THRESHOLD VALUE DETERMINATION FOR CLOUD MASKING PROCESS USING LANDSAT 8 IMAGERY

Fithrotul Aulia[1], Totok Wahyu Wibowo[2]
[1]Diploma of Remote Sensing and Geographic Information Students, Vocational School, Universitas Gadjah Mada
[2]Faculty of Geography, Universitas Gadjah Mada

fithaulia@gmail.com, totok.wahyu@ugm.ac.id

Abstract. Optical remote sensing is inevitable from cloud cover problems, especially in tropical country. Cloud cover could reduce the potential usage of the imagery, for instance it would be impractical to use it for land cover classification. The existance of cloud and it's shadow in an imagery could hinder image analysis such as image transformation and image classification. Furthermore, cloud cover could reduce the obervable area of an imagery. Automatic or semi-automatic cloud masking is considered an effective means of removing the cloud cover. Threshold value in the cloud masking process is essential to provide a clean cloud removal.Cloud masking was conducted using Landsat 8 imagery, which has been through pre-processing such as Top of Atmospheric (ToA) correction, Bidirectional Reflectance Distribution Function (BDRF) correction, and radiometric terrain correction. LAPAN's (National Institute of Aeronautics and Space) cloud removal algorithm was used since it provide a semi-automatic procedure. Four threshold value options were chosen based on pixel's statictics in cloud, cloud shadow, water, and other objects that are potentially identify as cloud or cloud's shadow. Cloud was defined based on albedo value using visible channels, whereas the cloud's shadow is defined based on Near Infrared and Short Wave Infrared bands.Four threshold combinations were successfully made based on three different acquisition date of Landsat 8 imagery. The best threshold value should be able to identify cloud and it's shadow, but shows minimal effect to the objects resemble to cloud or it's shadow. The result shows that the most effective threshold is, 1650 and 3600 for cloud, 11000 for lower limit of shadow, 12000 for upper limit of shadow, -1350 for a set of cloud shadow, and 900 for water.

Keywords: cloud masking, semi-automatic cloud masking, pixel of cloud/shadow, threshold

1. Introduction

Remote sensing data has been used extensively in many fields related to space on earth. The quality of imagery determines the potential usage, which differ from one to another remote sensing system. Some remote sensing system may experience a lot of noise, especially the optical remote sensing data inequatorial area. Sensor quality, platform's position, topographical condition, and atmospheric condition are several factors that contribute to image quality (Purwadhi 2002).

Landsat is one of many optical remote sensing data that is freely distributed. In a tropical country, such as Indonesia, cloud cover is the main obstacle of digital image processing (Asner 2001). The brightness of the cloud and the darkness of its shadow could cause inaccuracies in data analysis, such as image transformation, landcover classification, and landuse detection. The area that covered by cloud couldn't be included in the analysis, since it would only disrupt the result. Therefore, the detection of cloud and its shadow is the first step for further analysis (Arvidson et al.2001; Irish 2000; Simpson and Stitt 1998).

Generally, there are two types of cloud, ie thick cloud (opaque) and thin cloud (semi-transparent). The thick cloud is relatively easier to identifiy, because of it's high reflectance value in the visible bands. In the other hand, the thin cloud is relatively harder to identify, because there are a mixing between thin cloud and the objects below it (Gao et al. 2002). Cloud masking is a method that can be employed to remove cloud cover in an imagery, which focussed on cloud and shadow detection (Wang et al. 1999; Andre 2009; Orepoulus 2011, Kustiyo et al. 2012). Automatic Cloud Cover Assessment (ACCA) is acloud cover algorithm developed by Irish et al. (2006). ACCA can work well to estimate the percentage of overall coverage in a scene of Landsaat 8 imagery, but it can't work effectively to determine the edge of the cloud, which could provide an automatic cloud masking process. Thus, the determination of threshold value of unwanted objects (cloud, shadow, and water body) plays an important role in cloud masking process.

Cloud masking process using threshold value was proposed by Kustiyo et al. (2012), which utilized Visible, Near Infrared, Shortwave Infrared, and Thermal bands of Landsat 7. However, the method hasn't been optimally impemented using Landsat 8 imagery. This paper discusses the optimal value of cloud cover masking algorithm on Landsat 8 by utilizing the algorithm by Kustiyo et al. (2012).

2. Methods

2.1. Data

Three Landsat 8 path/row 124/63 with different acquisition dates were used in this study in order to get a variation of clouds. The data were captured at (a) 2 Februari 2015, (b) 7 April 2015, and (c) 9 May 2015 respectively. The first and second data have a larger cloud cover, ie 40.63% and 55.46%, than the third data which has 11.09% cloud cover.

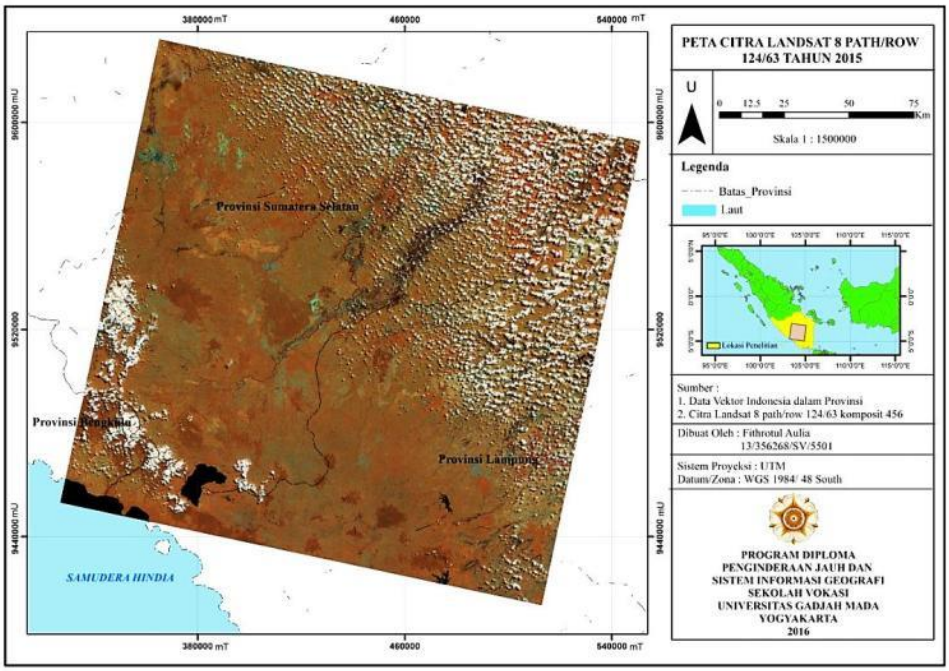


Figure 1. Map of the study area

2.2. Pre-Processing

Image pre-processing were done using three steps, which are (1) file format conversion to suit the software needs (*.TIF to *.ERS), (2) pixel resampling using cubic covolution method, and (3) radiometric correction using ToA-BDRF and terrain correction.

