

POLITECNICO DI TORINO

Master's Degree in Environmental and Land Engineering

Master Thesis

Land Cover Classification of Deception Island with SAR and Optical Images



Supervisors

Prof. Piero Boccardo

Prof. Rogelio De La Vega Panizo

Candidate

Vittoria Marolo

March 2021

INDEX

1. OBJECTIVES AND INTRODUCTION	5
1.1 OBJECTIVE OF THE WORK.....	5
2. CASE STUDY: DECEPTION ISLAND	13
2.1 LOCATION, GEOGRAPHY AND TOPOGRAPHY	13
2.2 GEOLOGY AND GEOMORPHOLOGY	16
2.3 METEOROLOGY	19
3. MATERIAL.....	21
4. METHODS.....	31
4.1 SAR.....	42
4.1.1 PREPARATION OF THE IMAGE	44
4.1.2 SEGMENTATION	50
4.1.3 UNSUPERVISED CLASSIFICATION	52
4.1.4 SUPERVISED CLASSIFICATION	56
4.2 OPTICAL	59
4.2.1 PREPARATION OF THE IMAGE	60
4.2.2 UNSUPERVISED CLASSIFICATION	63
4.2.3 SUPERVISED CLASSIFICATION	65
4.2.4 INDEX CLASSIFICATION	67
4.3 PREPARATION OF THE GROUND TRUTH.....	78
4.4 REALIZATION OF THE CONFUSION MATRICES.....	87
5. RESULTS AND DISCUSSION.....	93
5.1 SAR.....	94
5.1.1 UNSUPERVISED CLASSIFICATION	94
5.1.2 SUPERVISED CLASSIFICATION	97
5.2 OPTICAL	103
5.2.1 UNSUPERVISED CLASSIFICATION	103
5.2.2 SUPERVISED CLASSIFICATION	105
5.2.3 INDEX CLASSIFICATION	110
6. CONCLUSIONS	113
7. BIBLIOGRAPHY.....	115
8. APPENDIX	120
8.1 GEOLOGIC MAP	121
8.2 GEOMORPHOLOGIC MAP	123
8.3 TOPOGRAPHIC MAP	125
8.4 IMAGES USED.....	127

8.5 SAR.....	129
8.5.1 EFFECT OF THE FILTERS.....	130
8.5.2 SEGMENTED IMAGES.....	132
8.5.3 UNSUPERVISED CLASSIFICATION	134
8.5.4 SUPERVISED CLASSIFICATION	136
8.6 OPTICAL	145
8.6.1 UNSUPERVISED CLASSIFICATION	146
8.6.2 SUPERVISED CLASSIFICATION	148
8.6.3 INDEX CLASSIFICATION	150

1. OBJECTIVES AND INTRODUCTION

In this introductory chapter, the purpose of the work will be clarified and its approach will be explained in general, from a theoretical point of view (a more detailed and practical explanation will be given in Chapter 4).

1.1 OBJECTIVE OF THE WORK

Optical images and SAR images have different strengths and weaknesses. Optical images are easier to interpret visually, as the human eye is more accustomed to such views, but they are dependent on solar illumination and weather conditions. SAR images, on the other hand, actively illuminate the scene with microwaves, so the presence of clouds is no longer a problem and even scenes captured at night will be visible. The difficulty, however, of SAR images lies in the more difficult interpretation, due to the presence of a single band in the case of single polarisation, but above all due to the noises which affect them: since the image is acquired from the side it is subject to geometric distortion and presents shadows, and furthermore, the interference between the various backscattering of point elements in the same pixel causes a "salt and pepper" effect typical of these images, called speckle.

In addition, there are several studies of SAR images that show their goodness of classification in urban contexts, where there are geometric elements with strong contrasts, such as streets or buildings, but there is a lack of studies regarding non-urbanised areas, where the contrasts between ground covers are less clear and classification is consequently more difficult.

The objective of this study is therefore to classify an optical image and a SAR image of the same non-urbanised area, acquired on different days, in order to compare the results. Specifically, images of Deception Island acquired one day after the other were analysed: the optical image on 19.01.2020 and the SAR image on 20.01.2020, and the covers that will be classified are bare ground/rock, snow/ice and water.

The specific reason for choosing Deception Island as the study area lies in the fact that it is subject to seismic and volcanological hazards, which are constantly monitored by two scientific bases located on it: it is therefore a necessity to have maps of the ground cover which are present in order to produce risk maps, and it would be appropriate to have techniques capable of providing such maps quickly (the covers are in fact not constant, but subject to continuous change in the case of snow or temporary lakes).

The work will be structured as follows:

- Chapter 1: this chapter gives a general description of the processes followed and their theoretical basis;
- Chapter 2: an overview of the study area, Deception Island, is given;

- Chapter 3: the images used for the study are explained in detail: those to be analysed and the sources used as ground truth;
- Chapter 4: the procedure followed for the processing and classification of the images is described in detail, as well as the creation of the ground truth and thanks to it the quantification of the goodness of the classifications;
- Chapter 5: the best results obtained are presented and discussed, that is the classified images and the associated confusion matrices;
- Chapter 6: the conclusions of the work are drawn.

1.2 STRUCTURE OF THE WORK

In order to classify the optical and the SAR images, the work was carried out on two parallel lines, and consisted of a first preparation phase, which was different in the two cases, and a second classification phase, firstly unsupervised and then supervised, also, an index-based classification was added in the case of the optical image. The goodness of the classifications was then quantified through the use of confusion matrices, whereby the images produced were compared with a ground truth, derived from two separate sources: the topographic map produced in 2005 by the Spanish Land Army and the Pixel QA band of a Landsat image acquired at a date close to those analyzed. The qualitative and theoretical introduction to the procedure carried out will now be given, while the process will be described in detail in chapter 4. To have more details about the images used and their satellites, please refer to chapter 3.

PREPARATION OF THE IMAGES

Before classifying the images, it is necessary to carry out some preparation operations, firstly because the downloaded images include areas that will not be studied (Figures 3.2 and 3.6), and continuing with an excessively large image will only lead to an unnecessary burdening of operations, secondly because, while the optical image of Level A2 is already ready for processing, being characterised by a so-called Bottom of Atmosphere reflectance, from which the effect of the atmosphere has been removed, the SAR image requires a series of operations aimed at correcting errors typically present in these images, such as speckle and geometric distortion caused by lateral acquisition.

SAR Image

For the preparation of the SAR image, the preparation operations were:

- cropping of the target area, as mentioned above, because the original image included a much larger area than the area of interest;
- calibration, in order to normalize the present values; and to do this it uses the ground area, the illuminated area in the look direction or the tangent of the local angle of incidence;

- filtering, to reduce the speckle effect, granular noise that causes a "salt and pepper" effect in SAR images, due to the interference of the waves reflected by many point scatterers;
- georeferencing and geometric correction of the image, using an appropriate DTM,
- conversion of the band into decibels, to obtain values that are easier to handle.

After these operations, the clipping was also refined with the coastline.

As far as filtering is concerned, four different filters were tested: an initial application of the Median filter was made to the image, then, as suggested by Lee (Lee et al., 1989), a multiple application of the Lee filter was carried out, of which the two modified versions present in SNAP were also tested: Lee Sigma and Refined Lee. These four filters will now be briefly described.

Median

The Median filter works by moving a scan window that replaces the center pixel with the median of the pixels in the window. It is effective in reducing point and isolated noise such as speckle, suppressing or even removing it, and for this reason it is often used as an alternative to the Mean filter, which instead causes blurring (Kupidura, 2016). Its drawback is that it can lead to the deletion of small objects or subtle elements; however, this may be the case in an urban setting, but not in the image studied, where elements of this type were not present.

Lee

Many variations of this filter have been proposed over the years, but the original Lee filter dates back to 1980 (Lee, 1980). It is considered to be the first model-based speckle removal filter and it is a local statistics filter that operates linearly and reduces speckle by minimizing the mean square error of the mean intensity within the scan window.

Lee Sigma

The Lee Sigma filter, devised by Lee in 1983 (Lee, 1983) assumes that the distribution of reflected signals is Gaussian and calculates the sigma probability, then considers as outliers the pixels that fall outside the two-sigma interval starting from the central pixel of the scan window and averages the others and then substitutes the value of the central pixel with this one.

This filter is able to bring out the parts of the image that have a different contrast, but does not remove the dark spots.

Refined Lee

The Refined Lee filter was designed by Lee in 1981 (Lee, 1981) to overcome the problem of noisy edge boundaries produced by the Lee filter. It operates using a 7x7 moving window that, when it encounters an edge, estimates its orientation using the local gradient: inside the window there are in fact eight non-square windows for eight different edge orientations. The algorithm considers the window that best fits the orientation of the edge and with it

calculates the local mean and variance (if no edge is found, the calculations are performed taking into account the whole area of the 7x7 window).

This filter works particularly well on edges and high contrast areas, but the same cannot be said if the image is excessively segmented.

Optical Image

For the preparation of the optical image, the operations were:

- cropping of the interested area, as for the SAR image;
- resampling of the bands to make them all have the same cell size and thus be able to carry out the operations necessary to calculate the indices.

As for the SAR image, the clipping was refined with the coastline.

The indices that were considered were the following: NDWI, MNDWI, Mu-WIR, NDSI, NDGI, NDSII. They will now be briefly illustrated, and the bands indicated in brackets in the formulas refer to the Sentinel 2 satellite, which is the one of interest.

NDWI

The Normalized Difference Water Index (NDWI) is an index used to detect the presence of water, of which there are at least two versions: one, introduced by Gao in 1996 (Gao, 1996), which uses the near-infrared band (NIR) and the short-wave infrared band (SWIR) and is sensitive to changes in the water content of vegetation, useful because it is less sensitive to atmospheric effects, which instead condition the NDVI (Gao, 1996), an index with a similar purpose. Its formula is as follows:

$$NDWI_{Gao} = \frac{NIR (B8) - SWIR (B11)}{NIR (B8) + SWIR (B11)}$$

The other version, which is the one used in this case, was introduced by McFeeters in 1996 (McFeeters, 1996) and is instead used to detect water bodies, being able to enhance them while leaving out coverages such as soil and vegetation. Its formula exploits the green and near-infrared bands and is given below:

$$NDWI_{McFeeters} = \frac{Green (B3) - NIR (B8)}{Green (B3) + NIR (B8)}$$

MNDWI

The Modified Normalized Difference Water Index (MNDWI) was first introduced by Xu in 2006 (Xu, 2006) and replaces the near-infrared band used in McFeeters' NDWI with the mid-infrared band (MIR), with the advantage of increasing the positive value of water, since it absorbs more light in the MIR than in the NIR, and having negative values for buildings, soil and vegetation, which reflect more in the MIR than in the NIR, so as to have a greater contrast between water and soil, facilitate water extraction and obtain a more accurate classification.

