Remote Sensing of Japanese WWII airstrips in the Papua Province Republic of Indonesia

Classification of the area surrounding three WWII airstrips (Mongosah, Otawiri and Sagan)

D. H. van der Valk, BSc.
Student number: 4154150
Delft University of Technology
June 2017

TUDelft

Remote Sensing of Japanese WWII airstrips in the Papua Province Republic of Indonesia

Classification of the area surrounding three WWII airstrips (Mongosah, Otawiri and Sagan)

By

D. H. van der Valk, BSc

in partial fulfilment of the requirements for the degree of

Master of Science
in Geoscience and Remote Sensing

at the Delft University of Technology,

Supervisor:	Dr. R.C. Lindenbergh,	TU Delft
Thesis committee:	Prof.dr.ir. R.F. Hanssen,	TU Delft
	Drs. B. Kreuger,	Kreuger *in* Kultuur
	Drs. T. de Groot,	Universiteit Leiden

TUDelft

Preface

In front of you lies the additional thesis: Remote Sensing of Japanese WWII airstrips in the Papua Province Republic of Indonesia. I choose this additional thesis as a voluntary part of my Master Degree Geoscience and Remote Sensing at the Delft University of Technology. I have been approached
by Drs. B. Kreuger (Kreuger *in* Kultuur) and Drs. T. de Groot, whether I was interested in writing this report. He provided me with the necessary information and background so I could accomplish this project. I would like to thank him for his help.

Together with Dr. R.C. Lindenbergh and Prof.dr.ir. R.F. Hanssen, I came up with the research questions and tried to answer them in this thesis. I would like to thank both of them for their support and help. I hope I answered the research questions adequately and that you will enjoy reading my thesis.

D. H. van der Valk, BSc
June 2017, Delft

Abstract

In the Second World War Dutch New Guinea was a strategic battle front for both the Japanese and the Allied forces in the Pacific War. A lot of airstrips were constructed and bombed during this time, of which at least three (Mongosah, Otowari and Sagan) have never been visited after the war. This provided a great opportunity to find potential war heritage and airstrip equipment. Later this year an additional research team will go on an in-situ exploration to potentially find those objects. To do so, they needed a classification map giving information on the type and location of the vegetation. This map helps to know where to land with a helicopter, to setup base camp, to find travel ways, etc. Thus, the main objective of this thesis is to check whether it is possible to create a proper classification image with the available data. I used data obtained from the Sentinel 2 Mission (Optical data), the ALOS PALSAR Mission (L-Band Radar data) and the SRTM Mission (Digital elevation data). I pre-processed the data and used the supervised classification method, "Maximum Likelihood Classification" (MLC). I masked clouds via three different cloud masking methods, MLC Method, Threshold Method and Sen2cor (scene classification) Method. I compared the three different methods with each other and there is no significant difference between them. The classifications have been cross-validated with a reference validation dataset and the classified pixels are on average about 90% correctly classified.

Table of Contents

1 Introduction

At the end of the Second World War, the Japanese army abandoned several airstrips in the Papua Province Republic of Indonesia (at the time called Netherlands New Guinea). These airstrips were constructed between 1943 and 1944 and were heavily bombed by the allied forces. After abandoning these airstrips, no one used them for civil aviation or for any other purposes. This means that, because the area is very remote, these airstrips lay there untouched, which provides a great opportunity to find heritage and potential airstrip equipment, such as crashed or shot-down aircrafts, trucks, mess halls, etc. The location of the region of interest is shown in figure 1.
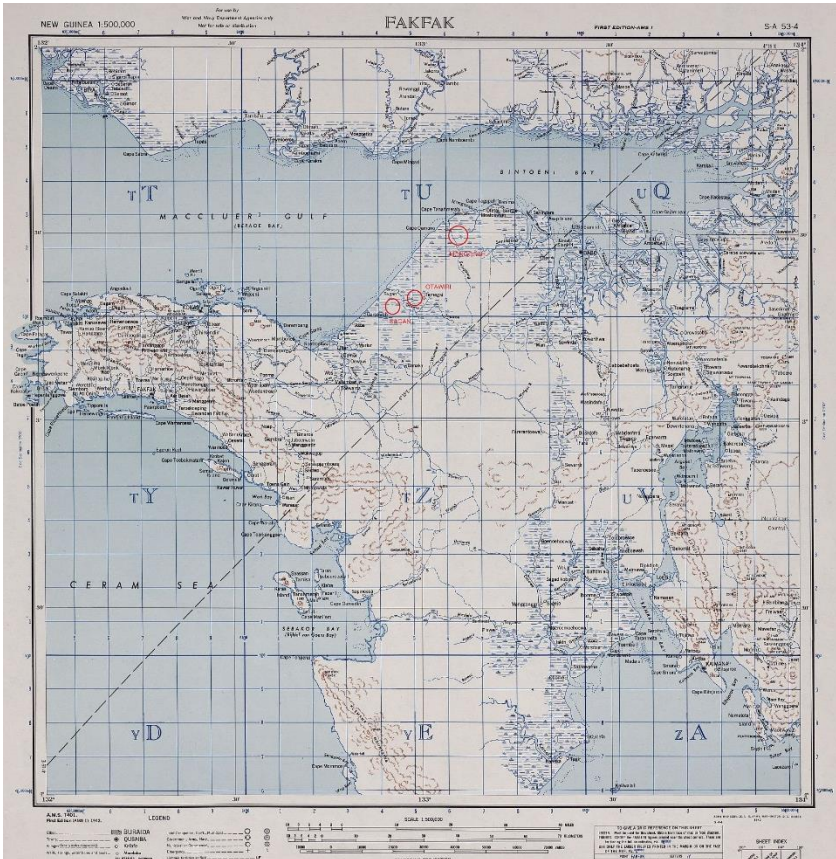

Figure 1: Region of interest. The red circles on the map indicate the location of the three airstrips

1.1   Thesis objective

The main objective of this thesis is to run a classification of the region surrounding the three airstrips (Mongosah, Otowari and Sagan) in Papua Province Republic of Indonesia. The reason why I want to create such a classification image is because, later this year an additional research team will go on an in-situ exploration to find potential war heritage and/or airstrip equipment and for them it is important to know where and what kind of vegetation is present in the area. For example, this information can help them to know where they can setup a basecamp, land their helicopter, know the best way to travel through the rainforest, etc.

1.2   Historical background

First a summary will be given concerning the invasion and liberation of Dutch New Guinea and after that I will provide some equipment usage, because that is what the ground research team eventually is after; war heritage.

1.2.1   Invasion by Japan [1]

Dutch New Guinea (DNG) became involved in the WWII battle scene in the year 1942. The Japanese invasion did not last long (November 1941 to April 1942) due to the absence of battalions of the Royal Dutch Army in the East Indies (KNIL) in DNG. This was because the Netherlands were occupied by Nazi Germany at the time and thus was quite weakened.
The Japanese army attacked DNG so they could use it as a strategic base towards the Solomon Islands to cut of Australia from the United States. The Japanese advanced from the north by first taking over FakFak, a harbour city just south of the Bird's Head Peninsula in DNG on the 1st of April 1942. After they invaded this harbour city, they proceeded quickly to Manakwari (12 April) and Hollandia (20 April) but their invasion stopped at the Frederik Hendrik Island.

1.2.2   Liberation of New Guinea [1]

In January 1943, the US landed their troops on DNG and start launching attacks on the Japanese army. There were two main reasons why the Americans were interested in DNG. The first one was because it was a strategic location on the march from Taiwan towards Japan. And the second one was about the Japanese oil reserves. Japan was heavily depended on the oil reservoirs of Dutch East Indies. This was because the US cut of their oil supplies due to the aggressive expansion of Japan at the time. So DNG was a perfect starting point for the US to eventually cut of the Japanese oil supply to the Dutch East Indies.

In the meantime, the Japanese had relocated their headquarters in Buma on the north coast of DNG and the allied forces setup theirs in Port Moresby. The Australian troops were able to hold off the Japanese force while they tried to invade Port Moresby.

In March 1943, the 5th US air force was able to sink a troop transport ship in the Bismarck Sea which meant a turning point in the battle of DNG. The liberation was leaded by the army general Donald MacArthur and lasted from 1943 to 1944. They started at Hollandia in April 1944, after that allied forces liberated Wakde (17th may 1944) and Biak (27th may 1944). They advanced quickly to Noemfoer (2nd July 2944), Sansapor (29th June 1944) and Mar (30th June 2944). The last step before the allied forces could proceed to The Philippines was the invasion of Morotai in September 1944.

