accuracy still needed to be explored further.

As the composition and characteristics of land surface element, land cover is the key environmental index and is important to the ecosystem assessment and management with political purposes (Cihlar, 2000). Recently, the succession of land cover and climatic anomaly by global change have highlighted the role of ecosystems in soil retention, water flow regulation, water-quality security and the mitigation and adaption of climate change. Therefore, the landscape analysis and ecological modeling of regional area representing the internships between natural conditions and human impacts across time and space are particularly important, which are based on the data of land cover categories frequently. The use of land cover data to answer ecological questions could be greatly increased by the effective removal of haze from satellite images (Hughes and Hayes, 2014).

Whether the haze removal process can effectively improve classification accuracy and how to improve it further is still needed to be explored in the remote sensing community. Thus, the objective of this study was to assess the effectiveness of two different haze removal algorithms (Improved HF and VCP) and their impacts on land cover classification.

2. Overview of the haze removal

2.1 HF algorithm

Images usually consist of light reflected from objects. The basic nature of an image can be characterized by two components based on the illumination-reflectance model (Liu and Hunt, 1984): the amount of light coming from the incident source on the scene or the illumination components and the amount of light reflected by objects in the scene or the reflectance components (Sousa et al., 2011). When surfaces are covered by haze or thin clouds, the illuminated components represent the solar radiation reflected by haze and the reflected components represent the solar radiation reflected by the planetary surface. The simplified concept is shown as follow (Fan and Zhang, 2011; Adelman, 1998).

$$f(x, y) = f_i(x, y)f_r(x, y) \quad (1)$$

Where $f_i(x, y)$ is the illumination component, whose spectrum is concentrated in the low-frequency part, so $f_i(x, y)$ can be regarded as the cloud distribution function; $f_r(x, y)$ is the reflection component, which is a function of surface features, and its spectrum is concentrated in the high-frequency part. The process of removing cloud contamination is to remove $f_i(x, y)$, and find $f_r(x, y)$. Using this fact, high-pass filter technique aims to reduce the significance of $f_i(x, y)$ by reducing the low frequency components of the image. This can be achieved by executing the filtering process in frequency domain. In order to process an image in frequency domain, the image need to be transformed from spatial domain to frequency domain by using Fourier transform. After that, the Butterworth high-pass filter chosen as the HF (homomorphic filter) is used to extract the high-frequency component, which can yield information on surface features,

and to filter out the low-frequency component, which contains the cloud information (Cai et al., 2011).

$$H(u, v) = \frac{1}{1 + k \left[\frac{D_0}{D(u, v)} \right]^{2n}} \quad (2)$$

Where D_0 is the cut-off frequency and, traditionally, D_0 is a fixed value. By taking into account the fact that different thicknesses of cloud can result in different components of illumination and reflection, D_0 should vary with the thickness of cloud cover. k is obtained in the condition $D(u, v) = D_0$ and $H(u, v) = \sqrt{2}/2$, $k = 0.414$; $n = 1$.

Using high-pass filtering to reduce the brightness value is the basic principle of haze removal with HF. But the DN value in cloudless areas has been modified as well. A simple processing is to use the DN value of the original image as the judgment condition for per-pixel replacement. The DN value of the original image and the processed image is expressed in $f(i, j)$ and $g(i, j)$, respectively, and the DN value after dealing with the pixel replacement, is expressed in $F(i, j)$.

$$F(i, j) = \begin{cases} f(i, j), & f(i, j) < g(i, j) (\text{clear area}) \\ g(i, j), & f(i, j) \geq g(i, j) (\text{cloud area}) \end{cases} \quad (3)$$

2.2 VCP algorithm

The visible bands for most land surfaces under clear-sky conditions are highly correlated, and the spectral response to haze is different between the blue and red wavelengths. Thus, based on these differences, the Haze-Optimized Transformation (HOT) for Landsat data has been proposed by Zhang et al. (2002). However, HOT has a limitation in that it is based on the assumption that the blue and red bands are highly correlated. HOT does not seem particularly robust because blue and red bands are not always highly correlated (Liu et al., 2011). Therefore, an ideal index named the 'Background Suppressed Haze Thickness Index' (BSHTI) used for haze removal was proposed by Liu et al. (2011), which not only reflect cloud thickness but also suppressed background noise.

BSHTI derived from an optimal statistical analysis reflects haze thickness well in accord with visual perception based on experience. However, there are still shortcomings since some spurious land cover types may cause severe bias that is statistically abnormal. Therefore some haze perfection algorithms were designed to correct the spurious values (the overestimated or underestimated haze thickness index) resulting from haze detection caused by the spatial information of some land cover types (He et al., 2010). The robust algorithms proposed by Planchon and Darboux (2002) to detect and fill sinks, which can solve the low HOT problem. The mathematical morphological operations (Liu et al., 2011) were used to solve the high BSHTI problem.

There are three algorithms with good performance in haze removal including DOS, histogram matching (HM), and VCP algorithms (He et al., 2010; Liu et al., 2011). All of them are implemented on each band separately, after the density slicing of haze image. Considering our image characteristics (different land cover type compositions and cloud situations), we selected the VCP to remove haze. Based on the detailed VCP algorithm, the point of intersection

(BSHTIVCP, DNVCP) for each band can be found. Then, projecting all pixels (BSHTI, DN) onto the vertical line (BSHTI = 0) to get the dehazed image, and it can be calculated as follows:

$$y_r = (y \cdot x_{VCP} - x \cdot y_{VCP}) / (x_{VCP} - x) \quad (4)$$

where y_r is the DN value of a pixel after haze removal, y is the original DN value of the pixel before haze removal, x is the BSHTI value of the pixel, and x_{VCP} and y_{VCP} are the BSHTI value and DN value of VCP, respectively.

3. Materials and methods

3.1 Materials

The study area located between 104°55' - 110°73' E and 27°96' - 32°69' N, which is a component part of the Sichuan Basin, covering the entire Chongqing region. The Sichuan Basin in southwestern China is a topographically well-defined rhomboid-shaped basin, bounded on all sides by mountains and drained by the Yangtze River and a number of major tributaries (figure. 1). The area for most of the year experiences very humid conditions with a subtropical monsoon climate. With the low altitude and dense river net, the area is influenced by strong evaporation and air humidity. Due to the high mountains block, the water vapor can't spread out and continue to be condensed into cloud and rain within the basin.

