



Clouds Detection Algorithms for Proba-V

Proba-V Clouds Detection Round Robin Experiment

Algorithm Theoretical Basis Document

Cloud Masking ATBD

L. Gómez-Chova, G. Mateo-García, J. Muñoz-Marí, and G. Camps-Valls

Universidad de Valencia, Valencia (Spain)



VNIVERSITAT DE VALÈNCIA

December 15, 2016

Contents

Acronyms and Abbreviations	4
Abstract	6
1 Introduction	6
1.1 Review of Cloud Masking Methods	6
2 Proba-V Data	8
2.1 Manually Labelled Training Set	9
3 Feature Extraction, Feature Selection, and Sample Selection	10
3.1 Feature Extraction	10
3.2 Feature Selection	14
3.3 Sample Selection	17
4 Cloud Masking Algorithm	18
4.1 Supervised Classification Algorithms	18
4.2 Cloud Classification Results	20
5 Summary and Conclusions	24
Bibliography	25
Appendix	29

Acronyms and Abbreviations

ATBD Algorithm Theoretical Basis Document

BC Brockmann Consult

BEAM Basic ERS & Envisat (A)ATSR and MERIS toolbox

BELSPO Belgian Federal Science Policy Office

CART Classification And Regression Trees

CCI Climate Change Initiative

CE Commission Error

COT Cloud Optical Thickness

CTP Cloud Top Pressure

DEM Digital Elevation Model

EO Earth Observation

ESA European Space Agency

ESRIN European Space Research Institute

FAR False Alarm Rate

FOV Field of View

HDF5 Hierarchical Data Format 5

HR Hit Rate

IDEAS+ Instrument Data quality Evaluation and Analysis Service

IR InfraRed

MEP Mission Exploitation Platform

MLP MultiLayer Perceptron

NDVI Normalized Difference Vegetation Index

NetCDF Network Common Data Form

NIR Near InfraRed

NN Neural Network

OA Overall Accuracy

OE Omission Error

PA Producer's Accuracy

PDF Portable Document Format

POD Probability of detection

PROBA-V Project for on-board Autonomy-Vegetation

PV-CDRR Proba-V Cloud Detection Round Robin

QA Quality Assessment

QC Quality Control

SNAP Sentinel Application Platform

SVM Support Vector Machine

SWIR ShortWave InfraRed

SZA Solar Zenith Angle

TOA Top Of Atmosphere

UA User's Accuracy

UV University of Valencia

VIS Visible

VITO Flemish institute for technological research

VZA View Zenith Angle

Abstract

This document describes the theoretical basis of the cloud detection algorithm developed by the University of Valencia for the study ‘*Clouds Detection Algorithms for Proba-V*’ in the framework of the Proba-V Clouds Detection Round Robin (PV-CDRR) experiment. Accurate and automatic detection of clouds in satellite scenes is a key issue for a wide range of remote sensing applications. With no accurate cloud masking, undetected clouds are one of the most significant sources of error in both sea and land cover biophysical parameter retrieval. The objective of the algorithms presented in this document is to detect clouds accurately providing a cloud flag per pixel. For this purpose, the method exploits the information of Proba-V using statistical machine learning techniques to identify the location of clouds present in Proba-V products.

1 Introduction

The main objective of this document is to define and implement a cloud masking scheme for Proba-V. Images acquired by Proba-V instrument (Dierckx et al., 2014), which works in the visible and infrared (VIS-IR) ranges of the electromagnetic spectrum, may be affected by the presence of clouds. The corresponding cloud influence depends on the cloud type, cloud cover, cloud height, and cloud distribution in the sky; e.g. thick opaque clouds impede the incoming radiation reaching the surface, while thin transparent clouds contaminate the data by photons scattered in the observation direction, or attenuate the signal by the removal of photons in their travel to the sensor. As a result, non-masked clouds would lead to significant errors in retrieved bio-physical parameters.

1.1 Review of Cloud Masking Methods

Cloud masking approaches, also referred to as cloud masking or detection, are generally based on the assumption that clouds present some useful characteristics for its identification: clouds are usually brighter and colder than the underlying surface, the spectral response is different from that of the surface covers, and cloud height produces a shorter optical path thus lowering atmospheric absorptions. The simplest approach to cloud detection in a scene is the use of a set of static thresholds (e.g. over reflectance or temperature) applied to every pixel in the image, which provides a cloud flag (binary classification) (Wang and Shi, 2006). However, these methods can fail for several reasons, such as subpixel clouds, high reflectance surfaces, illumination and observation geometry, sensor calibration, variation of the spectral response of clouds with cloud type and height, transparency of clouds, etc.

- On the one hand, the thickest clouds should be easily detected and masked out from visible and near-infrared Proba-V bands, but this is not true for thin clouds, which are semitransparent to solar radiation. Moreover, bright pixels, such as ice and snow in the surface, can be misclassified as clouds. Bright land covers and clouds have a similar reflectance behavior, thus thresholds on reflectance values do not solve the problem.
- On the other hand, signal coming from optically-thin semitransparent clouds is mostly affected by surface contribution, and it ranges from very low to extremely high values depending on whether the cloud is over water or ice, respectively. Therefore, they are extremely difficult to detect from reflectance properties in VNIR data.

These problems preclude the use of simple approaches based on static thresholds and suggest the use of more advanced cloud masking methods (Gómez-Chova et al., 2007a,b).

In the literature, the simplest approach to mask clouds in a particular scene is the use of a set of static thresholds (e.g. over features such as albedo or temperature) applied to every pixel in the image and ultimately providing a binary flag (Wang and Shi, 2006). These methods can fail for several reasons, such as subpixel clouds, high reflectance surfaces, illumination and observation geometry, sensor calibration, variation of the spectral response of clouds with cloud type and height, etc. (Simpson et al., 1998). Spatial coherence methods have an advantage over static threshold methods because they use the local spatial structure to determine cloud free and cloud covered pixels. Usually, these algorithms are based on extracted textural features (Tian et al., 1999; Christodoulou et al., 2003), contextual approaches (Papin et al., 2002), or simple thresholds applied to the spatial variability of spectral bands (Martins et al., 2002), which is mainly applicable over the ocean where the surface background is sufficiently homogeneous. However, spatial coherence methods can fail (Jones et al., 1996) when the cloud system is multi-layered (which is often the case), the clouds over the scene are smaller than the instrument spatial resolution, or the scene presents cirrus clouds (which are not opaque). As a consequence, researchers have turned to developing adaptive threshold cloud-masking algorithms (Simpson et al., 1998; Di Vittorio and Emery, 2002; Yang et al., 2007). Some other algorithms take advantage of the multi-angular (Yang et al., 2007; Mazzoni et al., 2007) or the multi-temporal (Saitwal et al., 2003) information depending on the instrument characteristics and the application constraints. Given the extreme complexity of cloud masking, most of the operational cloud masking applications of current satellite multispectral systems, such as Moderate Resolution Imaging Spectroradiometer (MODIS) (Ackerman et al., 1998; Platnick et al., 2003), Advanced Very High Resolution Radiometer (AVHRR) (Saunders and Kriebel, 1988), Along Track Scanning Radiometer (ATSR) (Simpson et al., 1998), POLarization and Directionality of the Earth's Reflectances (POLDER) (Buriez et al., 1997), or MERIS (Santer et al., 1997),

consist in a series of cloud detection threshold tests, which can vary depending on surface type, solar illumination, geographic location, or climatological criteria. In particular, current Proba-V cloud detection uses multiple thresholds applied to the blue and the SWIR spectral bands (Lisens et al., 2000), but the definition of global thresholds is practically impossible. Hence, for next Proba-V reprocessing (Wolters et al., 2015), monthly composites of cloud-free reflectance in the blue band are used to define dynamic thresholds depending on the land cover type.

In this context, few works using more sophisticated machine learning tools have been presented so far, such as Bayesian methods (Murtagh et al., 2003; Li et al., 2003; Merchant et al., 2005), fuzzy logic (Ghosh et al., 2006), artificial neural networks (Yhann and Simpson, 1995; Tian et al., 1999; McIntire and Simpson, 2002; Torres Arriaza et al., 2003), or recently kernel methods (Srivastava and Stroeve, 2003; Lee et al., 2004; Mazzoni et al., 2007; Gómez-Chova et al., 2007b). The proposed cloud masking method might take the best elements of these algorithms, but basically it must be an advanced non-linear method capable of exploiting the information of Proba-V features in order to improve the cloud masking products.

2 Proba-V Data

Proba-V instrument presents a limited number of spectral bands (Blue, Red, NIR and SWIR) which makes cloud detection particularly challenging since it does not present thermal channels or a dedicated cirrus band.

The main characteristics of the Proba-V products are the following:

- Swath 2250 km
- Resolution 100 m at nadir, 350 m across full field of view
- Spectral bands:
 - Blue (438-486nm)
 - Red (615-696nm)
 - NIR (772-914nm)
 - SWIR (1564-1634nm)

In particular, for this Round Robin exercise, we consider as input Proba-V Level 2A products with TOA reflectance in the four Proba-V bands, radiometrically and geometrically corrected and resampled at 333m.

The available data set consists of 331 products acquired in four days covering the four seasons: 21/03/2014, 21/06/2014, 21/09/2014, and 21/12/2014. From this dataset, a reduced set of Proba-V data for a number of representative sites worldwide (Fig. 1) is used to train the algorithm and validate its performance.

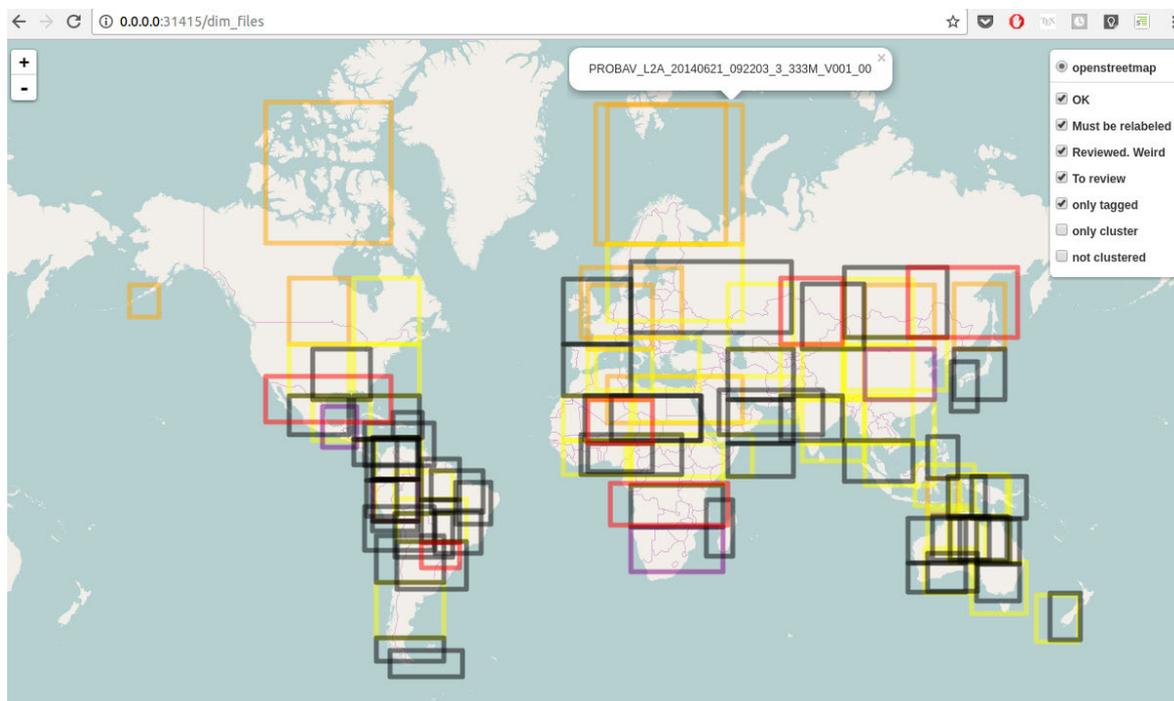


Figure 1: Proba-V images selected for validation of the cloud masking algorithm: geographic location and dates.

2.1 Manually Labelled Training Set

In order to train statistical machine learning models from real data, a representative number of samples have to be labelled as *cloud-contaminated* or *cloud-free* samples. To label in a semi-automatic way a sufficient number of pixels from the Proba-V images, we have adapted the user-driven methodology proposed in [Gómez-Chova et al. \(2007a\)](#) for MERIS to the Proba-V images, where the labeling of cloud clusters found in the image is done by the user.

Figure 2 shows the interactive interface used to generate the Ground Truth by an expert user. In addition to label the clusters as ‘cloudy’ or ‘cloud-free’, it allows to label the pixels as ‘bright clouds’, ‘clouds’, ‘cirrus’, ‘shadows’, ‘vegetation’, ‘soil’, ‘water’, ‘ice/snow’, ‘sand’; which can be a useful information for the generation of a comprehensive training set with samples of all relevant land-covers.

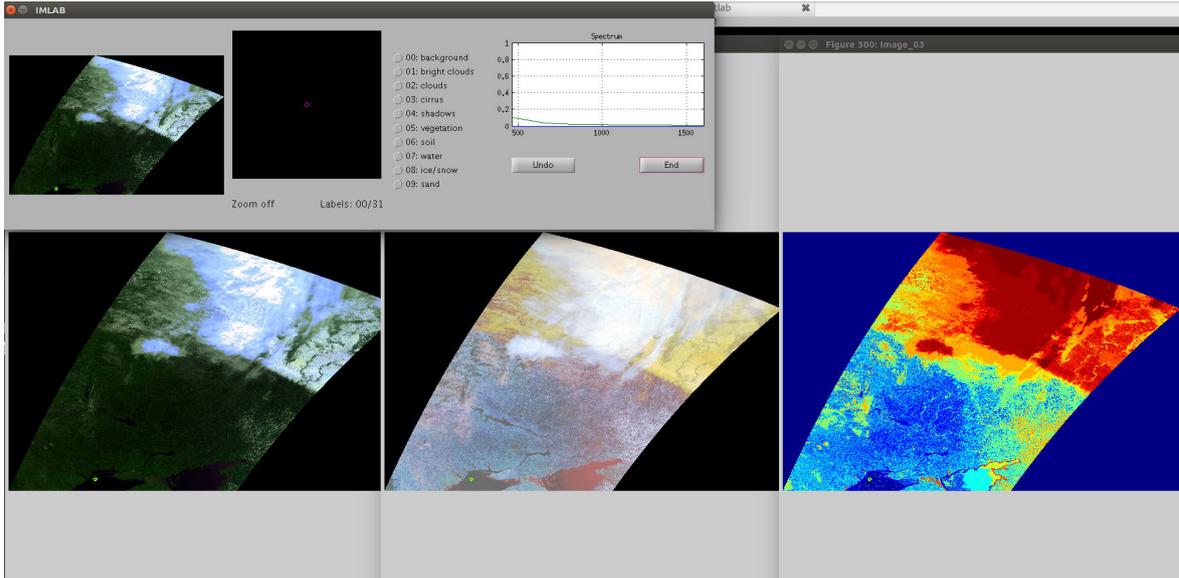


Figure 2: Manual Ground Truth generation: cloudy/cloud-free labels

Figure 3 shows an example of the resulting products generated using this user-driven methodology. First row shows a) the RGB false color composite and b) the class map of the labels of the clusters. The second row shows c) the membership (posterior probability map) of each pixel to the cloud clusters, which is also useful to include different types of cloudy pixels in the training set, and d) the spatial distribution of all clusters.

3 Feature Extraction, Feature Selection, and Sample Selection

The following sections describe the feature extraction, feature selection, and sample selection approaches applied to the dataset in order to provide a high quality training set to the machine learning algorithms.

3.1 Feature Extraction

Several physically-inspired features can be extracted from the spectrum before applying the classification methods in order to improve their performance.

The measured spectral signature depends on the illumination, the atmosphere, and the surface. Figure 4 shows Proba-V channel locations compared with the spectral curve of healthy vegetation and bare soil.

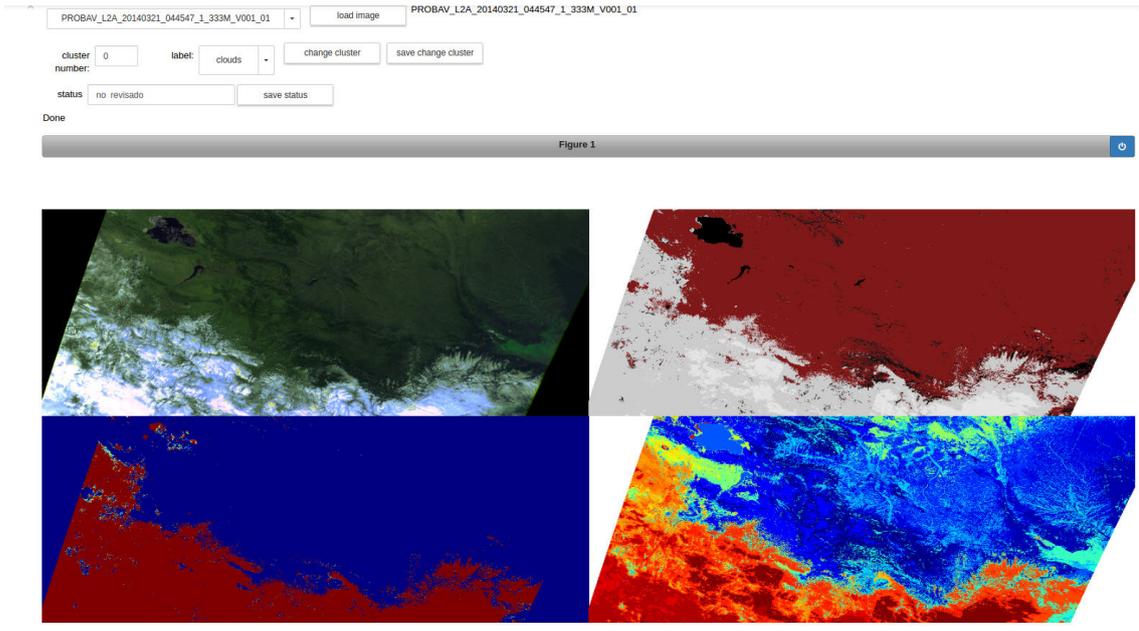


Figure 3: Manual Ground Truth validation: visual inspection and clustering

In this project, we take advantage of previous research and, rather than working with the spectral reflectance only, physically-inspired features are extracted in order to increase the separability of clouds and surface covers. A detailed analysis of the most important extracted features follows.