1.2.3   Used equipment of both the Japanese air force as the Allied forces air force [2]

The most commonly used airplanes by the Japanese air force were:

- Fighter planes: Nakajima Ki-43 Hayabusa (Oscar), Kawasaki Ki-61 Hien (Tony)
- Bomber planes: Mitsubishi Ki-21 (Sally), Kawasaki Ki-48 (Lily), Nakajima Ki-49 (Helen)

The most commonly used airplanes by the Allied forces air force were:

- Fighter planes: P-40 Kittyhawk, P-38 Lightning, P-47 Thunderbolt
- Middle heavy bomber planes: A-20 Havoc, B-25 Mitchell, B-26 Maurader
- Heavy bomber planes: B-17 Flying Fort, B-24 Liberator

2 Remote sensing datasets

The goal of this thesis is to perform a classification of the area surrounding 3 WWII airstrips in the Papua Province Republic of Indonesia. To do that, I will use multiple different data sets acquired from different types of satellites

2.1 Sentinel 2 (Optical Imagery)

The Sentinel 2 mission is a mission that consists of two different satellites, Sentinel 2A and 2B. These two are land monitoring satellites, which were launched at the 23rd of June 2015 and 7th of March 2017 respectively by the European Space Agency (ESA). Both satellites carry a multi-spectral instrument (MSI) that will sample 13 spectral bands (ranging from visible to short wave infrared light). Four spectral bands have a spatial resolution of 10 meters, six of them have a resolution of 20 meters and three bands have a resolution of 60 meters. The satellites operate simultaneously in the same orbit at 180° apart from each other at an altitude of 786 km. Due to the fact that the two satellites are identical and operate exactly opposite to each other, the revisiting time is five days.

Since I am going to do a classification of the area I decided to use as much bands I could but I do have computational problems and therefore I decided to use the bands of 10m resolution (figure 2) and the bands of 20m resolution (figure 3). I rescaled the 20m resolution to a 10m resolution by use of a bilinear interpolation method.
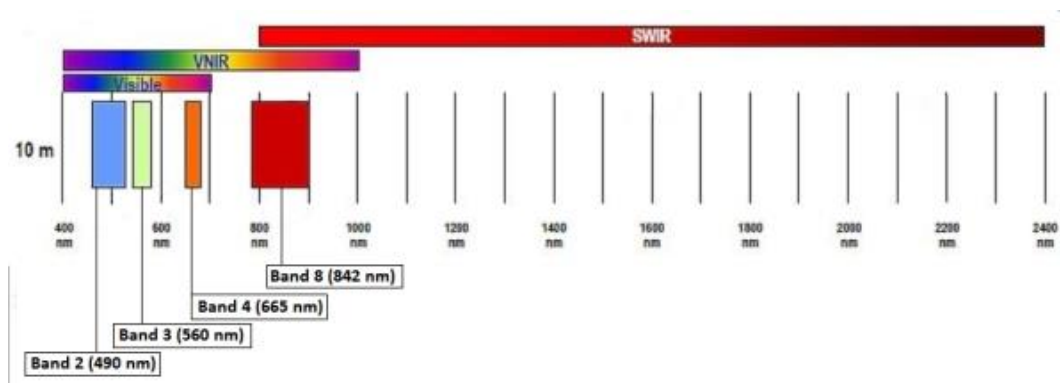

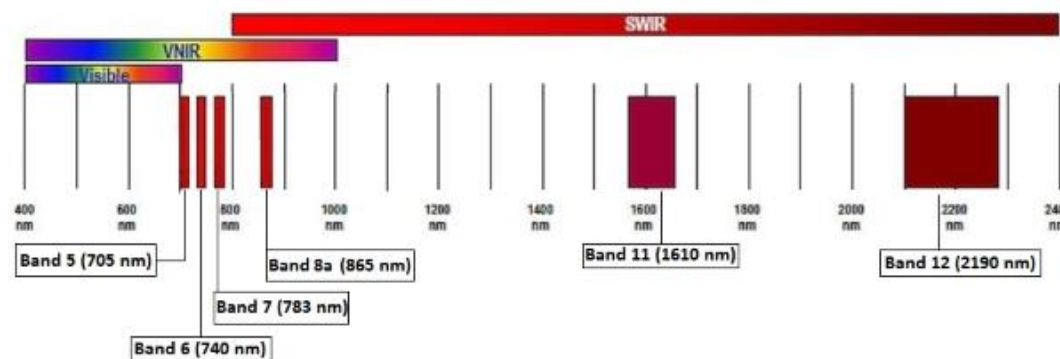*Figure 2: Sentinel 2; Spectral bands 10m resolution [3]*

Figure 3: Sentinel 2; Spectral bands 20m resolution [3]

2.2   ALOS PALSAR (SAR)

The Advanced Land Observation Satellite (ALOS), renamed Daichi, is a Japanese satellite that has been launched 24th of January 2005. It is a satellite that has three different types of measurement equipment on board (AVNIR-2, PRISM and PALSAR). ALOS PALSAR is a Synthetic Aperture Radar (SAR) satellite. An SAR satellite is an active radar satellite, which means it sends out a pulse signal to the ground surface and measures the time that it takes for an echo from an object to come back. This 2-way travel time is a measurement for distance.

An SAR satellite's resolution is splits in two different types of resolution. Range resolution and Azimuth resolution. Range resolution is the resolution in the direction perpendicular to the flying direction and Azimuth resolution is the resolution in the flying direction. Both resolutions are dependent on different parameters.

The range resolution is determined by equation 1.

$$Resolution_{Range} = \frac{1}{2} * c * \frac{1}{B} \ [m] \qquad eq. 1$$

Where:
$c = speed\ of\ light = 3.0 * 10^8$
$B = Bandwidth\ of\ the\ frequency\ of\ the\ transmitted\ signal$
$factor\ 0.5\ is\ due\ to\ the\ 2-way\ travel\ time$

The resolution is Azimuth direction for a SAR satellite is shown in equation 2.

$$Resolution_{Azimuth} = \frac{L}{2} \ [m] \qquad eq. 2$$

Where:
$L = length\ in\ Azimuth\ direction\ of\ the\ antenna$

For ALOS PALSAR, which sends out a 28 MHz or 14 MHz bandwidth pulse and has an antenna with dimension ~8.9 m x 3.1m, this means that the resolution in range direction is ~5.36 m or ~10.71 m respectively and the azimuth resolution ~4.5m.

I chose the ALOS PALSAR satellite because it is an L-Band satellite, which means that the satellite sends out a pulse with centre frequency of 1270 MHz. The advantage of such a long wavelength is that, the signal travels through clouds and the canopy of the rainforest, which is necessary if I want to (potentially) see objects on the surface of the rainforest, like possible (military) WWII heritage [4].

2.3   Shuttle Radar Topography Mission (SRTM)

Shuttle Radar Topography Mission (SRTM) was an instrument aboard the Space Shuttle Endeavour from the National Aeronautics and Space Administration (NASA). This space shuttle launched at 11th of February 2000 and landed 11 days later at 22nd of February 2000. In those 11 days the instrument collected data of 80% of the earth's topography. The resolution of the SRTM data is 1 arc-second, which corresponds to ~30 meters [5].

3 Pre-processing

After acquiring the desired data, I had to pre-process the data, since the downloaded data are raw data. The methods used for pre-processing of the data are different for each type of data that has been used.

3.1   Pre-processing data Sentinel 2

Sentinel 2 data can be downloaded from the Copernicus data hub managed by the ESA. You can download individual granule (of size 100x100 km) and select specific dates that the satellite measured the optical reflectance. I did this for eventually for one granule at four different timespans.

3.1.1   Bottom-of-atmosphere

The data downloaded is called level 1C data, which means that it is Top-Of-Atmosphere (TOA) data. This means that de measured reflectance still holds errors that arise due to the fact that the signal has to travel through the atmosphere. These TOA data have to be corrected for those atmospheric errors. This process is called atmospheric correction and transforms the Level 1A data to level 2A data (Bottom-Of-Atmosphere, BOA).

This atmospheric correction has been done using the plugin sen2cor of the program SNAP. This a Python based plugin that is able to directly transform the TOA data to BOA data. This has been done for all the four timespans, which took about 4.5 hours [6]. The result is shown in figure 4.

Figure 4: Left: Level 1C Input, Bands 2-4; right: Level 2A Output, Bands 2-4, RGB composite images

3.1.2 Normal-Difference Vegetation Index (NDVI)

The normal difference vegetation index (NDVI) is an index which is widely used in optical imagery to discriminate difference vegetation. To calculate is the NDVI equation 3 is used. Lower index means less dense vegetation. If the index is smaller than zero, water is present. The NDVI map is used as an extra feature space (= the total amount of datasets that will be used to as input for the classification) and is shown in figure 5. This is the NDVI map of the Sentinel 2 dataset which is acquired at the 6th of March 2017 and thus with clouds. The clouds in figure 5 are white.

