

# A NEW CLOUD SHADOW DETECTION ALGORITHM BASED ON PRIOR LAND TYPE DATABASE SUPPORT

X Y.Zhou<sup>1</sup>, L. Sun<sup>2,\*</sup>

<sup>1</sup> School of Remote Sensing and Information Engineering, Wuhan University, Wuhan; 430079, China – zhouxueying666@hotmail.com

<sup>2</sup> Geomatics College, Shandong University of Science and Technology, Qingdao, China – sunlin6@126.com

**Commission VI, WG VI/4**

**KEY WORDS:** Cloud Shadow Detection, Cloud Shadow Probability, Landsat 8, Land Type Database, GlobeLand30, Pixel Database

## ABSTRACT:

Cloud shadow detection is one of the basic step for remote sensing image processing. The threshold method is commonly used currently because of its easy implementation and good accuracy. Aiming at the problems that common threshold settings are difficult to meet the complex surface conditions and the results is two-value, this paper proposes a detection method of cloud shadow pixels based on land cover data support by calculating shadow probabilities using Landsat 8 data as an example. Then, a validation with visual interpretation is used to verify the accuracy. The results show that the method can achieve high cloud shadow detection results.

## 1. INTRODUCTION

The existence of cloud shadows on remote sensing images can cause wrong expression of the original surface information, which makes the qualitative and quantitative results inaccurate. Extracting cloud shadow areas is the basis for remote sensing image processing (Li et al., 2016).

Threshold method of shadow detection is commonly used because of its easy implementation and good accuracy. It is based on the difference between cloud shadow and typical underlying surface spectrum to identify the cloud shadow pixels (Shahtahmassebi et al., 2013). The traditional threshold method (Song, Civco, 2002; Boright, Sluder, 2007; Zhu, Woodcock, 2012) finds the threshold with the spectral features based on the band value, band ratio or difference. However, the surface condition is complicated. If the threshold is applied to various types of surface, it is difficult to achieve comprehensive and accurate cloud shadow recognition (Sun et al., 2018a). Moreover, and the result of traditional methods is two-value result (cloud shadow or clear land). However, the support of prior data can better improve the accuracy (Tian et al., 2018). To solve these problems, this paper proposes a detection method of cloud shadow pixels based on land cover data support using spectral characteristic differences between cloud shadow and clear land. This paper takes Landsat 8 data as an example and uses GlobeLand30 surface type database as a support to obtain the threshold of cloud shadows under different surface types. Then, the cloud shadow probability is calculated based on the database of land types and Cloud Shadow & Clear Land (CSCL) pixels. Cloud shadow detection experiments are carried out in areas with cultivated land, forest, grassland, shrubland, wetland, artificial surface and bareland. Moreover, the validation with visual interpretation is used to verify the accuracy. The results show that the method can achieve high cloud shadow detection results.

## 2. METHOD

### 2.1 Data source

**2.1.1 GlobeLand30:** The 30-meter Global Land Cover Dataset (GlobeLand30) is a global land cover product generated from 30-meter multispectral images, including Landsat TM5, ETM+ multispectral images and multispectral images from the Chinese Environmental Disaster Alleviation Satellite (HJ-1). It also combines a lot of auxiliary data and reference materials at the same time. The overall accuracy of GlobalLand30-2010 can reach 83.51% (Chen et al., 2017). The relevant Kappa indicator is 0.78. Thus, this land cover type classification product achieves high accuracy, which provides an accuracy guarantee for its application in cloud detection in this study (Sun et al., 2018b).

### 2.1.2 Pixel database building:

The pixel database consists of cloud shadow and clear land utilizing Landsat 8 images. In order to ensure the diversity and availability of samples, the sample points are selected from 40 images in the world. It not only samples the CSCL pixels of different types of surface, but also samples the cloud shadow of different degrees, which makes the shadow features more complete and provides guarantee for the detection results.

### 2.2 Optimal bands selection

According to the built prior CSCL pixel database, the Shadow Correct (True Negative) Rate (SCR) and Clear-land Error (False Positive) Rate (CER) is calculated at each threshold which is reflectance changing from 0 to 1 at intervals of 0.01. Figure 1 shows the simulation trend of shadow accuracy rate and clear land error rate. When the threshold is very low, SCR and CER are all 0. Then, some shadows can be identified correctly which results in an increasing SCR. Meanwhile, CER is low due to the

\* Corresponding author: sunlin6@126.com

accurate recognition of clear land. Big difference between SCR and CER start to appear as the threshold increasing. Finally, both the two rates end to 1. This simulation is performed on each surface type and each band. When there is a band that SCR is greater than 0.95 and CER is less than 0.1, then this band is included in one of the optimal bands for one land type.