Its formula is as follows:

$$MNDWI = \frac{Green (B3) - MIR (B8A)}{Green (B3) + MIR (B8A)}$$

MuWI-R

The MuWI-R, or Multi-spectral Water Index Revised (MuWI-R), is a reduced version of another index, the Multi-spectral Water Index Complete (MuWI-C), to which it removes some non-fundamental terms, making calculation easier. Both indices were developed in 2018 by Wang et al (Wang et al., 2018) and were specifically designed for Sentinel 2, with the aim of generating water maps at 10 m resolution, thus having higher classification accuracies. Its formula consists of a combination of normalized differences, thus ensuring threshold stability:

$$MuWIR = -4 * \frac{Blue (B2) - Green (B3)}{Blue (B2) + Green (B3)} + 2 * \frac{Green (B3) - NIR (B8)}{Green (B3) + NIR (B8)} + 2 * \frac{Green (B3) - SWIR (B12)}{Green (B3) + SWIR (B12)} - \frac{Green (B3) - SWIR (B11)}{Green (B3) + SWIR (B11)}$$

NDSI

The term Normalized Difference Snow Index (NDSI) was first coined by Hall et al. in 1995 (Hall et al., 2010). Although methods similar to NDSI, which used visible and near-infrared bands with the intention of mapping snow and separating it from clouds, have been tested since the mid-1970s. It is an index that takes advantage of the high reflectance of snow and ice in the visible bands, particularly in the green band, and their strong absorption in the near-infrared or shortwave areas to map and separate them from clouds, which have high reflectance in both parts of the spectrum mentioned above. Its formula is as follows:

$$NDSI = \frac{Green (B3) - SWIR (B11)}{Green (B3) + SWIR (B11)}$$

NDGI

The Normalized Difference Glacier Index (NDGI) was formulated by Keshri in 2009 (Keshri et al., 2009), with the intention of separating pure snow and ice from ice mixed with debris. The bands used are green and red, as the spectral characteristics of the two mentioned categories are different in these two bands and the two histograms can be easily divided. Its formula is as follows:

$$NDGI = \frac{Green (B3) - Red (B4)}{Green (B3) + Red (B4)}$$

NDSII

Like the NDGI, the Normalized Difference Snow Ice Index (NDSII) was formulated by Keshri et al. in 2009 (Keshri et al., 2009), and its purpose is to separate snow from ice using the green and near-infrared bands, as their reflectance is similar in the green but different in the near-infrared, where the reflectance of the ice decreases. The index formula is as follows:

$$NDSII = \frac{Green (B3) - NIR (B8)}{Green (B3) + NIR (B8)}$$

Once these preparation operations had been carried out, it was possible to proceed with the classification of the images.

CLASSIFICATION

Classifying an image means grouping its cells into groups, called classes, based on spectral or texture characteristics. Classification methods are divided into supervised and unsupervised, whereby in the first case it is the user who selects regions of interest in the various classes to provide the algorithm with training data on which to base itself in order to group the cells in the defined classes, while the second is an automatic method that does not require training data.

The classifiers used were those present in the ArcGIS software and will be briefly illustrated below: the Iterative Self-Organizing (ISO) cluster classifier for unsupervised classification and the Maximum Likelihood (ML), Random Trees (RT) and Support Vector Machine (SVM) classifiers for supervised classification.

ISO cluster

The ISO cluster involves subdividing cells into the number of classes specified at the start in an automatic and iterative manner using a technique called migrating means, based on the calculation of the Euclidean distance in feature space (each cell is assigned to the cluster for which the distance is minimum) and on the recalculation of the cluster mean at each iteration [4].

Maximum Likelihood

The modern version of the Maximum Likelihood classification method was created by Fisher, between 1912 and 1922 (Hald, 1999), although rudimentary forms already existed earlier under different names. The algorithm that operates the Maximum Likelihood Classification tool in ArcGIS is based on Bayes' decision theorem and on the fact that the sample cells of each class are normally distributed in multidimensional space. Under these assumptions, each cell is assigned to the class to which it is most likely to belong based on mean and covariance [5]. It is one of the most used classifiers in remote sensing and provides very good results

when working with unimodal distributions, but has difficulty working with multimodal distributions since it assumes that the distribution is normal (Liu et al., 2011).

Random Trees

The Random Trees classifier, also known as Random Forests, was first introduced in 2001 by Breiman (Breiman, 2001) and works with multiple decision trees, each based on a set of random samples that are independent of those in the other trees (there is thus a possibility that the same sample is selected from several trees). Each tree assigns a probability of belonging to a certain class for a certain cell and the final assignment is given by the average of the various probabilities calculated by each tree. This classifier has gained much fame in the remote sensing world due to its speed and ability to handle large amounts of data, it is however very sensitive to sampling design and this should be taken into account in order to avoid misclassification as much as possible (Belgiu et al., 2016).

Support Vector Machine

The Support Vector Machine classifier was initially developed by Vapnik in 1979 (Vapnik, 1979) and is based on the representation of elements in a feature space, whereby different classes are located in different regions of the space and the assignment of a cell to one class rather than another depends on its proximity to it in the said space. The drawback of this method is that the assumption that different classes are located in completely distinct regions of feature space is theoretical when in reality point data of different classes may overlap and this leads to misclassification, in addition to the fact that this classifier is not optimized to work with data affected by noise and even a small percentage of training samples assigned to the wrong class can drastically reduce the goodness of its performance. It does, however, have the advantage that it does not require a large amount of training data and does not assume any particular distribution for the data representation, which is a good thing since data acquired by remote sensing usually have unknown distributions (Mountrakis, 2011).

These classifications can also be pixel-based or object-based. The first one bases the assignment of a pixel to a given class only on its spectral characteristics, without taking into account how neighboring pixels behave; the second one instead involves grouping neighboring pixels on the basis of spectral characteristics and textures into structures called segments, and then using these for classification, taking into account several of their attributes. For the formation of segments ArcGIS uses the Mean Shift algorithm, conceived in 1975 by Fukunaga (Fukunaga, 1975) to find the modes of a probability density function. In particular, the command works in such a way that a window is scrolled to calculate the mean in order to understand which pixels to include in the segment, so that the latter will be the mean value; continuing to scroll on the image, it will iteratively recalculate the mean so as to verify that each segment has been formed well [7].

Once the classification is complete, it is necessary to evaluate its goodness and this is done by comparing the classified image with a reference one called ground truth and, quantitatively, by making confusion matrices. These are constructed selecting a series of control points, independent from those used as training samples in the case of supervised classification, in which one compares which value is assumed by the pixel in the classified image and which in the reference one (Stehman, 1997). On the basis of this, it is possible to quantify the accuracy of the classifications, using measures such as the producer's accuracy, relative to the errors of omission, for which a class is found to be lacking in points because the points that should have been classified as belonging to it have been inserted in another, and the user's accuracy, relative to the errors of commission, for which instead the pixels that should have belonged to a certain class are classified as belonging to another, which consequently will include more points than it should have [1].

Another value used to estimate the goodness of the classification is the coefficient of agreement k , which represents the overall accuracy of the classification.

User's Accuracy, Producer's Accuracy and k move between 0 and 1, where 0 represents very bad accuracy and 1 represents very good accuracy.

2. CASE STUDY: DECEPTION ISLAND

This chapter will provide an overview of the study area: Deception Island. In particular, its geography and geomorphology will be briefly explained, and the meteorology of the period characterizing the acquisition of the images used will be analysed.

2.1 LOCATION, GEOGRAPHY and TOPOGRAPHY

Deception Island is part of the South Shetland Islands archipelago, located approximately 120 km from the Antarctic Peninsula (Figure 2.1.1), and more precisely it is situated between latitudes 62°53'30S and 63°01'20 "S and longitudes 60°29'20 "W and 60°45'10 "W. It was discovered by the explorer Nathaniel Palmer in 1820 (Smellie et al., 2002), who gave it this name because what at first seemed to be an island, when approaching Neptune's Bellows, the 500 m strait that allows access to its interior, turns out to be a volcanic caldera [8]. In fact, the island has a horseshoe shape, produced by the flooding of the caldera of a still active volcano, whose last eruption dates back to 1970 (Shultz, 1972). This bay is about 6 x 10 km in size (similar to Krakatoa or Santorini) and is called Port Foster. It is about 190 m deep (Luzón et al., 2011) and is embraced by mostly gravelly and sandy shores (except near Mount Pond), which are affected by several small lava deltas, as well as a shell of slag and thin lavas close to the edge of the caldera (Smellie, 2001). These beaches then give way to steep hills in the central part of the island (Smellie, 2001). In contrast to this, the outer coast is characterised by ice cliffs that are 30-70 m high, but can reach 300 m in some places (Smellie, 2001).

The island has a total area of 98.5 km² [11], including the islands under its jurisdiction, of which 57% is covered by permanent glaciers, mainly affecting Mount Kirkwood and Mount Pond and reaching as far as the north-eastern part of Kendall Terrace, where the surface layer consists of ash (Smellie, 2001). The remaining 43% of the island consists of pyroclastic deposits and sediments, with the exception of a few areas where the rock is exposed (Luzón et al., 2011), as well as glacial moraines, mostly evident at Baily Head and along the outer coast between Vapour Col and Entrance Point (Smellie, 2001). In addition, there are many craters, often filled with lakes.

The diameter of the island is 15 km and its highest peak is 539 m above sea level (Smellie et al., 2002), reached by Mount Pond, but its base diameter is in reality of 30 km and the island rises 1400 m above the seabed (Smellie, 1990).

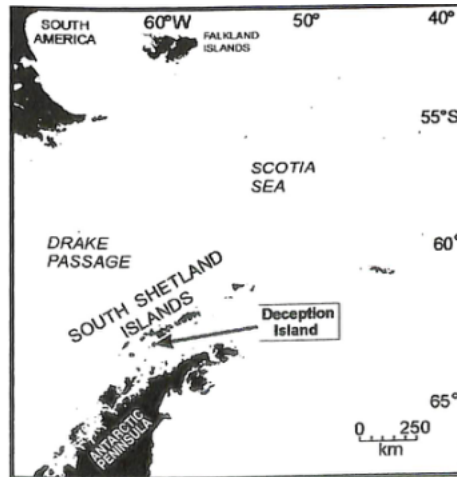


Figure 2.1.1 Location of Deception Island (source: Smellie et al., 2002)

The island is the most active volcano in the South Shetland Islands, and over the years has erupted a large quantity of tephra, estimated to have reached as far as the South Pole (Palais et al., 1989); Moreton (Moreton, 1999) analyzed ash samples in the sediments of the Scotch and Weddell Seas and found that Deception Island was the unequivocal source; however, more recent eruptions have been smaller in volume and have occurred near the inner coast (Figure 2.1.2). Nowadays, evidence of its activity is provided by fumarole emissions (Fumarole Bay owes its name to this very reason), hydrothermal springs, especially active in Pendulum Cove and Whalers Bay, seismicity, monitored since 1986 by several teams of researchers from the Spanish Antarctic Program, as well as uplift of the north floor of Port Foster (Luzón et al., 2011).

