A New Cloud Detection Algorithm for PROBA-V

Marian-Daniel Iordache, Luc Bertels, Else Swinnen, Wouter Dierckx, Erwin Wolters, Iskander Behadj, Dennis Clarijs, Carolien Toté, Tom Van Roey

Flemish Institute for Technological Research (VITO), Mol, Belgium

Abstract

This poster illustrates a new cloud detection algorithm designed for the multispectral PROBA-V instrument. The methodology to generate the algorithm, auxiliary input data, algorithmic steps, and examples of obtained cloud masks are shown. The presented figures show high quality of the obtained cloud masks in heterogeneous scenes and for different cloud types.

Context

Challenges to tackle for an accurate cloud detection algorithm for PROBA-V:

- The reduced **number of spectral bands**; PROBA-V has 4 spectral bands (central wavelengths): 464 nm (**B**lue), 655 nm (**R**ed), 837 nm (**N**ear-**I**nfra**R**ed) and 1603 nm (**S**hort-**W**ave **I**nfra**R**ed)
- No **thermal** spectral band
- **Temporal shift** of 12 seconds between the first three bands and the **SWIR** band
- The **heterogeneity** of Earth surface (*e.g.*, land, water, wetland, ice/snow); **customized** cloud detection for different surface types is needed

The newly designed scheme relies on auxiliary data to guide the detection towards specific sets of tests based on band thresholding, reflectance ratios, and similarity indices. The classification scheme is trained on a large database of pixels and yields binary classification maps in which the cloudy pixels are flagged [1].

Proposed Algorithm

Input/output data

The input data are: PROBA-V *image*, *status* map and reflectance *reference map* for the Blue spectral band. The output is the cloud cover mask.

For each pixel of the image: status map = surface type (land, water, snow/ice, unknown type) [3]; reference map = reflectance value characterizing the pixel Blue reflectance with high probability at the respective time of the year (monthly) (see [4]). Both input maps are derived from MERIS satellite data. Missing data in the reflectance reference map are replaced with albedo values delivered by the GlobAlbedo product [4,5].

Training data

Training database: 80,000+ pixels, manually built by VITO. Heterogeneous subsets of 20,000 pixels, containing cloudy, semi-cloudy and clear pixels of different land covers were randomly selected for supervised training [2]. The proposed per-pixel classification scheme is shown in Figure 1.

Proposed per-pixel classification scheme

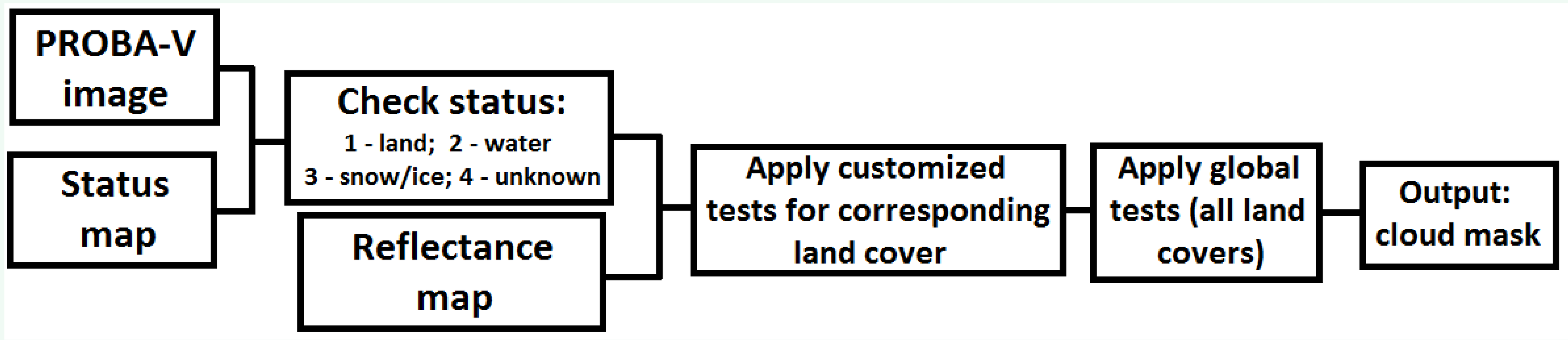


Figure 1: Proposed cloud detection scheme

The **customized tests** include:

- Band thresholding
- Difference between observed and reference reflectance values
- Band ratios thresholding
- Similarity checks using the Spectral Angle Distance (SAD) to reference spectra

The **global tests** apply to all land covers: test on the difference in amplitude between reference and observed spectra.