ToA-BDRF was intended to obtain a reflectan value, which is more objective value for multidate imagery. C-Correction method was applied in terrain correction as described in Wu et al. (2004).

Terrain correction was applied to address terrain illuminati effect, in which the object located on slope that facing the sun will receive more energy than the object in the other side of the slope (Furby 2010).

2.3. Sampling

The pixel value of cloud, cloud's shadow, and other similiar objects (such as bareland and water body) were sampled in each scene. The variation of cloud based on the thickness was also sampled (Figure 1). Mean value of each sample was used to determine the threshold value of each object in cloud masking process.

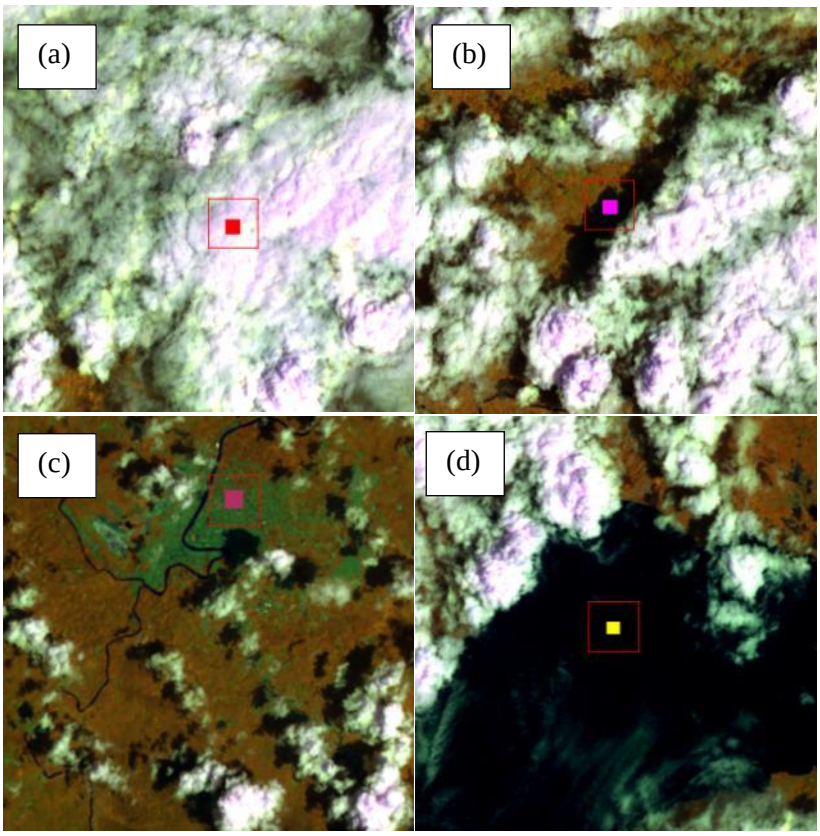


Figure 2. Example of sample taken on (a) thick cloud, (b) cloud's shadow, (c) urban, dan (d) water body.

2.4. Threshold Value Determination

The algorithm developed by Kustiyo et al. (2012) needs several threshold for each unwanted object in image. In this study, thermal band was excluded from the algothm, the calculation of the clouds temperature was not done, and also to maximize the usage of visible amd infrared bands. The following formulas were applied in the image:

- Cloud Threshold. Similar object, such as bareland, waas taken as a consideration for determining cloud threshold.
  $X_2 > DNa$.....(1)
  $(2*X_1-X_3) > DNb$.....(2)

  X_i : Pixel value of *band* –i on Landsat 8*
  DNa : threshold value for albeldo cloud
  DNb : threshold value for bareland and cloud
- Cloud's Shadow Threshold. The following formula was used to determine the cloud's shadow threshold:

$(3*X_4 + X_5) < DNs$.....(1)

$(X_1-X_4) > DNt$.....(2)

X_i : Pixel value of *band* –i on Landsat 8*
DNs : threshold value for the upper limit and the lower limit of cloud's shadow
DNt : threshold value for shadow-cloud edge

- Water Body Threshold. The water body threshold was needed to separate water and shadow, by means of using the following formula:
  $X_4 < DNw$
  Xi : Pixel value of *band* –i on Landsat 8*
  DNw : threshold value for water body

**Band 1 is blue band, not aerosol band*

A command prompt's based software called L8INCASCSv4 was used to facilitate and ensure automation process. Though, the thresholds value from previous step were manually entered into the software.

3. Result and Discussion

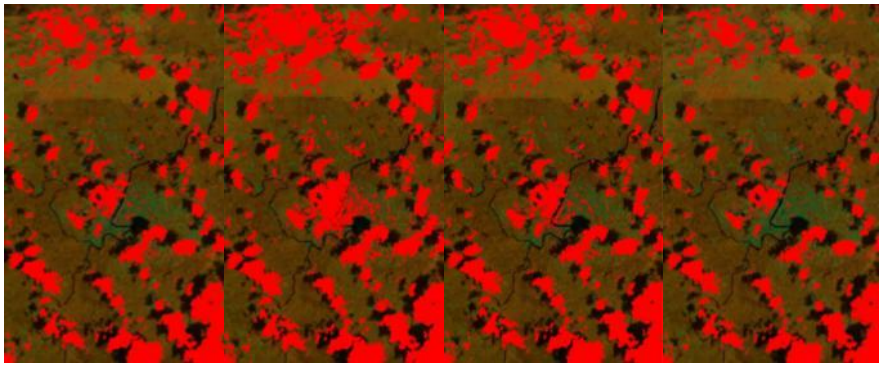
3.1. Threshold Value Determination

Table 1 presents four threshold scenario, which was combination of unwanted objects (cloud, shadow and water body). For instance, sample calculation was taken into account in determining the cloud GREEN/albedo threshold. Should the resulting value were 8600, 3165, 2290, and 5900, then the threshold is 2290 since the formula for albeldo is threshold < B2.