As one of the most important ecological function zones, the study area is located on the upper reaches of the Three Gorges Reservoir Area (TGRA), where breeds the largest water conservation project in the world. The status of the landscape structure and ecological function here has direct impacts on TGPA. Eight case areas (figure. 1) were selected for the case studies because of their certain haze contamination images with high quality of corresponding reference images (non-haze contamination). Moreover, these areas are regarded as the places which are hard to be accurately characterized in the world with the heterogeneity and fragment of landscape. Thus, these areas served as important case studies for testing algorithms to monitor haze removal and their impacts on classification.

With a long history of widespread use, Landsat data has become one of the most valuable datasets that are available for mapping land cover information (Cohen et al., 1998; Watmougha et al., 2011). Although selected image data are now available for free, which was one of the constraints highlighted by Foody (2002), users of Landsat data are still suffer from the problem that there may be no option but to select clouds affected image and search for a proper method to identify and remove this contamination when an image from a specific date is required for subsequent data analysis (Annemarie, 2012). Therefore, how to remove or reduce the influence of thin clouds from the remote sensing image in the preprocessing process becomes the barrier of related research. In this paper, 16-scene Landsat TM images (8-scene original and reference images) were collected for the current study, which were listed in Table 1.

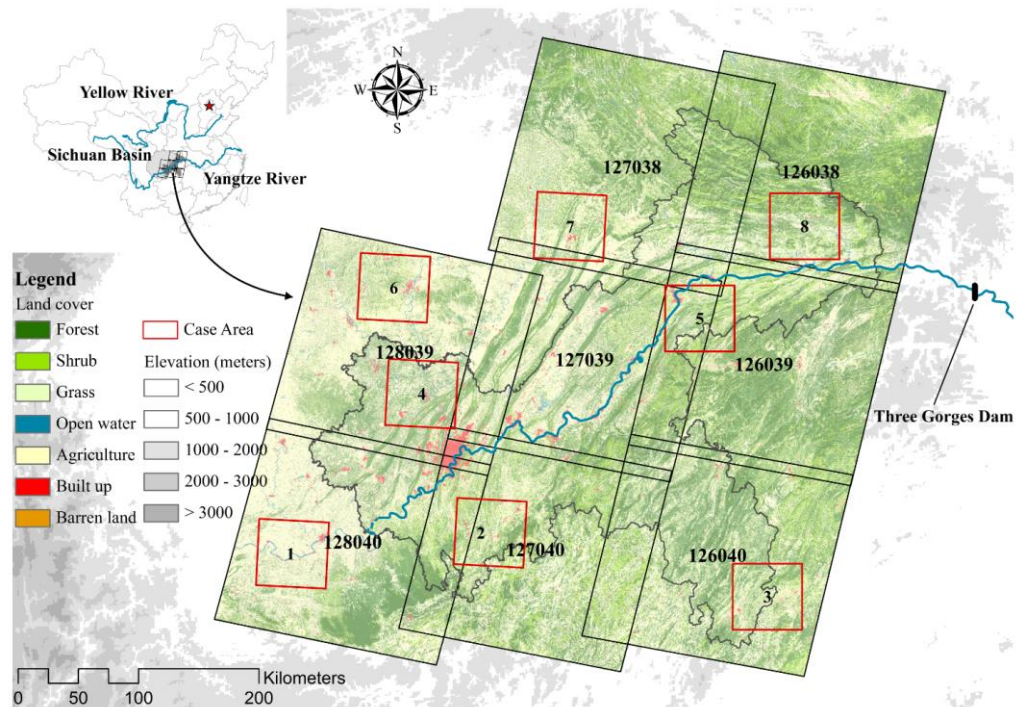


Figure 1: The location of each case study area.

Table 1: List of Landsat TM images used for the current study

Application	Acquisition date	Path/row	Cell center lat/long	Data source		
Original image	2009/7/25	126/038	31.7/109.5	United States	Geological Survey	Geological
	2009/4/20	126/039	30.3/109.2	United States	Geological Survey	Geological
	2011/4/26	126/040	28.9/108.8	Chinese Academy of Sciences		
	2009/11/5	127/038	31.7/108.0	United States	Geological Survey	Geological
	2009/11/5	127/039	30.3/107.6	United States	Geological Survey	Geological
	2009/11/5	127/040	28.9/107.2	United States	Geological Survey	Geological
	2010/8/11	128/039	30.3/106.1	United States	Geological Survey	Geological
	2011/8/30	128/040	28.9/105.7	Chinese Academy of Sciences		
Reference image	2011/4/26	126/038	31.7/109.5	Chinese Academy of Sciences		
	2011/4/26	126/039	30.3/109.2	Chinese Academy of Sciences		
	2009/4/20	126/040	28.9/108.8	United States	Geological Survey	Geological
	2010/8/4	127/038	31.7/108.0	United States	Geological Survey	Geological
	2011/5/19	127/039	30.3/107.6	Chinese Academy of Sciences		
	2011/5/19	127/040	28.9/107.2	Chinese Academy of Sciences		
	2011/8/30	128/039	30.3/106.1	Chinese Academy of Sciences		
	2010/8/11	128/040	28.9/105.7	United States	Geological Survey	Geological

*Note: The original image as the haze contaminated input data in each study area were used for haze removal algorithm; Reference image with the clear comparable data were used for evaluating the effect of haze removal algorithm. All Landsat TM images were radiometrically and geometrically corrected. Landsat data collections from United States Geological Survey (USGS) can be available by order at <http://glovis.usgs.gov/>.

The study area consisted of diverse land cover classes including man-made and natural types. Referring to the land use and land cover classification system

(Anderson et al., 1976), a total of seven land cover classes were identified in the study area to facilitate analysis for accuracy assessment (see Table 2 for class descriptions) (Center for Advanced Spatial Technologies, 2007). Choice of these classes was guided by: i) the objective of the research and ii) the expected certain degree of accuracy in image classification.