Spectral Brightness and Whiteness Regarding the reflectance of the surface, one of the main characteristics of clouds is that they present *bright* and *white* spectra (Fig. 5). We can exploit Proba-V channels for extracting information about the target reflectance, i.e. cloud brightness and cloud whiteness for cloudy pixels:

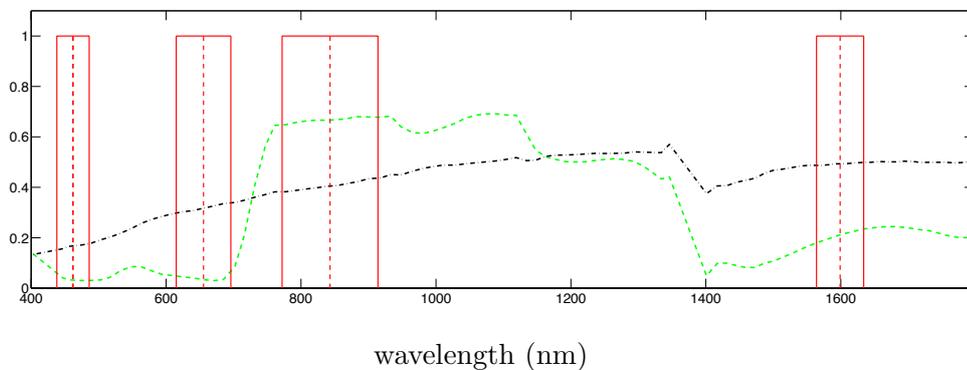


Figure 4: Proba-V channel locations (boxes) superimposed to a reflectance spectra of healthy vegetation and bare soil (dash-dotted line).

-
- A *bright* spectrum means that the intensity of the spectral curve (related to the albedo) should present relatively high values. Therefore, cloud brightness is calculated for each pixel as the integral of spectrum, $f_{Br} = \int \rho(\lambda)d\lambda$, which is approximated through trapezoidal numerical integration:

$$\hat{f}_{Br} = \frac{1}{\lambda_{max} - \lambda_{min}} \sum_{\lambda_i} \frac{\rho(\lambda_{i+1}) + \rho(\lambda_i)}{2} (\lambda_{i+1} - \lambda_i), \quad (3.1)$$

which has the same units as $\rho(\lambda)$, and differs from the average of the spectral channels since it takes into account the distribution of the energy along the spectrum.

- A *white* spectrum means that the spectral signature must be flat along the spectrum. The first derivative of the spectral curve should present low values, but noise and calibration errors may reduce the accuracy in the estimation of the spectrum flatness when computing the spectral derivative in channels with similar wavelengths. Therefore, we compute for each pixel the deviation from the flatness as the (trapezoidal approximate) integral of $e(\lambda) = |\rho(\lambda) - \hat{f}_{Br}|$:

$$f_{Wh} = \frac{1}{\lambda_{max} - \lambda_{min}} \sum_{\lambda_i} \frac{e(\lambda_{i+1}) + e(\lambda_i)}{2} (\lambda_{i+1} - \lambda_i) \quad (3.2)$$

Further surface features can be obtained by considering independently the VIS and NIR spectral ranges, where surface covers present different reflectance properties. Therefore, instead of working with f_{Br} and f_{Wh} , we can obtain 2 + 2 features from (3.1) and (3.2) respectively: $f_{Br,VIS}$ and $f_{Wh,VIS}$; and $f_{Br,NIR}$ and $f_{Wh,NIR}$. For example, clouds over land should be better recognized in $f_{Br,VIS}$ than in $f_{Br,NIR}$ since land covers have less reflectance in the VIS range, while the opposite is true for clouds over sea.

Spatial Features Together with the spectral features, spatial features are extracted at different scales: the mean (μ) and standard deviation (σ) are computed for each pixel-based feature at two different scales in 3x3 and 5x5 windows.

List of Extracted Features

Previous sections have introduced most of the characteristics that are useful to discriminate clouds from surface. However, these are only some of the features presented in the literature to identify clouds. The final set of features that are analyzed in the frame of this project are listed in Table 1.

Summarizing, we consider the four Proba-V spectral channels (4), the spectral features described in Table 1 (10), and the mean (μ) and standard deviation at two different scales,

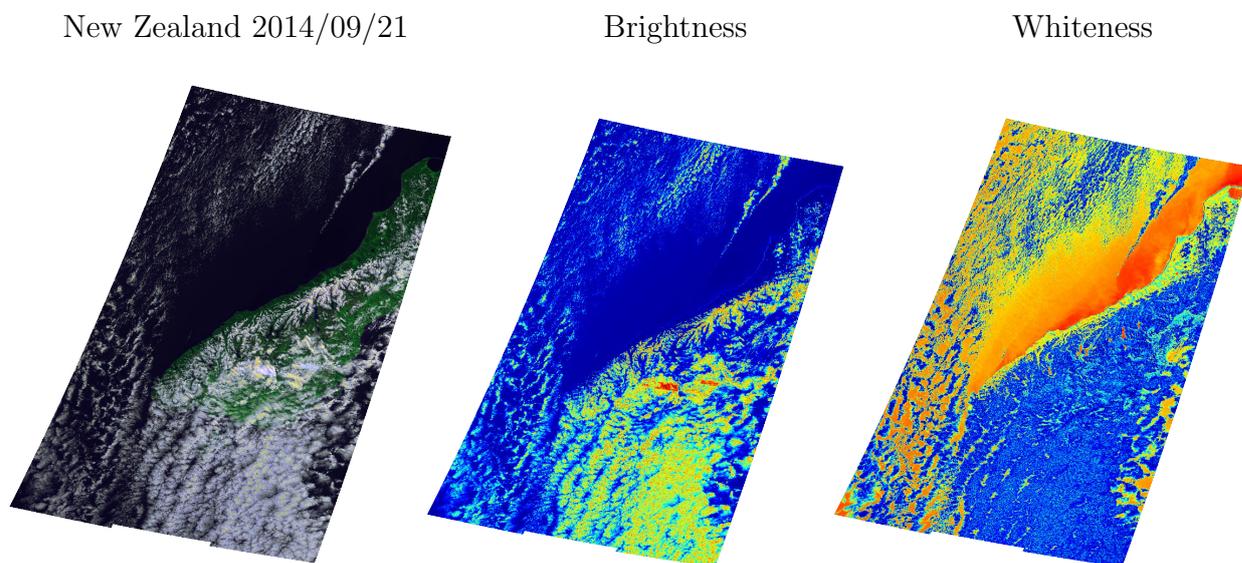


Figure 5: Cloud brightness and whiteness features extracted from the TOA reflectance of a Proba-V image.

Table 1: Cloud features extracted from Proba-V images.

Cloud Feature	Feature	Channels
Brightness	f_{Br}	(Blue,Red,NIR,SWIR)
Brightness VIS	$f_{Br,VIS}$	(Blue,Red)
Brightness NIR	$f_{Br,NIR}$	(NIR,SWIR)
Whiteness	f_{Wh}	(Blue,Red,NIR,SWIR)
Whiteness VIS	$f_{Wh,VIS}$	(Blue,Red)
Whiteness NIR	$f_{Wh,NIR}$	(NIR,SWIR)
Snow NDSI	$f_{(Blue-NIR)/(Blue+NIR)}$	(Blue,NIR)
Snow NDSI	$f_{(Blue-SWIR)/(Blue+SWIR)}$	(Blue,SWIR)
Red-SWIR ratio	$f_{Red/SWIR}$	(Red,SWIR)
NDVI	$f_{(NIR-Red)/(NIR+Red)}$	(Red,NIR)

which are computed for each pixel-based feature $((4 + 10) \times 4)$. That results in a total number of 70 possible input features (4 reflectance bands, 10 spectral features, 56 spatio-spectral features).

3.2 Feature Selection

Machine learning methods have proven to be effective techniques in classification of remote sensing data. However, one of the main problems with multispectral image processing is the huge amount of data involved. Certainly, this is a major problem for pattern recognition methods, since they are sensitive to problems associated to high-dimensional feature spaces - known as Hughes phenomenon or curse of dimensionality (Hughes, 1968). The close relationship between the complexity of the classifier and the size of the training set suggests the idea of a prior reduction of the input space in order to avoid wrong estimations of the classifier parameters (Fukunaga and Hayes, 1989). This objective can be achieved by two different ways. The first one is to identify those variables that do not contribute to the classification task and to ignore them (feature selection). The second approach is to find a transformation to a lower dimension feature space while retaining the useful information for the classification problem (feature extraction).

Feature selection methods have several advantages compared to feature extraction: data transmission and storage (only selected features are required); interpretability of results (the selected features are spectral bands with physical meaning); and extrapolation of results to other spectrometers with slightly different spectral bands. We propose a dimensionality reduction strategy that eliminates redundant information, by means of local correlation criterion between contiguous spectral bands, and a subsequent selection of the most discriminative features based on combination of several feature ranking methods.

The Feature Selection Problem

The Feature Selection Problem (FSP) in a ‘learning from samples’ approach can be defined as choosing a subset of features that achieves the lowest error according to a certain loss functional. Following a general taxonomy, the FSP can be tackled using *filter* (Blum and Langley, 1998) and *wrapper* (Kohavi and John, 1997) methods. Filter methods use an indirect measure of the quality of the selected features, e.g. evaluating the correlation function between each input feature and the observed output. A faster convergence of the algorithm is thus obtained. On the other hand, wrapper methods use as selection criteria the goodness-of-fit between the inputs and the output provided by the learning machine under consideration, e.g. a neural network. This approach guarantees that, in each step of the algorithm, the selected subset improves performance of the previous one. Filter methods

might fail to select the right subset of features if the used criterium deviates from the one used for training the learning machine, whereas wrapper methods can be computationally intensive since the learning machine has to be retrained for each new set of features.

In the first part of the study, we use filter methods to obtain a ranking of the most relevant important bands. With these methods we quantitatively measure the relationship between the spectral bands (input features) and the presence vs. absence of clouds.

Filter Methods: Feature Ranking Methods

Several additional methods are used for feature ranking. They quantitatively measure the relationship between each band and the desired output and rank key features by class separability criteria (independent evaluation criterion for binary classification). Since they work under different learning paradigms (multivariate statistics, information theory, and detection and estimation frameworks), the obtained pooled of ranking methods may be potentially useful.

- *The t -test estimate.* Absolute value two-sample t -test with pooled variance estimate (Liu and Motoda, 1998).
- *The classification Chernoff bound.* This is the well-known Bhattacharyya distance. Minimum attainable classification error, also known as the Chernoff bound (Theodoridis and Koutroumbas, 1999).
- *The Wilcoxon test.* This method is basically the absolute value of the u -statistic of a two-sample unpaired Wilcoxon test, also known as the Mann-Whitney test (Liu and Motoda, 1998).
- *The ROC curve.* This is the area under the empirical receiver operating characteristic (ROC) curve (Theodoridis and Koutroumbas, 1999).
- *The Kullback-Leibler divergence.* This is the relative *entropy*, also known as Kullback-Liebr distance or divergence (KLD). The measure, which is based on solid information theory foundation, yields an estimate of shared information content between each band and the output (Liu and Motoda, 1998; Peng et al., 2005).
- *The minimum redundancy and maximal relevance (mRMR) criterion.* The method jointly optimizes the *maximal relevance* of the features and the *minimum redundancy* among them. The Max-Relevance is approximated with the mean value of all mutual information values between individual features and the observed class label. When

two features highly depend on each other, the respective class-discriminative expressive power should not change if one were removed. Therefore, a Min-Redundancy condition can be added to select mutually exclusive features (Peng et al., 2005).

Note that ‘t-test’, ‘entropy’, and ‘Bhattacharyya’ methods assume normal distributed classes while ‘ROC’ and ‘Wilcoxon’ are non-parametric tests. Also, we should stress that most tests are feature independent.

Wrapper Methods: Sequential Feature Selection

Most of the presented ranking methods in the previous section analyze each band independently and using as selection criteria measures of class separability instead of using the classification accuracy directly. In this section, more attention is paid to wrapper methods for feature selection (that jointly analyze the feature space) in the results. Therefore, to conclude the analysis about the spectral bands relevance, we explore a wrapper feature selection method that uses as selection criteria the classification accuracy of the classification method for all the possible number of selected bands and most band combinations.

The Sequential Forward Selection (SFS) algorithm identifies those bands that better discriminate among classes in a two-stage iterative feature selection process. Firstly, a search strategy for feature group selection is carried out, and secondly, the objective function that evaluates the different subgroups is calculated. In our case, we use as objective function the classification accuracy (kappa statistic) of the classifier trained with the spectral bands and features under study. The SFS method will produce a hierarchy of the input bands by selecting first the best individual band and then adding the next best band (given the already selected bands) for each dimension. The SFS algorithm presents the following useful characteristics:

- Sequential: Sequential algorithms present better computational requirements than the algorithms that perform an exhaustive search. The latter are infeasible working with a high number of input features (exponential growth).
- Forward: The forward search begins with one variable and continues adding features. It presents a couple of practical advantages with regard to the backward search. Firstly, if the initial dimension is very high, the calculation of the criterion can be complex or impossible (for example the estimation of the covariance matrix from few samples). Secondly, if the process is long and interrupted, the best subsets for small dimensions are already obtained.
- Selection: Feature selection is well-suited to our problem since results can be directly interpreted and extrapolated to other sensors or instrument configuration changes.

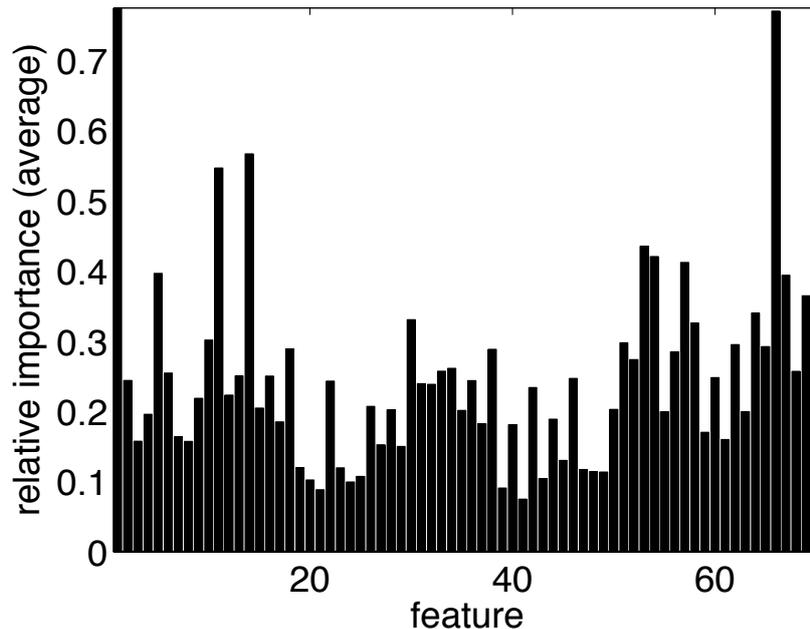


Figure 6: Ranking of the extracted features: relative importance between 0 and 1.

Feature Selection Results

Attending to the ranking of the features, one can evaluate the performance of different classifiers for different subsets of relevant features, which are selected using both filter and wrapper approaches. The set of features analyzed consist of the 4 Proba-V TOA reflectance bands, the 10 cloud features extracted from spectral bands (pixel-based), and the corresponding 14×4 spatial features (see Table 1 for details).

Figure 6 shows the relative importance ($[0,1]$) of the extracted features. Using this ranking we define two sets with the first 20 and 40 best features that will be compared in the classification experiments. The ten most relevant features are: $\rho_{TOA,Blue}$, $\sigma_{Wh,NIR}^{5 \times 5}$, $f_{Wh,VIS}$, $f_{Red/SWIR}$, $\mu_{(NDVI)}^{5 \times 5}$, $\sigma_{(NDVI)}^{5 \times 5}$, $\mu_{Red/SWIR}^{5 \times 5}$, $f_{(Blue-NIR)/(Blue+NIR)}$, $\mu_{Wh,VIS}^{3 \times 3}$, $\mu_{Wh,VIS}^{5 \times 5}$.

3.3 Sample Selection

The samples of the dataset can be used to train a supervised classifier taking into account the sample labels ('cloudy' or 'cloud-free'). However, when the number of training samples is high, the training of the model can be computationally time consuming. In this cases, the number of samples used in the training is reduced by means of subsampling or sample selection techniques. For this purpose, we reduce the number of samples while preserving the original data distribution and complexity.

4 Cloud Masking Algorithm

The method should be capable of detecting clouds accurately. The cloud masking process relies on the extraction of meaningful physical features (e.g. brightness, whiteness, temperature) that are combined with spatial features to increase the cloud detection accuracy. The supervised classifiers to be developed and tested on this data should allow the use of a high number of input features and allow an easy integration of heterogeneous sources of information.

Firstly, a feature extraction based on meaningful physical facts is carried out (clouds are bright, white, high, etc.). Then, a supervised pixel-based classification, based on the TOA reflectance and on a manually labeled training set, is applied to these features providing the pixel label (*cloud* or *cloud free*).

4.1 Supervised Classification Algorithms

The extracted features and the original spectral bands are used as inputs of advanced supervised classification algorithms, which are required to solve complex classification problems such as cloud masking.