Tabel 1. Threshold combination value for each unwanted object

	Cloud "GREEN"	Cloud "HOT"	Lower limit Shadow	Upper limit Shadow	Shadow-cloud edge	Water
Threshold (1)	1650	3600	10000	11000	-1350	900
Threshold (2)	1550	3200	11000	12000	-1200	800
Threshold (3)	1450	4000	9000	10000	-1000	700
Threshold (4)	1750	2900	8000	9000	-1500	600

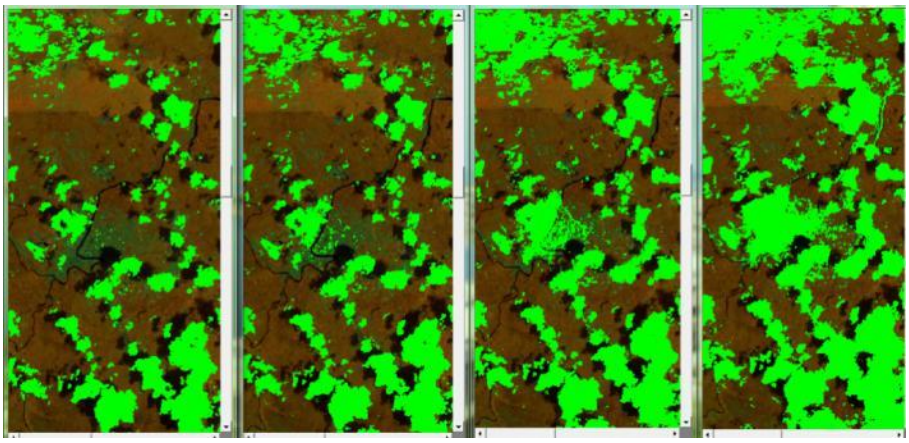
Figure 3 shows explicitly that threshold value 1450 is the best threshold for cloud albedo, without regarding other object such as urban features. The 1450 threshold was also the best for thin cloud identification among other thershold, especially 1750. However the selection of threshold value should be equal, not only good at detecting the cloud but also maintain other objects as well. By means of visual qualitative assessment, it is assigned that 1650 value gives a balance result.



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

Figure 3. Cloud Albedo detection on Landsat 8 with (a) 1650, (b) 1550, (c) 1450, and(d) 1750 thresholds

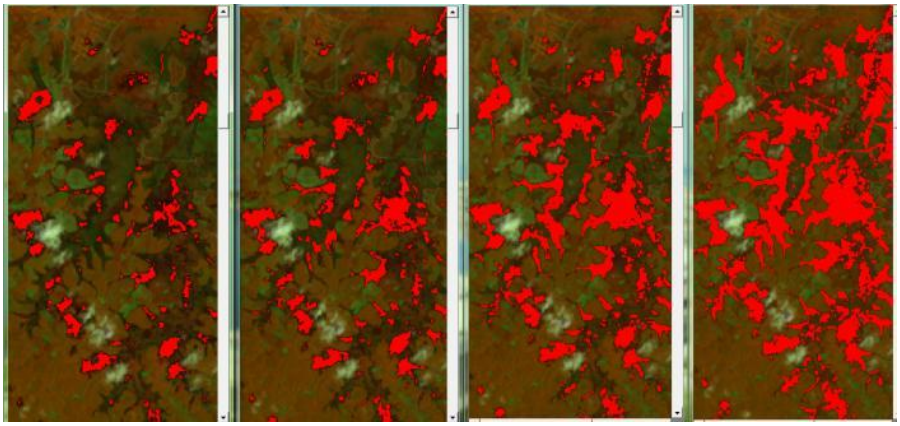
HOT cloud detection as shown in Figure 4 could be effectively identified using 3600 threshold value, due to it's capabilities to detect thin cloud and other non-cloud objects. Although the image is still contains haze, but not as overestimate as other threshold value.



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

Figure 4. HOT cloud detection on Landsat 8 with(a) 4000, (b) 3600, (c) 3200, and (d) 2900 thresholds

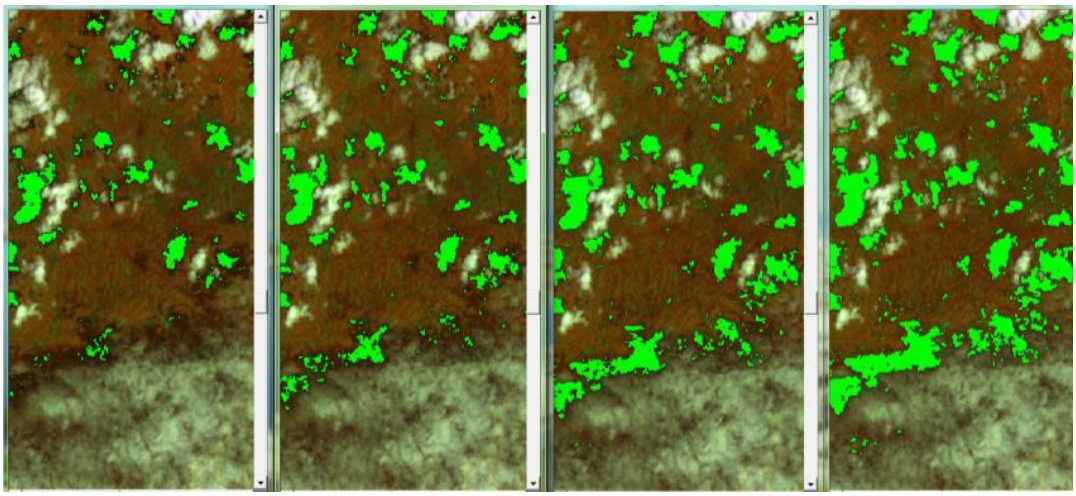
The detection of cloud's shadow distracted by the exitance of water body, since their pixel value is similarly low. The higher the threshold then the more potential of non cloud's shadow object can also be detected. Based on Figure 5, the threshold value 8000 gives an underestimation of cloud's shadow detection, but in the same time could separate water object and cloud's shadow.