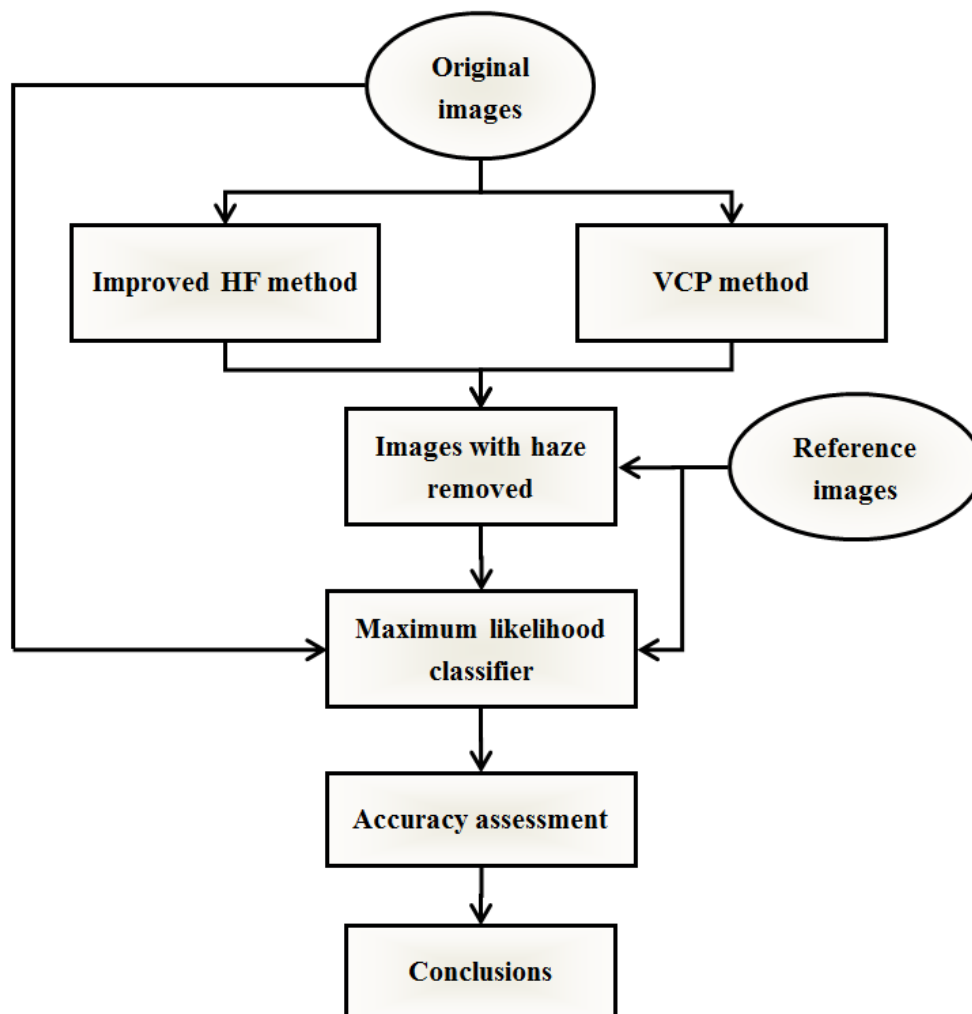
Table 2: Description of land cover classes

Land cover class	Description
Forest	Coniferous trees, deciduous trees and mixed trees
Shrub	Sparse shrubs growing naturally together with some grass
Grassland	Predominant grasses, grasslike plants, forbs, or low shrubs
Water	Rivers, lakes, and some small reservoirs for agricultural and consumption purposes
Agriculture	Lowland and highland fields with or without crops
Built-up	Cities, villages and other artificial infrastructures such as bridges and roads
Barren land	River banks, dry river beds, bare soil and rock outcrop

Sampling collection is essential for classification training and validation, since all further exploration is based on the sample data. In this research, selection of sampled data was performed in-lab for each case study area through visual interpretation of Landsat data and Google Earth VHR image, and through on-the-ground visits to each location during June 2010 and August 2011. All the sampled points were double checked visually using both Google Earth and the Landsat TM images. All the training samples distributed in clear region (non-haze contaminated area through visual interpretation of original image) were used as the input data for classification.

3.2 Methods

Flow chart of the method is shown in figure 2. Due to the fact that the weakened quality of the image caused by haze contamination mainly was distributed on the visible and near-infrared band (1, 2, 3 and 4) with low impacts of haze on mid-infrared bands (5 and 7), band 1, 2, 3 and 4 of Landsat data were used as the input for removing haze contamination. A spectral match was performed between each corresponding band in original and reference images to minimize the negative effects by the inter-annual variability of the landscape.

Figure 2: Flow chart of methodology.

To successfully classify all seven land cover classes, a MLC method was used, which is acknowledged as one of the most efficient parametric methods for image classification (Kozak et al., 2008; Bayarsaikhan et al., 2009). In this research, pixel-based supervised maximum likelihood image classification was performed within ENVI v.4.8 software. A post-classification smoothing process was executed to improve the accuracy of image classification by operating a 3×3 moving window across the images (Krishna Bahadur, 2009).

The complexity of the real Chinese landscape with intra annual variability is needed to be explored by a large number of exemplars, thus any collected sample sites should therefore be used during processing. The impact of training set variability could be minimized by the use of stratified random sampling method. To assess the accuracy of the classification, a confusion matrix (or error matrix), which had been the core of the analysis and estimation procedures for the accuracy assessment (Stehman and Raymond, 1998), was applied here. Important accuracy indexes such as the overall accuracy and kappa coefficient were calculated. Moreover, the Kappa coefficient was calculated as an improvement to this overall accuracy assessment metric, and expresses the proportionate reduction in error generated by a classifier compared with the error of a completely random classification.