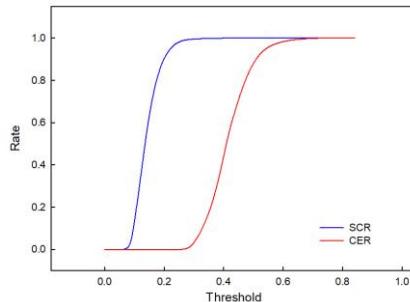


Figure 1. Simulation trend of SCR and CER

### 2.3 Cloud shadow probability calculation

The optimal bands for each land type are obtained by the above simulation.  $MIN'$  and  $MAX'$  are obtain by following equations:

$$MIN' = T_{\max\{SCR=0, CER=0\}} \quad (1)$$

$$MAX' = T_{\min\{SCR=1, CER=1\}} \quad (2)$$

where  $T$ = threshold value

The probability of cloud shadow for each band is calculated in which interval that traversed in turn at an interval of 0.001 when  $T \in [MIN', MAX']$ . The value at the right end of each interval is taken as the threshold representing the interval, so the cloud shadow probability corresponding to each threshold is obtained. The cloud shadow probability of the  $i$ th interval is calculated as follows:

$$P_i = N_{\text{shadow}} / N_{\text{total}} \quad (3)$$

where  $N_{\text{shadow}}$  = cloud shadow pixel number in  $[MIN' + (i-1) * 0.001, MIN' + i * 0.001]$

$N_{\text{total}}$  = total pixel number in this interval

So the probability of cloud shadow at a certain threshold is obtained. In the calculation,  $MIN$  is the maximum value in the thresholds of the cloud shadow probability of 1 and  $MAX$  is the minimum threshold value when probability is 0. Therefore, when reflectance is less than  $MIN$ , the cloud shadow probability is 1, and when the pixel reflectance is greater than  $MAX$ , the probability is 0.

When the threshold is between  $MIN$  and  $MAX$ , according to statistics, the cloud shadow probability corresponding to each threshold will present a certain trend. The function that best fits this curve is the Sigmoid function (S-shaped curve). The function is defined as equation (4), and Figure 2 shows an example of three optimal bands for grassland.

$$f(T) = \frac{a}{1+e^{-(x-x_0)/b}} \quad (4)$$

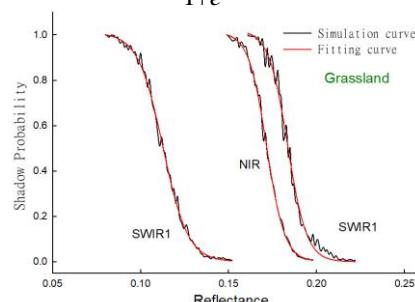


Figure 2. Fitting curves for grassland

### 2.4 Cloud shadow probability generation algorithm

Each land type usually has 2~3 optimal bands, so these bands need to be combined to get the final probability results.

Based on the CSCL pixel database, the average difference between the clear land and cloud shadow reflectance is calculated for each band, and then the corresponding weight is given by the distance weight method. The difference bigger, the weight larger. It is because more distinguishable. The final algorithm is as follows:

$$P_{i,j} = \begin{cases} 1 & \rho_{i,j} \leq MIN \\ f(\rho_{i,j}) = \frac{a}{1+e^{-(\rho_{i,j}-x_0)/b}} & MIN \leq \rho_{i,j} \leq MAX \\ 0 & \rho_{i,j} \geq MAX \end{cases} \quad (5)$$

$$P = \sum_{j=1}^n P_{i,j} \times W_{i,j} \quad (6)$$

where  $i$ = land type

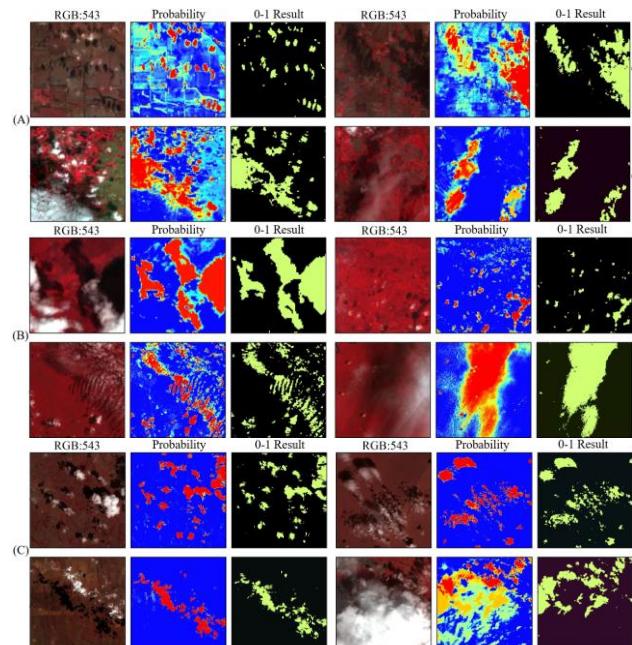
$n$  = number of optimal bands

$W_{i,j}$  = weight of band  $j$  on land type  $i$

$\rho$  = apparent reflectance

## 3. RESULT

Cloud shadow detection results can be obtained by selecting an appropriate probability threshold according to user needs. Generally speaking, a pixel whose probability is greater than 0.75 can be considered as cloud shadow in all probability. Thus, in this paper, 0.75 is used as the probability threshold to the experiments. Some results are shown in Figure 3. It shows the good detection effect of this algorithm in different land types and shadows of different degrees.