$$NDVI = \frac{Band\ 8 - Band\ 4}{Band\ 8 + Band\ 4} \qquad eq.3$$
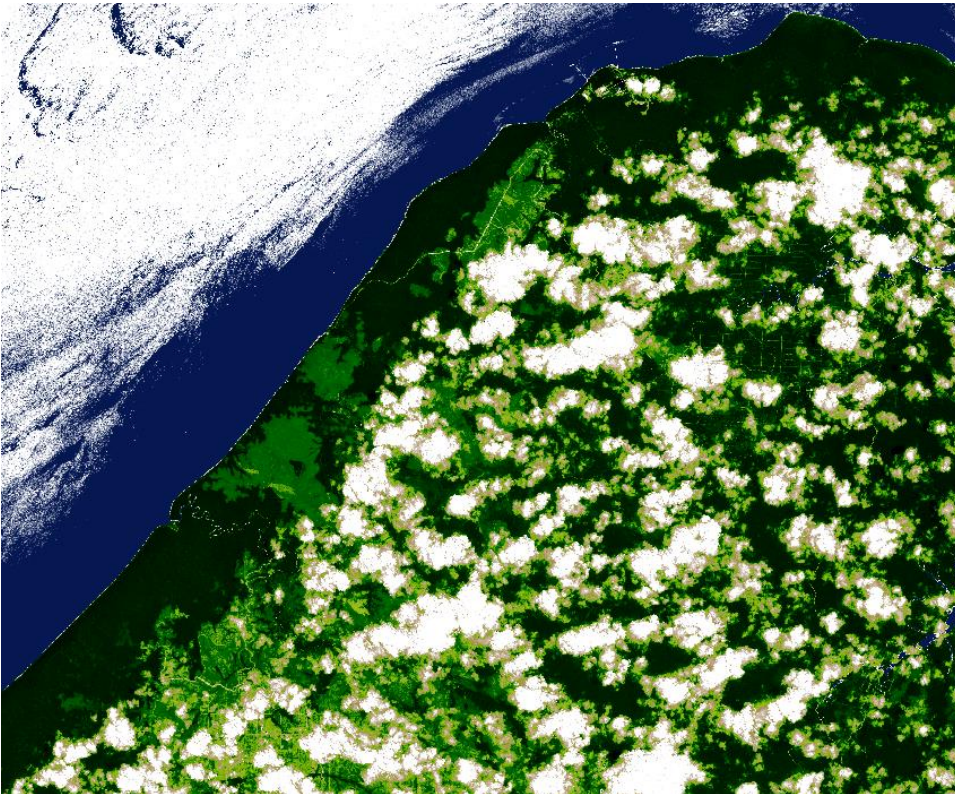

Figure 5: NDVI Map, 6th of March 2017

3.2   Pre-processing data ALOS PALSAR

ALOS PALSAR data can be downloaded from the Alaska Satellite Facility (ASF). I downloaded three different timespans of one single granule. They provide you either with raw radar data, which means that you have to do a lot of pre-processing of the data before you can actually use it, or they provide you with high resolution terrain corrected data. I decided to download the second dataset, which saved me a lot of time pre-processing the data.

3.2.1   Speckle

A well-known issue with radar is the phenomena called speckle. Speckle is noise that inherently exists and decreases the quality of a SAR image. This problem can be solved by either spatially averaging the pixels or by time averaging the pixels. The first option decreases the resolution of the image, which is in this case undesirable, so I decided to time average the pixels. The result is shown in figure 6.
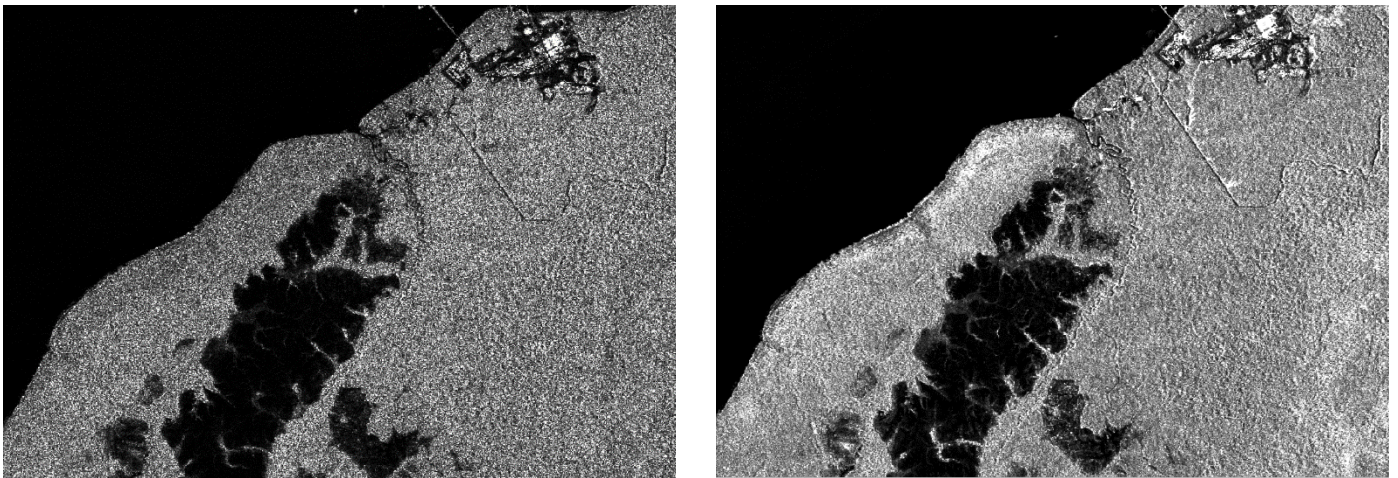

Figure 6: Left: SAR image with speckle; right: SAR image speckle removed by time averaging the pixels

3.2.2 False colour image HH, HV polarization

The radar data from the ALOS PALSAR satellite sends out a pulse with a horizontal polarization and is able to receive horizontal but also vertical. Figure 7 functions to clarify the principle behind polarization. I decided to add a false colour image (figure 8) of these HH and HV polarization to this chapter (<u>not</u> to the feature space). The reason to do so, is that eventually (as explained in the introduction) a ground team is interested in the possible war heritage that might still be there. An advantage of radar (L-band; wavelength: ~24 cm) is that the signal penetrates the dense forest and theoretically I should be possible to find objects (like WWII planes, WWII trucks etc.). However, these objects are very small (not more than 1 or 2 pixels) and thus, I was not able to find any, but a trained eye might. A full-size can be found in appendix A.

Figure 7: Principle of polarization;
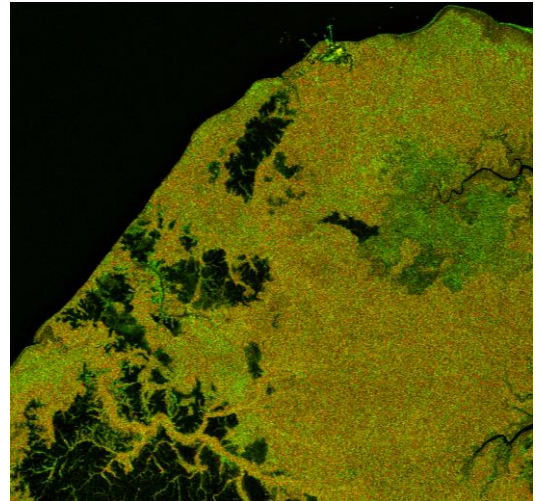*Left: horizontal polarization, Right: Vertical polarization [7]*

Figure 8: False colour image HH, HV polarization

3.3 Pre-processing data SRTM

The SRTM Data can be downloaded from the USGS Earth Explorer website. SRTM measures the topography of the earth's surface, so I am able to use this elevation data to calculate some other geometric attributes, such as the slope, hill shade, ruggedness and aspect.

SRTM measures, as said, the topography of the earth's surface, but is does not measure the location of water. Since there is water in my region of interest, I decided to manually add a value to the water pixels for each attribute. Even though that value itself is an unrealistic value, it ensures that the classification does not go wrong. The resulting images are shown in figure 9.
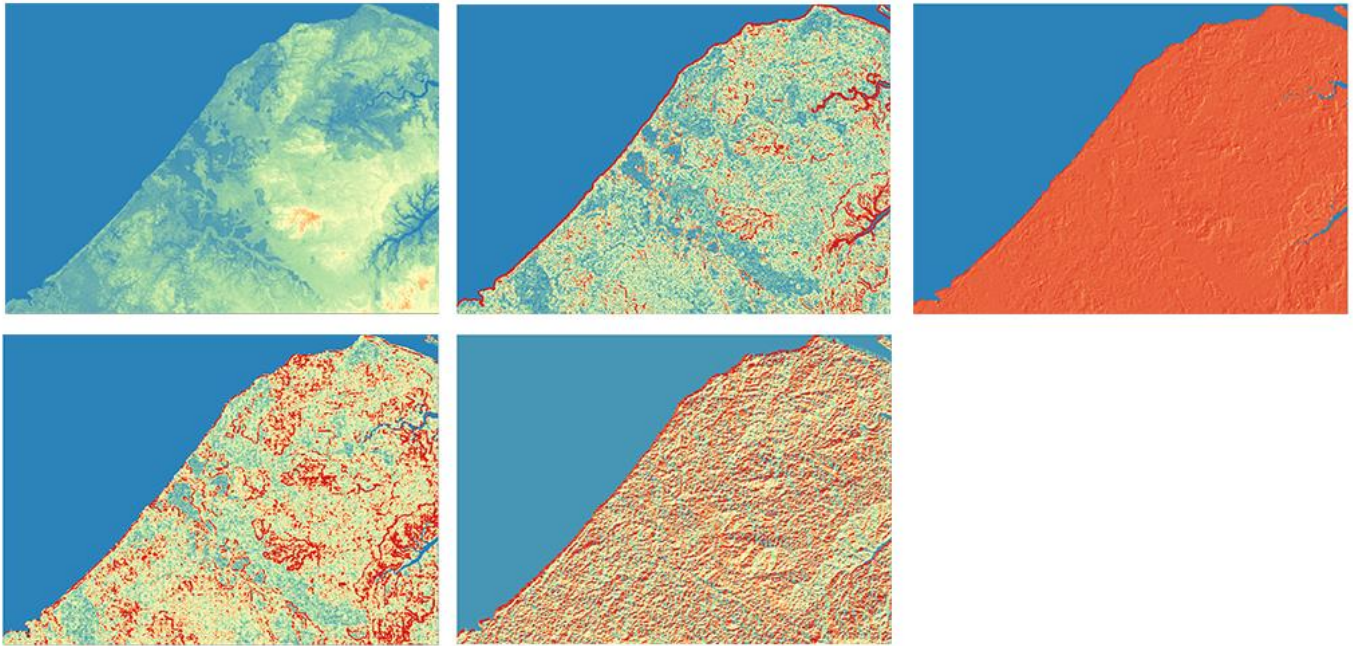

Figure 9: LRTB: Elevation, slope, hill shade, ruggedness and aspect

4 Methodology

This chapter will elaborate the methodology that will be used to create a classification image of the surrounding areas of the three airstrips in Papua Province Republic of Indonesia.