To evaluate the impact of haze removal on classification (both the addition of transformed features and reduction of inputs), a series of experiments were devised in which different

combinations of bands were tested as input to ML classifier. As being showed in Table 3, the first experiment used visible and near-infrared bands as inputs; this experiment was treated as the direct effects of images on classification after haze removal. In the second experiment, near-infrared and mid-infrared bands were used as inputs to figure out classification accuracy with the reduction of band1,2,3,4, which caused to lower image quality. In the third experiment, the impacts of combinative bands after haze removal and bands without processing were tested by using all Landsat bands. During the following sets of experiments, the Landsat data (with or without haze removal) were augmented with some additional features (normalized difference vegetation index (NDVI)) to produce an expanded dataset for each study area (cases 4–6).

Table 3: The objectives for each combination of Landsat images and NDVI tested in each study area

Case	Input features	Objectives
1	Band 1, 2, 3, 4	Direct impacts of images after haze removal
2	Band 5, 7	Raw classification accuracy without haze removal
3	All Landsat data	Effects of combination bands after haze removal and non-processing bands
4	Band 1, 2, 3, 4, NDVI	Influence of images after haze removal with additional feature
5	Band 5, 7, NDVI	Influence of expanded dataset
6	All Landsat data, NDVI	Effects of combination bands after haze removal and non-processing bands with additional feature

4. Results and Discussion

4.1 Haze removal visual examination

A simple assessment of the haze reduction was implemented by comparing the images before and after haze removal visually with reference images. In this research, we selected four typical areas to illustrate the performance of haze removal.

From an inspection of the results (figure 3), most of haze, including some thin clouds, had been well removed in the both processed results (HF and VCP results) and their differences were obvious, particularly over the area 3 and 6. Note that the color of the corrected image looked distorted from the original image because of the image enhancement processing for visualization to a great extent. By comparing the VCP results with HF results, the VCP visually appeared to remove more thin cloud and to be more similar to the reference images. The spectrum and texture features of VCP results mostly were in accordance with the reference data, despite of some noticeable excessive processing parts.

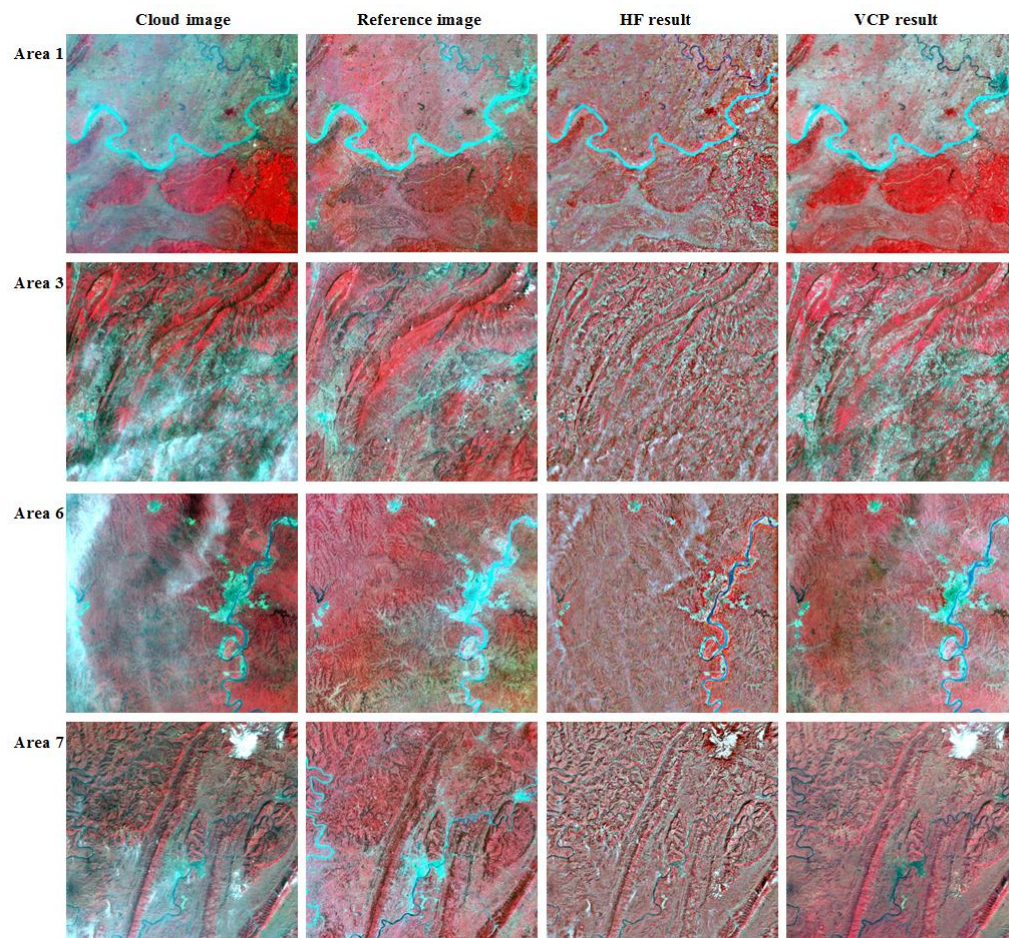


Figure 3: A series of panels showing the performance of the haze removal based on the HF and VCP for four typical areas (band 4, 3, 2 of Landsat TM): Area 1 (top row), Area 3 (second row), Area 6 (third row), and Area 7 (bottom row). For references, Landsat images after linear regression (second column) are shown.

4.2 Haze removal statistical evaluation

In order to assess the effectiveness of the algorithms in recovering the true spectral information under haze, and improving their accuracy rigorously, one common method used to quantitatively evaluate the reliability of the results is the comparison of the image from overlapped region with clear image (spectral matched reference image). The statistical results for the haze image (before and after haze removal) against the reference image are shown in figure 4 and figure 5. After haze removal, the squared correlation coefficient (R^2) for each band between the de-hazed image and clear image became greater than 0.25 (HF) and 0.40 (VCP), respectively, which is much better than that before haze removal. It is clear that de-hazed technique can ameliorate the hazy Landsat TM images significantly. However, there exist differences between the two algorithms. The R^2 for each band based on the VCP are greater than the results by the HF, implying that the VCP is superior to the HF.