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

Figure 5. Cloud's shadow (lower limit) detection on Landsat 8 with (a) 8000, (b) 9000, (c) 10000, and (d) 11000 thresholds

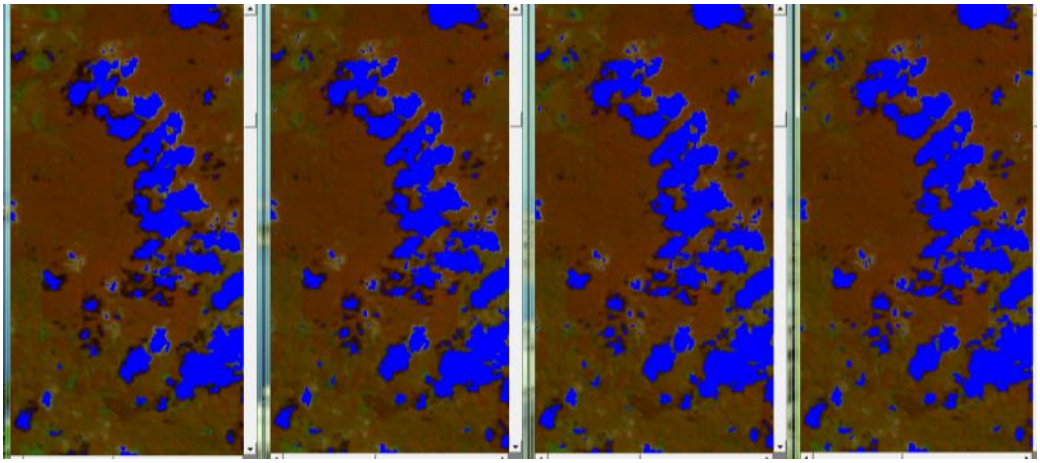
The optimum threshold value for upper shadow detection should also preserve other non-upper shadow object. Figure 6 shows several threhold value, and the optimum value is 11000 since the cloud can be identified correctly without neglecting of other objects.



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

Figure 6. Upper shadow detection on Landsat 8 with (a) 9000, (b) 10000, (c) 11000, and (d) 12000 thresholds

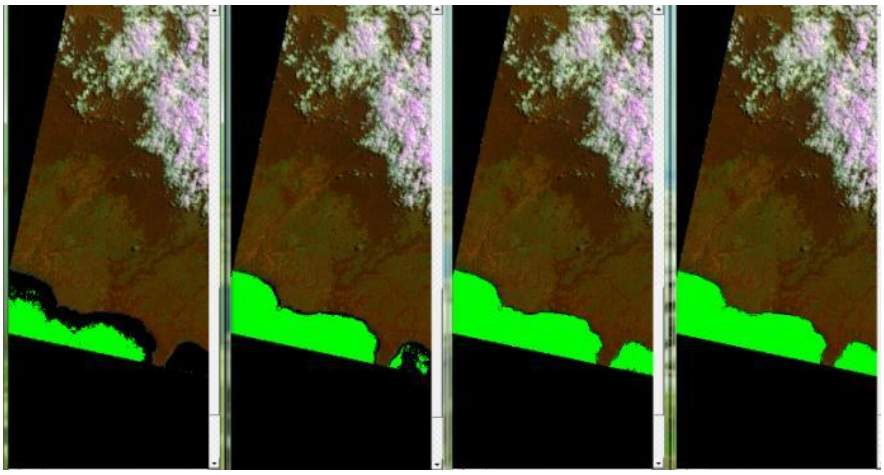
Shadow detection were succesfully detected using -1350 threshold value. The higher threshold value would not cover all of cloud, while the lower threshold would gives an overestimate identification since it also detect non-shadow objects (Figure 7).



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

Figure 7. Shadow detection on Landsat 8 with (a)-1000, (b) -1200, (c) -1350, and (d) -1500 thresholds

The optimum threshold value for water identification is 900, since the result shows that not only deep water but also shallow water can be detected by the threshold. Though normally, deep water has a pixel value arround 500 in band 4 (near infrared band). Figure 8 presents the gradual result of threshold value 600, 700, 800, and 90 respectively.



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

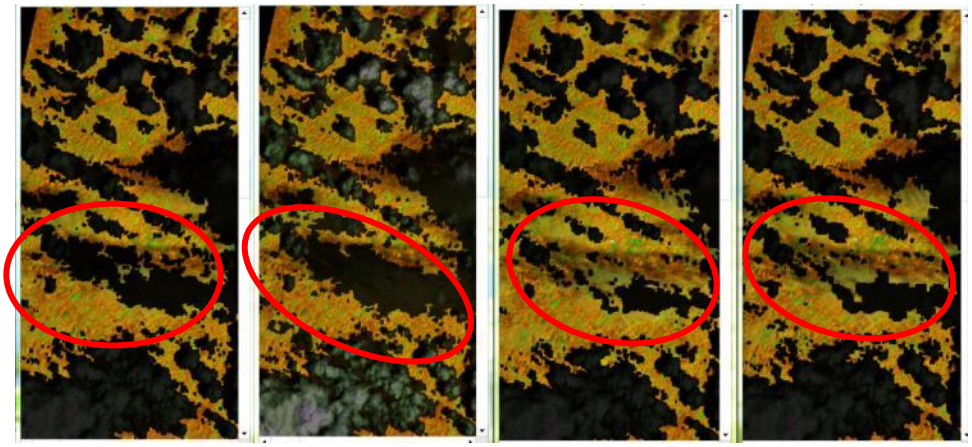
Figure 8. Water detection on Landsat 8 with(a) 600, (b) 700), (c) 800), and (d) 900 thresholds

3.2. Automatic Cloud Masking

There are 4 combination of thresholds (as shown in Table 1) which have been sorted based on visual properties. There are several consideration in determining the threshold value, due to pixel value similarities among particular objects. For instance, the similarities between cloud and bareland area and the similarities between water body and cloud's shadow. Therefore, the threshold value must be carefully examined.