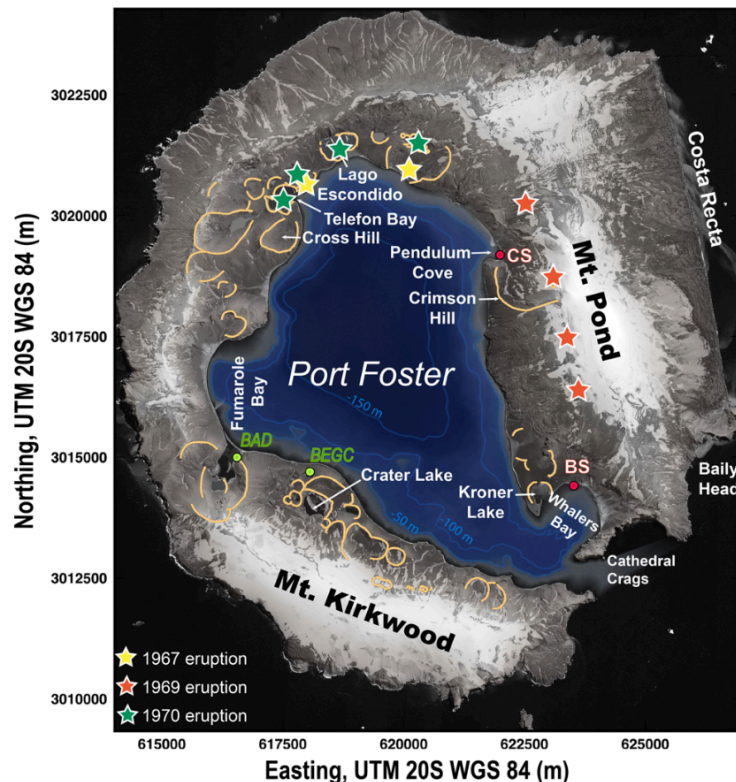


Figure 2.1.2: Most recent eruptions (source: Geyer et al., 2018). BAD: Argentinian Base “Decepción”; BEGC: Spanish Base “Gabriel de Castilla”; BS: British Base (destroyed); CS: Chilean Base (destroyed)

The topographic map which was used as a reference for the study is the 2005 map, drawn up by the Spanish Army Geographical Survey; for more information about it, please refer to Chapter 3, a simplified one is shown in Figure 2.1.3 for greater clarity, whose source is [11]. The island hosts two scientific stations and is one of the main touristic destinations in Antarctica.



Figure 2.1.3: Topographic map (source: [11])

2.2 GEOLOGY AND GEOMORPHOLOGY

Deception Island is a broad Quaternary-era shield volcano, composed of basaltic andesite and located along the Bransfield Strait, a 15-20 km wide back-arc basin located between the Antarctic Peninsula and the South Shetland Islands (Figure 2.2.1) (Smellie, 2001). The cause of the formation of the Bransfield Trench was the subduction of the Phoenix Plate under the Antarctic Plate during the Mesozoic-Cenozoic period (Dalziel, 1984); this movement is still ongoing, although at a very low rate (between 2.5 and 7.5 mm/y over the last 2 million years, according to Henriot (Henriot et al., 1992)). Although the relative ages of the island's individual volcanic units are known, there is a lack of information on their absolute ages. However, an analysis of paleomagnetic data shows that the rocks that have emerged have normal magnetic polarity, which means they are less than 780000 years old, when the last magnetic pole reversal occurred (Valencio et al., 1979, Baraldo et al., 2003).

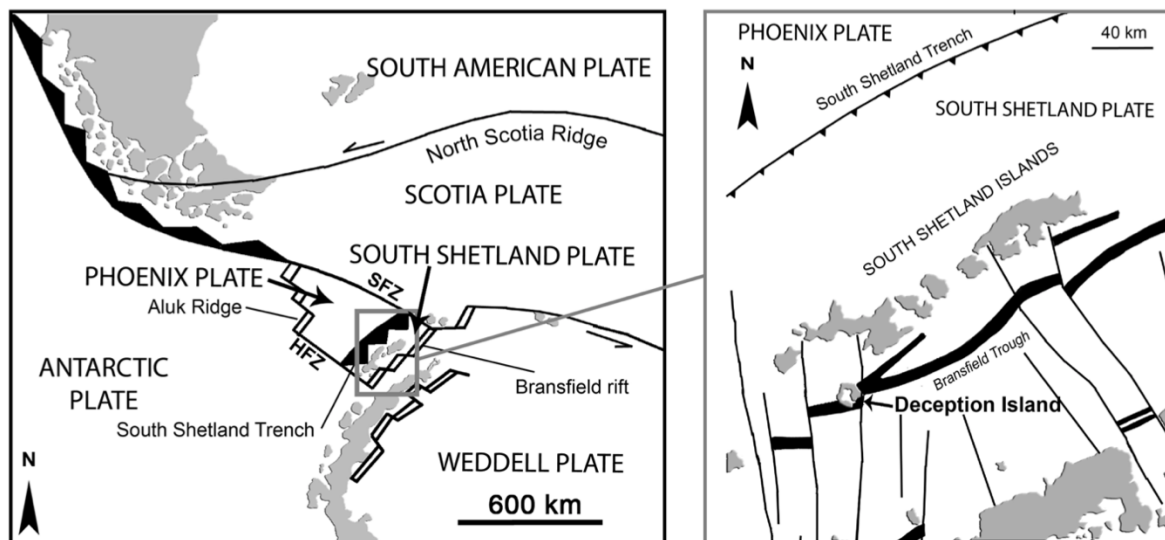


Figure 2.2.1: Bransfield Strait (source: Geyer et al., 2008)

Over the years, the island has experienced many eruptions, including one that is estimated to have erupted 60 km³ of magma (Geyer et al., 2008) and led to the collapse of the caldera (in 8300 BC, as assessed by Olivia Urcia et al. (Oliva-Urcia et al., 2016)); for this reason, its eruptions are divided into pre-caldera and post-caldera.

The pre-caldera period began with the emission of tephra from multiple centers, possibly due to subaqueous fire fountains during shoaling; after which the activity became subaerial effusive or weakly pyroclastic (Strombolian and Hawaiian), leading to the formation of the basaltic shield and subsequently resulting in the great eruption that formed the caldera, then followed by smaller successive collapses (Smellie, 2001).

The post-caldera period comprises the most recent eruptions and is characterized by small-volume mainly phreatomagmatic eruptions, with varying degrees of explosiveness depending

on the amount of water interacting with the magma and its source (aquifer, sea, ice melting) (Geyer et al., 2008), located within the caldera but also along its structural boundaries. The magmas that the island erupts are sodic, while the erupted rocks have a wide range of composition, ranging from basalt to dacite, but they are mainly basaltic andesites (Smellie, 2001).

Speaking of the geology of the island, according to a study conducted by Luzón et al. (Luzón et al., 2011), the volcano is composed of layers of pyroclastic deposits and loose sediments that extend to a depth of 400 m, with different degrees of compaction. In particular, two layers with thicknesses of approximately 100 and 300 m can be differentiated on the basis of the different propagation velocities of s-waves within them, of which the second one is the most consolidated. These are deposits resulting from pre- and post-caldera eruptions that have then been subjected to erosion, sedimentation and snow cycles.

On this basis, the deposits on the island are divided into two main groups: The Port Foster Group, which comprises pre-caldera deposits, and the Mount Pond Group, which comprises post-caldera deposits (Smellie et al., 2002). The Port Foster Group includes the Fumarole Bay, Basaltic Shield and Outer Coast Tuff formations, which are only visible in the outer coast cliffs and in some parts of the inner wall of the caldera, because they are covered by post-caldera materials, and consist mainly of mafic rocks formed by submarine Strombolian deposits, submarine lavas and deposits from currents of pyroclastic density. More specifically, Smellie (Smellie et al., 2002) states that Fumarole Bay is the result of a predominantly subaqueous eruptive activity, generated by several centers, the Basaltic Shield is composed of subaerial effusive and minor pyroclastic material, and the Outer Coast Tuff is the result of a major eruption immediately preceding, and probably related to, the caldera collapse. Other sources, such as Martí et al. (Martí et al., 2013), state that the caldera collapse occurred prior to the deposition of the Outer Coast Tuff, which they classify as a syn-caldera collapse deposit. These units would therefore testify the progressive evolution from subaqueous to subaerial volcanism (Oliva-Urcia et al., 2016).

The Mount Pond Group comprises the Baily Head, Pendulum Cove and Stonethrow Ridge formations and consists of small volume phreatomagmatic eruption deposits, such as hydrovolcanic tephra, lavas, tuff cones and maar deposits, and deposits from magmatic eruptions, such as strombolian scoriae and lavas, of basaltic to dacitic composition, erupted from different centers (Smellie et al., 2002).

Figure 2.2.2 shows a simplified geological map of the island, drawn up by Smellie (Smellie, 2001), for a more complete map please refer to the appendix, where the geological and geomorphological maps drawn up by López-Martínez et al. in 2000 are shown.

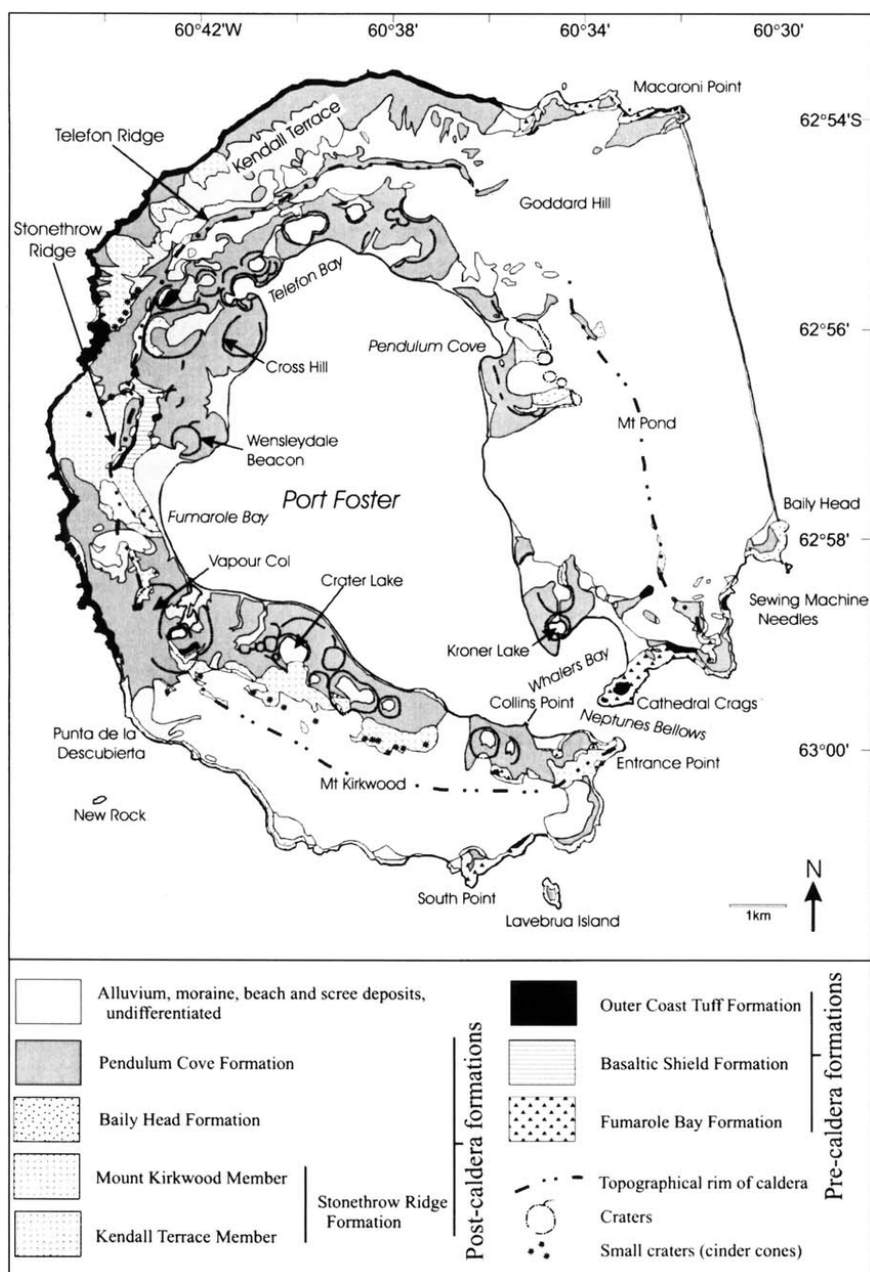


Figure 2.2.2: Geologic map (source: Smellie, 2001)

2.3 METEOROLOGY

The climate of Deception Island is polar maritime, and annual temperatures move in the range -28 °C +11 °C, with an average of -3 °C, although locally 70 °C can be reached in proximity of fumaroles and hot springs [10].