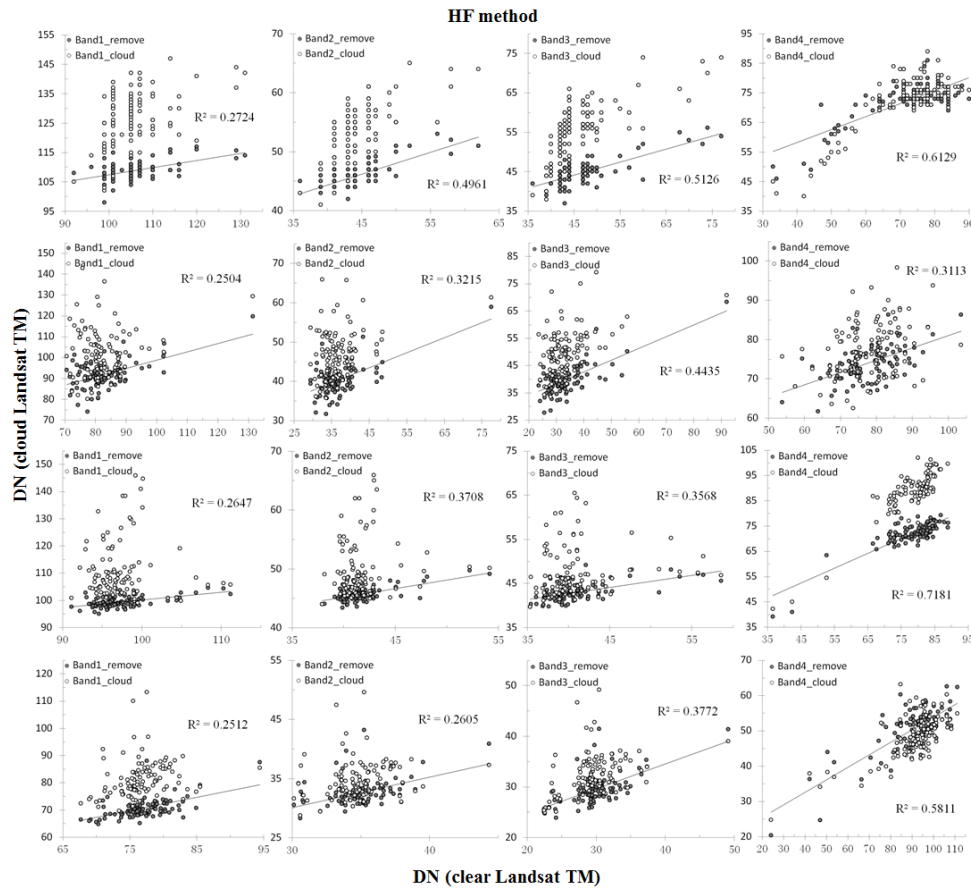


Figure 4: Scatter plot for the cloud image (before and after haze removal based on HF) versus the reference image using 100 paired polygon samples in the overlapping region. DN, Digital Number.

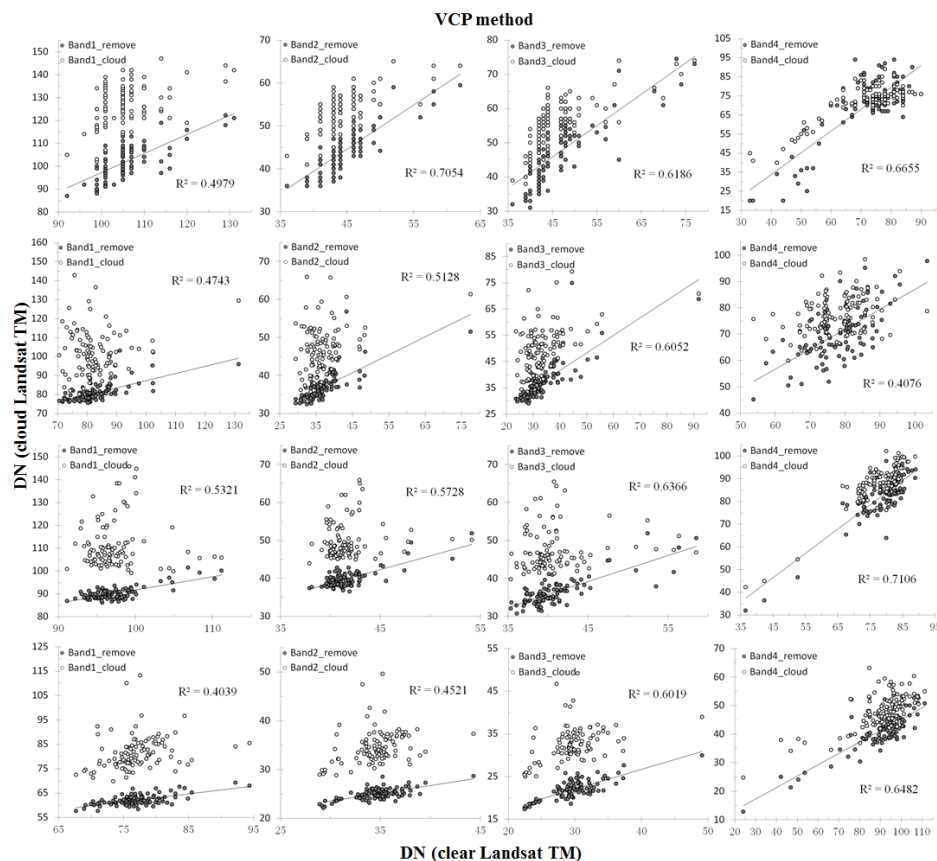


Figure 5: Scatter plot for the cloud image (before and after haze removal based on VCP) versus the reference image using 100 paired polygon samples in the overlapping region. DN, Digital Number.

4.3 Classification accuracy assessment

In the past, it was common that developments in the haze removal algorithm could not reach the practical applications. A further step beyond a new method development is to demonstrate the benefits of such improvement, especially in land cover classification, which has recently been a hot research topic for a variety of applications (Townshend, 1994). In order to assess and judge the effect of haze removal algorithm on classification accuracy, and the impact of data quality and quantity on classifier performance, classification accuracy before and after haze removal were evaluated with multiple criteria through a series of experiments. Meanwhile, classifications of the clear image were also provided references for comparison.