The quality of each cloud masking threshold combination were examined through the resulting image, whether the result is over removal or under removal. In this case, the best threshold combination gives the balance between the preserved objects and neglected objects.Figure 9 shows

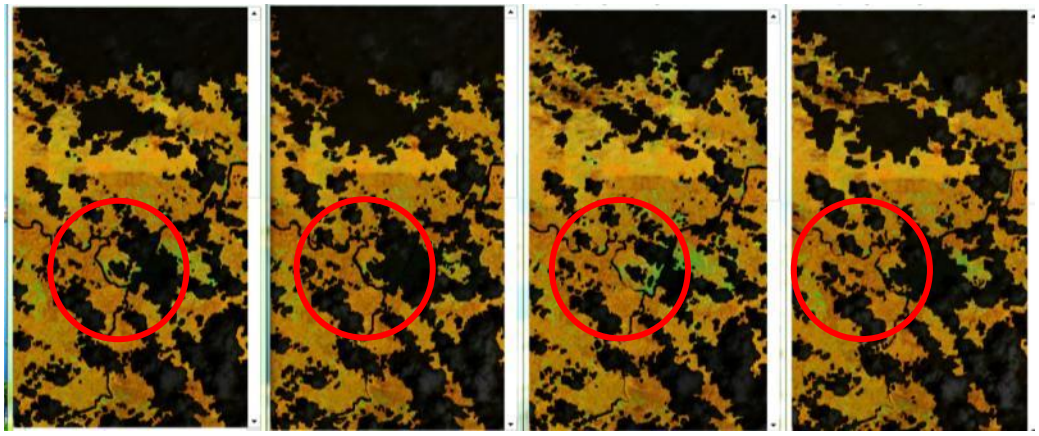
that thin cloud was very well removed using threshold 1 and threshold 2. Nevertheless, the four threshold combination were not optimum to remove thin cloud and it's shadow.



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

Figure 9. Thin cloud removal using (a) threshold1, (b) threshold 2, (c) threshold3, and (d) threshold4

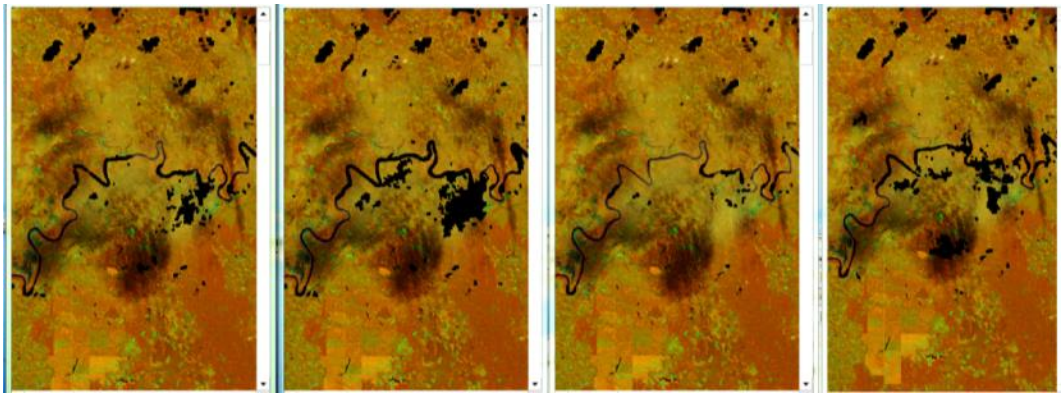
Urban objects, which mostly composed of roof tile, has a similar pixel value to bareland. Moreover particular roof material has a high reflectance value, such as zinc roof, concrete roof and asbestos roof. This condition needs to be consider to determine which threshold is best suited. Threshold 3 can preserve urban object better than other threshold (Figure 10). Overall in the first Landsat (02/02/2015) threshold 1 gives the most satisfactory result, both to remove the unwanted object and preserve essential pixels.The second Landsat 8 image (07/04/2015) is quite the same with the first data. The most effective threshold combination is threshold 1.



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

Figure 10. Urban objects preservation in (a) threshold1, (b) threshold 2, (c) threshold3, and (d) threshold4

The third Landsat 8 (09/05/2015) image, which has the smallest cloud cover, was proved more difficult to remove thin cloud than two previous data (Figure 11). In this case, when the imagery used is relatively has a small cloud cover, then the threshold value could become ineffective. In other words, the well proven threshold of cloud masking in one image, not necessarily has the same result in other image condition or other image.

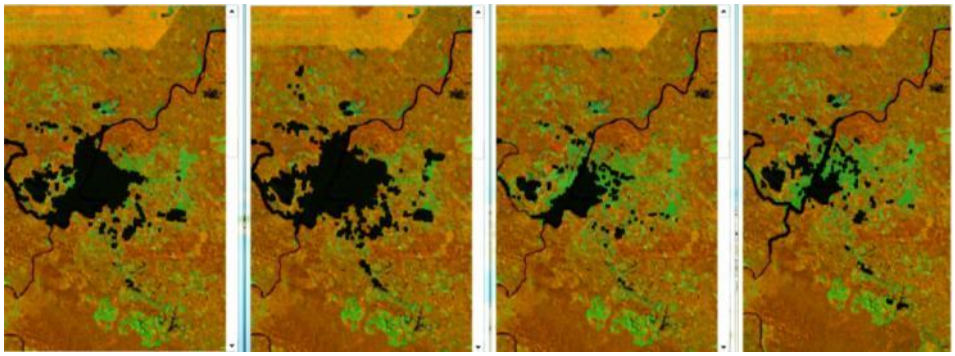


(a) (b) (c) (d)