For the study of the meteorology, the reference used was the Antarctic Campaign of the Spanish Land Army (Campaña Antártica del Ejército de Tierra, 2019-2020). In particular, the period 12.01.2020-09.02.2020 was considered, which includes all the dates of the downloaded images, from the first one of Sentinel 2, on 19.01.2020, to the one of Landsat 8 used as ground truth, on 09.02.2020, and also covers the week before the first image, for a more complete analysis.

Examining the data provided, it emerges that the temperature was always above 1 °C, occurring seven times, while on the majority of days (sixteen) the minimum temperature was 2 °C, reaching 3 °C (six days) on the hottest days.

Speaking of the maximum temperatures, in most cases they reached 4 °C (sixteen days), peaking at 5 °C (eight days) on the hottest days and falling to 2 °C on the two coldest days.

There was no snowfall throughout the period, and the days were always cloudy, often accompanied by rainfalls and more frequently by sunshine.

DATE	Weather conditions	T min (°C)	T max (°C)
12/01/2020	Cloudy with light rain during a time slot	3	5
13/01/2020	Cloudy	3	5
14/01/2020	Cloudy	3	5
15/01/2020	Cloudy and partly sunny	2	4
16/01/2020	Cloudy and partly sunny	1	4
17/01/2020	Cloudy	1	3
18/01/2020	Mostly sunny	1	3
19/01/2020	Partly cloudy and partly sunny	1	3
20/01/2020	Mostly sunny	1	4
21/01/2020	Cloudy and partly sunny	2	4
22/01/2020	Cloudy and partly rainy	2	4
23/01/2020	Cloudy, with light rain during two time slots	3	5
24/01/2020	Cloudy and rainy	3	5
25/01/2020	Cloudy and partly rainy	2	4
26/01/2020	Cloudy and rainy	2	4
27/01/2020	Cloudy but mostly rainy	2	5
28/01/2020	Cloudy but mostly rainy	2	5
29/01/2020	Cloudy and rainy	2	4
30/01/2020	Cloudy with light rain during a time slot	2	4
31/01/2020	Cloudy with light rain during a time slot	2	4
01/02/2020	Mostly rainy	2	4

02/02/2020	Cloudy and rainy	2	4
03/02/2020	Partly cloudy and partly sunny, with light rain during a time slot	2	4
04/02/2020	Mostly sunny	1	2
05/02/2020	Mostly sunny	1	2
06/02/2020	Cloudy and partly sunny	3	5
07/02/2020	Partly sunny	2	4
08/02/2020	Partly sunny	2	4
09/02/2020	Mostly sunny	2	4

Table 2.3.1: metereological data

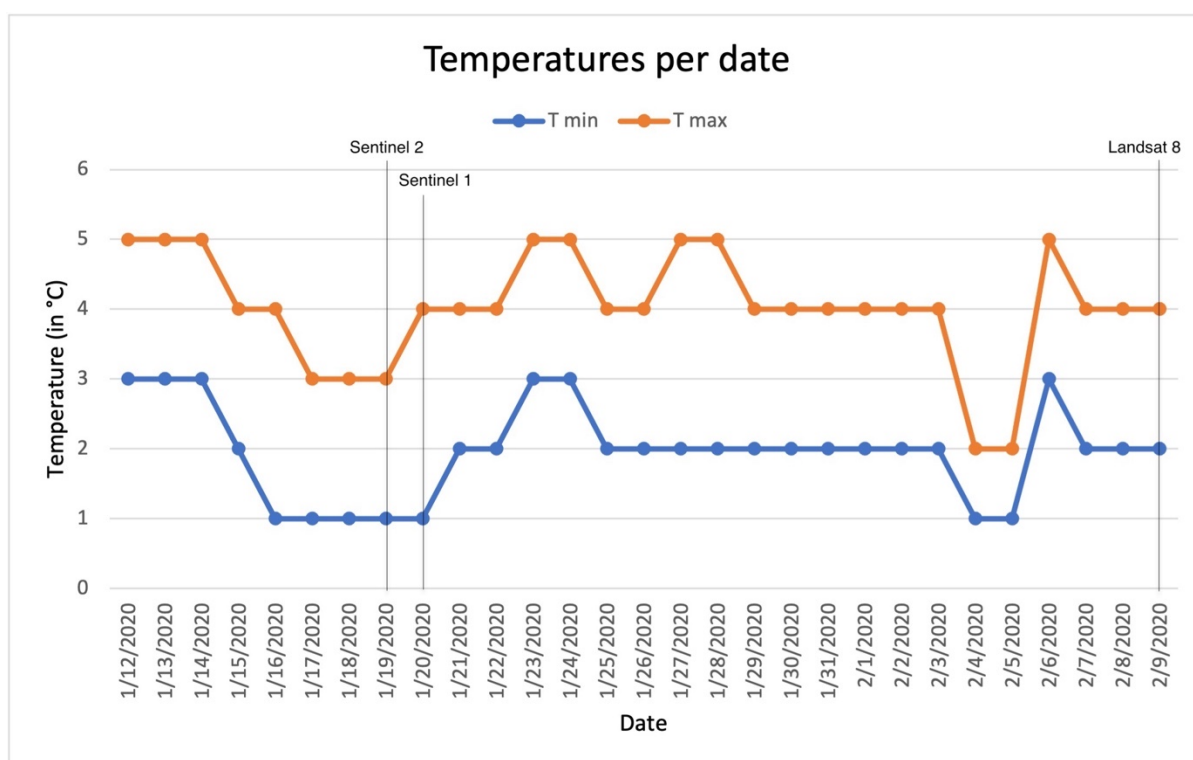


Figure 2.3.1: temperatures graph with lines in correspondence with the dates in which the images used were acquired

To consult the data divided by time slot, reference can be made to the tables provided by the Diary of Operations of the Spanish Land Army's Antarctic Campaign (Campaña Antártica del Ejército de Tierra, 2019-2020), shown in the Appendix.

3. MATERIAL

The images analyzed for the study belong to the period of the Antarctic summer and were acquired by the Sentinel 1 and Sentinel 2 satellites of the European Space Agency (ESA), on the dates of the 19th (optical image) and 20th (SAR image) of January 2020, respectively. The sources used as ground truth were an image acquired by Landsat 8, a joint NASA/ United States Geological Survey (USGS) satellite, for snow/ ice and the 2005 topographic map developed by the Geographic Centre of the Spanish Army for lakes and bare soil/ rock. A brief description of the 3 mentioned satellites and the topographic map is given below.

The Sentinel images were downloaded from ESA's Copernicus Open Access Hub (formerly known as Sentinels Scientific Data Hub) [9] and the Landsat images from the USGS Earth Explorer [23].

A brief description of the three satellites is given below and the used images are also specified with more detail.

Sentinel 1

Sentinel 1 refers to a mission comprising two satellites, Sentinel 1A and Sentinel 1B, both launched by the European Space Agency on the 3rd of April 2014 and the 25th of April 2016, respectively, for the Copernicus initiative, taken jointly with the European Commission (EC). These are two C-band synthetic aperture radars (5.6 cm wavelength) capable of providing a ground track repetition of 12 days for the single satellite and 6 for the two. They can provide single or dual polarization and have four possible acquisition modes, shown in Figure 3.1: Stripmap (SM), in continuity with the ERS and Envisat missions, Interferometric Wide swath (IW), which is the default mode for ground-based acquisitions, Extra-Wide swath (EW), used instead for maritime, glacial and polar areas, and Wave mode (WV), which aims to determine the direction, wavelength and height of ocean waves.

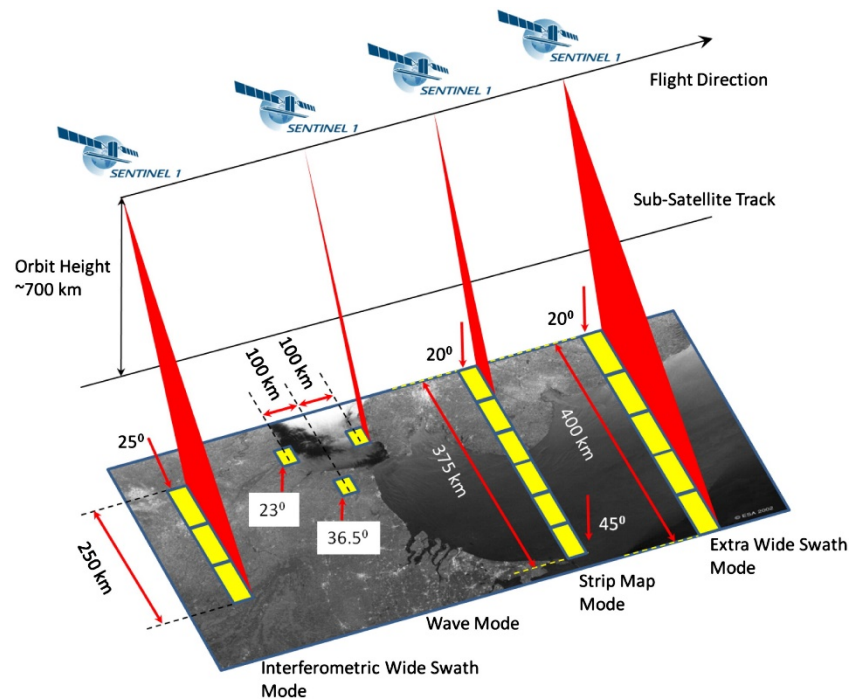


Figure 3.1: Sentinel 1 acquisition modes. Source: [12]

In particular, the image used was acquired in IW mode, which consists of a data acquisition with a swath of 250 km operated by the TOPSAR (Terrain Observation with Progressive Scans SAR), an instrument that works by moving the acquisition beam cyclically between several adjacent sub-swaths. This allows to provide large swaths with a reduced scalloping effect and the result is one image for each sub-swath and one for each polarisation channel, for a total of three images in single polarisation or six in dual polarisation, which can then be merged into a single image covering the entire area. To guarantee this possibility, each sub-swath has a small overlap area of about 2km.