The detection of clouds can be considered as a two-class classification problem. In these problems, we are given a set of ℓ labeled (training) samples $\{\mathbf{x}_i, y_i\}_{i=1}^{\ell}$, where $\mathbf{x}_i \in \mathbb{R}^d$ is defined in an input space \mathcal{X} , and $y_i \in \{0, 1\}$ belongs to the observation (output) space ('cloudy' or 'cloud free').

In this project, different classification and regression methods are analyzed: decision trees (Breiman et al., 1984), support vector machines (SVMs) (Vapnik, 1998; Schölkopf and Smola, 2002), and multilayer perceptron (MLP) neural networks (Duda et al., 1998; Haykin, 1999).

Decision Trees (TREE) Decision trees are simple, yet powerful, non-parametric classification techniques. There are different types of decision trees, such as Classification and Regression Trees (CART) (Breiman et al., 1984), Interactive Dichotomizer Version 3 (ID3) (Quinlan, 1986), and C4.5 algorithm (Quinlan, 1993). In this work, CART is used to classify cloudy pixels.

CART algorithm is a tree graph structure with a sequence of nodes that are partitioned or split into two branches by means of decision rules and each terminal node (leaf) is classified with the predicted value for that node. For each branch node, the left child node corresponds to the points that satisfy the condition, and the right child node corresponds to the points that do not satisfy such condition. CART presents some advantages over other

classification or regression techniques as high generalization, capability to divide a complex problem into a collection of simpler decisions making tree models simple to understand and interpret. A common problem of classification algorithms is the overfitting problem (Hughes' phenomenon) that produces poor generalization. Pruning and cross-validation methods are usually employed on CART algorithms to avoid over-fitting. In this work, a 10-fold cross-validation procedure is used to find the minimum-cost tree and to estimate the best level of pruning. Tree is pruned based on an optimal pruning scheme that first prunes branches giving less improvement in error cost.

Support Vector Machines (SVM) The Support Vector Machine (SVM) binary classifier is a statistical learning algorithm based on constructing a maximum margin separating hyperplane in a reproducing kernel Hilbert space (Vapnik, 1995). SVMs allow the use of a high number of input features as it combats the curse of dimensionality efficiently. In addition, they provide simple (sparse) solutions that can be interpreted (1) by inspecting the weights associated to the most relevant examples, and (2) by applying traditional saliency techniques.

Notationally, be \mathbf{V} a set of N observed and labeled data, $\mathbf{V} = \{(\mathbf{x}_1, y_1), \dots, (\mathbf{x}_N, y_N)\}$, where $\mathbf{x}_i \in \mathbb{R}^n$ and $y_i \in \{-1, +1\}$. Be $\phi(\mathbf{x}_i)$ a nonlinear transformation to a (generally unknown) higher dimensional space \mathbb{R}^B , where a separating hyperplane is given by $\langle \phi(\mathbf{x}_i), \mathbf{w} \rangle + b = 0$. The key in non-linear SVM is that both the solution and the solving procedure can be expressed as a function of dot products among samples, and thus, these products can be computed in the mapped space without even knowing the mapping ϕ . The dot products are known as kernel, $K(\mathbf{x}_i, \mathbf{x}_j) = \langle \phi(\mathbf{x}_i), \phi(\mathbf{x}_j) \rangle$ and has to fulfill Mercer's conditions, that is, the matrix has to be positive (semi-)definite for the formulation to be valid.

In this work, we use a Gaussian RBF kernel, given by $K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\frac{\|\mathbf{x}_i - \mathbf{x}_j\|^2}{2\sigma^2})$; and free SVM parameters are selected by following an 8-fold cross-validation procedure.

Multilayer Perceptron (MLP) The well-known MLP neural network Duda et al. (1998); Haykin (1999), which has been a traditional approach for supervised cloud classification (Yhann and Simpson, 1995; Tian et al., 1999; McIntire and Simpson, 2002; Torres Arriaza et al., 2003; Preusker et al., 2006), is also included in the comparison. The MLP is a feed-forward artificial neural network model that produces flexible input-output mappings. This artificial neural network (ANN) attempts to separate samples belonging to different classes by mapping samples \mathbf{x}_i to a one-dimensional output space o_i , through a highly non-linear mapping function consisting in a net of interconnected neurons with a given activation function arranged in layers and connected by synaptic weights, where $\hat{y}_i = 1$

(cloudy) when o_i is over a given threshold $b = 0.5$ (for our two-class problem). The MLP are trained using the Levenberg-Marquardt algorithm, which is more efficient in terms of computational cost than the standard back-propagation algorithm. The Levenberg-Marquardt algorithm approaches the Hessian matrix following a Newton-like updating, and thus it is no longer necessary to compute the whole Hessian matrix. This method is specially fast for training moderate-sized feedforward neural networks (Hagan and Menhaj, 1994). Different number of hidden neurons in the hidden layer are tested (2 to 30) and the best architecture is selected evaluating the averaged 3-fold cross-validation kappa statistic. In all the cases, the neurons of the hidden layer present the hyperbolic tangent sigmoid activation function while the neuron of the output layer presents a linear output function in order to better analyze the distribution of the output values, o_i , and to directly combine the outputs of several ANN.

4.2 Cloud Classification Results

In this section, the performance of different classifiers is evaluated for different subsets of spectral bands and extracted features.

Classification methods comparison

In order to decide which features are more relevant, the information of the spectral channels and the extracted features for cloud detection is analyzed in terms of classification accuracy of classification trees (TREE), multilayer perceptron (MLP) ANN, and support vector machines (SVM) in our problem. In the following experiments, classifiers are trained with different numbers of training samples and different combinations of features. Figure 7 shows the Overall Accuracy (OA) for each number of features, i.e. the classification accuracy of the TREES, MLPs and SVMs for the sets of selected features.

Several conclusions can be extracted from Fig. 7. Classification trees are very efficient classification algorithms but provide the less accurate detection results for all cases. SVMs provide excellent results when few training samples are available but present a huge computational cost when the number of samples increases (no SVM models have been trained for more than 15000 training samples per class). Finally, MLP neural networks provide excellent cloud detection accuracy and the extracted spatio-spectral cloud features drastically improve results, obtaining the best results with the top 40 selected features (sec. 3.2).

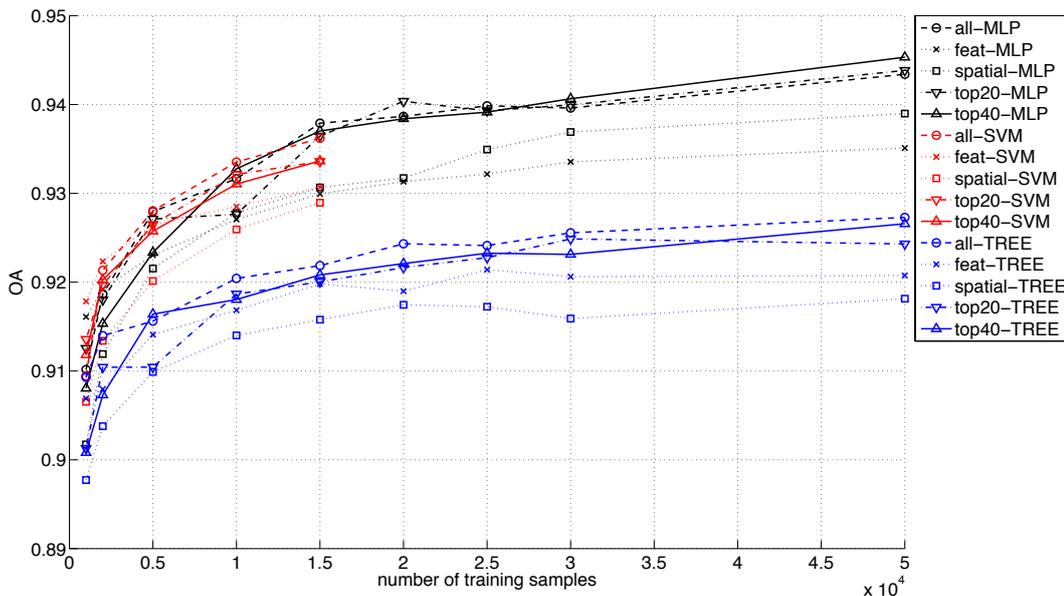


Figure 7: Overall Accuracy (OA%) over the test sets for the analyzed methods (TREE, SVM and MLP). The number of input features (spectral, spatial, and all features) and training samples per class (1000, 2000, 5000, 10000, 15000, 20000, 25000, 30000, 50000) vary for each test set.

Selected classifier performance

In the following, we focus only on MLPs, which have offered improved performance, using the top 40 selected features for Proba-V. Also, and given that accuracy is not improved too much with an increasing number of samples, we concentrate on using 50×10^3 training samples per class in all cases and classification accuracy is computed using 380×10^3 test samples.

Figure 8 shows the Overall Accuracy (OA%) and Cohen’s Kappa Statistics (κ) for the 54 validation images that have been manually labeled in order to be used as reference (ground truth). In this plot, one can observe that in most images the cloud detection accuracy is higher than 90%. This confirms that the trained MLP provides an excellent generalization over the analyzed images. However, it is important to remark that we are using as reference cloud mask a ‘ground truth’ that has been manually generated. Hence, the risk of learning and reproduce the errors present in the ground truth there always exists.

Finally, in order to better analyze the type of errors that we are committing, Table 2 shows the overall Confusion Matrix for the 54 validation images manually labeled. One can observe that the proposed method provides a balanced number of false negatives (FN) and

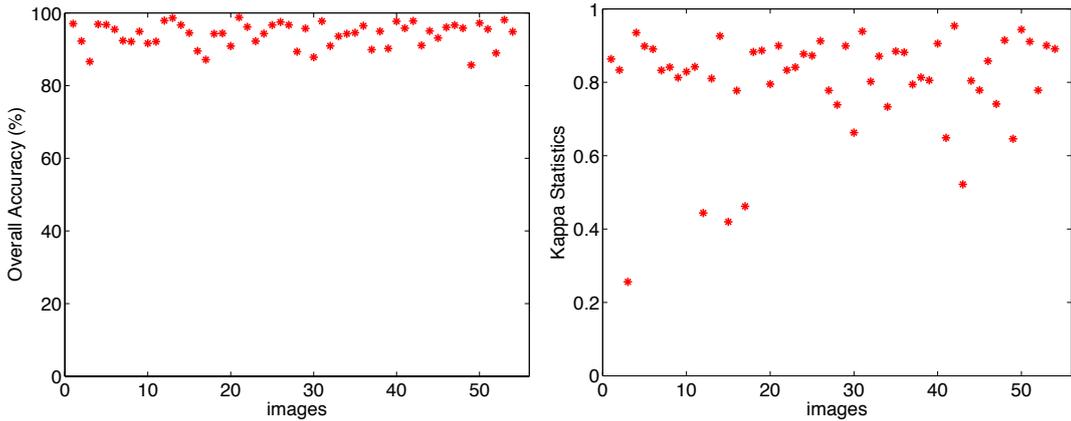


Figure 8: Overall Accuracy (OA%) and Kappa Statistics (κ) for the 54 validation images that have been manually labeled in order to be used as reference (ground truth).

false positives (FP), although the number of FN is relatively higher. However, the average Overall Accuracy for all images is 93% and we can consider that the agreement between the predicted cloud masks and the generated ground truth is high enough.

Table 2: Confusion Matrix for the 54 validation images that have been manually labeled in order to be used as reference (ground truth). True Negatives (TN), False Negatives (FN), True Positives (TP), False Positives (FP), Producer Accuracy (PA), User Accuracy (UA), and Overall Accuracy (OA).

	Manual Labels		Total	
	Cloud-free	Cloudy		
Predicted Cloud-free	TN:345308804	FN:25205511	370514315	PA:93%
Predicted Cloudy	FP:9133061	TP:191024152	200157213	PA:95%
Total	354441865	216229663	536332956	
	UA:97%	UA:88%		OA:93%

Cloud Product Validation Examples

An example of the results is shown in Fig. 9. In this figure, the proposed cloud mask is benchmarked against the manually generated ‘ground truth’ for a Proba-V product. This image has been selected because it illustrates most common cloud detection problems (e.g. cloud borders, thin clouds, ice/snow covers). In these plots, pixels detected as cloudy pixels by the cloud detection method but labeled as cloud free in the ‘ground truth’ are plotted in *blue*, while discrepancy pixels classified as cloud free but marked as clouds in the ‘ground truth’ are shown in *orange*.

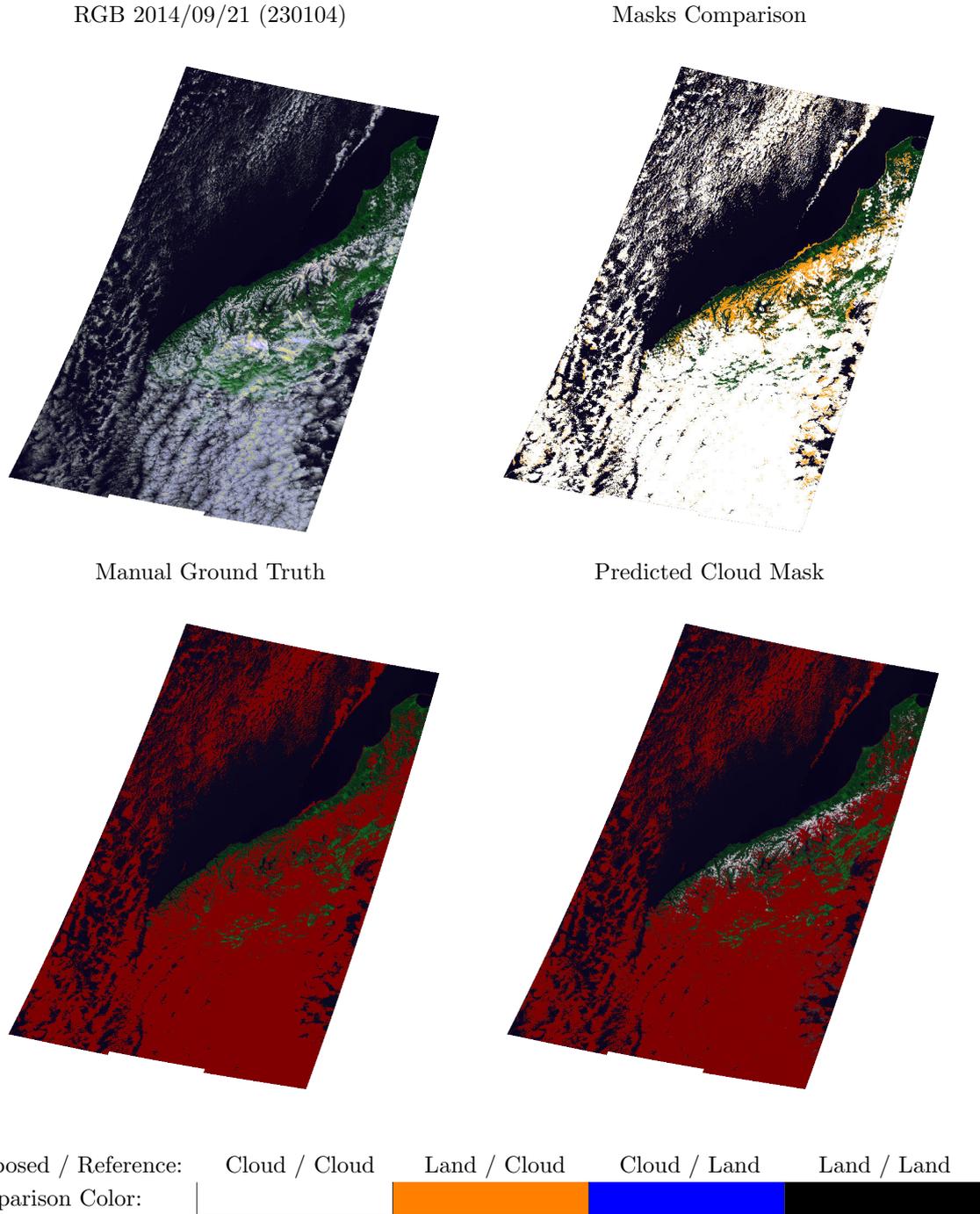


Figure 9: Cloud detection example showing the RGB false color composite, the manually generated ground truth, the cloud mask obtained with the final MLP classifier, and the comparison of the ‘manual ground truth’ with the ‘Proba-V Cloud Flag’: discrepancies are shown in blue when proposed method detects cloud and in orange when pixels are classified as cloud-free.

The pixels detected as cloud-free by the algorithm that are labeled as cloudy in the ‘ground truth’ (*orange*) show a good agreement with snow and glaciers over mountains in the South Island of New Zealand, and with some parts of the coastline. Therefore, one can assume that these differences are not errors but cloudy pixels, and that proposed method provides an accurate recognition in these difficult cases, although it is true that some thin cloud borders are also detected as cloud-free. However, in these cases a spatial growing of the cloud mask should easily improve the results. Validation results for the complete list of manually labeled Proba-V images can be found in the Appendix.

Cloud Mask Product

The output cloud product for the developed cloud masking method is provided in a HDF5 file that consists of a new band with the cloud flags. This band has the same spatial coverage than the original image and provides a prediction (‘cloudy’ or ‘cloud-free’) only over pixels with valid values for the four Proba-V spectral channels. Information about the corresponding flags coding is provided in Table 3.

Table 3: Bands of the generated cloud product.

Band	Type	Unit	Description
cloud mask	int16	1	Cloud flags (1: cloudy pixel, 0: cloud-free pixel, -1: invalid pixel).