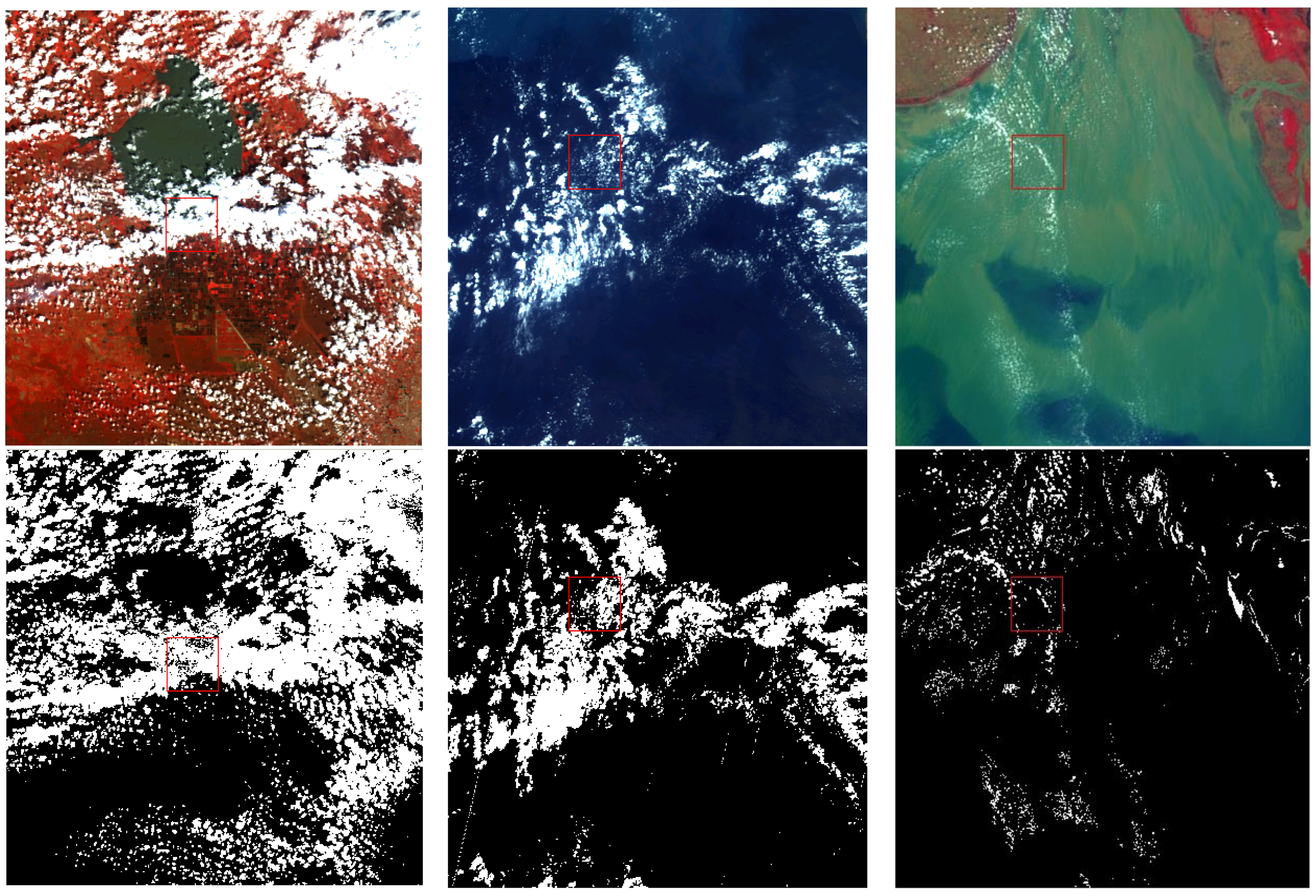
Parameter tuning All parameters in the classification scheme were fine-tuned to achieve a minimum classification error in the training datasets. **Table 1** shows the overall average classification errors over 5 training sets and 5 test sets (40,000 pixels each) selected from the initial training database. **Figure 2** displays examples of cloud masks obtained in PROBA-V data.

Data	Underdetection [%]	Overdetection [%]	Overall [%]
Training	8.74	10.17	8.93
Testing	8.96	10.19	9.04

Table 1 Classification errors of the cloud detection algorithm

Cloud detection: Examples

CLOUDS OVER LAND | CLOUDS OVER WATER | CLOUDS OVER TURBID WATER



Conclusions and validation plan

A new cloud detection algorithm for PROBA-V multispectral data was presented. The algorithm was trained following a supervised scheme using training pixels manually selected and labeled from the satellite imagery. The presented quantitative performances show that the overall errors are lower than 10% in the testing data. The cloud masks obtained in several PROBA-V images confirm the high quality of the cloud mask. Ongoing work at VITO includes parameter refining and in-depth testing on newly acquired PROBA-V imagery.

The obtained cloud masks will be thoroughly validated with various datasets, in particular with the Meteosat-SEVIRI High Resolution Visible (HRV) cloud mask [6] for 2014. Using SEVIRI's Rapid Scan Service, this cloud mask is available every 5 minutes and will thus give minimum observation time differences for the comparison.

References

[1] Jedlovec, G. (2009), *Automated Detection of Clouds in Satellite Imagery, Advances in Geoscience and Remote Sensing*, Gary Jedlovec (Ed.), InTech, DOI: 10.5772/8326., available from: http://www.intechopen.com/books/advances-in-geoscience-and-remote-sensing/automated-detection-of-clouds-in-satellite-imagery

[2] Mohri, M., Rostamizadeh, A. and Talwalkar, A. (2012), *Foundations of Machine Learning*, The MIT Press, ISBN: 9780262018258

[3] European Space Agency, *GlobCover Project*, available online: http://due.esrin.esa.int/page_globcover.php (accessed on 18 April 2016)

[3] European Space Agency (ESA), *MERIS Global Land Surface Albedo Project,* available online: http://www.brockmann-consult.de/albedomap/index.html (accessed on 18 April 2016)

[4] European Space Agency (ESA), *GlobAlbedo Project,* available online: http://www.globalbedo.org (accessed on 18 April 2016)

[5] Muller, J.-P., et al. (2012), The ESA GlobAlbedo Project for mapping the Earth's land surface albedo for 15 Years from European Sensors., IEEE Geoscience and Remote Sensing Symposium (IGARSS) 2012, IEEE, Munich, Germany, 22-27.7.12.

[6] Bley, S., & Deneke, H. (2013). A threshold-based cloud mask for the high-resolution visible channel of Meteosat Second Generation SEVIRI. *Atmos. Meas. Tech.*, *6*(10), 2713-2723.