The results of classification reveal a major trend regarding classifier performance after haze removal (figure 6). Both the HF and VCP outperformed the cloud image in terms of overall accuracy in the whole area. The results for the two haze removal approaches (62.5%–74.3%) are higher than those for the cloud image (62.0–69.5%). However, the classification after VCP preceded HF approach in the production of higher accuracy. Among the six classification experiments for VCP, the band combinations that include all Landsat data and NDVI (cases 6) showed optimistic results, and achieve an overall accuracy of 74.3% across the eight study areas, which is similar to the result for reference image (74.7%). The setup of experiments containing NDVI achieved more accuracy in regards to vegetation because of the NDVI responded the highest for vegetation and the lowest for bare soil and built-up (Krishna Bahadur, 2009). These

results showed that higher rates of accuracy could be achieved by haze removal approaches. Given the high degree of variability in the scope and extent of cloud effect, and the complexity of land cover in each study area, the low accuracy rates for the cloud image are expected.

The classification results were assessed further to understand the principle how each algorithm performed in the cloud contaminated area. The assumption here is that the haze contaminated bands provided added variability that causes these areas more difficult to be classified. Evaluating these results therefore offers a more specific viewpoint of classification performance than those in the whole area. Not surprisingly, the results from the cloud area corroborate the results in the whole area: the HF and VCP performed better than the results for cloud image. The overall accuracies in cloud area for the HF and VCP (52.0%–69.0%) were markedly higher than those for the cloud image (49.4%–62.3%), with some added variability indicated by the increase in the error bars for all feature combinations.

There was also a similar pattern for the overall Kappa statistic in the whole area: the haze removal image classification based on the VCP had overall Kappa statistic between 0.55–0.64, the clear image classification was 0.58–0.64, and the haze removal image classification based on the HF had an overall Kappa between 0.49–0.60, while the overall Kappa statistic of original image classification was 0.50–0.58. These results revealed that, in general, the Landsat TM image with VCP haze removal can improve the classification accuracy effectively in comparison to the image without haze removal, especially in cloud contaminated area. While the results for HF performed slightly better than those for the cloud image. The VCP processed images with combination of all Landsat data and NDVI resulted in the best classification with overall accuracy 74.3%, and Kappa coefficient 0.64 for the whole of study areas.

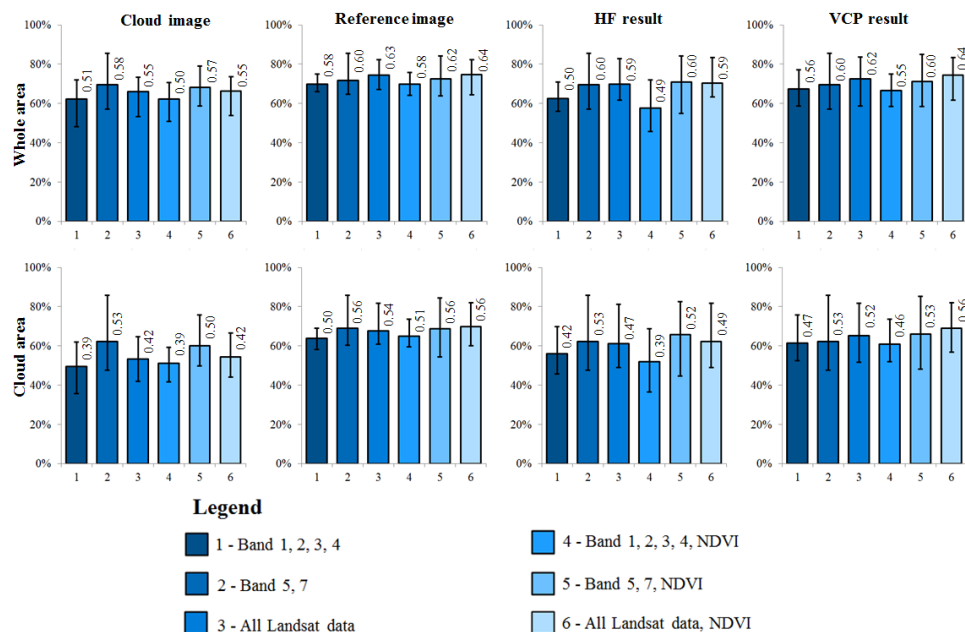


Figure 6: Classification accuracy assessment for each combination of Landsat imagery and transformed features (NDVI) using the maximum likelihood classifier. Classification accuracy in whole area (top row), and classification accuracy in cloud contaminated area (bottom row). For these tests, the results from the areas of eight case studies were then averaged to estimate mean overall accuracy and class accuracy. The value portrayed on each pillar was Kappa coefficient.

4.4 Classification visual examination

The classification results (case 6) of the image before and after the application of haze removal algorithm are shown in figure 7. A visual comparison of the resultant land cover images shows some differences between the classifications. The results indicate that VCP algorithm provide the best results: the size and shape of each land cover features are in good agreement with the morphology of reference image of each area. Moreover, some noticeable misclassification could be found from the results for original image (cloud image). Due to the presence of heavy haze that reducing image quality, the land cover were incorrectly classified using the original image without haze removal by visual comparison with the classifications of reference image.

On the other hand, after haze removal based on HF and VCP, the land cover types were correctly classified in the same area. The classified results after VCP haze removal were similar to the results of reference image, although there still existed some differences between the results of reference image and VCP. The different acquisition time of the reference image may be the major reason for this discrepancy. figure 6 also reveals that the HF did not achieve useable map results compared to VCP, but it still improved the accuracy of classification.