The images acquired are open source and the processing levels provided are the followings:

- Level 0, this is the raw and compressed SAR product, in order to be used it needs to be decompressed and processed using a SAR processor;
- Level 1, which in turn is divided into two types of products: Single Look Complex (SLC), which consist of complex samples that preserve phase information and are characterised by the presence of a single look in each dimension that exploits the full bandwidth of the transmission signal, and Ground Range Detected (GRD), consisting of SAR data detected, multi-looked and projected to ground level using a ground ellipsoid model, but which lose phase information. GRD products can have three different resolutions, depending on the number of multi-looks performed: Full Resolution (FR), High Resolution (HR) and Medium Resolution (MR).
- Level 2, which contains three components: the Ocean Swell spectra (OSW), which is a two-dimensional spectrum of waves on the ocean surface, and includes an estimate of speed and wind, the Ocean Wind Fields (OWI), which is a grid with estimates of wind speed and direction at 10 m above the ground, derived from GRD images, and the Surface Radial Velocities (RVL),

which indicates the difference between the Doppler grid measured in level 2 and the geometric Doppler calculated in level 1.

For more information on this subject, please refer to [13], the official source from which the above information was taken.

The image used for the study, shown in Figure 3.2, is the “S1A_IW_GRDH_1SSH_20200120T081742_20200120T081807_030881_038B41_9A9B”, was acquired by the Sentinel 1A (S1A) satellite on 20-01-2020 from 08:17:42 to 08:18:07, with an Interferometric Wide Swath Mode (IW), in a downward direction along orbit 9. The product is level 1 and Ground Range, Multi-Look, Detected (GRD), high resolution (H) and has single HH polarisation. This information was retrieved from the metadata of the product and from the interpretation of its nomenclature with the (ESA, 2016).

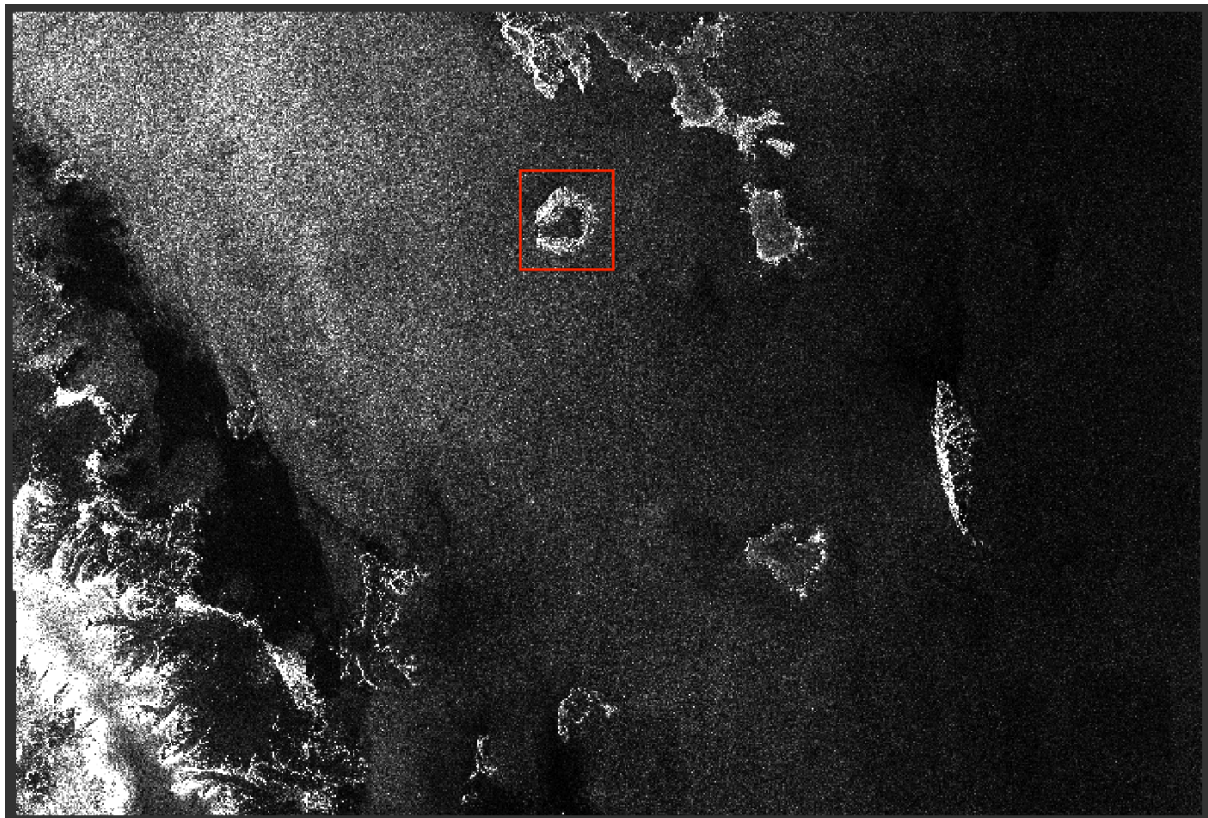


Figure 3.2: Original Sentinel 1 image, the study area is circumscribed by the red rectangle

Sentinel 2

Sentinel 2 is the mission comprising the Sentinel 2A and Sentinel 2B satellites, launched by the European Space Agency on the 23rd of June 2015 and on the 7th of March 2017, respectively, for the Copernicus initiative. The two satellites travel on the same polar and sun-synchronous orbit, with a relative phase difference of 180° the one from the other. The revisit time that they are able to provide is 10 days at the equator for the single satellite and 5 days for the two. Each of the two satellites is equipped with a Multi-Spectral Instrument for the

acquisition of the images which is able to acquire 13 spectral bands, characterized by three possible spatial resolutions: 10 m (bands B2, B3, B4 and B8), 20 m (bands B5, B6, B7, B8A, B11 and B12) and 60 m (bands B1, B9 and B10), and by the wavelengths reported in Figures 3.3, 3.4 and 3.5 (source: [15]).

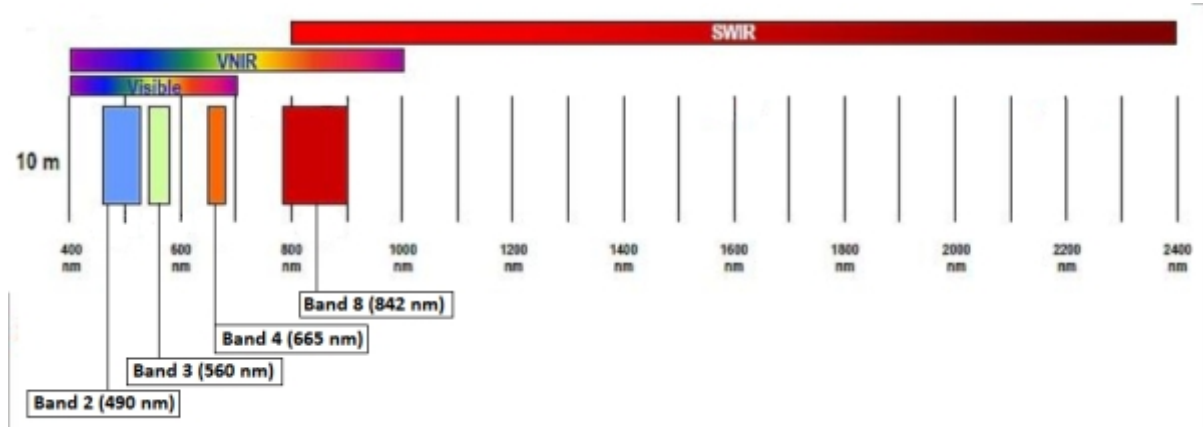


Figure 3.3: Sentinel 2 10 m spatial resolution bands: B2 (490 nm), B3 (560 nm), B4 (665 nm) and B8 (842 nm). Source: [15]

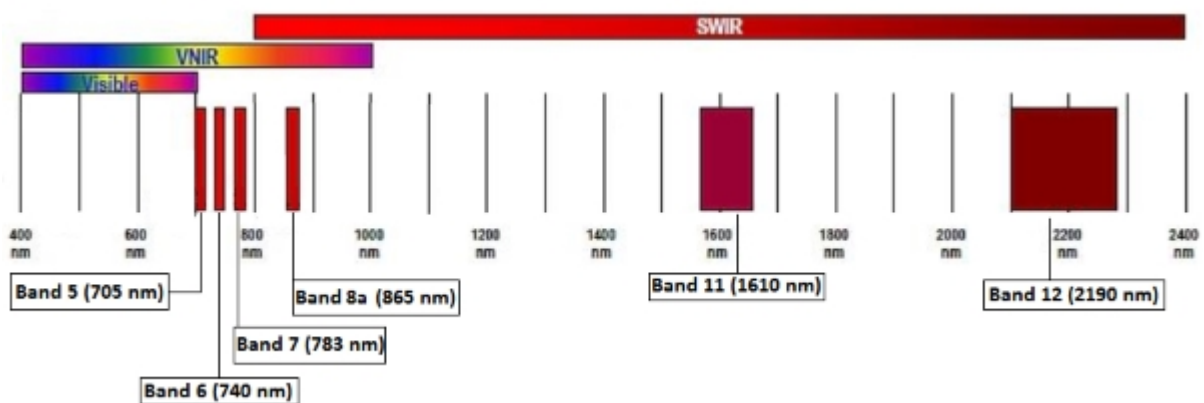


Figure 3.4: Sentinel 2 20 m spatial resolution bands: B5 (705 nm), B6 (740 nm), B7 (783 nm), B8a (865 nm), B11 (1610 nm) and B12 (2190 nm). Source: [15]

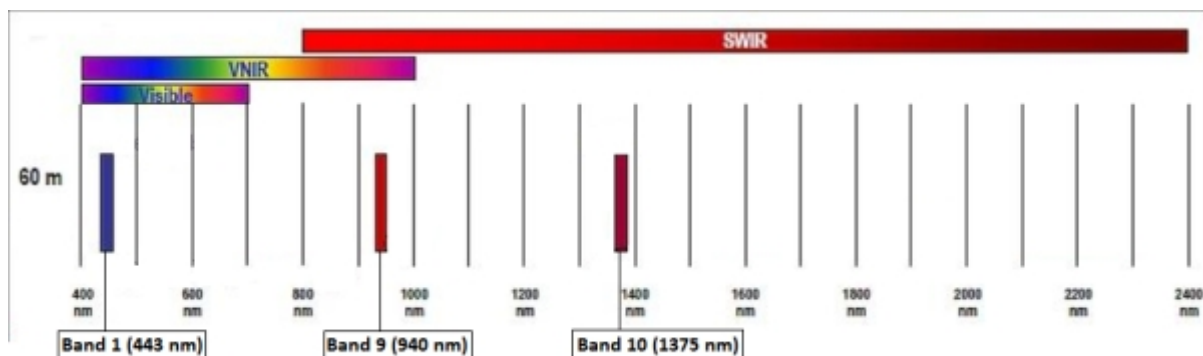


Figure 3.5: Sentinel 2 60 m spatial resolution bands: B1 (443 nm), B9 (940 nm) and B10 (1375 nm). Source: [15]

For the purposes of this study, not all bands were useful and in particular bands 1, 9 and 10 were excluded, because the first two bands both contain measurements that are not in our interest: the first is related to aerosol monitoring (ESA, 2015) and the second indicates an absorption value related to the water vapour content in the atmosphere [14]. As for band 10, it is the one related to cirrus and is already excluded from level 2A products as it does not contain any information related to surfaces.