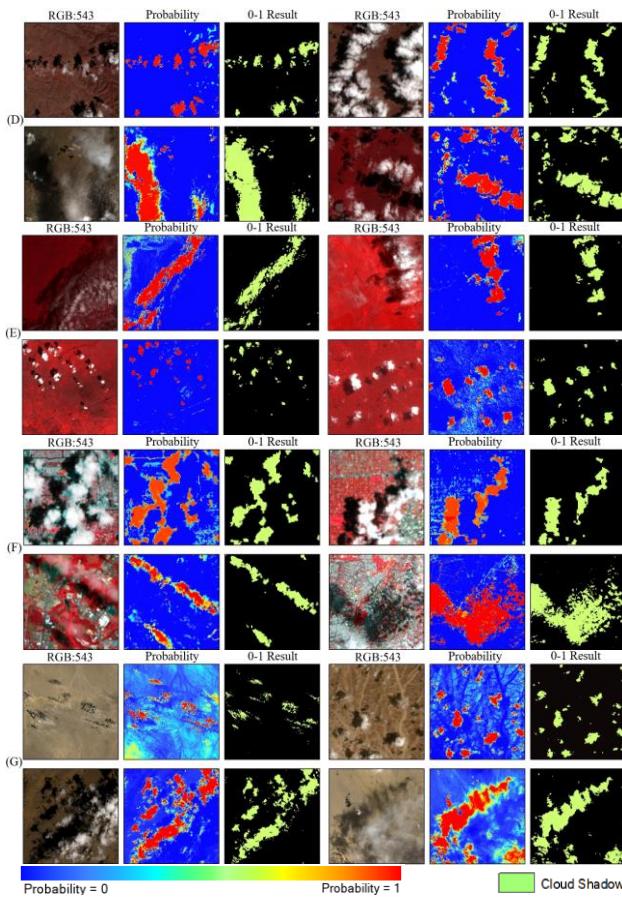


Figure 3. Cloud shadow detection results ( $P>0.75$ )

(A): cultivated land (B): forest (C): grassland (D): shrubland (E): wetland (F): artificial surface (G): bareland

#### 4. CONCLUSION

Aiming at the problem that traditional threshold method cannot adapt to complex surface conditions, land type data and CSCL pixel database are added as references in this paper, so that the threshold can be set according to different situations. In addition, the detection result of the threshold method is usually a 0-1 result. This paper proposes a shadow probability generation method based on the pixel database, which is better than the two-value result. The results show that the algorithm can achieve good accuracy. This paper provides a new idea for cloud shadow detection.

#### REFERENCES

- Boright, A. L., & Sluder, J. C. 2007. U.S. Patent No. 7,184,890. Washington, DC: U.S. Patent and Trademark Office.
- Chen, J., Liao, A., Chen, J., Peng, S., Chen, L., & Zhang, H. 2017. 30-meter global land cover data product-GlobeLand30. *Geomatics World*, 24(1), 1-8.
- Li, Z. , Shen, H. , Li, H. , Xia, G. , Gamba, P. , & Zhang, L., 2016. Multi-feature combined cloud and cloud shadow detection in gf-1 wfv imagery. *Remote Sensing of Environment*, 191, 342-358.
- Shahtahmassebi, A., Yang, N., Wang, K., Moore, N., & Shen, Z. 2013. Review of shadow detection and de-shadowing

methods in remote sensing. *Chinese Geographical Science*, 23(4), 403-420.

Song, M., & Civco, D. L. 2002,. A knowledge-based approach for reducing cloud and shadow. In *Proc. of* (pp. 22-26).

Sun, L., Zhou, X., Wei, J., Wang, Q., Liu, X., Shu, M., ... & Zhang, W. 2018. A New Cloud Detection Method Supported by GlobeLand30 Data Set. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(10), 3628-3645.

Sun, L., Wang, Q., Zhou, X., Wei, J., Yang, X., Zhang, W., & Ma, N. 2018b. A Priori Surface Reflectance-Based Cloud Shadow Detection Algorithm for Landsat 8 OLI. *IEEE Geoscience and Remote Sensing Letters*, 15(10), 1610-1614.

Tian, X., Liu, Q., Song, Z., Dou, B., & Li, X. 2018. Aerosol optical depth retrieval from Landsat 8 OLI images over urban areas supported by MODIS BRDF/Albedo Data. *IEEE Geoscience and Remote Sensing Letters*, 15(7), 976-980.

Zhu, Z., & Woodcock, C. E. 2012. Object-based cloud and cloud shadow detection in Landsat imagery. *Remote sensing of environment*, 118, 83-94.

*Revised January 2020*