5 Summary and Conclusions

In this document, a methodology that faces the problem of accurately identifying the location of clouds in Proba-V images is described. The cloud masking algorithm is based on simple spatio-spectral physical features, which are intended to increase separability between clouds and ground covers, and are extracted from the converted top of atmosphere reflectance in order to reduce dependence on illumination and geometric acquisition conditions. An supervised classification is performed based on the selected extracted features and selected training samples covering most relevant image conditions, background surfaces, and cloud types. In particular, several machine learning methods have been trained using different sets of input features and different sets of training samples in order to select the best empirical model. The final implemented method is based on artificial neural networks trained with manually labeled real data.

The performance of the method has been tested on a large number of real images and on scenes presenting most critical cloud detection problems, and results show an accurate discrimination of thin clouds, cloud borders, and bright surfaces.

Bibliography

- Ackerman, S., Strabala, K., Menzel, W., Frey, R., Moeller, C., and Gumley, L. (1998). Discriminating clear sky from clouds with MODIS. *Journal of Geophysical Research*, 103(D24):32141–32157.
- Blum, A. and Langley, P. (1998). Selection of relevant features and examples in machine learning. *Artificial Intelligence*, 97:245–271.
- Breiman, L., Friedman, J., Stone, C. J., and Olshen, R. A. (1984). *Classification and Regression Trees*. Chapman & Hall/CRC, Monterey, CA.
- Buriez, J. C., Vanbauce, C., Parol, F., Goloub, P., Herman, M., Bonnel, B., Fouquart, Y., Couvert, P., and Seze, G. (1997). Cloud detection and derivation of cloud properties from POLDER. *Int. Journal of Remote Sensing*, 18(13):2785–2813.
- Christodoulou, C. I., Michaelides, S. C., and Pattichis, C. S. (2003). Multifeature Texture Analysis for the Classification of Clouds in Satellite Imagery. *IEEE Trans. on Geoscience and Remote Sensing*, 41(11):2662–2668.
- Di Vittorio, A. and Emery, W. (2002). An automated, dynamic threshold cloud-masking algorithm for daytime AVHRR images over land. *IEEE Trans. on Geoscience and Remote Sensing*, 40(8):1682–1694.
- Dierckx, W., Sterckx, S., Benhadj, I., Livens, S., Duhoux, G., Achteren, T. V., Francois, M., Mellab, K., and Saint, G. (2014). Proba-v mission for global vegetation monitoring: standard products and image quality. *International Journal of Remote Sensing*, 35(7):2589–2614.
- Duda, R. O., Hart, P. E., and Stork, D. G. (1998). *Pattern Classification and Scene Analysis: Part I Pattern Classification*. John Wiley & Sons, 2nd edition.
- Fukunaga, K. and Hayes, R. (1989). Effects of sample size in classifier design. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 11(8):873–885.
- Ghosh, A., Pal, N., and Das, J. (2006). A fuzzy rule based approach to cloud cover estimation. *Remote Sensing of Environment*, 100:531–549.
- Gómez-Chova, L., Camps-Valls, G., Calpe, J., Guanter, L., and Moreno, J. (2007a). Cloud-screening algorithm for ENVISAT/MERIS multispectral images. *IEEE Trans. on Geoscience and Remote Sensing*, 45(12, Part 2):4105–4118. (DOI:[10.1109/TGRS.2007.905312](https://doi.org/10.1109/TGRS.2007.905312)).

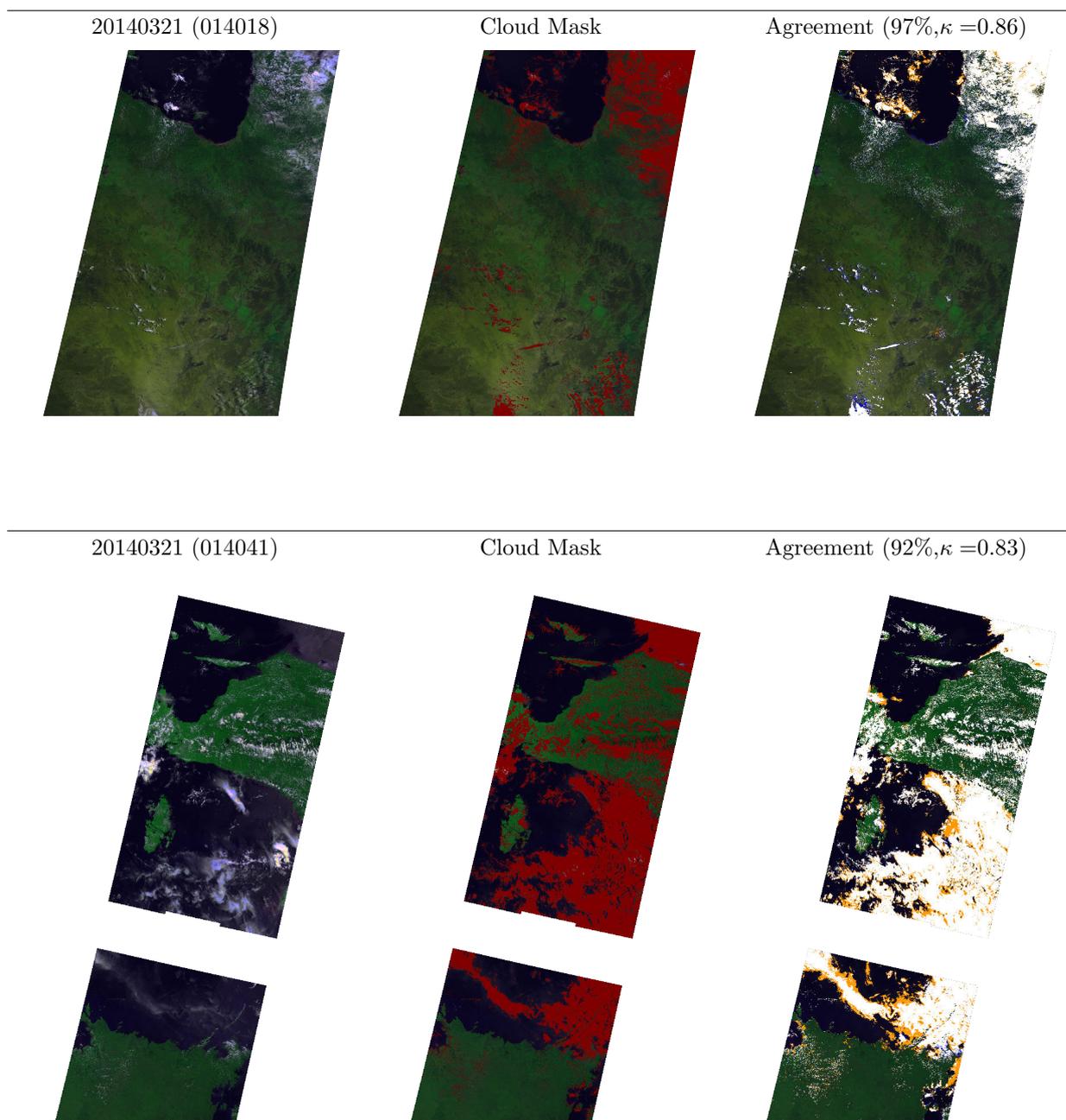
- Gómez-Chova, L., Camps-Valls, G., Muñoz-Marí, J., and Calpe-Maravilla, J. (2007b). Semi-supervised Cloud Screening with Laplacian SVM. In *IEEE Int. Geoscience and Remote Sensing Symposium, IGARSS'2007*, pages 1521–1524, Barcelona, Spain. (DOI:[10.1109/IGARSS.2007.4423098](https://doi.org/10.1109/IGARSS.2007.4423098)).
- Hagan, M. and Menhaj, M. (1994). Training feed-forward networks with the Marquardt algorithm. *IEEE Trans. on Neural Networks*, 5(6):989–993.
- Haykin, S. (1999). *Neural Networks: A Comprehensive Foundation*. Prentice Hall, Englewood Cliffs, NJ, USA.
- Hughes, G. F. (1968). On the mean accuracy of statistical pattern recognizers. *IEEE Trans. on Information Theory*, 14(1):55–63.
- Jones, M. S., Saunders, M. A., and Guymer, T. H. (1996). Global remnant cloud contamination in the along-track scanning radiometer data: Source and removal. *Journal of Geophysical Research*, 101(12):141–147.
- Kohavi, R. and John, G. H. (1997). Wrappers for features subset selection. *Int J Digit Libr*, 1:108–121.
- Lee, Y., Wahba, G., and Ackerman, S. (2004). Cloud Classification of Satellite Radiance Data by Multicategory Support Vector Machines. *Journal of Atmospheric and Oceanic Technology*, 21(2):159–169.
- Li, J., Menzel, W., Yang, Z., Frey, R., and Ackerman, S. (2003). High-Spatial-Resolution Surface and Cloud-Type Classification from MODIS Multispectral Band Measurements. *Journal of Applied Meteorology*, 42:204–226.
- Lisens, G., Kempencers, P., Fierens, F., and Rensbergen, J. V. (2000). Development of cloud, snow, and shadow masking algorithms for VEGETATION imagery. In *IGARSS 2000. IEEE 2000 International Geoscience and Remote Sensing Symposium.*, volume 2, pages 834–836.
- Liu, H. and Motoda, H. (1998). *Feature Selection for Knowledge Discovery and Data Mining*. Kluwer Academic Publishers, Boston, USA.
- Martins, J., Tanré, D., Remer, L. Kaufman, Y., Mattoo, S., and Levy, R. (2002). MODIS cloud screening for remote sensing of aerosols over oceans using spatial variability. *Geophysical Research Letters*, 29(12):8009.
- Mazzoni, D., Garay, M., Davies, R., and Nelson, D. (2007). An operational MISR pixel classifier using support vector machines. *Remote Sensing of Environment*, 107(1/2):149–158.

- McIntire, T. and Simpson, J. (2002). Arctic sea ice, cloud, water, and lead classification using neural networks and 1.6 μm data. *IEEE Trans. on Geoscience and Remote Sensing*, 40(9):1956–1972.
- Merchant, C., Harris, A., Maturi, E., and MacCallum, S. (2005). Probabilistic physically-based cloud screening of satellite infra-red imagery for operational sea surface temperature retrieval. *Quart. J. Royal Met. Soc.*, 131:2735–2755.
- Murtagh, F., Barreto, D., and Marcello, J. (2003). Decision Boundaries Using Bayes Factors: The Case of Cloud Masks. *IEEE Trans. on Geoscience and Remote Sensing*, 41(12):2952–2958.
- Papin, C., Bouthemy, P., and Rochard, G. (2002). Unsupervised segmentation of low clouds from infrared METEOSAT images based on a contextual spatio-temporal labeling approach. *IEEE Trans. on Geoscience and Remote Sensing*, 40(1):104–114.
- Peng, H., Long, F., and Ding, C. (2005). Feature selection based on mutual information: Criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Trans. Pattern Analysis and Machine Intelligence*, 27(8):1226–1238.
- Platnick, S., King, M., Ackerman, S., Menzel, W., Baum, B., Riédi, J., and Frey, R. (2003). The MODIS Cloud Products: Algorithms and Examples From Terra. *IEEE Trans. on Geoscience and Remote Sensing*, 41(2):459–473.
- Preusker, R., Huenerbein, A., and Fischer, J. (2006). Cloud detection with MERIS using oxygen absorption measurements. *Geophysical Research Abstracts*, 8:09956.
- Quinlan, J. R. (1986). Induction of decision trees. *Machine Learning*, 1(1):81–106.
- Quinlan, R. J. (1993). *"C4.5: programs for machine learning"*. Morgan Kaufmann Publishers Inc.
- Saitwal, K., Azimi-Sadjadi, M., and Reinke, D. (2003). A multichannel temporally adaptive system for continuous cloud classification from satellite imagery. *IEEE Trans. on Geoscience and Remote Sensing*, 41(5):1098–1104.
- Santer, R., Carrère, V., Dessailly, D., Dubuisson, P., and Roger, J. (1997). MERIS Algorithm Theoretical Basis Document (ATBD 2.17). Pixel Identification. Technical Report issue 4, European Space Agency. <http://envisat.esa.int/instruments/meris/atbd/>.
- Saunders, R. W. and Kriebel, K. T. (1988). An improved method for detecting clear sky and cloudy sky radiances from AVHRR data. *Int. Journal of Remote Sensing*, 9:123–150. errata - IJRS 9, 1393-1394.

- Schölkopf, B. and Smola, A. (2002). *Learning with Kernels – Support Vector Machines, Regularization, Optimization and Beyond*. MIT Press Series, Cambridge, MA, USA.
- Simpson, J., Schmidt, A., and Harris, A. (1998). Improved Cloud Detection in Along Track Scanning Radiometer (ATSR) Data over the Ocean. *Remote Sensing of Environment*, 65(1):1–24.
- Srivastava, A. N. and Stroeve, J. (2003). Onboard detection of snow, ice, clouds and other geophysical processes using kernel methods. In *12th Int. Conference on Machine Learning (ICML): Machine Learning Technologies for Autonomous Space Applications Workshop*, Washington, DC, USA.
- Theodoridis, S. and Koutroumbas, K. (1999). *Pattern Recognition*. Academic Press, San Diego, California, USA.
- Tian, B., Shaikh, M., Azimi, M., Haar, T., and Reinke, D. (1999). A study of cloud classification with neural networks using spectral and textural features. *IEEE Trans. on Neural Networks*, 10(1):138–151.
- Torres Arriaza, J. A., Guindos Rojas, F., Peralta López, M., and Cantón, M. (2003). An Automatic Cloud-Masking System Using Backpro. Neural Nets for AVHRR Scenes. *IEEE Trans. on Geoscience and Remote Sensing*, 41(4):826–831.
- Vapnik, V. N. (1995). *The nature of statistical learning theory*. Springer-Verlag, New York, NY, USA.
- Vapnik, V. N. (1998). *Statistical Learning Theory*. John Wiley & Sons, New York, USA.
- Wang, M. and Shi, W. (2006). Cloud Masking for Ocean Color Data Processing in the Coastal Regions. *IEEE Trans. on Geoscience and Remote Sensing*, 44(11):3196–3105.
- Wolters, E., Swinnen, E., Benhadj, I., and Dierckx, W. (2015). PROBA-V cloud detection evaluation and proposed modification. Technical Report Technical Note, 17/7/2015, QWG.
- Yang, Y., Di Girolamo, L., and Mazzoni, D. (2007). Selection of the automated thresholding algorithm for the Multi-angle Imaging SpectroRadiometer Radiometric Camera-by-Camera Cloud Mask over land. *Remote Sensing of Environment*, 107(1/2):159–171.
- Yhann, S. R. and Simpson, J. J. (1995). Application of neural networks to AVHRR cloud segmentation. *IEEE Trans. on Geoscience and Remote Sensing*, 33(3):590–604.

Appendix A. Validation for all manually labeled images

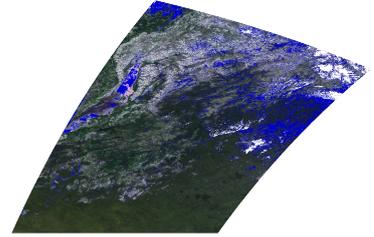
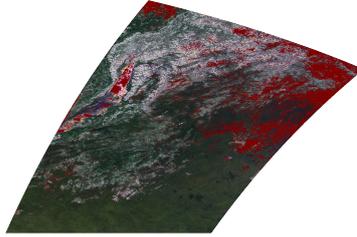
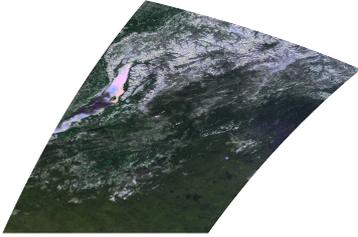
This Appendix presents the validation results for all manually labeled Proba-V images. *First column* shows the RGB composite. *Second column* shows the cloud mask obtained with the final MLP classifier. *Third column* shows the comparison of the ‘manual ground truth’ with the ‘Proba-V Cloud Flag’. Discrepancies are shown in blue when proposed method detects cloud and in orange when pixels are classified as cloud-free.