Regarding the processing levels of the provided products, there are two levels which are open to the public: level 1C, which is characterised by Top-of-Atmosphere reflectance values, and level 2A, which provides Bottom-of-Atmosphere reflectance values.

The official source for the above information is [16].

The image used for the study, shown in Figure 3.6, is the “S2B_MSIL2A_20200119T131859_N0213_R095_T20EPR_20200119T145349” and was acquired by the Sentinel 2B (S2B) satellite with the Multi-Spectral Instrument (MSI) on 19-01-2020 at 13:18:59 in a downward direction along orbit 95 in tile number 20. The product is level 2A and was generated on the same date at 14:53:49, and has a cloud coverage of 38.88%. This information was retrieved from the metadata of the product and from the interpretation of its nomenclature with the Products Specification Document of Sentinel 2 (ESA, 2018).

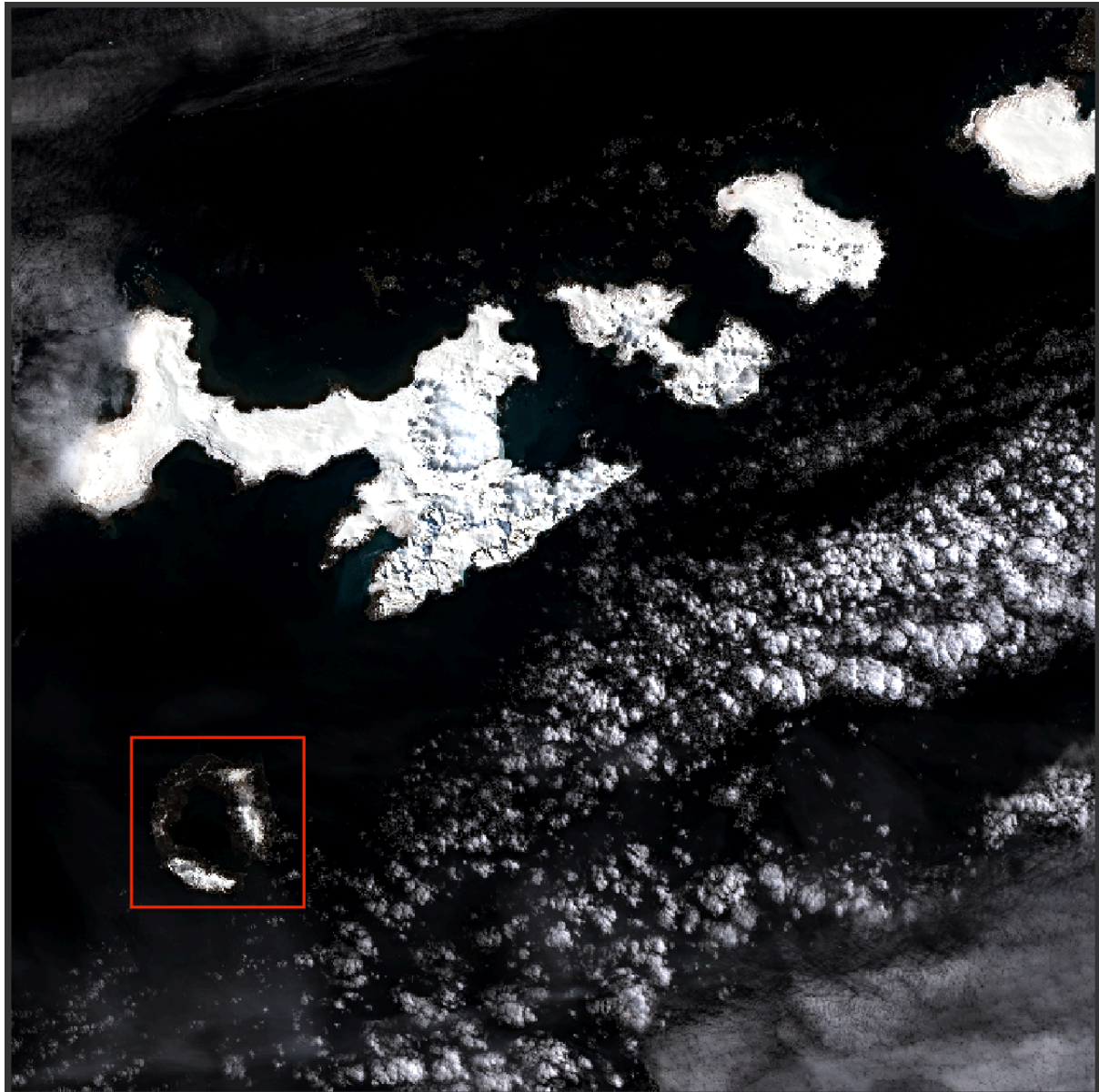


Figure 3.6: Sentinel 2 image (RGB combination: B4-B3-B2), the study area is framed with a red rectangle

Landsat 8

Landsat 8 was launched on the 11th of February 2013 from NASA's Vandenberg Air Force Base in California and moves in a near-polar orbit, so its temporal resolution is of 16 days. It has two instruments on board: the Operational Land Imager (OLI), which measures the visible, near-infrared and short-wave infrared in nine different bands, including one panchromatic band, and the Thermal Infrared Sensor (TIRS), which measures the temperature of the Earth's surface in two thermal bands. Speaking of the spatial resolution, shown in Figure 3.8, it is equal to 30 for all of the bands except for the panchromatic band (band 8), where it is equal to 15 m, and for the two thermal infrared bands, acquired by the TIRS, where it is equal to 100. Figure 3.7 shows the spectral resolutions of the different bands of Landsat 8 compared with those of Landsat 7.

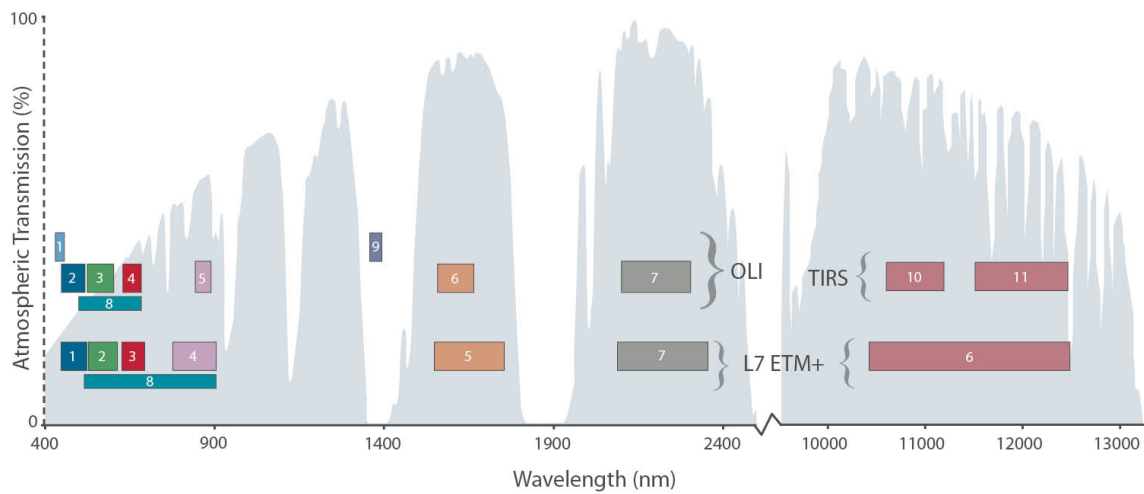


Figure 3.7: Comparison between Landsat 8 and Landsat 7 spectral resolutions. Source: [20]

Bands	Wavelength (micrometers)	Resolution (meters)
Band 1 - Coastal aerosol	0.43-0.45	30
Band 2 - Blue	0.45-0.51	30
Band 3 - Green	0.53-0.59	30
Band 4 - Red	0.64-0.67	30
Band 5 - Near Infrared (NIR)	0.85-0.88	30
Band 6 - SWIR 1	1.57-1.65	30
Band 7 - SWIR 2	2.11-2.29	30
Band 8 - Panchromatic	0.50-0.68	15
Band 9 - Cirrus	1.36-1.38	30
Band 10 - Thermal Infrared (TIRS) 1	10.6-11.19	100
Band 11 - Thermal Infrared (TIRS) 2	11.50-12.51	100

Figure 3.8: Landsat 8 bands spatial and spectral resolution. Source: [22]

Again, there are two processing levels available for public use: one subjected to atmospheric correction (Level 2) and one not (Level 1).

The data acquired by Landsat are then divided into two collections: Collection 1, which contains all the Level 1 data acquired from 1972 to the present from Landsat 1-8, and Collection 2, which is a reprocessing of the Level 1 images that aims to bring improvements such as greater accuracy in geolocation, and includes the Level 2 products, which provide the reflectance and surface temperature of the scenes acquired from 1982 to the present.

The sources from which the above information was drawn were [21] and [24].

For the purposes of the study, the ground truth of the snow was taken from the Pixel Quality Assessment band of the image “LC08_L2SR_219104_20200209_20201016_02_T2”, acquired by the Landsat 8 satellite on 09-02-2020 from 13:15:19 to 13:15:51 along orbit 219 referred to the Worldwide Reference System-2 (WRS-2) Path of the product and orbit 104 referred to the WRS-2 row of the product. The image processed at level 2 on 16-10-2020 was chosen and the scene is 20.73% covered by clouds but only 4.65% of it is on dry land. These data were retrieved from the product metadata and interpreted using the Landsat 8 Collection 2 (C2) Level 2 Science Product (L2SP) Guide (USGS, 2020). The image is shown in Figure 3.9, while, in order to see its Pixel QA band, please refer to chapter 4.3 (Figure 4.3.5).