4.1 Classification method

I am going to use the maximum likelihood classification (MLC) method to classify the area. This method is based on statistics and it is a supervised method, which means that the classification will be using training-data. These are data which are already been assigned by hand to a certain class (trees, grass, buildings, etc.), which is possible because the class of those pixels is already known. The training-data will then be used to train the classification method to assign other unseen pixels (pixels that still must be assigned to a class) to the correct class. Better and more training-data will result in a better classification.

The formula that the MLC uses to determine the best class for each individual pixel is originally based on the multivariate probability density function, multi PDF (equation 3).

$$P(x|B_i) = \frac{1}{(2\pi)^{\frac{k}{2}} Tr(C_i)} \cdot \exp\left(-\frac{1}{2}(X - \mu_i)^T C_i^{-1}(X - \mu_i)\right) \qquad eq.\,3$$

Where:
$x = Unseen\ pixel$
$B_i = Class\ i$
$C_i = The\ k * k\ experimental\ covariance\ matrix$
$\mu_i = Mean\ of\ traingsdata\ of\ class\ B_i$

To decide which pixel has to be assigned to a certain class, it is more convenient to rewrite equation 3 (the multi PDF) to the so-called discriminant function (equation 4) which is in fact nothing more than just taking the natural logarithm of $P(x|B_i) \cdot P(B_i)$.

$$\ln\big(P(x|B_i) \cdot P(B_i)\big) = -\frac{k}{2} \cdot \ln(2\pi) - \frac{1}{2} \cdot \ln\big(Tr(C_i)\big) - \frac{1}{2}(x - \mu_i)^T C_i^{-1}(x - \mu_i) + \ln\big(P(B_i)\big) \qquad eq.\,4$$

To finalize this discriminant function, I assume a couple of things.
1. Since I am going to calculate and compare the probability of an unseen pixel to be in each class, constants can be neglected. Thus, the constant $-\frac{k}{2} \cdot \ln(2\pi)$ can be skipped.
2. I assume that each class that the same chance to be the correct class for an unseen pixel, thus $\ln\big(P(B_i)\big)$ can also be skipped.

By applying these two assumptions on equation 4, equation 5 is born. The final equation that the MLC will use to classify the surrounding [8].

$$g(x, B_i) := -\ln\big(Tr(C_i)\big) - (x - \mu_i)^T C_i^{-1}(x - \mu_i) \qquad eq.\,5$$

Where:
$x = Unseen\ pixel$
$B_i = Class\ i$
$g(x, B_i) = Probability\ of\ x\ in\ class\ B_i$
$C_i = The\ k * k\ experimental\ covariance\ matrix$
$\mu_i = Mean\ of\ traingsdata\ of\ class\ B_i$

4.2   Training-data

As said in the previous chapter, the MLC is a supervised classification method, which means that training-data will be used. This is done by hand drawing a lot of polygons from each land cover class and uses that as input to the MLC classification. Figure 10 shows how this is done.
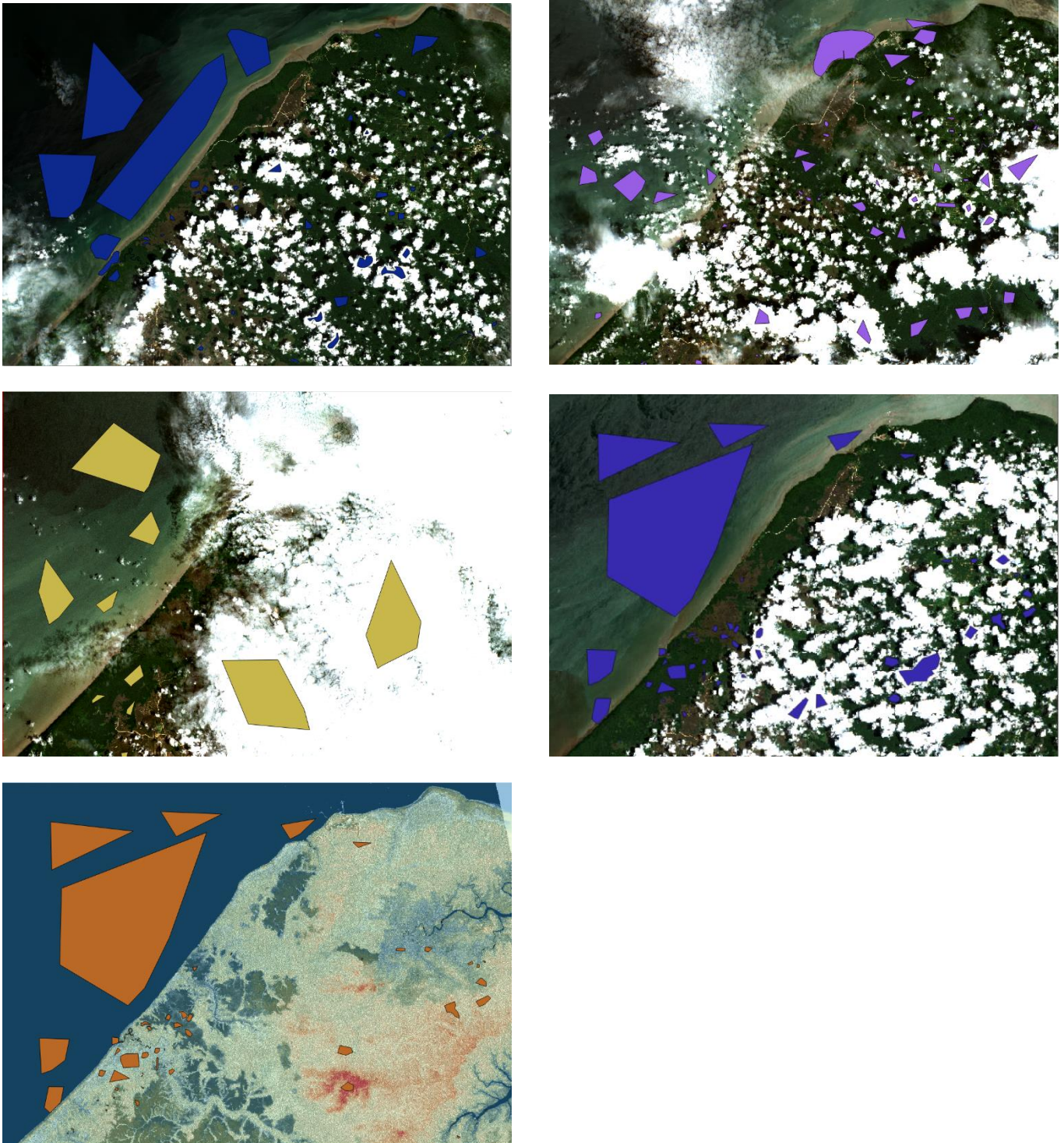

Figure 10: LRTB: Training-data at 26th December 2016, 15th January 2017, 14th February 2017, 6th March 2017 and SRTM and ALOS PALSAR

5 Cloud Masking

One of the main issues om this region is the constant presence of clouds. These clouds will have a big influence on the classification, since the system is not able to look through the clouds and therefore it will be unable to correctly classify the landcover underneath the clouds. I came up with three different methods to mask and eliminate the clouds.

5.1   Most likelihood classification (MLC) Method

The first method to mask and eliminate clouds is called the Most likelihood classification method (MLC). This method is based on the fact that clouds themselves can be a classifiable group. What I did is that I simply classified all four different timespan images with four different training-data (they are shown in chapter 4.2). In these four training-data I included the class clouds and after these four were classified I started with a master image (6th of March 2017) and wherever clouds were classified I filled in the classified data of a different timespan classification. I continued this process until almost no clouds were present any longer. To eliminate the last clouds, I used the classification which was done by only the SRTM and ALOS PALSAR data and filled their classified data at the spots where the last clouds were still present. This process resulted in a cloud-free classified image, which will be shown in in chapter 6.

5.2   Threshold Method

The Threshold method has a different approach then the MLC Method. The threshold method masks clouds before the classification has been done and mosaics each spectral band into each other, via the same approach which has been used in the MLC method to get a cloud-free image. The lower Threshold is there so that cloud shadows are also masked. The threshold values that have been used for this method are shown in table 1 and the resulting mosaiced RGB composite image is shown in figure 11.