20140321 (044547)

Cloud Mask

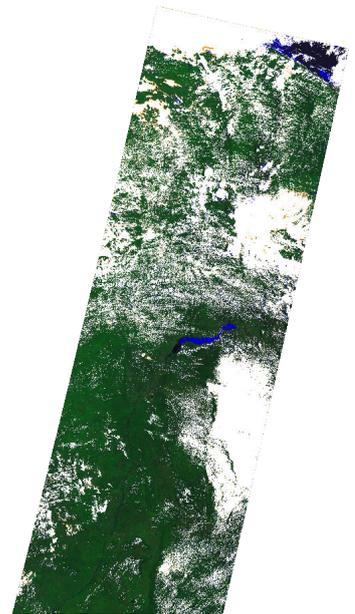
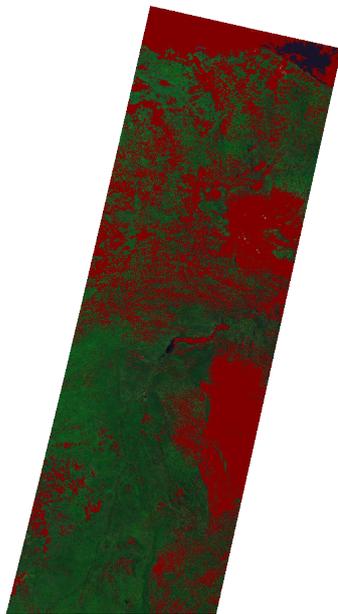
Agreement (86%, $\kappa = 0.26$)



20140321 (133006)

Cloud Mask

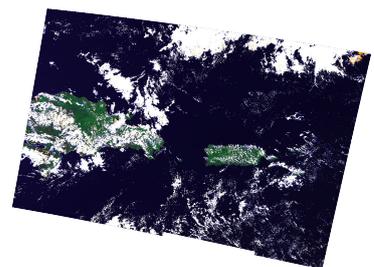
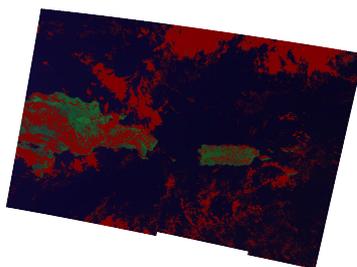
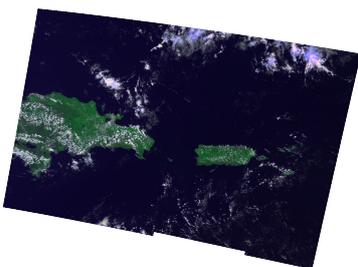
Agreement (96%, $\kappa = 0.94$)



20140321 (150455)

Cloud Mask

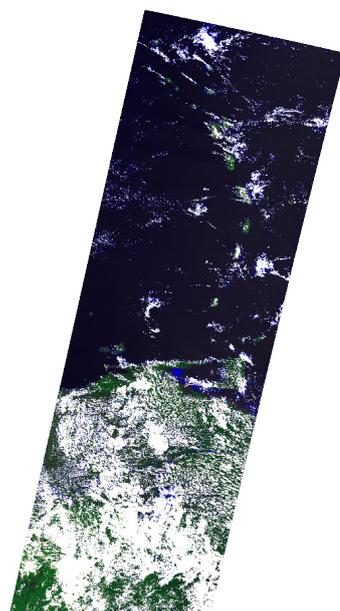
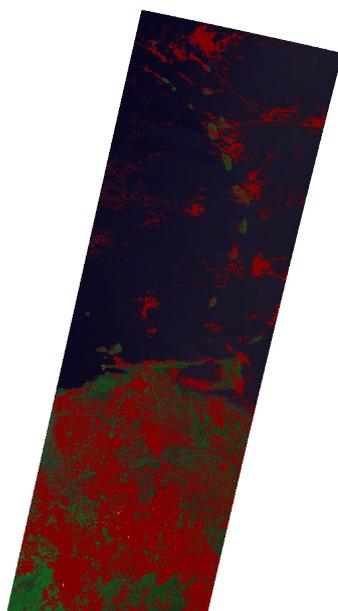
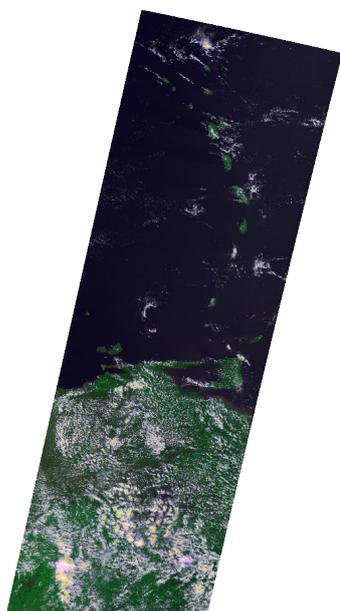
Agreement (96%, $\kappa = 0.9$)



20140321 (150528)

Cloud Mask

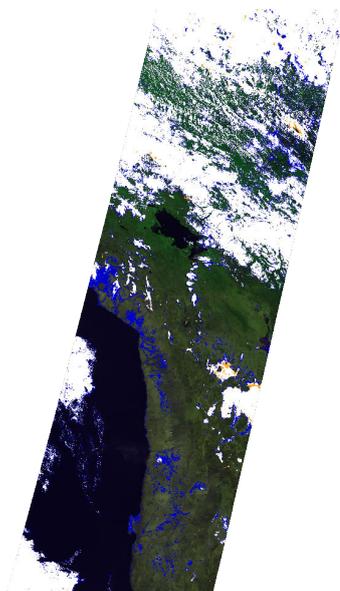
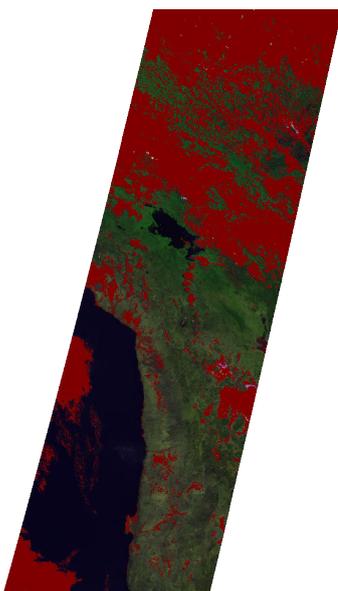
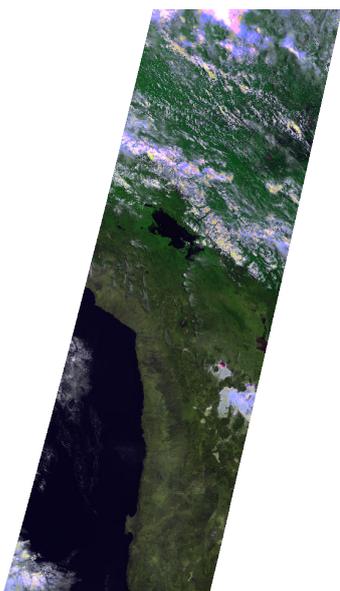
Agreement (95%, $\kappa = 0.89$)



20140321 (150528)

Cloud Mask

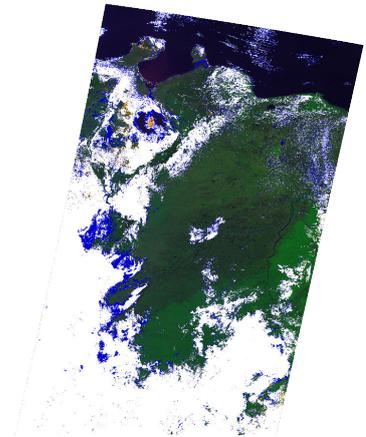
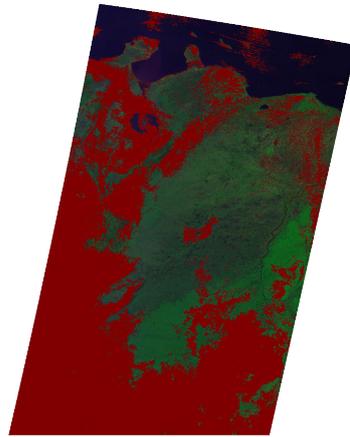
Agreement (92%, $\kappa = 0.83$)



20140321 (150724)

Cloud Mask

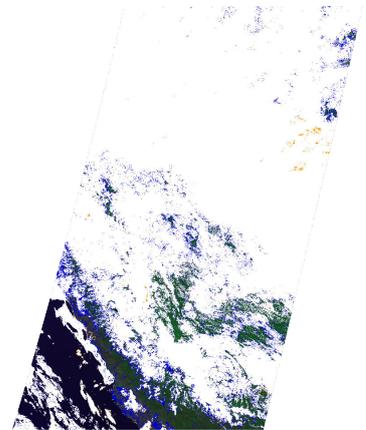
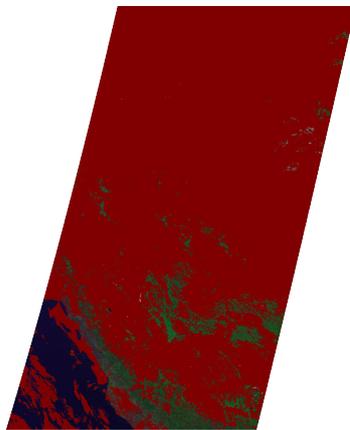
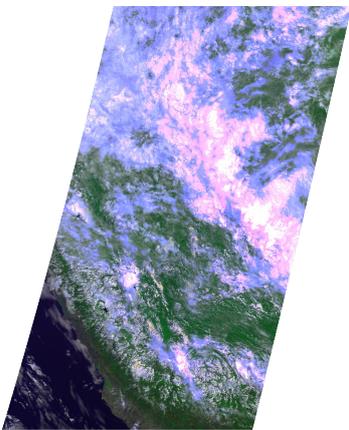
Agreement (92%, $\kappa=0.84$)



20140321 (150724)

Cloud Mask

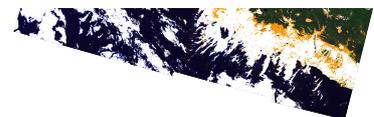
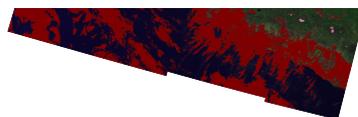
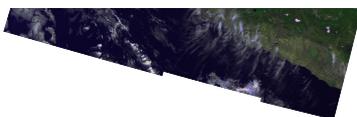
Agreement (94%, $\kappa=0.81$)



20140321 (150724)

Cloud Mask

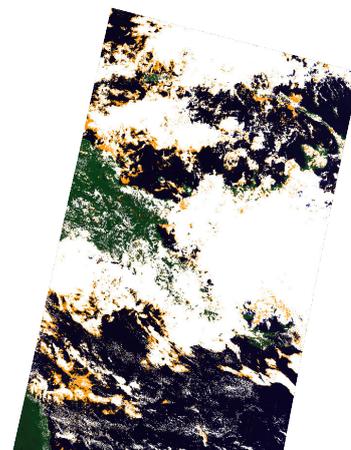
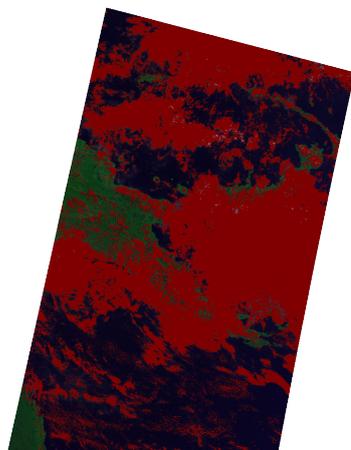
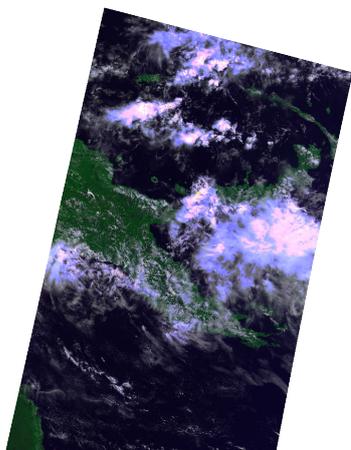
Agreement (91%, $\kappa=0.83$)



20140621 (011731)

Cloud Mask

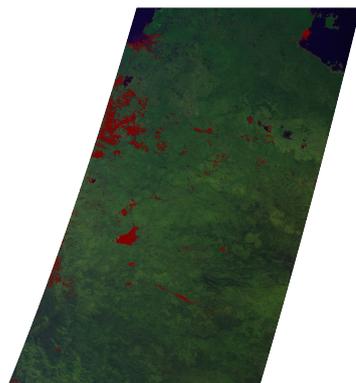
Agreement (92%, $\kappa = 0.84$)



20140621 (011810)

Cloud Mask

Agreement (97%, $\kappa = 0.44$)



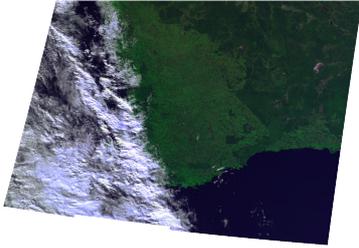
20140621 (025437)

Cloud Mask

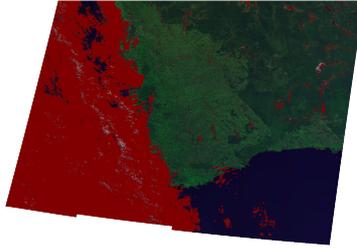
Agreement (98%, $\kappa = 0.81$)



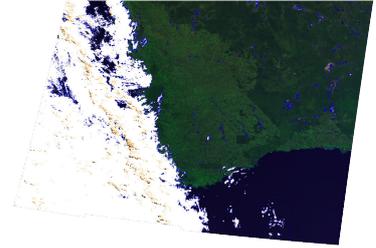
20140621 (025437)



Cloud Mask



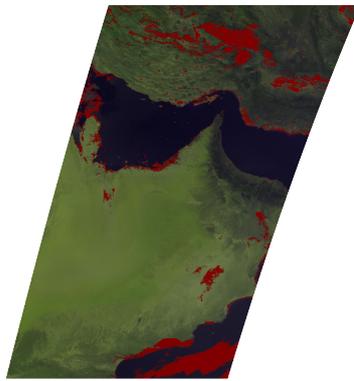
Agreement (96%, $\kappa = 0.93$)



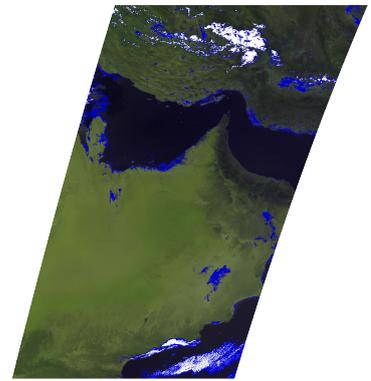
20140621 (074527)



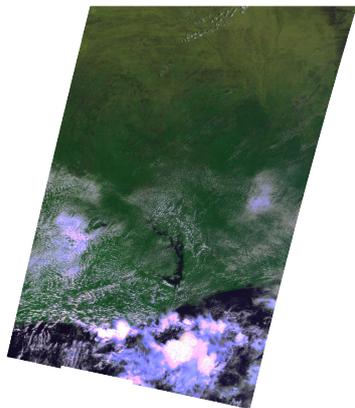
Cloud Mask



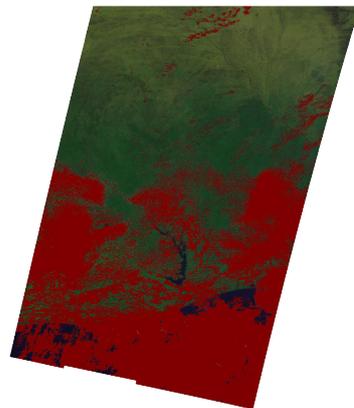
Agreement (94%, $\kappa = 0.42$)



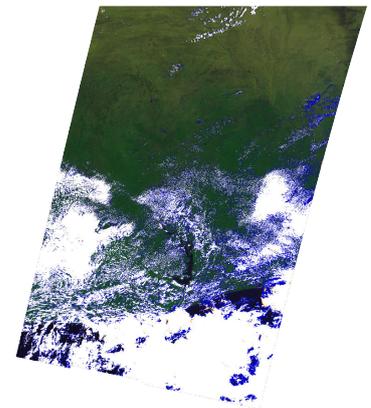
20140621 (110758)



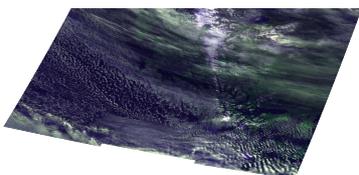
Cloud Mask



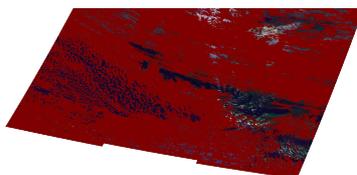
Agreement (89%, $\kappa = 0.78$)



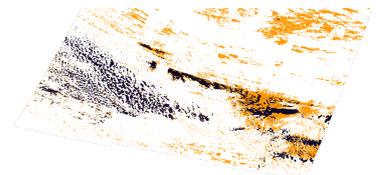
20140621 (144525)



Cloud Mask



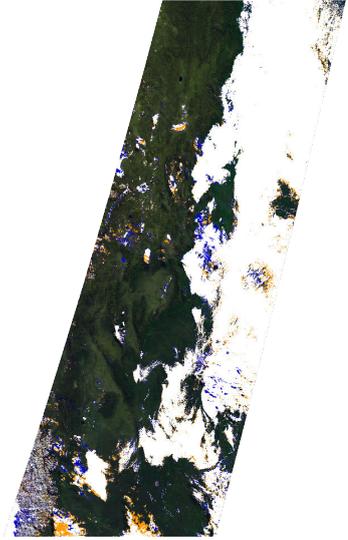
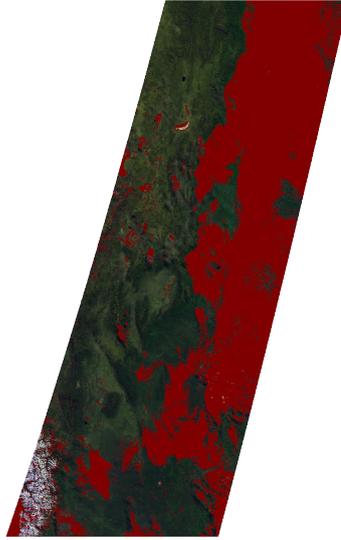
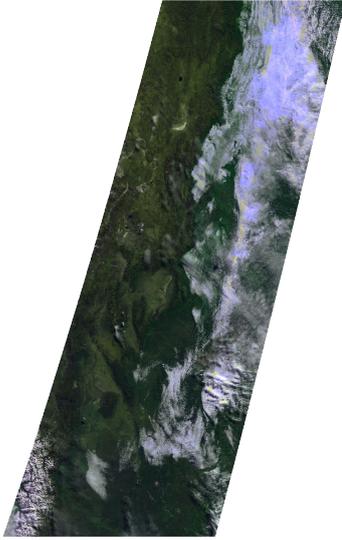
Agreement (87%, $\kappa = 0.46$)



20140621 (144525)