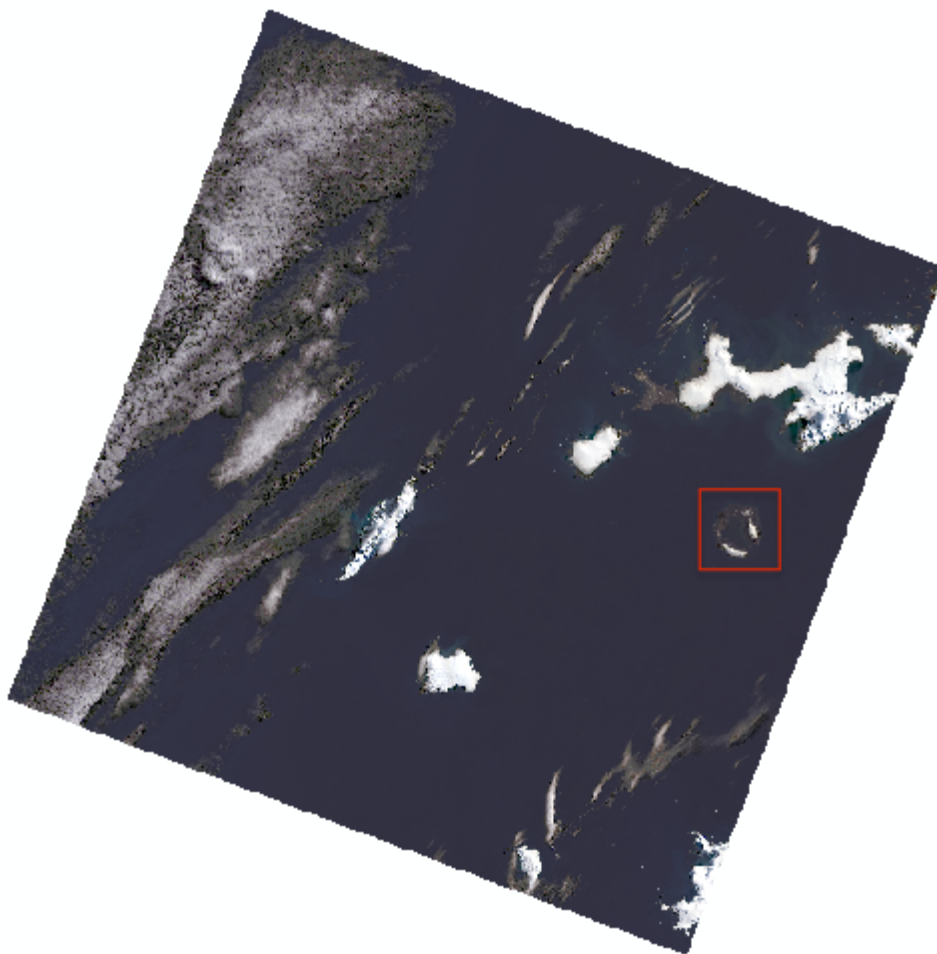


Figure 3.9: Landsat 8 image, the study area is framed with the red rectangle

Topographic map

The topographic map used as reference was produced by the Geographic Centre of the Spanish Army in collaboration with the University of Cádiz, and the edition used is the second, made in 2005, shown in Figure 3.10. The map is at a scale of 1:25000 and is based on field surveys carried out by the Geographic Centre of the Army in the Antarctic summer of 1992-1993, thus updating the aerial survey carried out in 1986, and on the coastline of an image

acquired by the QuickBird satellite in 2003. For more information on the other sources that underlie the map (for example bathymetric data, location of fauna, etc.), please refer to the map, which is included in the Appendix.

This map was then redrawn with Microstation software by a research group from the University of Cádiz and it is from these dgn files that the lakes were extracted when preparing the ground truth, as will be further explained in chapter 4.3.

Furthermore, the contour lines indicated in the map were used by the same university to form the DTM of the island, which will be used for the geometric correction of the SAR image, as explained in chapter 4.1.1. Finally, the contour line 0 is the one they used to create the shapefile of the coastline, which was later used to clip the images.

Softwares used

For an initial processing, such as coarsely cropping the area of interest, or, in the case of the SAR image, calibrating, georeferencing, applying speckle reduction filters and converting the bandwidth to decibels, ESA's SNAP software, version 7.0 [17], was used. Then the images were exported in GeoTIFF format

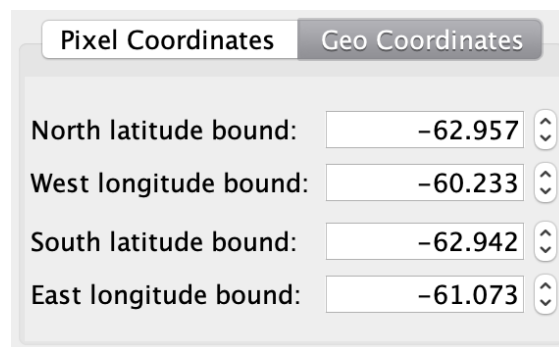
The work continued with the software ArcGIS, version 10.7 [2], and ArcGIS Pro, version 2.4 [6], where a more precise clip of the contour of the island was applied, thanks to a shapefile of the coastline, and it was also possible to carry out the different operations between the bands to generate the necessary indices, as well as the classification and segmentation operations and all of the steps that they involve. In the specific case of the classification of the optical image based on the application of subsequent indices, moreover, the limits capable of dividing their histograms in such a way as to separate the different land covers were identified using the ENVI software, version 5.1 [18].



4. METHODS

The aim of the study is to compare the classification of a SAR image with that of an optical image taken in close dates (so the results should be similar); to do this, therefore, the work was divided into three main phases: preparation of the images, classification, quantification of the goodness of the classified images (through comparison with a ground truth by means of confusion matrices).

As far as the preparation phase is concerned, it began for both of the images with the software SNAP, where they were subjected to a first crop, by selecting a rectangle (with the coordinates shown in Figure 4.1) that included the area of the island: the original images in fact included a much larger area and cropping was the first necessary operation, in order not to overload the next ones.



Pixel Coordinates	Geo Coordinates
North latitude bound:	-62.957
West longitude bound:	-60.233
South latitude bound:	-62.942
East longitude bound:	-61.073

Figure 4.1: coordinates used to crop the area

At this point, the optical image was ready for the export of each band in GeoTIFF format, while the SAR image had to firstly undergo the operations of calibration, speckle filtering and geometric correction and georeferencing, which will be described in more detail in sub-chapter 4.1.1.

Once the ArcGIS software was opened, the first thing to be done was to clip the various bands of the optical image and the SAR image with the shapefile of the island's coastline (Figure 4.2), obtained from the 2005 topographic map of the island, as explained in Chapter 3.

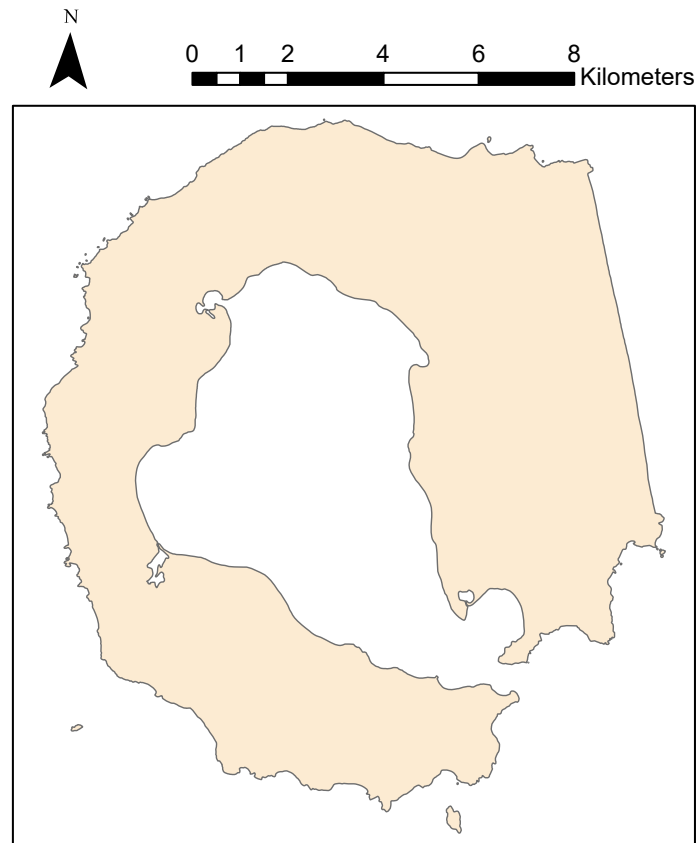


Figure 4.2: shapefile of the coastline

The bands of the optical image which had a resolution of less than 10x10 were then resampled to this resolution and put together with the others into a stack, as will be explained in more detail in sub-chapter 4.2.1.

At this point, both the optical and SAR images were ready to be subjected to classification, with the aim of identifying the different ground covers of the island.

The classifications made were both unsupervised and supervised for both, and in addition to these, a hierarchical classification based on the selection of limits in the histograms of different indices was also carried out for the optical image. A classification based on limits to be selected in the histograms was not possible for the SAR image, even after the repeated application of the Lee, Lee Sigma and Refined Lee filters, which will be discussed later, as there were no easily divisible distributions. Figure 4.3 shows the histograms of the SAR image with the different filter applications to demonstrate what has just been said.