Table 1: Threshold values for the Threshold method

Band number	Lower Threshold	Higher Threshold
B11	< 500	> 1100
B05	< 400	> 1200
RGB (composite image B4, B3, B2)	< 15	> 200

Figure 11: Mosaic RGB composite - Threshold method

5.3   Sen2cor scene classification Method

The third method is based on the same method as the Threshold method but does not use thresholds to mask clouds but the scene classification made by the sen2cor plugin which was used to generate bottom-of-atmosphere data (chapter 3.1). The scene classification is shown in figure 12 and the resulting mosaiced RGB composite image is shown in figure 13.
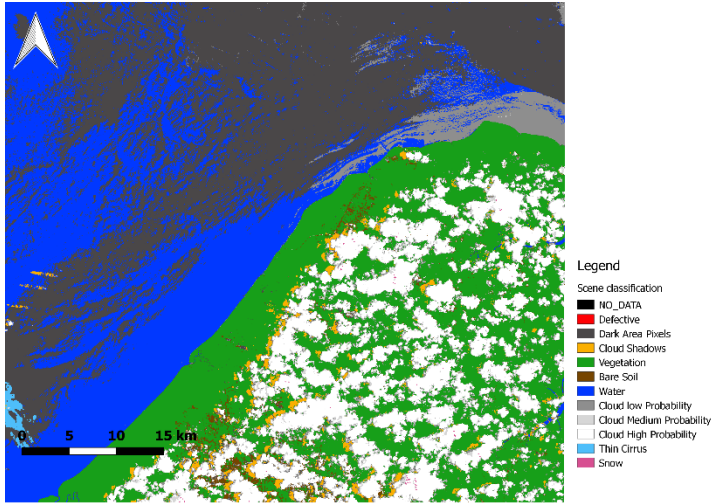

Figure 12: Scene classification Sen2cor

Figure 13: Mosaic RGB composite - Sen2cor Method

6 Results

The results of the classification are shown below. The first classification images are the four different timespan images containing each dataset and one cloud-free classification with only the SRTM geometric attributes and ALOS PALSAR datasets. I decided to not classify urban areas and the three airstrips, since their area is so small compared to the rest, that I was not able to get enough training-data. In the combined figure, however I manually added these areas. The full-size classification maps can be found in appendix A.

6.1   Most likelihood classification (MLC) Method

Figure 14 shows the classification maps of all the different times. I used the 4th classification (6th of March 2017) as a master image and filled in the data from respectively 1st, 2nd, 3rd and 5th image to get a cloud-free classification image (figure 15).
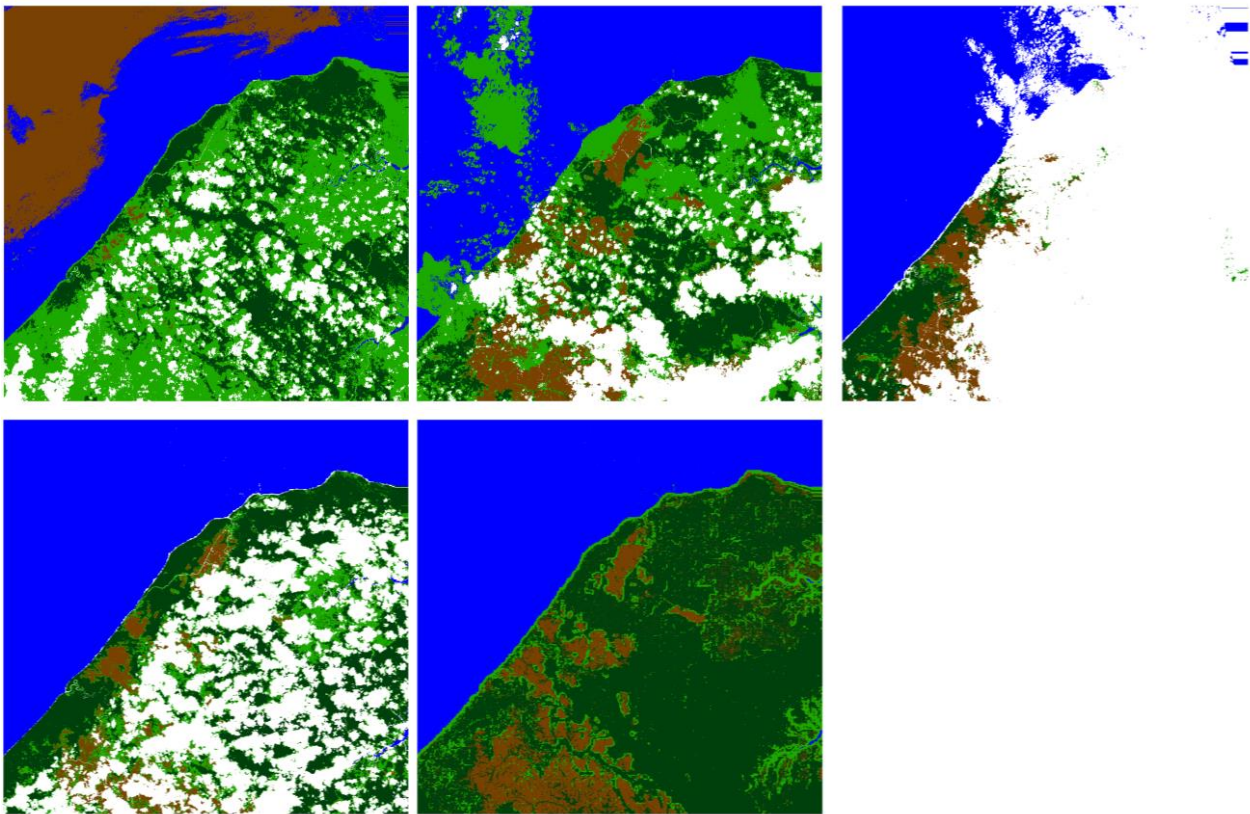

Figure 14: LRTB: Classification maps 25-12-16, 15-01-2017, 14-02-17, 06-03-2017 and SRTM/ALOS PALSAR
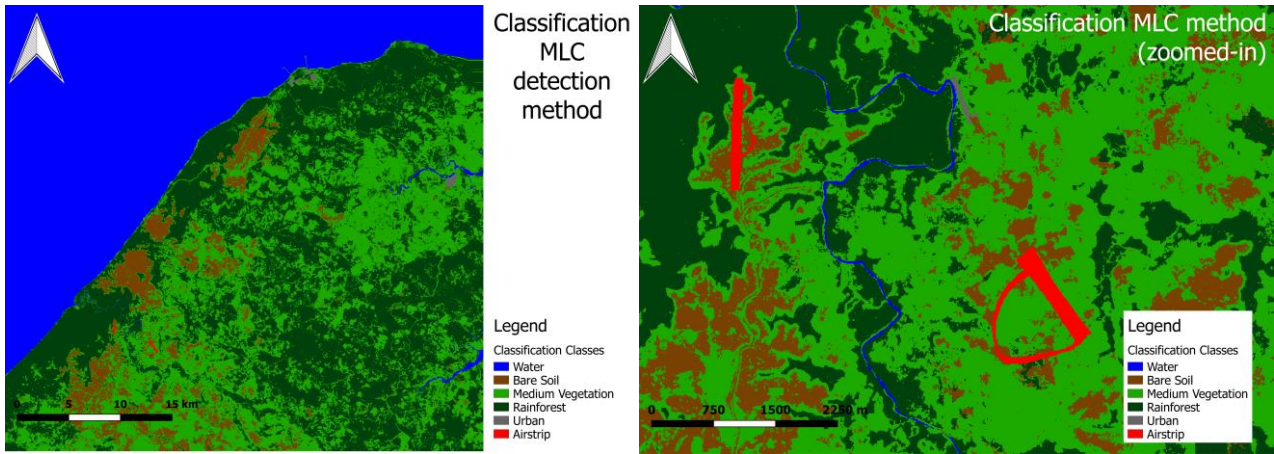
Figure 15: Classification MLC Method

6.2 Threshold Method

Figure 16 shows the classification map which was created with the Threshold method for cloud masking.
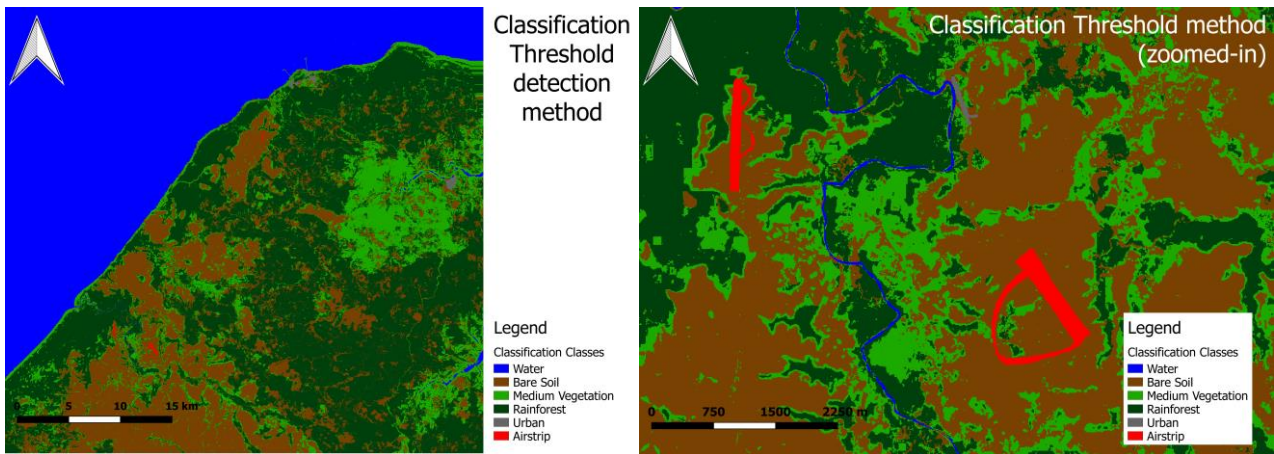