Cloud Mask

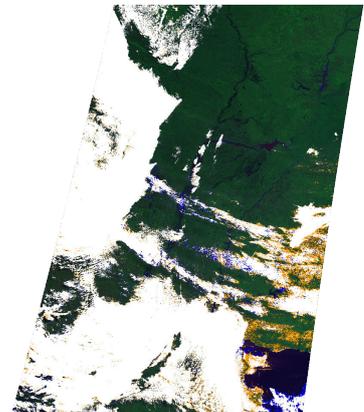
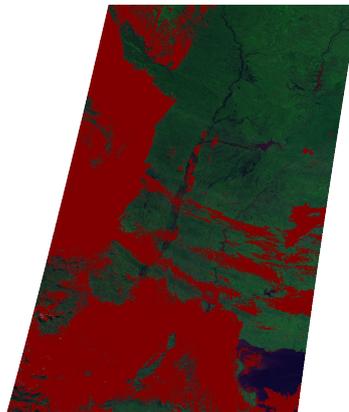
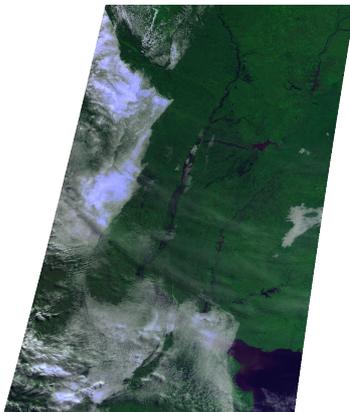
Agreement (94%, $\kappa=0.88$)



20140621 (144544)

Cloud Mask

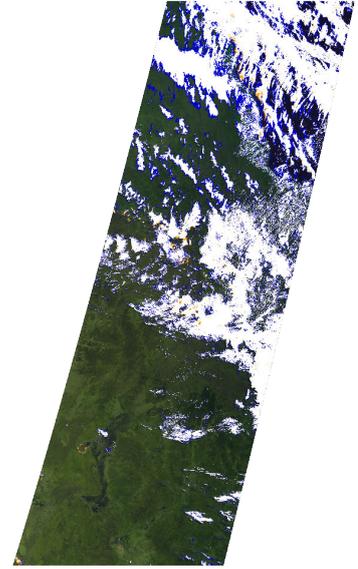
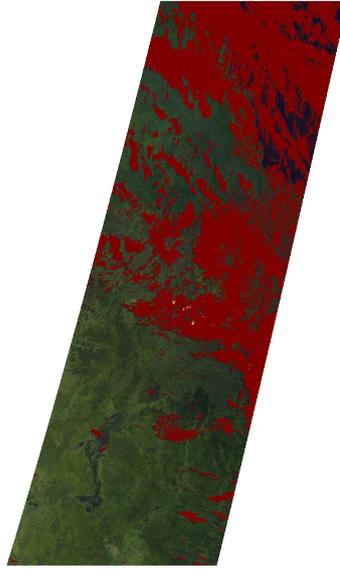
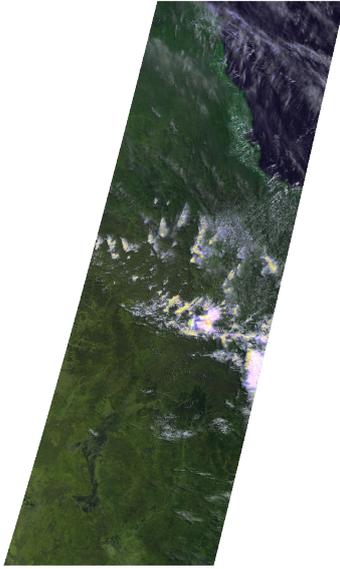
Agreement (94%, $\kappa=0.89$)



20140921 (005335)

Cloud Mask

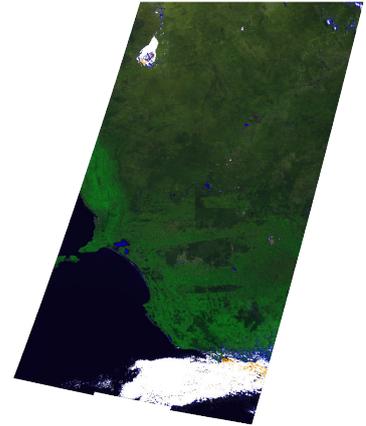
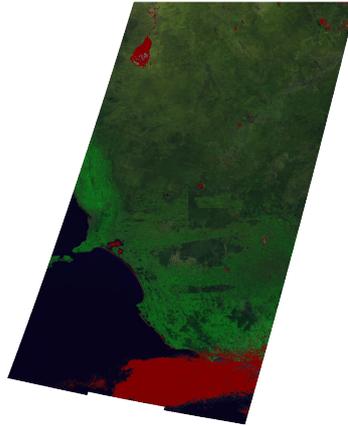
Agreement (90%, $\kappa = 0.8$)



20140921 (005335)

Cloud Mask

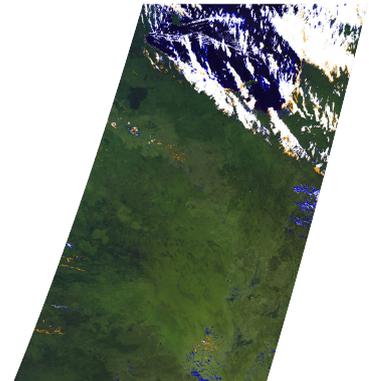
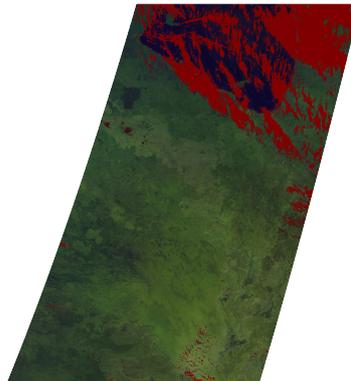
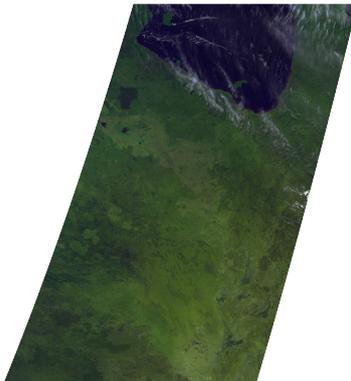
Agreement (98%, $\kappa = 0.9$)



20140921 (005350)

Cloud Mask

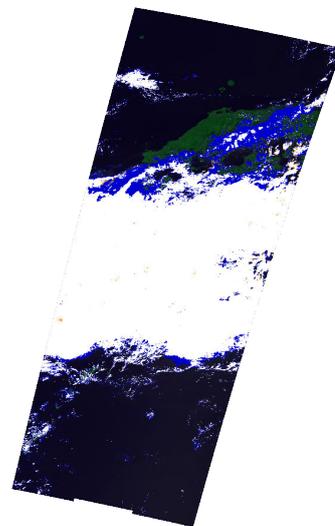
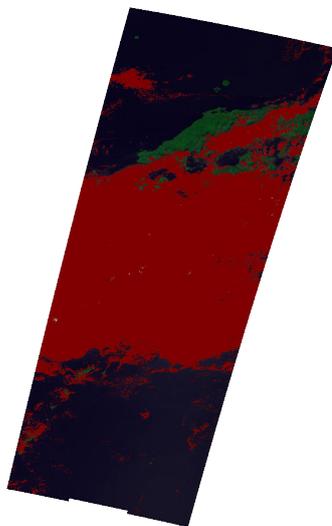
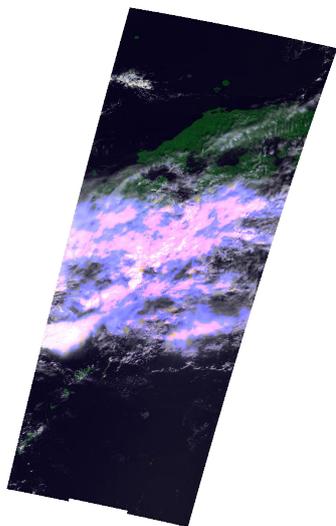
Agreement (96%, $\kappa = 0.83$)



20140921 (022355)

Cloud Mask

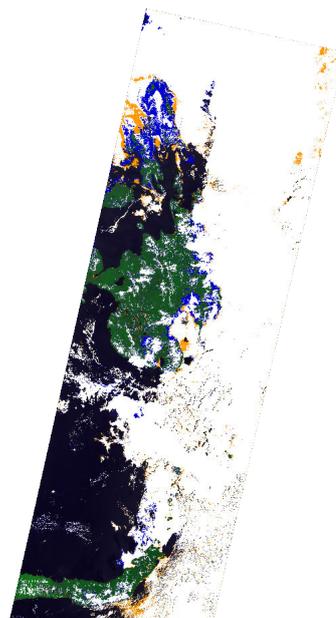
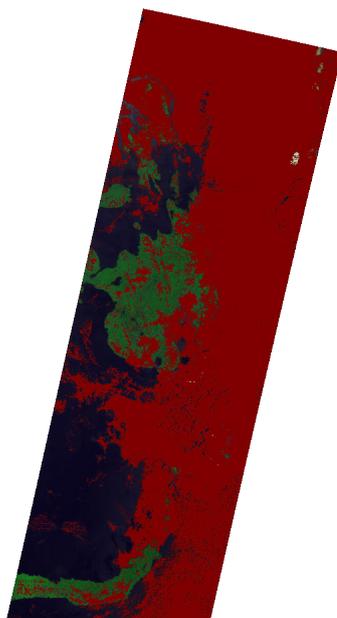
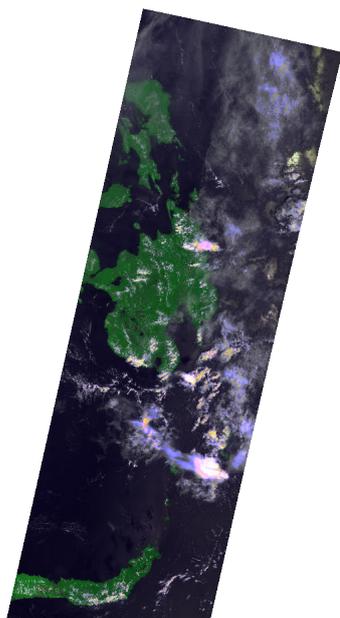
Agreement (92%, $\kappa = 0.84$)



20140921 (023045)

Cloud Mask

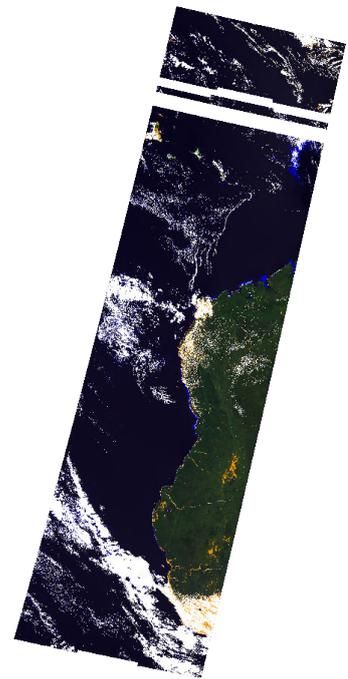
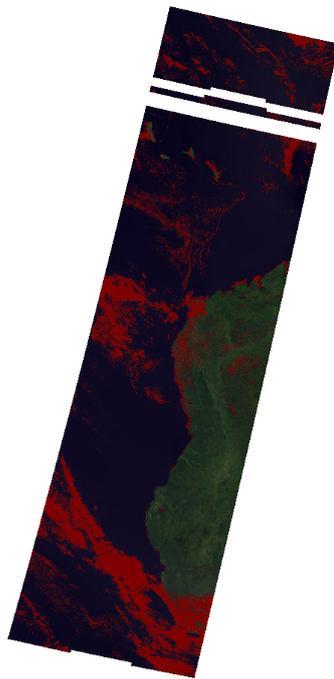
Agreement (94%, $\kappa = 0.88$)



20140921 (074055)

Cloud Mask

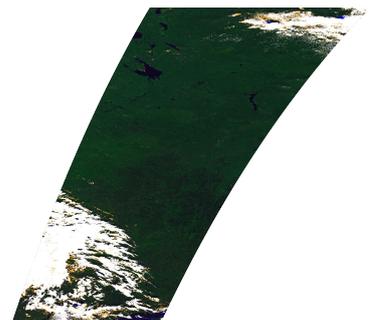
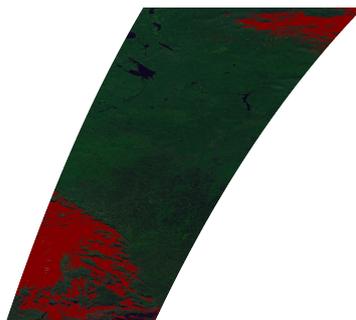
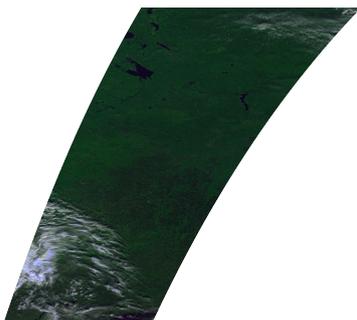
Agreement (96%, $\kappa=0.87$)



20140921 (085802)

Cloud Mask

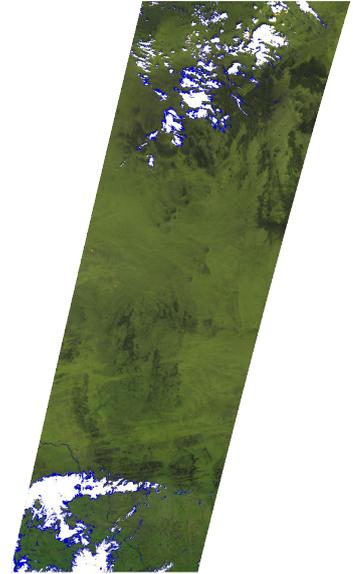
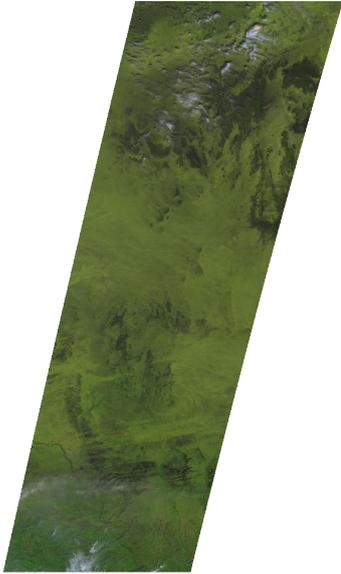
Agreement (97%, $\kappa=0.91$)



20140921 (104018)

Cloud Mask

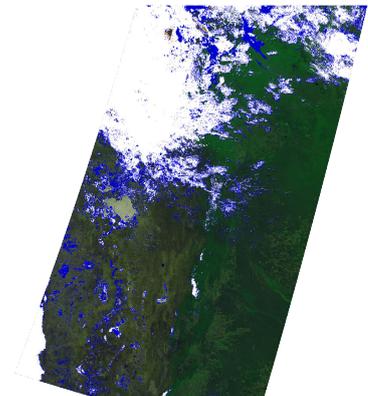
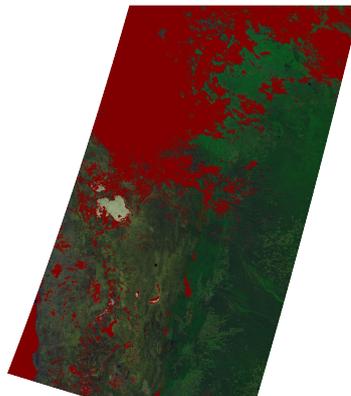
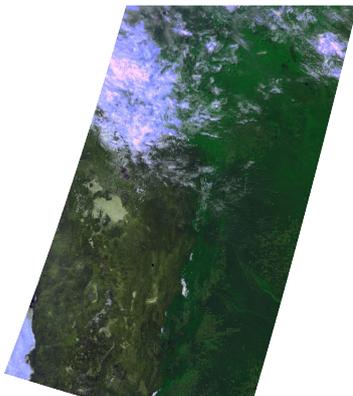
Agreement (96%, $\kappa=0.78$)



20140921 (141925)

Cloud Mask

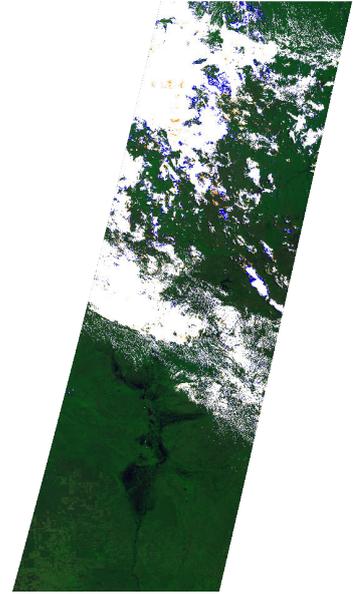
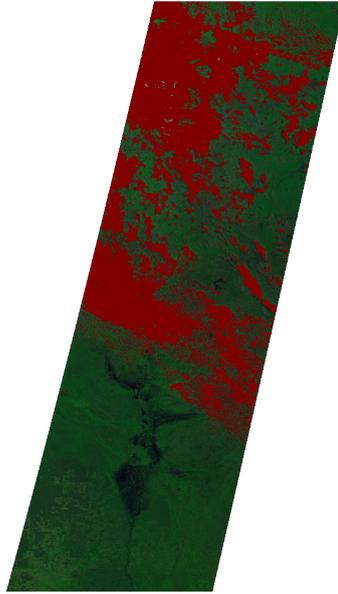
Agreement (89%, $\kappa=0.74$)



20140921 (142148)