Figure 11. Thin cloud removal on Landsat 8 (09/05/2015) with (a) threshold 1, (b) threshold2, (c) threshold3, and (d) threshold4

In term of removing cloud and cloud's shadow threshold 1 and 2 give a satisfactory result, but in the same time also eliminates a bit of several other objects such as wetland, urban area, and bareland. Figure 12a shows the urban area object when all threshold were applied to the third Landsat 8 image, while Figure 12b shows the effects of each thresholds towards wetland area.

a)



b)

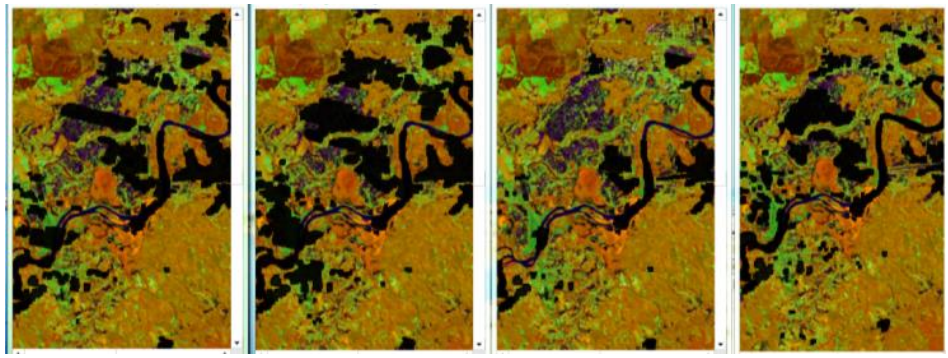


Figure 12. Thresholds effects on (a) urban area and (b) wetland area using Landsat 8 (08/09/2015)

Table 2. Presents the rank of qualitative satisfaction measure in three Landsat 8 image and four threshold combination. Meanwhile the complete result of cloud masking in each image can be seen in the closure.

	Threshold 1	Threshold 2	Threshold 3	Threshold 4
Landsat 8 (02/02/2015)	***	**	*	*
Landsat 8 (07/04/2015)	***	**	*	*
Landsat 8 (09/05/2015)	**	**	*	*

*** satisfactory, ** less satisfactory, * not satisfactory

5. Conclusion

Threshold value of particular object, ie cloud, cloud's shadow, and water, can be determined to achieve a satisfactory cloud masking result. Calculation of the threshold can be achieved using automatic cloud masking formula. All threshold were then can be combined in order to be applied in the imagery. From the experiment presented in this paper, two Landsat 8 image with cloud cover above 40% were succesfully executed using threshold combination 1, which consist of cloud albedo (1650), HOT cloud (3600), lower limit shadow (10000), upper limit shadow (11000), shadow-cloud edge (-1350), and water (900). Nevertheless, when the thershold combination was applied into the lower cloud cover image, the result was not so satisfactory. The later result bring assumption that the threshold value for cloud masking process could only be optimum to specific condition. Based on this experiment the differentiatiog factor is the percentage of cloud cover in an image.

References

AndréC and Hégarat-Mascle S L 2009 Use of markov random fields for automatic cloud/shadow detection on high resolution optical images. *ISPRS Journal of Photogrammetry and Remote Sensing***64** 351–366

Arvidson T, J Gasch, and S N Goward 2001 Landsat-7's long-term acquisition plan An innovative approach to building a global imagery archive, *Remote Sensing of Environment*, 78(1–2):13–26. Furby, S.L, Caccetta, P.A. and Wallace, J.F. 2010. Salinity Monitoring in WesternAustralia using Remotely Sensed and Other Spatial Data. J Environ. Quality 39 pp16-25

Asner G P 2001 Cloud cover in Landsat observation of the Brazilian Amazon. *International Journal of Remote Sensing***22** 18 3855–3862

Gao B C, P Yang, W Han, R R Li and W J Wiscombe 2002 An algorithm using visible and 1.38-μm channels to retrieve cirrus cloud reflectances from aircraft and satellite data, *IEEE Transactions on Geoscience and Remote Sensing***40** 8 1659–1668

Huang Chengquan, Thomas Nancy, Goward N Samuel, Masek G Jeffrey, Zhus Zhiliang, Townshend, G R John, and Vogelmann E James 2010*Automated Masking Of Cloud and Cloud Shadow for Forest Change Analysis Using alndsat Imagery. International Journal of Remote Sensing* 31 20 5449 - 5464

Irish R 2000 Landsat-7 automatic cloud cover assessment algorithms for multispectral, hyperspectral, and ultraspectral imagery. *The International Society for Optical Engineering*, 4049 348–355

Irish R, J L Barker, S N Goward and T Arvidson 2006 Characterization of the Landsat-7 ETM+ Automated Cloud-Cover Assessment (ACCA) algorithm. *Photogrammetric Engineering and Remote Sensing***72** 10 1179–1188

Kustiyo Dianovita, Hedi Ismaya, Mulia Inda Rahayu, Erna Sri Adiningsih2012*New Automated Cloud and Cloud-Shadow Detection uding landsat Imagery.International Jurnal of Remote Sensing and Earth Sciences* **9** 2

Oreopoulos L, M Wilson, and T Várnai 2011 Implementation on Landsat data of a simple cloud mask algorithm developed for MODIS land bands. *IEEE Transactionson Geoscience and Remote Sensing***8** 4 597–601

Purwadhi Hardiyanti dan Tjaturahono Budi S 2001*Pengantar Interpretasi Citra Penginderaan Jauh*. LAPAN: Jakarta

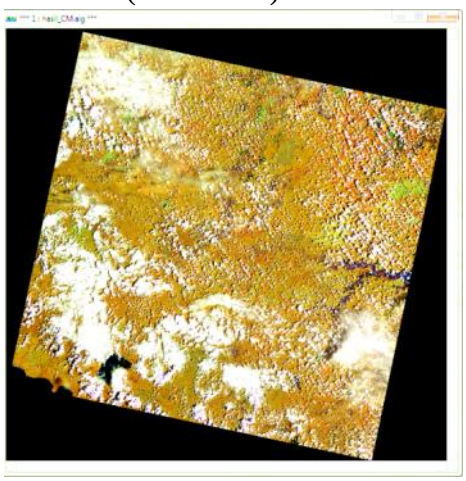
Simpson J J, and J R Stitt 1998 A procedure for the detection and removal of cloud shadowfrom AVHRR data over land. *Geoscience and Remote Sensing***36** 3 880–890