Figure 16: Classification Threshold Method

6.3 Sen2cor scene classification Method

Figure 17 shows the classification map which was created with the Sen2cor method for cloud masking.
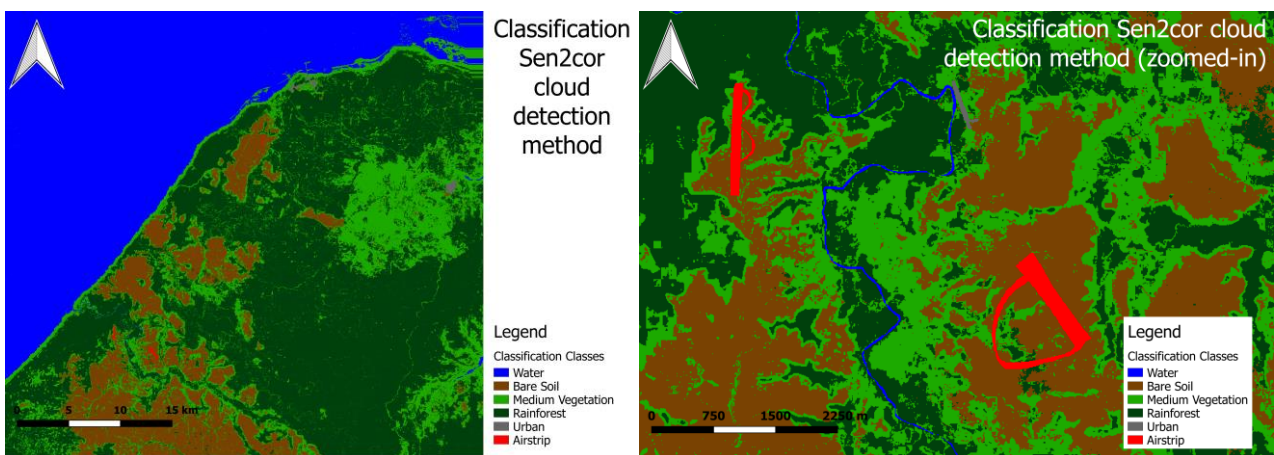

Figure 17: Classification Sen2cor Method

7 Validation

In this chapter, the classification maps of the three methods will be validated. First, I look into the rejection maps of the classification and secondly, I will cross-validate the results with a reference dataset.

7.1   Rejection maps

Most likelihood classification is, as explained in chapter 4, based on statistics. For each unseen pixel, the probability that a pixel belong to a class will be calculated and the class with the highest probability will then be selected to be the assigned class. If one class has a lot higher probability then all the other classes, you could say that the system is very sure that the assigned class is the correct class. I checked that for all the pixels and made a rejection map (figure 18), which is the opposite of a certainty map. A low rejection means that the algorithm is very sure that the assigned class is the correct class. In table 2 the average rejection percentages are shown.
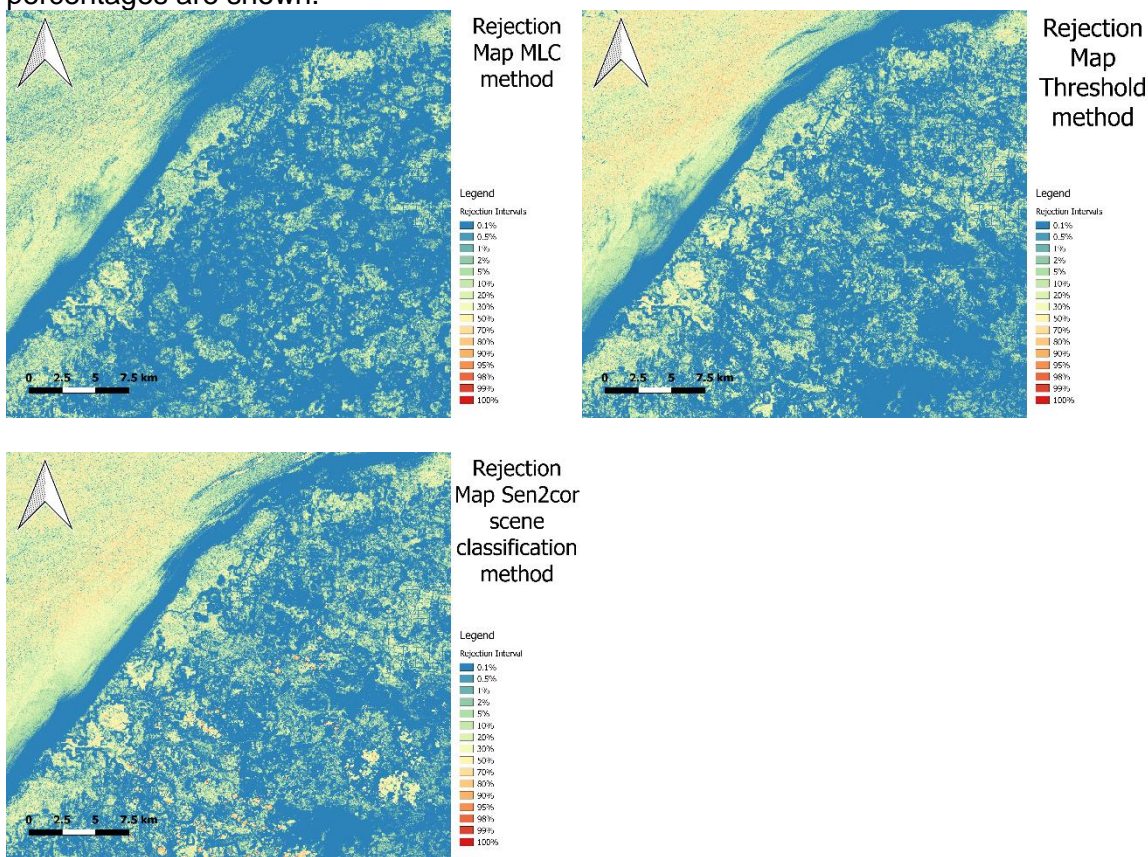

Figure 18: Rejection maps of the classification

Table 2: Rejection percentages

Method	Average rejection percentage
MLC	~ 2.60 %
Threshold	~ 4.10 %
Sen2cor	~ 3.98 %

The rejection maps can also be visualized in boxplots (figure 19). The MLC method has a slightly lower median, box and maximum rejection percentage. The two other methods seem to be almost the same.
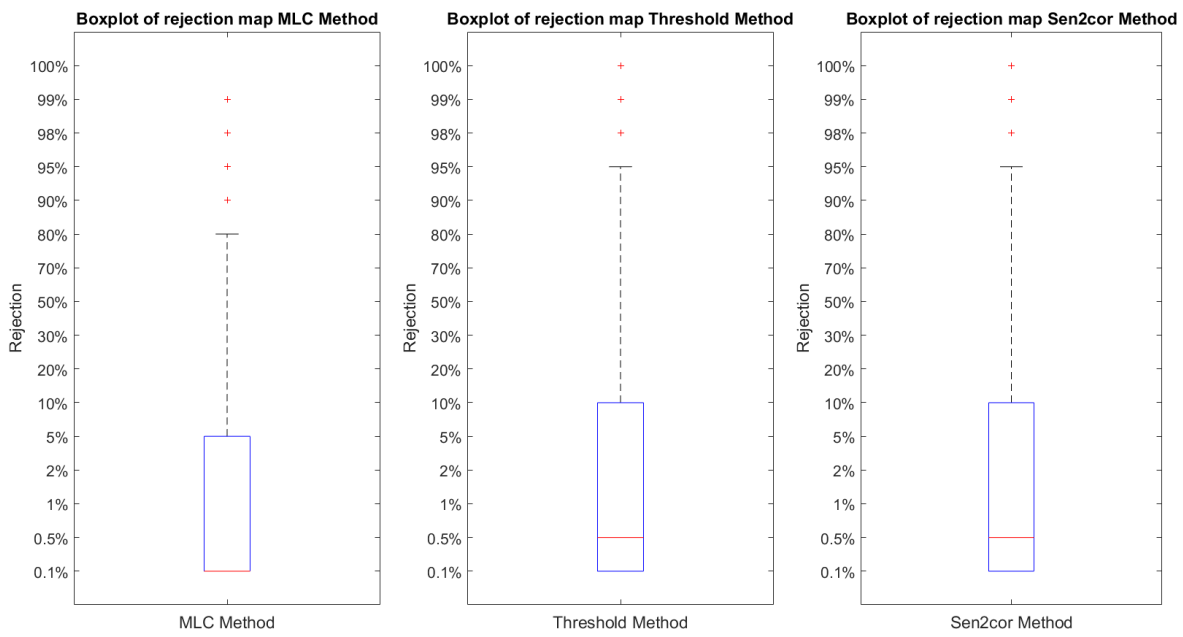