Cloud Mask

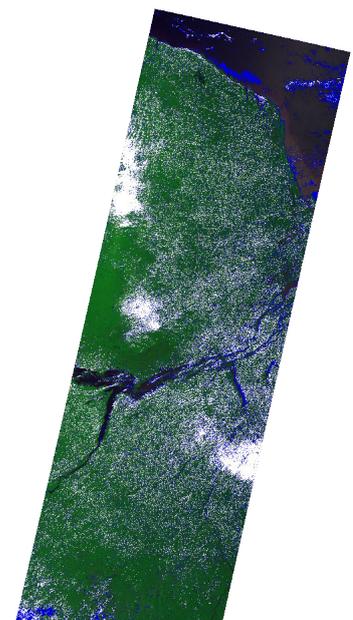
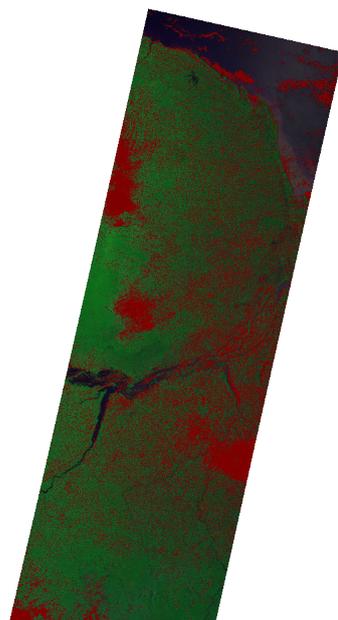
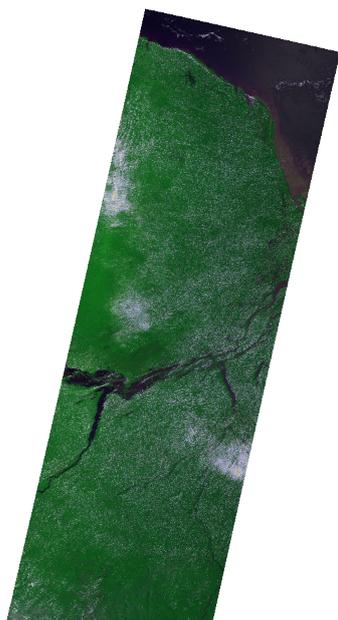
Agreement (95%, $\kappa = 0.9$)



20140921 (142148)

Cloud Mask

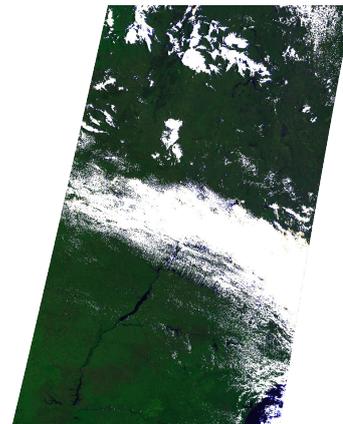
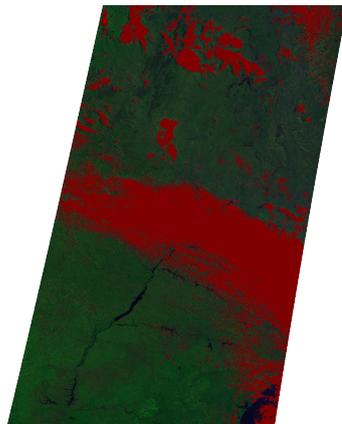
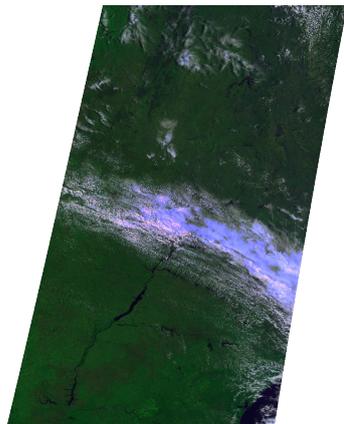
Agreement (87%, $\kappa = 0.66$)



20140921 (142232)

Cloud Mask

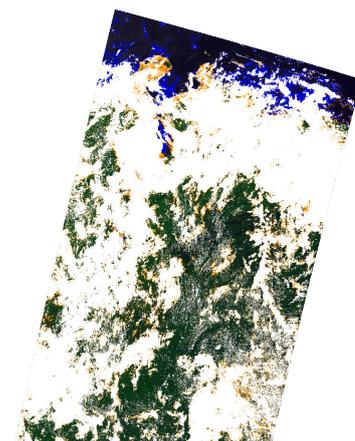
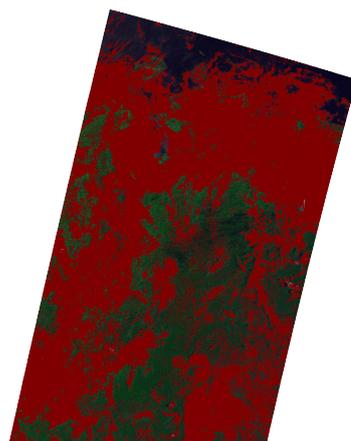
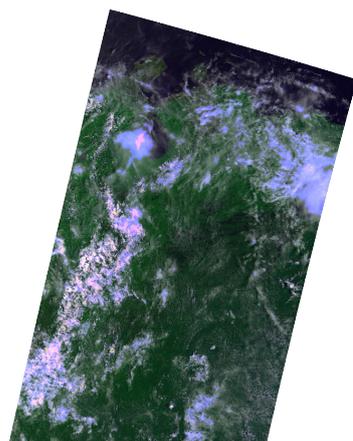
Agreement (97%, $\kappa = 0.94$)



20140921 (160041)

Cloud Mask

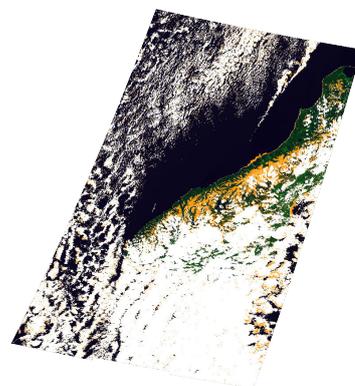
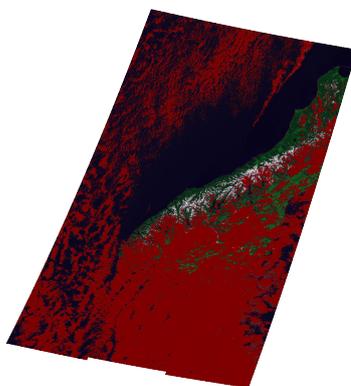
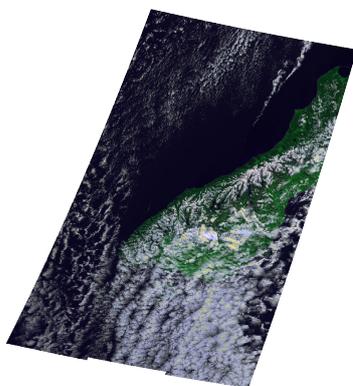
Agreement (90%, $\kappa = 0.8$)



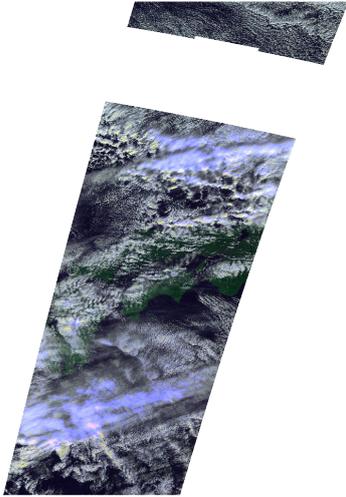
20140921 (230104)

Cloud Mask

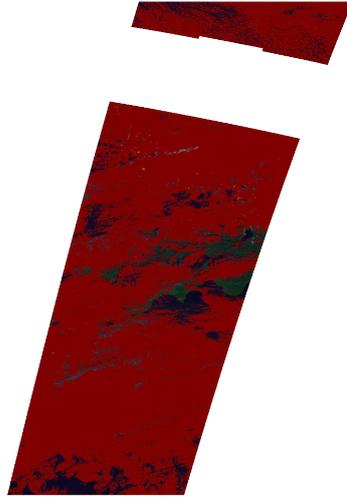
Agreement (93%, $\kappa = 0.87$)



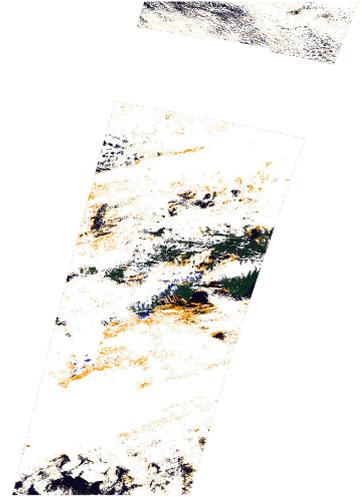
20141221 (021444)



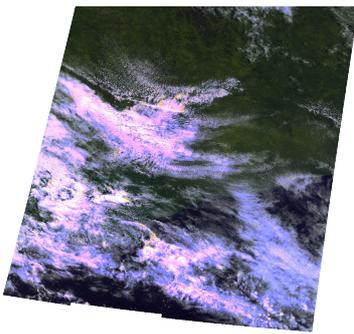
Cloud Mask



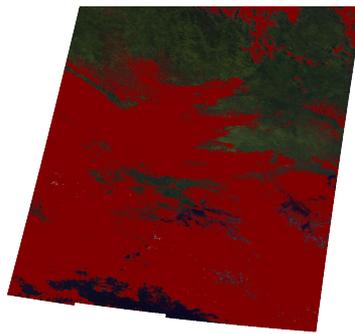
Agreement (94%, $\kappa = 0.73$)



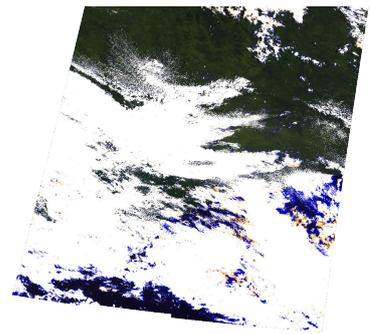
20141221 (022940)



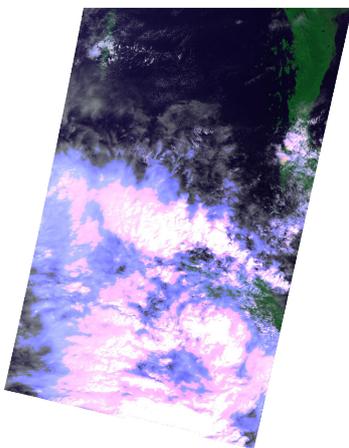
Cloud Mask



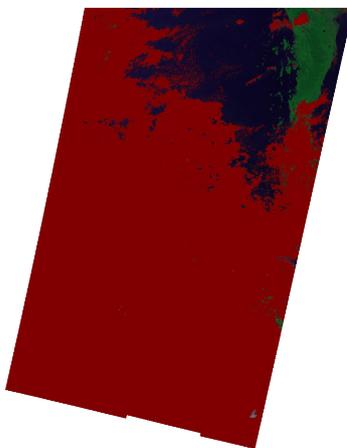
Agreement (94%, $\kappa = 0.88$)



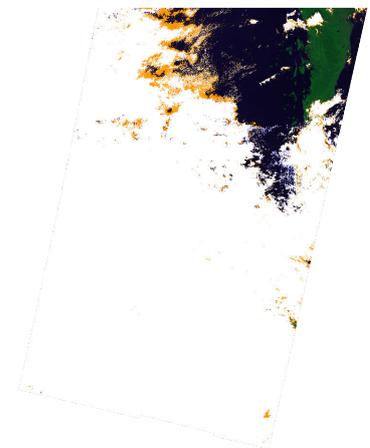
20141221 (035603)



Cloud Mask



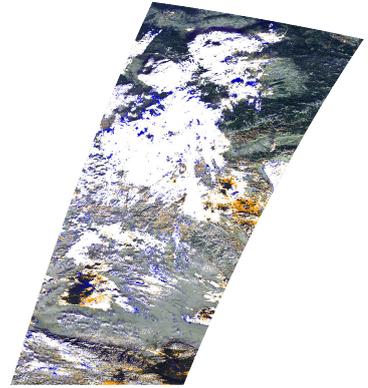
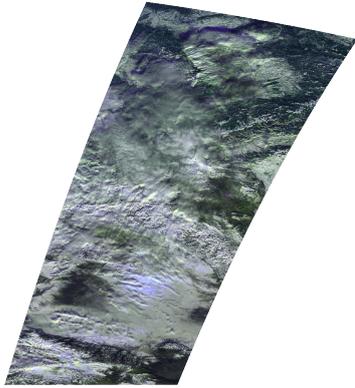
Agreement (96%, $\kappa = 0.88$)



20141221 (053716)

Cloud Mask

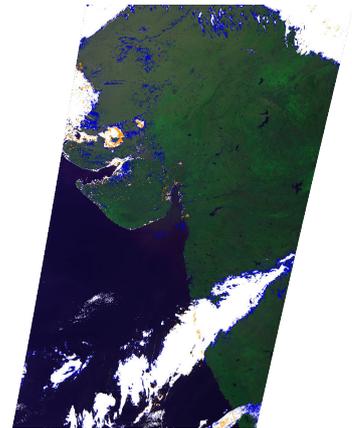
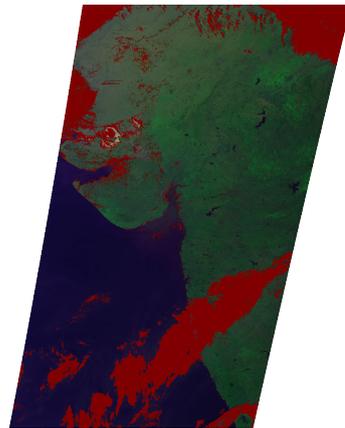
Agreement (89%, $\kappa=0.79$)



20141221 (053718)

Cloud Mask

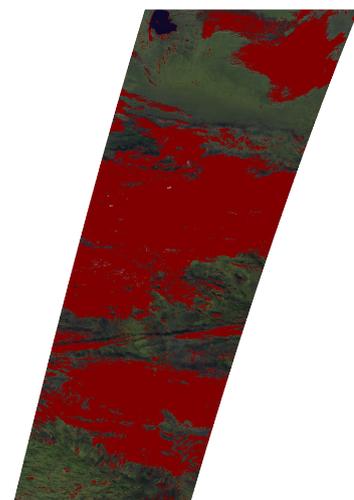
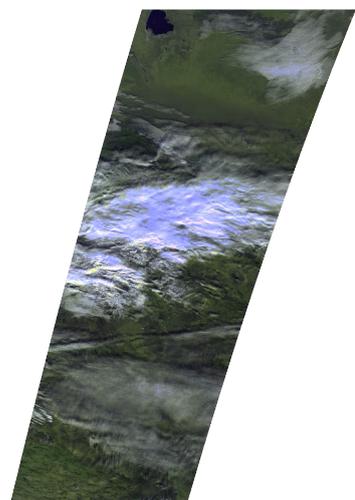
Agreement (94%, $\kappa=0.81$)



20141221 (071831)

Cloud Mask

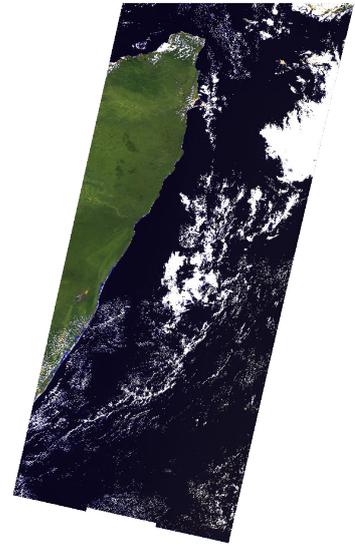
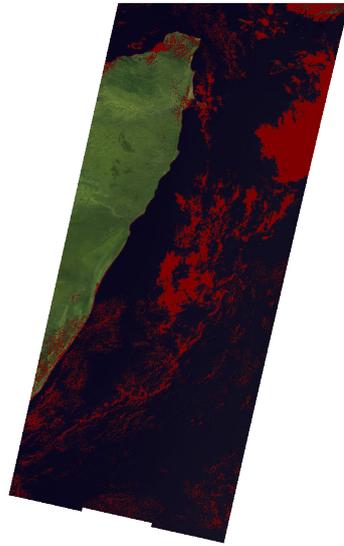
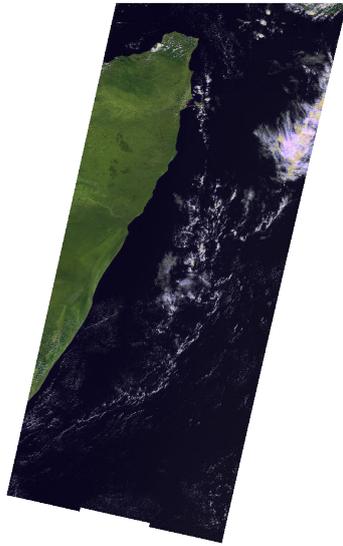
Agreement (90%, $\kappa=0.81$)



20141221 (071831)

Cloud Mask

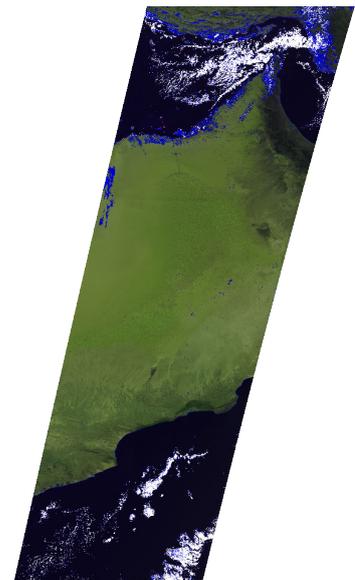
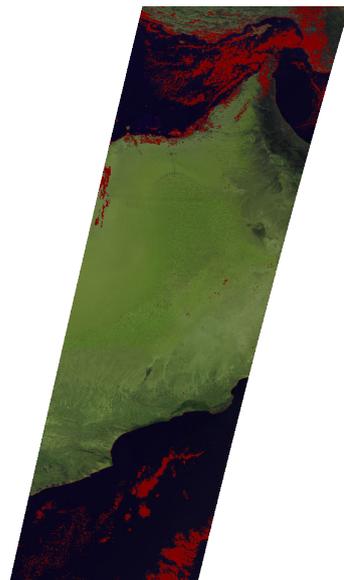
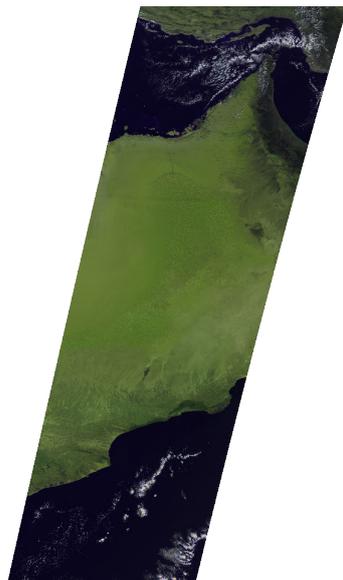
Agreement (97%, $\kappa=0.91$)