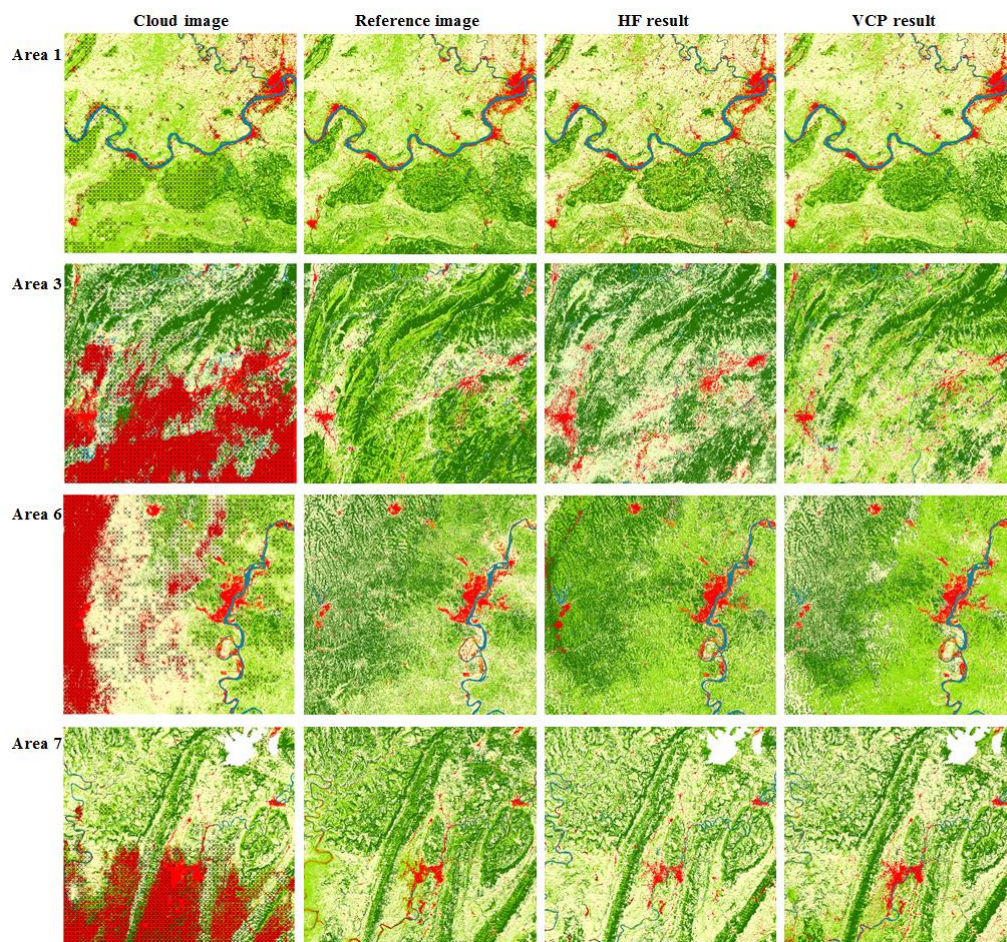


Figure 7: A series of panels showing the performance of land cover classification after removing the impacts of haze on four typical areas: Area 1 Thin cloud contaminated area in the west plain (top row), Area 3 Serious cloud contaminated area in the south mountain (second row), Area 6 Heavy cloud contaminated area in the west mountain (third row), Area 7 Moderate cloud contaminated area in the south hill (bottom row). These maps were produced using the all Landsat and NDVI (case 6) as the input to the maximum likelihood classifier.

*Note: the fork symbols in cloud images represent misclassification areas.

5. Conclusions

In this paper, we compared two major algorithms for haze removal (improved HF and VCP) to propose which is better at removing or weakening the masking influence of haze in humid areas, and find out their impacts on land cover classification. To evaluate the impacts of haze removal on classification, a series of experiments (case1-6) were devised in which different combinations of bands were tested as input to ML classifier. The results suggest that both of the haze removal algorithms perform well in processing Landsat images contaminated by haze although their results were different to some extent, due to the inconsistent in principle of processes. Furthermore, the classification results reveal that both the HF and VCP outperformed the cloud image in terms of overall accuracy. The spectral and textural signatures of haze area for individual classes showed quite satisfactory indication for separability of the classes after haze removal. It is concluded that the determination of land cover classes with different levels of haze contamination using the VCP before classification can lead to an improvement in the overall classification accuracy. The authors would like to point out that, as this research was aiming at the comparison of mapping results with using MLC before and after haze removal, there is no discussion of the classification with other classifier, so some parts of the study still need to be explored further. Nevertheless, it is expected that the processed methods developed in this study will provide theoretical support for other researches, and a stronger understanding of the haze impacts on classification in humid areas will improve classification accuracy, which will assist ecosystem assessment, management and policy purposes in China and other parts of the world.

Acknowledgement

This study was funded by National Key Technology Research and Development Program of China (No. 2011BAC09B08) and Special Issue of National Remote Sensing Survey and Assessment of Eco-Environment Change between 2000 and 2010 (No. STSN-04-01).

REFERENCES

- Adelmann H.G. "Butterworth equations for homomorphic filtering of images." *Computers in Biology and Medicine* 28 (1998): 169-181. doi:10.1016/S0010-4825(98)00004-3.
- Anderson J.R., Hardy E.E., Roach J.T. and Witmer R.E. "A land use and land cover classification system for use with remote sensor data." U.S. Geological Survey Professional Paper 964 (1976). USGS, Washington D.C. (<http://landcover.usgs.gov/pdf/anderson.pdf>. accessed on January 27th, 2014).
- Annemarie S. "Monitoring land cover change in urban and peri-urban areas using dense time stacks of Landsat satellite data and a data mining approach." *Remote Sensing of Environment* 124 (2012): 689-704. doi:10.1016/j.rse.2012.06.006.
- Asner G.P. "Cloud cover in Landsat observation of the Brazilian Amazon." *International Journal of Remote Sensing* 22 (2001): 3855-3862. doi:10.1080/01431160010006926
- Bayarsaikhan U., Boldgiy B., Kim K.R., Park K. and Lee D. "Change detection and classification of land cover at Hustai National Park in Mongolia." *International Journal of*