Figure 19: Boxplot rejection maps

7.2   Cross-validation

Even though rejection maps give an idea of how certain the MLC is about its own result, it does not contain a lot of value if de results are not validated. Therefore, I created a validation training-data, which will be assumed to be correct and each class will have approximately equal number of pixels selected, that I used as a reference dataset. I cross-validated the classification of all three methods with the reference dataset and checked how much of the pixels are assigned to the correct class. The pixels of the reference set have been randomly arbitrary selected so that no personal bias could occur. The reference set overlying (in this case the threshold method) a classification map is shown in figure 20 The result of this cross-validation is shown in a bar diagram (figure 21). The percentages represent the average of correct assigned pixels according to the reference dataset.
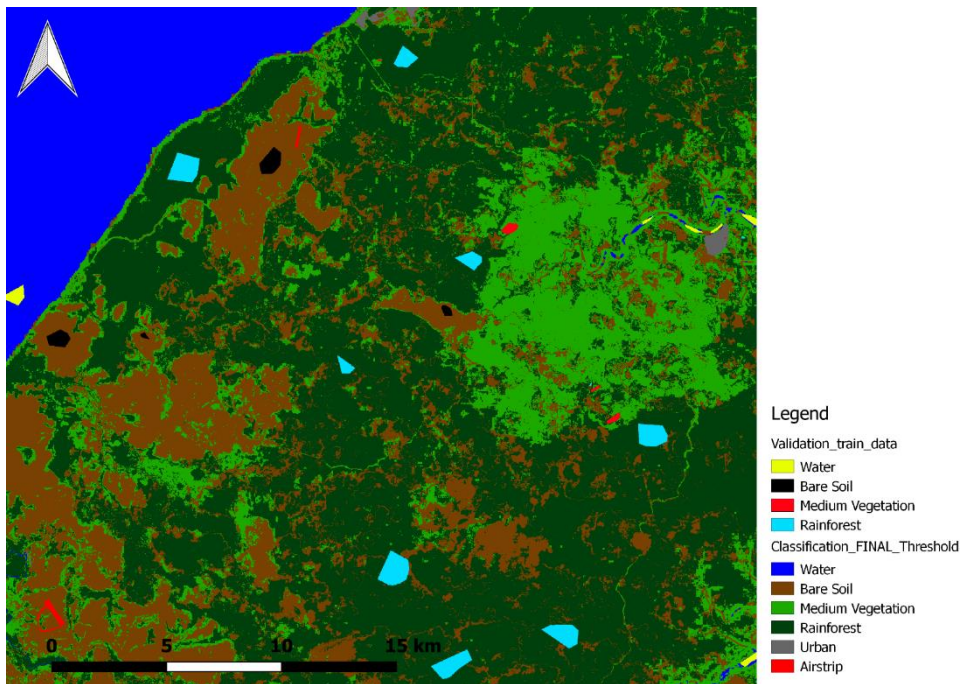

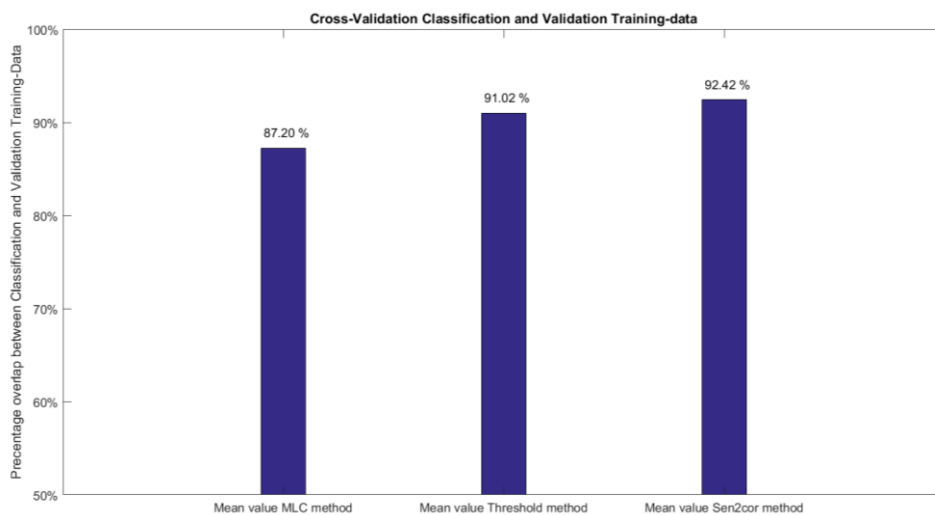*Figure 20: Reference dataset overlying classification (Threshold method)*

Figure 21: Bar diagram cross-validation results

8 Conclusion

Three different cloud masking methods (MLC method, Threshold method and, Sen2cor method) have been used to create a cloud-free classification. Visually, the three classifications seem to differ from each other as can be seen in figures 15, 16, and 17. To check whether this is actually true or this is just an optical illusion, I looked at the statistics of the rejection maps, in the form of boxplots. Even though, these boxplots (figure 19) show a slight difference between the three methods, the boxplots overlap each other quite severe which means that there is no significant difference in the certainty of the rejection maps. The main point you are able to notice about these boxplots is that the algorithm (the most likelihood classification) is very sure about its result. The average rejection percentages (the opposite of certainty) of table 2 are very low for all the methods. Again, a slight difference between the average rejection percentages is present, but this is so small that I concluded that between the methods no significant difference is present, but the overall the certainty of the algorithm is high.

The algorithm's certainty alone is not enough to see what the quality is of the classifications. Therefore, I created a validation training-data, which is assumed to be 100% correct, and used that dataset to cross-validate the classification results. The results (figure 21) show the percentages of pixels that are assigned to correct class. Unfortunately, the values are again very close to each other which means that between the methods there is no significant difference in the quality of the classification. The values (on average about 90%) are however very high, this means that the overall quality of the classifications is very good.

So, to conclude these comments, there is no significant difference in the classifications, when using different cloud masking methods. The three classifications do however have a high quality, proven by rejection maps and cross-validation, and thus I am confident enough to hand over the three classifications to be used as guidance maps for the research team to do their in-situ exploration in the winter of 2017/2018.

9 Discussion & Recommendations

Even though all three classifications went very good, there is room for improvement. One of the things that I could not do is creating absolute correct training-data. The best method to do so, is obviously undertaking a ground operation to several places in the area of interest and simply checking by hand which class is present at that pixel. If you would do this, you are a lot more certain that you have selected the correct class in your training-data. This applies not only for the training-data, but of course also for the reference dataset that has been used for the cross-validation.

Secondly, clouds were a big problem and I came up with three different methods for cloud detection and masking. The MLC method does assume that the ground cover of the area of interest does not change over time, which is not necessarily true. Due to the fact that with this method, I do not see the actual RGB composite image, I could not verify that there was no change. An imaginable effect of this might be that one pixel in the classification is classified as rainforest while the one directly next to it (which is in the master image a cloud) is classified as bare soil, since over time, there was a forest fire which burned the second pixel. The pixel is classified as bare soil, while (in the master classification) this was still rainforest or (if you select a different mater image, in this case after the forest fire) vice versa.
In this particular area, I did assume that the land cover was not changing since the area is very remote (thus little to no interference of humans) and I did not see any signs of natural disturbance, but if somebody wants to use the MLC cloud masking method, change over time should definitely be taken into account.

The Threshold cloud masking method could be improved by selecting better and more carefully the thresholds, since as you can see in figure 11 each mosaiced area still has some little white circles around it. These white circles are the edges of the clouds where they fade into the air and thus hard to discriminate and remove. The other thing you can see is that due to the Solar Zenith Angle (the angle between the sun and the land) the mosaiced areas have slightly different RGB values for the same landcover (as an example: in image 1 rainforest is dark green, in image 2 rainforest is just a little darker/lighter green). To improve the classification, one should adjust for this different Solar Zenith angle.

The Sen2cor scene classification cloud masking method, has the same issues as the Threshold method. Unfortunately, since the clouds are masked by the Sen2cor plugin, you cannot change improve the detection and masking in contrast to the Threshold method. However, the Sen2cor method also has the same problem of the Solar Zenith Angle and adjusting for that problem should increase the quality of the classification.

Next, I noticed that in 1st classification (26th of December 2016) in figure 14 is quite different than the other time classifications. As you can imagine, this has an impact on my final classification. So, the 1st classification might not be true. Further research should be done to examine whether this classification just simply went wrong or that the different classification is due to different reason (for instance, a lot of rain will most likely change bare soil into grass/small bushes, which is than classified as medium dense vegetation).

Fourthly, one should be careful by selecting the validation reference dataset. Some land covers are easier to classify than others. It is less obvious that the most likelihood classification mixes up rainforest with water than with medium dense vegetation. If you then select your reference data you should be careful that you do not select more pixels of the easier classified class and cross-validate them. If you do so, you will get an average overlap percentages, which is not true.

At last, there are a lot of aerial photographs of the three airstrips (and their surroundings), which have been made in WWII. A photograph specialist might be able to edit these photos to non-oblique pictures and use them as another method to cross-validate the results. A different advantage of these photos is that it might be possible to get knowledge of the military equipment which has been used and their location, so that it is easier to eventually find this military equipment. These WWII photos are available at Mr. Kreuger, B.

10 Bibliography

[1]   New-Guinea during World War II. (n.d.). Retrieved from
        http://www.papuaerfgoed.org/en/New-Guinea_during_World_War_II

[2]   Kreuger, B. (personal communication, May 22, 2017).