Wang B, A Ono, K Muramatsu, and N Fujiwara 1999 Automated detection and removal of clouds and their shadows from Landsat TM images. *IEICE Transactions on Information and Systems*, **82** 2 453-460

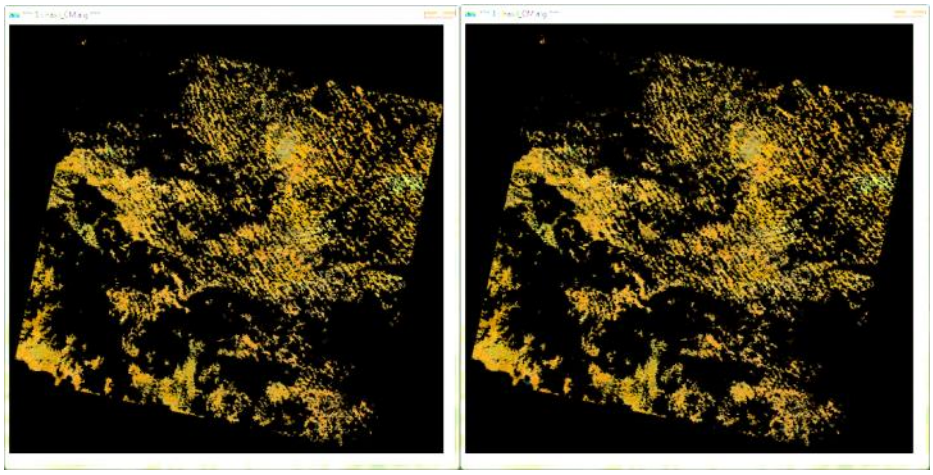
Wu X, Furby, S, and Wallace, JF 2004*An Approach Terrain Illumination Correction, In: Proceedings of the 12th Australian Photogrammetry and Remote Sensing Conference.* Fremantle Western Australia

Appendix

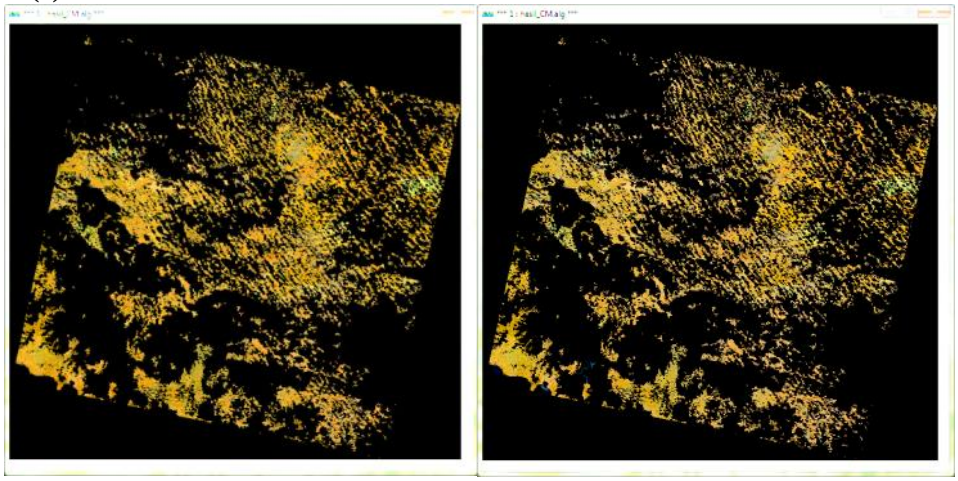
1. Landsat 8 (02/02/2015)



(a)



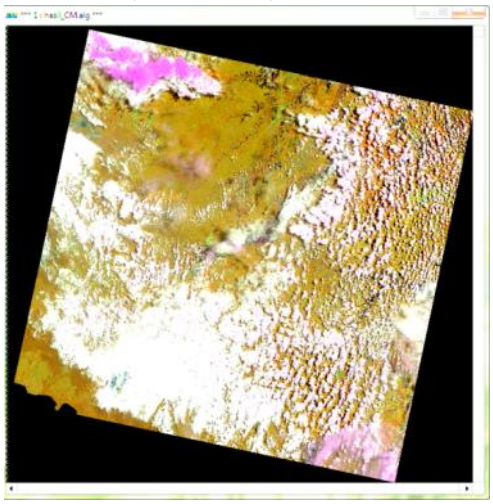
(b) (c)



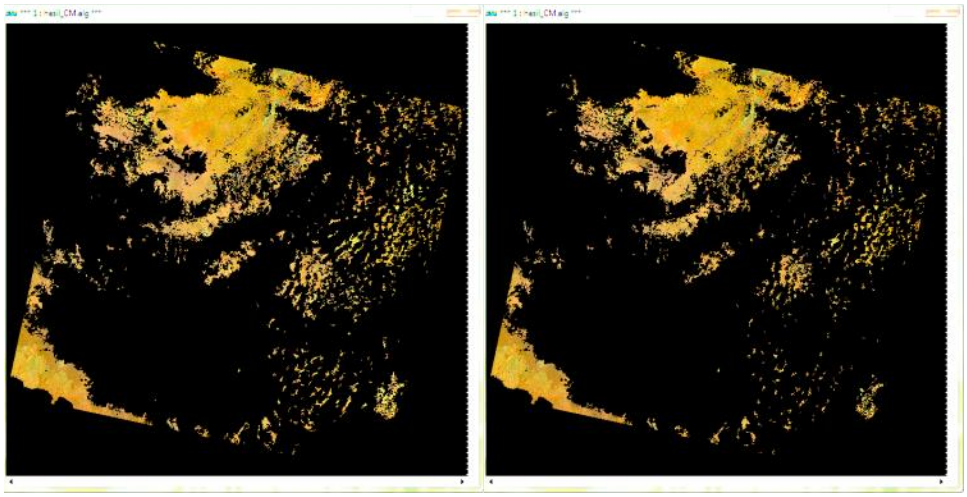
(d) (e)