20141221 (071831)

Cloud Mask

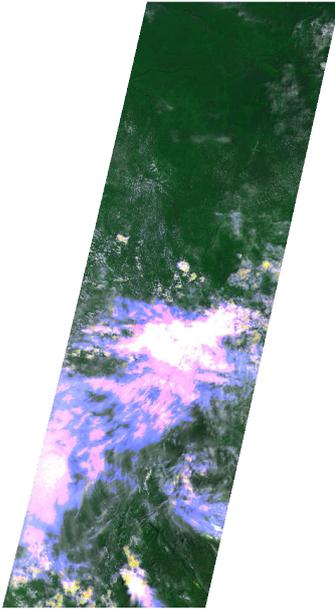
Agreement (95%, $\kappa=0.65$)



20141221 (085947)

Cloud Mask

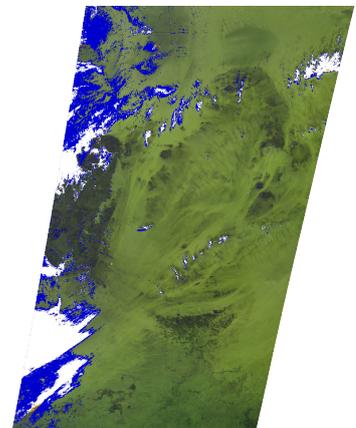
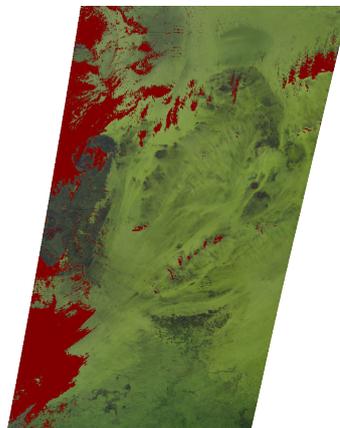
Agreement (97%, $\kappa = 0.95$)



20141221 (085949)

Cloud Mask

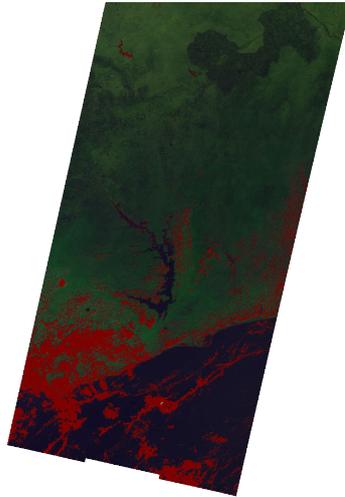
Agreement (91%, $\kappa = 0.52$)



20141221 (104103)



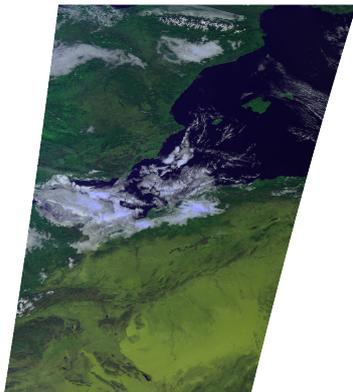
Cloud Mask



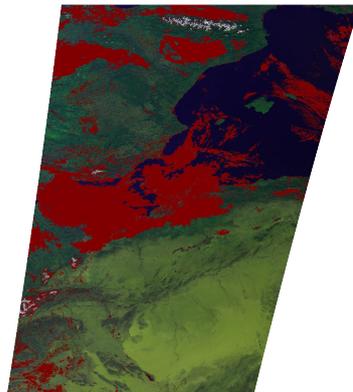
Agreement (95%, $\kappa = 0.8$)



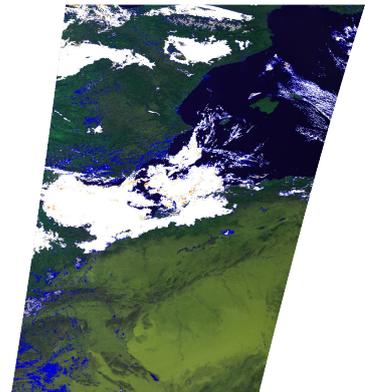
20141221 (104105)



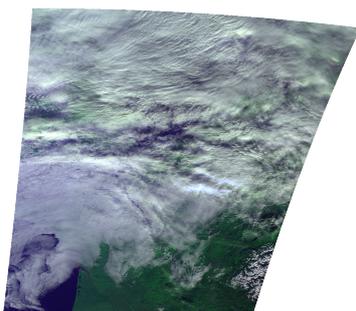
Cloud Mask



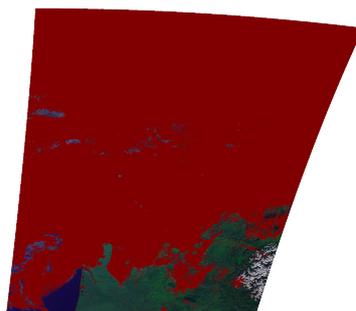
Agreement (93%, $\kappa = 0.78$)



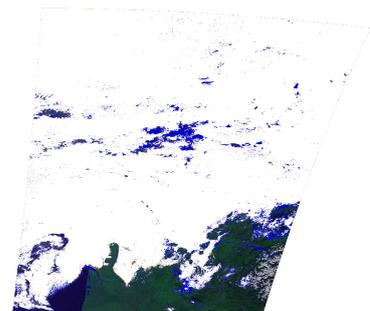
20141221 (104105)



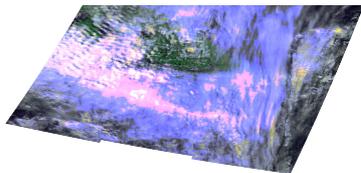
Cloud Mask



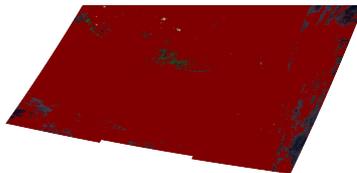
Agreement (96%, $\kappa = 0.86$)



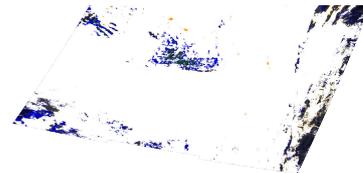
20141221 (141754)



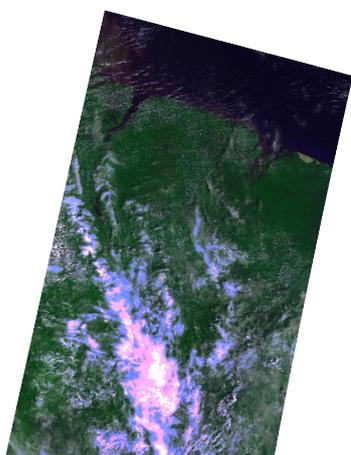
Cloud Mask



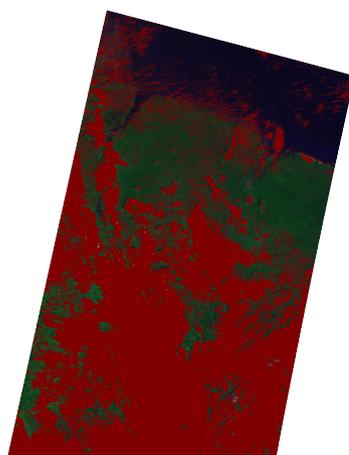
Agreement (96%, $\kappa = 0.74$)



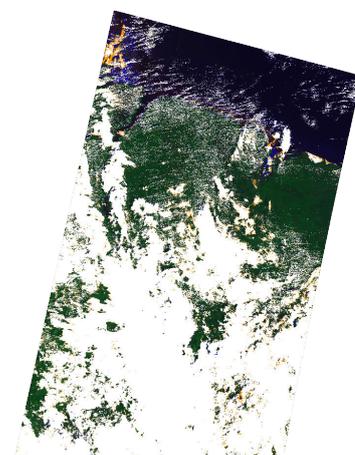
20141221 (141847)



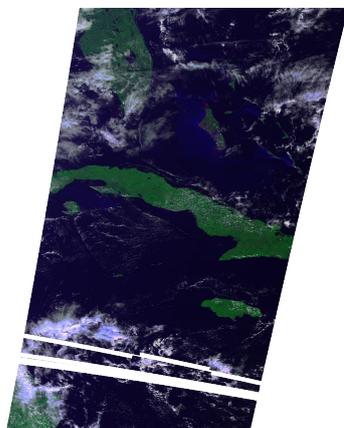
Cloud Mask



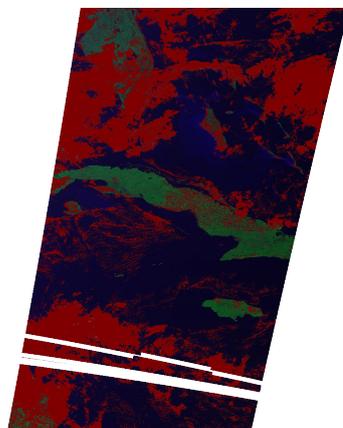
Agreement (95%, $\kappa = 0.91$)



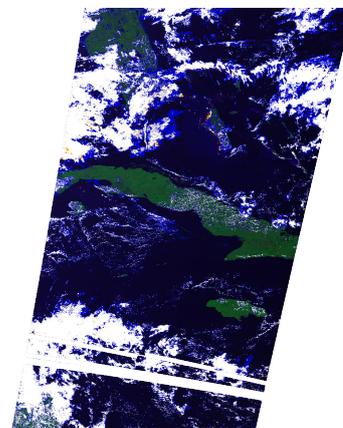
20141221 (154450)



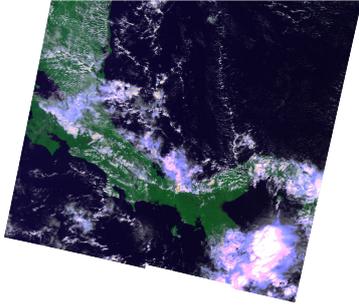
Cloud Mask



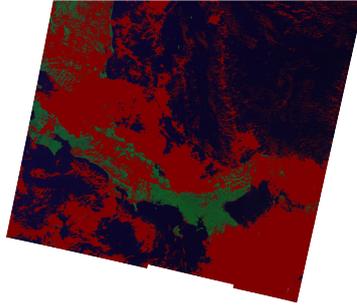
Agreement (85%, $\kappa = 0.65$)



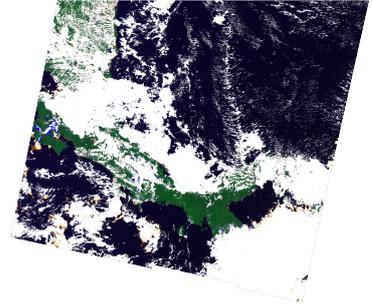
20141221 (154450)



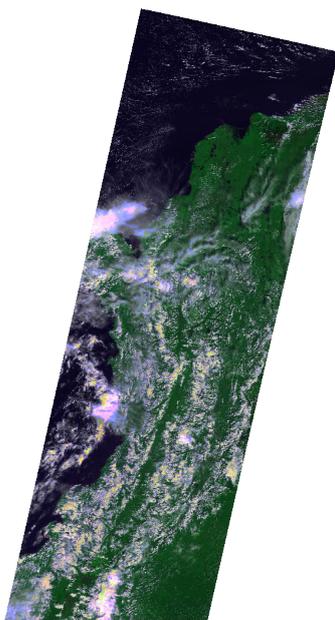
Cloud Mask



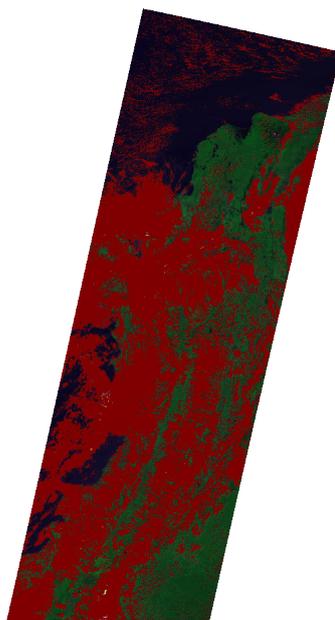
Agreement (97%, $\kappa = 0.94$)



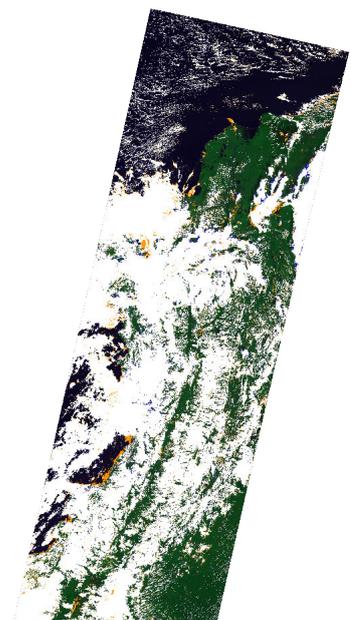
20141221 (155703)



Cloud Mask



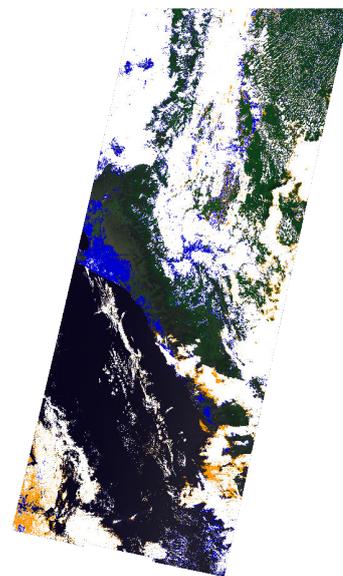
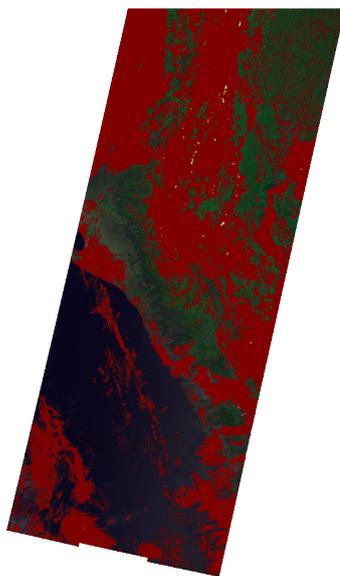
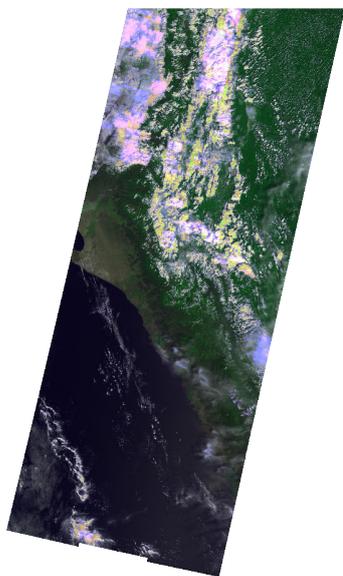
Agreement (95%, $\kappa = 0.91$)



20141221 (155703)

Cloud Mask

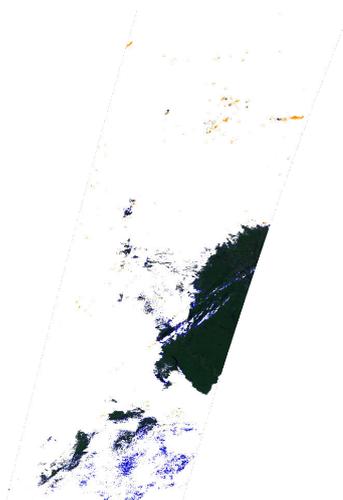
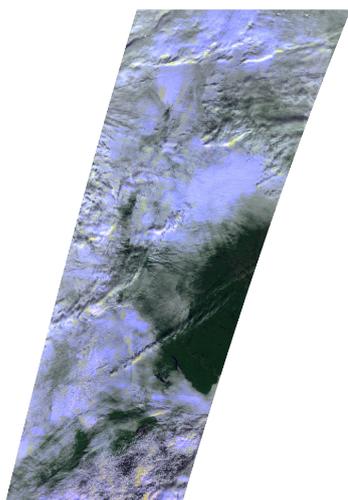
Agreement (88%, $\kappa=0.78$)



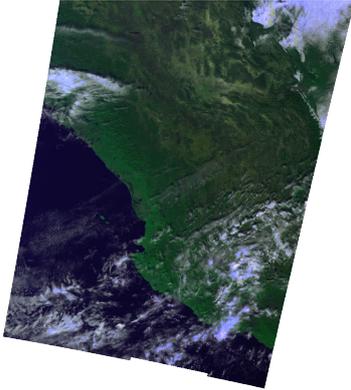
20141221 (172606)

Cloud Mask

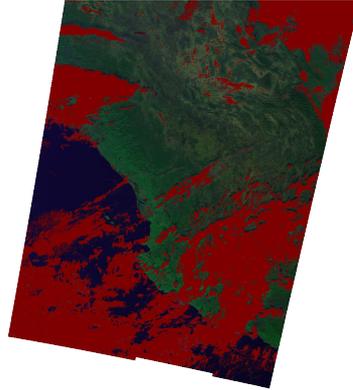
Agreement (98%, $\kappa=0.9$)



20141221 (172608)



Cloud Mask



Agreement (94%, $\kappa = 0.89$)