[3]   Sentinel 2: The operational Copernicus optical high resolution land mission.
        (2016, August 13). Retrieved from
        http://esamultimedia.esa.int/docs/S2-Data_Sheet.pdf

[4]   PALSAR Phased Array Type L-band Synthetic Aperture Radar. (n.d.). Retrieved from
        http://www.eorc.jaxa.jp/ALOS/en/about/palsar.htm

[5]   Shuttle Radar Topography Mission. (n.d.). Retrieved from
        https://www.jpl.nasa.gov/missions/shuttle-radar-topography-mission-srtm/

[6]   MuellerWilm, U. (2016, November 25). Sen2Cor Configuration and User Manual. Retrieved from
        http://step.esa.int/thirdparties/sen2cor/2.3.1/[L2A-SUM]%20S2-PDGS-MPC-L2A-SUM%20[2.3.0].pdf

[7]   Polarization. (n.d.). Retrieved from
        http://electron6.phys.utk.edu/light/6/polarization.htm

[8]   Lindenbergh, R.C. (2016, December 6). CIE4603, Geo-signal analysis. Retrieved from
        https://blackboard.tudelft.nl/bbcswebdav/pid-2898169-dt-content-rid-9999070_2/courses/38937-
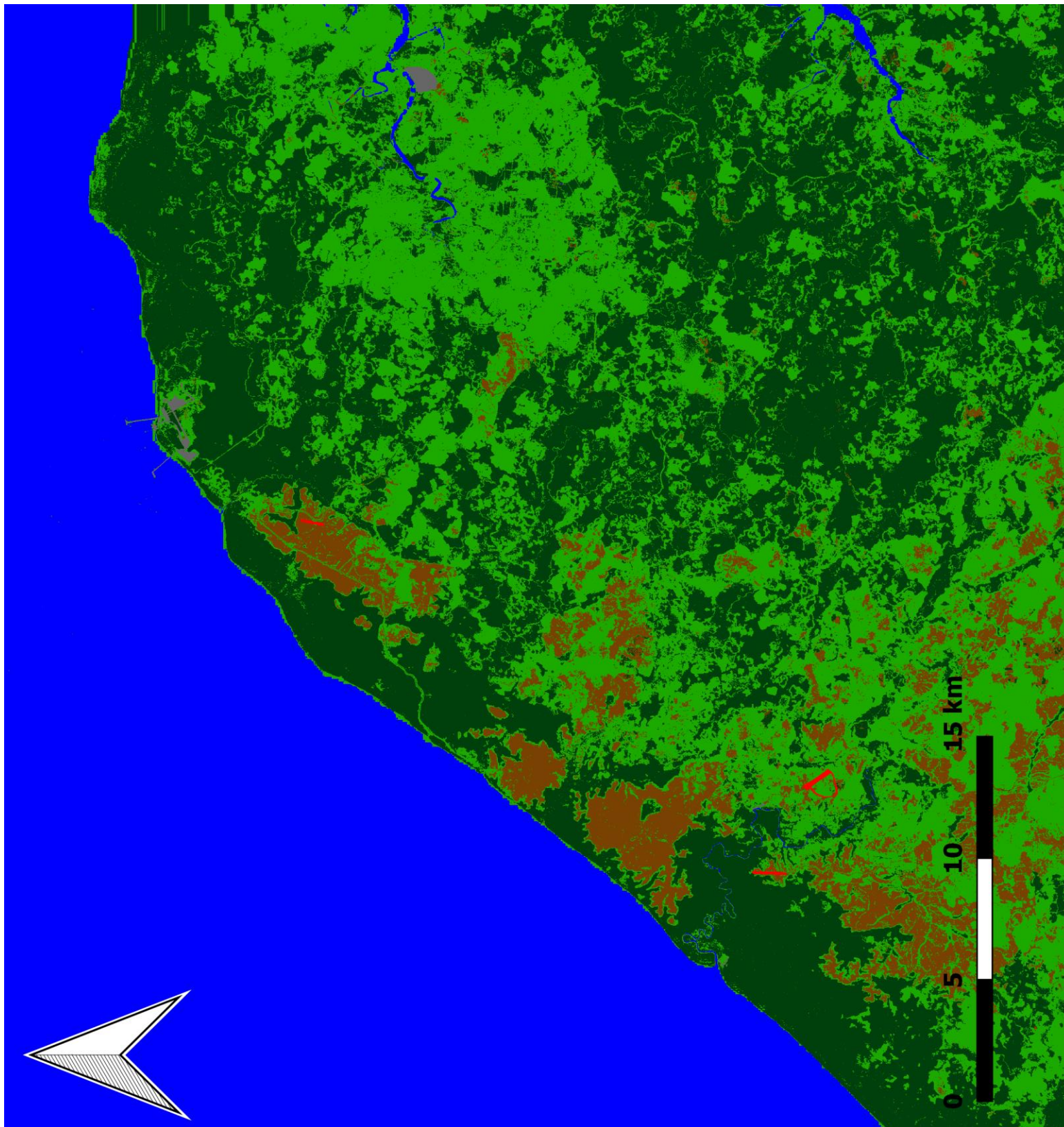   161702/GeoSig2016_4%281%29.pdf

11 Appendix A

11.1 False colour image HH, HV Polarization full size
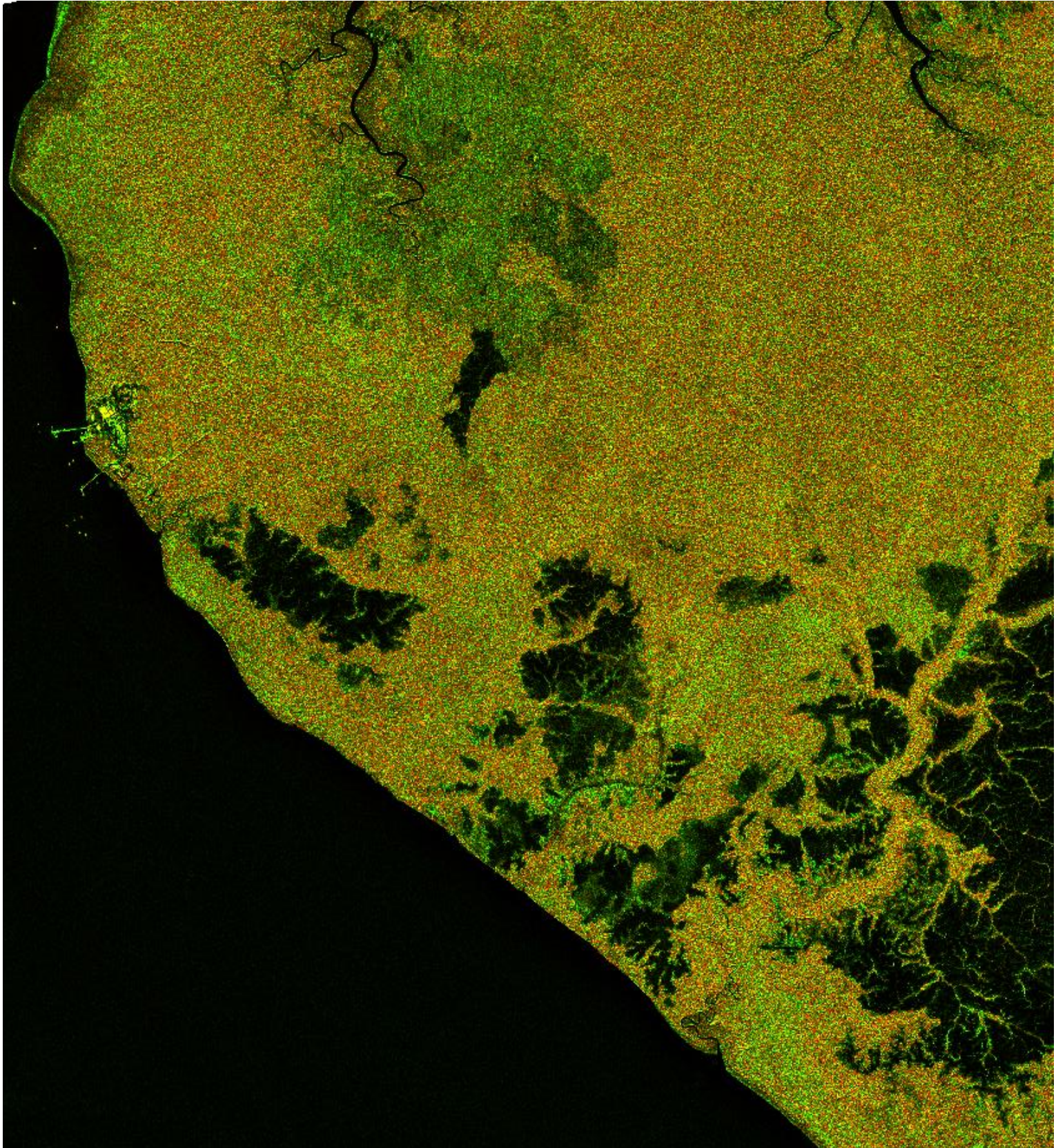
11.2 Classification maps full size

Classification MLC cloud masking Method. The legend can be found in figure (15).

Classification Threshold cloud masking Method. The legend can be found in figure (16).

Classification Sen2cor cloud masking Method. The legend can be found in figure (17).