Information :(a) Original data, (b) result of threshold1, (c) result of threshold 2, (d) result ofthreshold 3, and (e) result of threshold 4

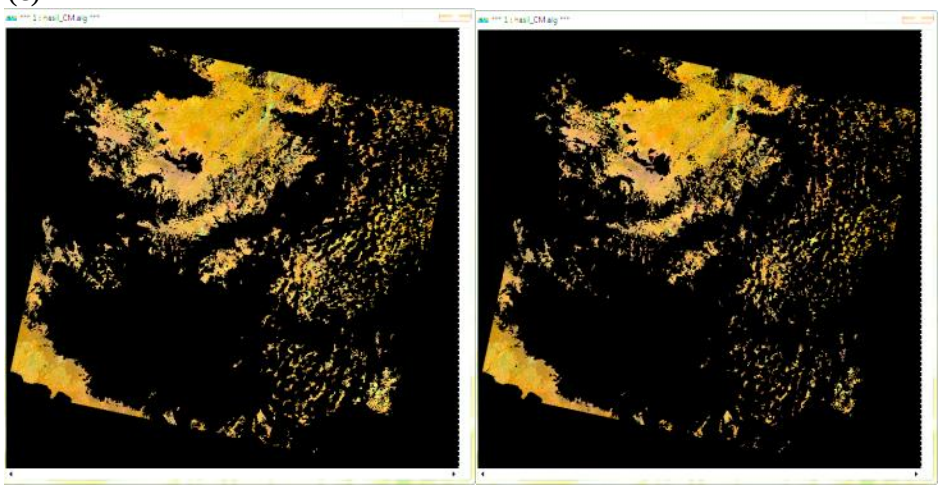
2. Landsat 8 (07/04/2015)



(a)



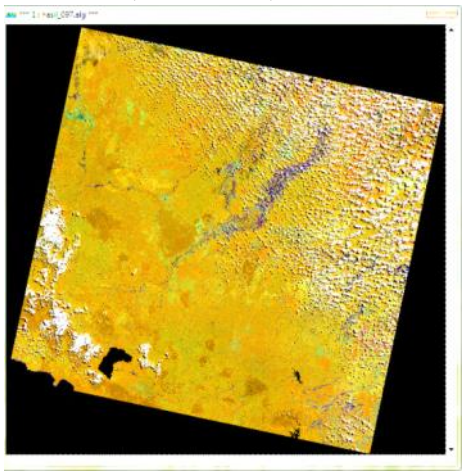
(b) (c)



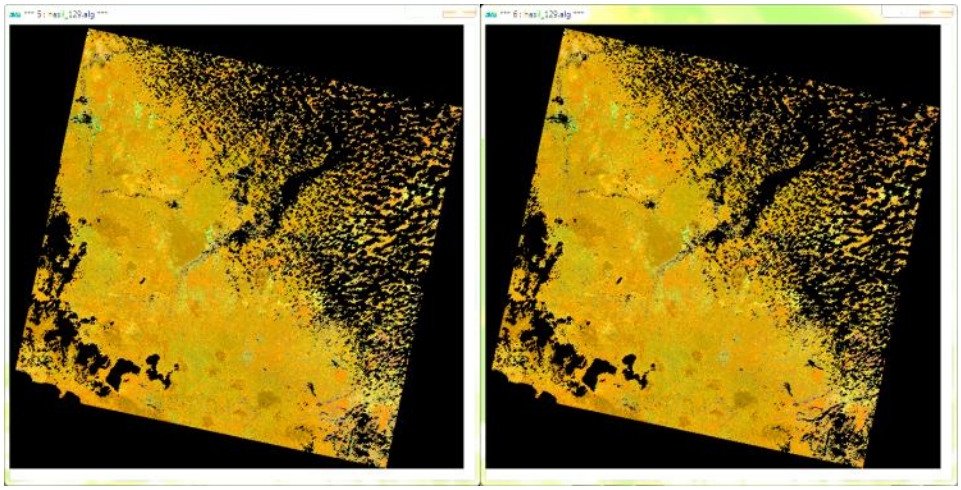
(d) (e)

Information: (a) Original data, (b) result of threshold 1, (c) result of threshold 2, (d) result of threshold 3, and (e) result of threshold 4

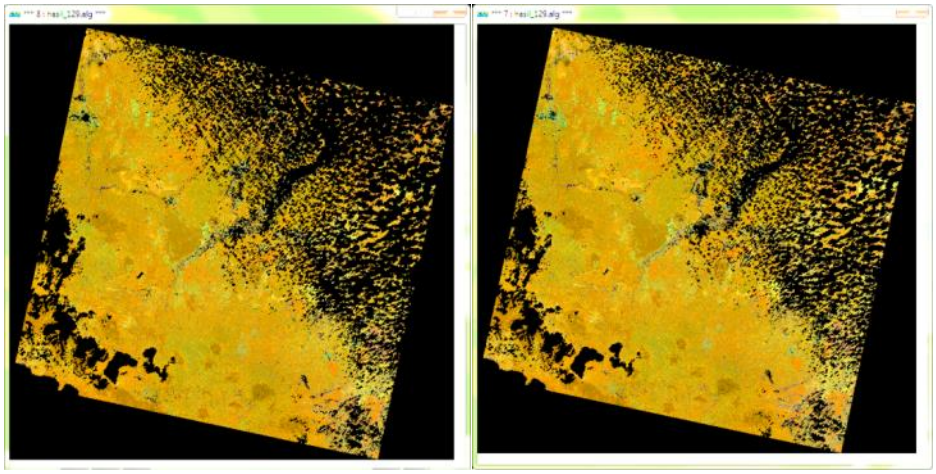
3. Landsat 8 (09/05/2015)



(a)



(b) (c)



(d) (e)

Information: (a) Original data, (b) result of threshold 1, (c) result of threshold 2, (d) result of threshold 3, and (e) result of threshold 4