- Applied Earth Observation and Geoinformation 11 (2009): 273-280. doi:10.1016/j.jag.2009.03.004.
- Cai W.T., Liu Y.X., Li M.C., Cheng L. and Zhang C.X. "A self-adaptive homomorphic filter method for removing thin cloud." 19th International Conference on Geoinformatics 1 (2011): 1-4. doi:10.1109/GeoInformatics.2011.5980963.
- Center for Advanced Spatial Technologies (CAST). "Final Reprot: 2006 Arkansas Land Use and Land Cover." Report prepared by Gorham B. E. and Tullis J. A. University of Arkansas, 2007. (<http://cast.uark.edu/cast-archived-page.php>). accessed on January 23th, 2014).
- Chavez P.S. "An improved dark-object subtraction technique for atmospheric scattering correction of multispectral data." Remote Sensing of Environment 24 (1988): 459-479. doi:10.1016/0034-4257(88) 90019-3.
- Cihlar J. "Land-cover mapping of large areas from satellites: Status and research priorities." International Journal of Remote Sensing 21 (2000): 1093-1114. doi:10.1080/014311600210092
- Cohen W.B., Fiorella M., Gray J., Helmer E. and Anderson K. "An efficient and accurate method for mapping forest clear cuts in the Pacific Northwest using Landsat imagery." Photogrammetric Engineering and Remote Sensing 64 (1998): 293-300.
- Du Y., Guindon B. and Cihlar J. "Haze detection and removal in high resolution satellite image with wavelet analysis." IEEE Transactions on Geoscience and Remote Sensing 40 (2002): 210-217. doi:10.1109/36.981363.
- Fan C.N. and Zhang F.Y. "Homomorphic filtering based illumination normalization method for face recognition." Pattern Recognition Letters 32 (2011): 1468-1479. doi:10.1016/j.patrec.2011.03.023.
- Foody G.M. "Status of land cover classification accuracy assessment." Remote Sensing of Environment 80 (2002): 185-201. doi:10.1016/S0034-4257(01)00295-4.
- He X.Y., Hu J.B., Chen W. and Li X.Y. "Haze removal based on advanced thin cloud-optimized transformation (AHOT) for multispectral imagery." International Journal of Remote Sensing 31 (2010): 5331-5348. doi:10.1080/01431160903369600.
- Hughes M.J. and Hayes D.J. "Automated Detection of Cloud and Cloud Shadow in Single-Date Landsat Imagery Using Neural Networks and Spatial Post-Processing." Remote Sensing 6 (2014): 4907-4926. doi:10.3390/rs6064907.
- Kozak J., Estreguil C. and Ostapowicz K. "European forest cover mapping with high resolution satellite data: the Carpathians case study." International Journal of Applied Earth Observation and Geoinformation 10 (2008): 44-55. doi:10.1016/j.jag.2007.04.003.
- Krishna Bahadur K.C. "Improving Landsat and IRS image classification: evaluation of unsupervised and supervised classification through band ratios and DEM in a mountainous landscape in Nepal." Remote Sensing 1 (2009): 1257-1272. doi:10.3390/rs1041257.
- Liang S.L., Fang H.L. and Chen M.Z. "Atmospheric correction of landsat ETM+ land surface imagery-Part I: methods." IEEE Transactions on Geoscience and Remote Sensing 39 (2001): 2490-2498. doi:10.1109/36.964986.
- Liang S.L., Fang H.L., Morissette J.T., Chen M.Z., Shuey C.J., Walthall C.L. and Daughtry C.S.T. "Atmospheric correction of landsat ETM plus land surface imagery-Part II: validation and applications." IEEE Transactions on Geoscience and Remote Sensing 40 (2002): 2736-2746. doi:10.1109/TGRS.2002.807579.

Liu C.B., Hu J.B., Lin Y., Wu S.H. and Huang W. "Haze detection, perfection and removal for high spatial resolution satellite imagery." *International Journal of Remote Sensing* 32 (2011): 8685-8697. doi:10.1080/01431161.2010.547884.

Liu Z.K. and Hunt B.R. "A new approach to removing cloud cover from satellite image." *Computer Vision, Graphics, and Image Processing* 25 (1984): 252-256. doi:10.1016/0734-189X(84)90107-5.

Planchon O. and Darboux F. "A fast, simple and versatile algorithm to fill the depressions of digital elevation models." *Catena* 46 (2002): 159-176. doi:10.1016/S0341-8162(01)00164-3.

Richter R. A spatially adaptive fast atmospheric correction algorithm. *International Journal of Remote Sensing* 17 (1996): 1201-1214. doi:10.1080/01431169608949077.

Sousa D., Siravenha A.C. and Pelaes E. "Comparing different high-pass filters to improve the accuracy of classification of satellite imagery obstructed by clouds and fog." *Signal Processing, Image Processing and Pattern Recognition, Communications in Computer and Information Science* 260 (2011): 66-73. doi:10.1007/978-3-642-27183-0_8.

Stehman S.V. and Raymond L.C. "Design and analysis for thematic map accuracy assessment: Fundamental principles." *Remote Sensing of Environment* 64 (1998): 331-344. doi:10.1016/S0034-4257(98)00010-8.

Townshend J.G.R. "Global data sets for land applications from the Advanced Very High Resolution Radiometer: an introduction." *International Journal of Remote Sensing* 15 (1994): 3319-3332. doi:10.1080/01431169408954333.

Watmough G.R., Atkinson P.M. and Hutton C.W. "A combined spectral and object-based approach to transparent cloud removal in an operational setting for Landsat ETM+." *International Journal of Applied Earth Observation and Geoinformation* 13 (2011): 220-227. doi:10.1016/j.jag.2010.11.006.

Zhang Y., Guindon B. and Cihlar J. "An image transform to change characterize and compensate for spatial variability in thin cloud contamination of Landsat images." *Remote Sensing of Environment* 82 (2002): 173-187. doi:10.1016/S0034-4257(02)00034-2.

Zhu Z. and Woodcock C.E. "Object-based cloud and cloud shadow detection in Landsat imagery." *Remote Sensing of Environment* 118 (2012): 83-94. doi:10.1016/j.rse.2011.10.028.

Recebido em 17 de novembro de 2015.

Aceito em 25 de maio de 2016.