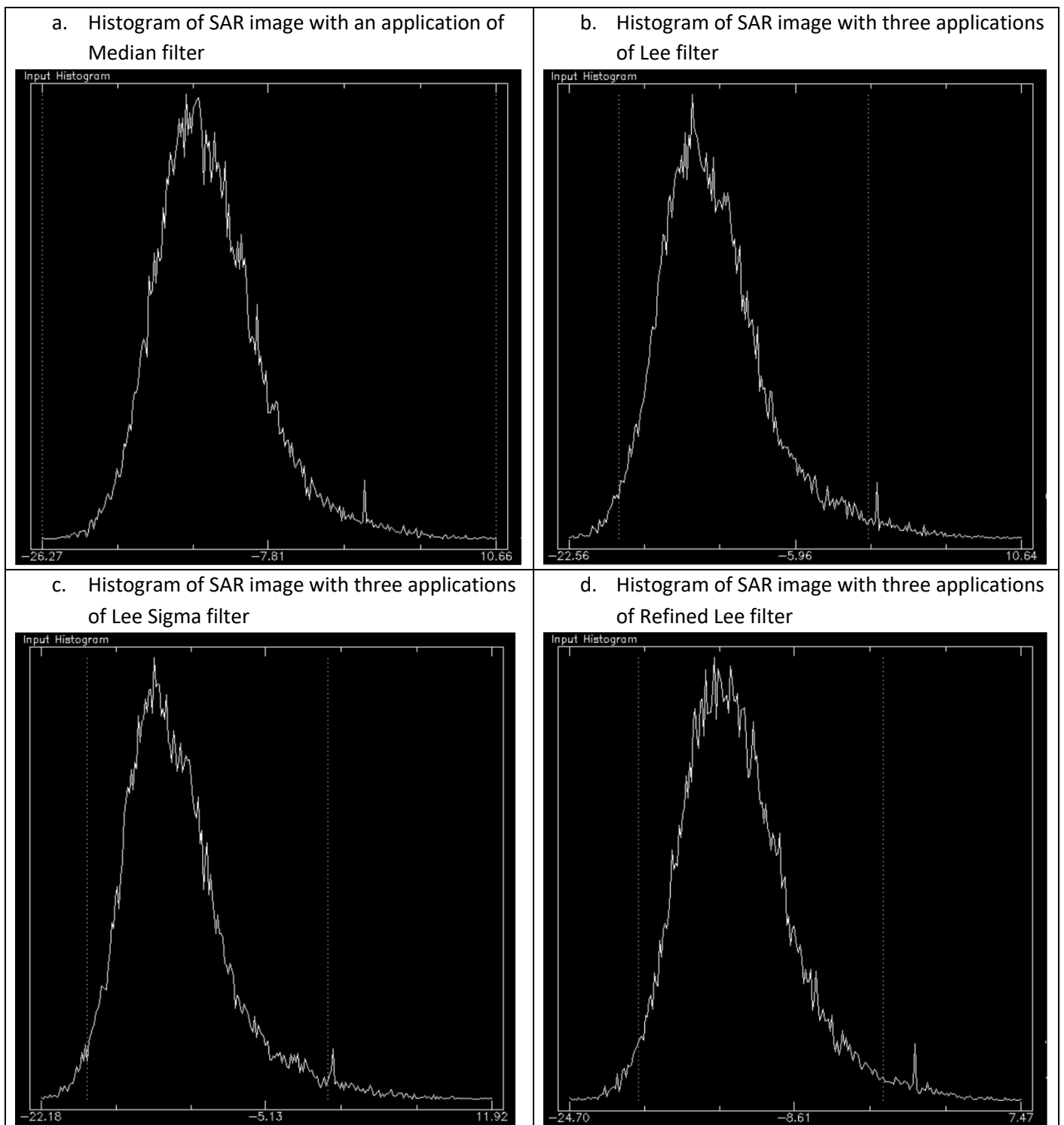
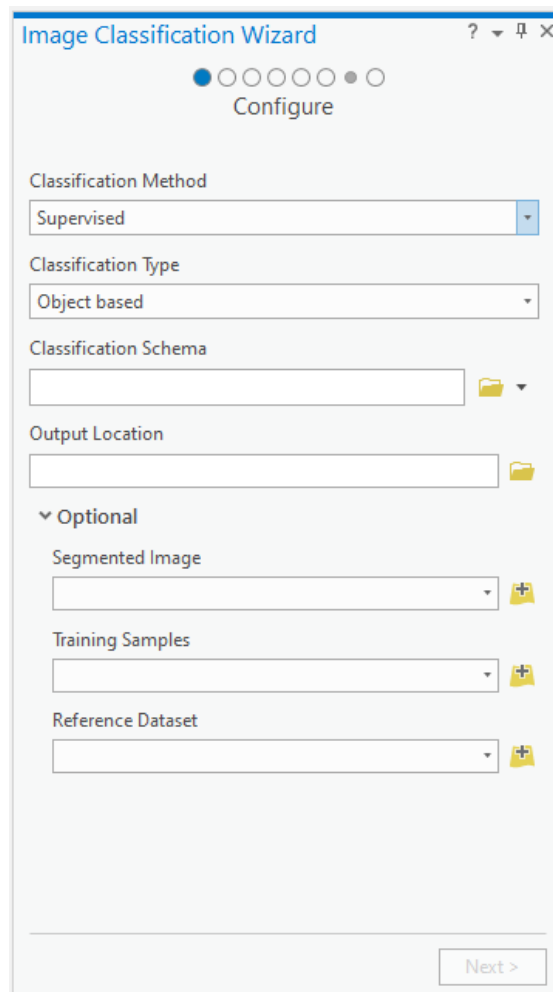


Figure 4.3: Histograms of SAR image with applications of different filters

In order to classify an image in ArcGIS Pro, the tool to use is the Classification Wizard, which proposes 8 guided steps that will be illustrated below; the individual decisions which were made will be specified in sub-chapters 4.1 and 4.2.

Step 1: Configure



The screenshot shows the 'Image Classification Wizard' window, specifically the 'Configure' step. The window has a title bar with a question mark, a maximize button, and a close button. Below the title bar is a progress indicator with seven circles; the first circle is filled blue, and the others are empty. The word 'Configure' is centered below the progress indicator. The main area contains several configuration options:

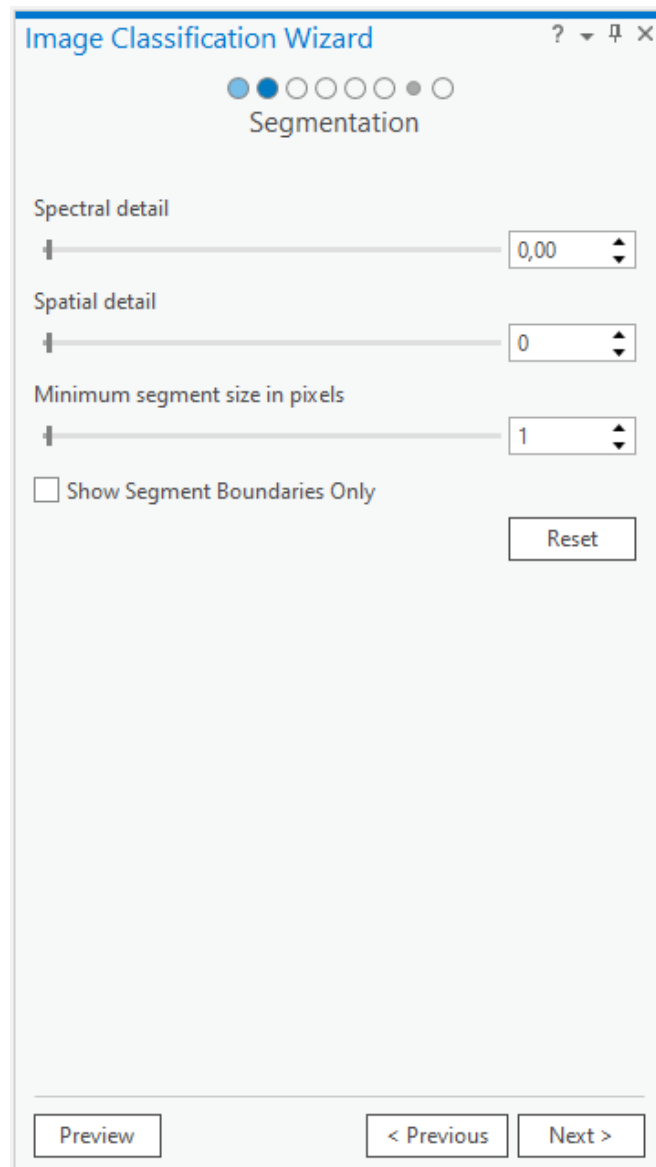
- Classification Method:** A dropdown menu with 'Supervised' selected.
- Classification Type:** A dropdown menu with 'Object based' selected.
- Classification Schema:** A text input field with a folder icon and a dropdown arrow to its right.
- Output Location:** A text input field with a folder icon to its right.
- Optional:** A section header with a downward arrow.
- Segmented Image:** A text input field with a plus icon to its right.
- Training Samples:** A text input field with a plus icon to its right.
- Reference Dataset:** A text input field with a plus icon to its right.

At the bottom right of the window is a 'Next >' button.

In this first step you set which type of classification we want to operate, object based or pixel based and supervised or unsupervised. You specify which schema you want to use for the classification, and the options are to use a schema that you have, to generate it from training samples (in the case of supervised classification) or from a classified image or finally to use the default schema provided by ArcGIS, based on the USA National Land Cover Database of 2011. The latter is the result of a mapping operation of the 48 southern United States, Hawaii, Alaska and Puerto Rico by the Multi-Resolution Land Characteristics (MRLC) Consortium, consisting of federal agencies cooperating to generate land cover information on a national scale (<https://www.mrlc.gov>).

Other optional inputs are a segmented image (if an object-based classification is to be performed), training samples (if a supervised classification is chosen) and a reference dataset, which is then compared to the classified image to check how good the latter is.

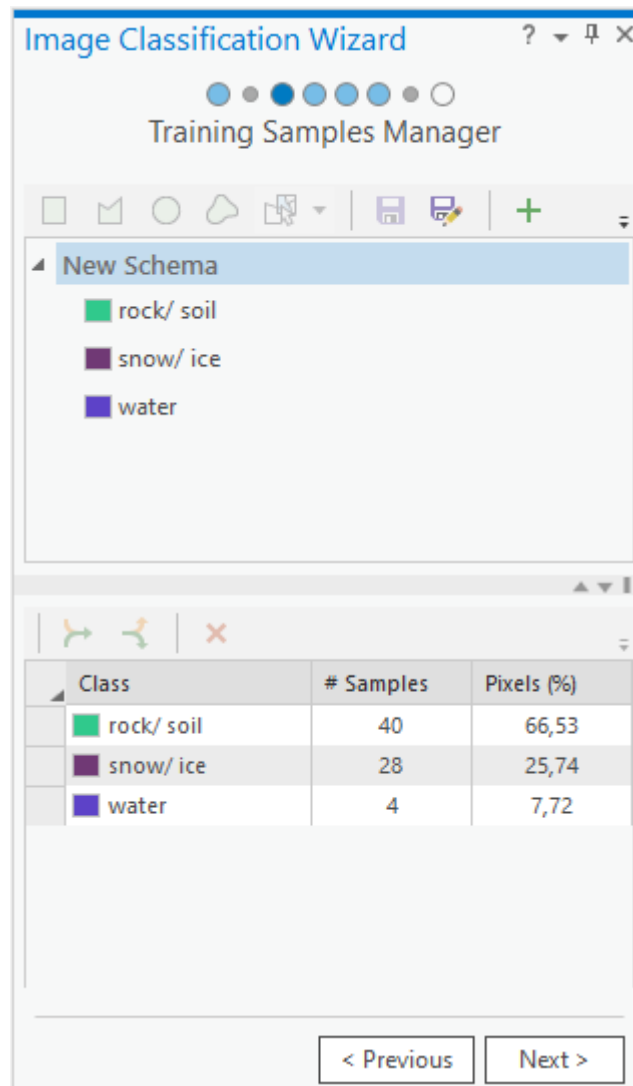
Step 2: Segmentation



The screenshot shows the 'Image Classification Wizard' window, specifically the 'Segmentation' step. The window has a title bar with a question mark, a maximize button, and a close button. Below the title bar, there are seven circular progress indicators; the second one from the left is filled with blue, and the others are empty. The word 'Segmentation' is centered below these indicators. The main area of the window contains three sliders and a checkbox. The first slider is labeled 'Spectral detail' and has a value of '0,00' displayed to its right. The second slider is labeled 'Spatial detail' and has a value of '0' displayed to its right. The third slider is labeled 'Minimum segment size in pixels' and has a value of '1' displayed to its right. Below these sliders is a checkbox labeled 'Show Segment Boundaries Only', which is currently unchecked. To the right of the checkbox is a 'Reset' button. At the bottom of the window, there are three buttons: 'Preview' on the left, '< Previous' in the middle, and 'Next >' on the right.

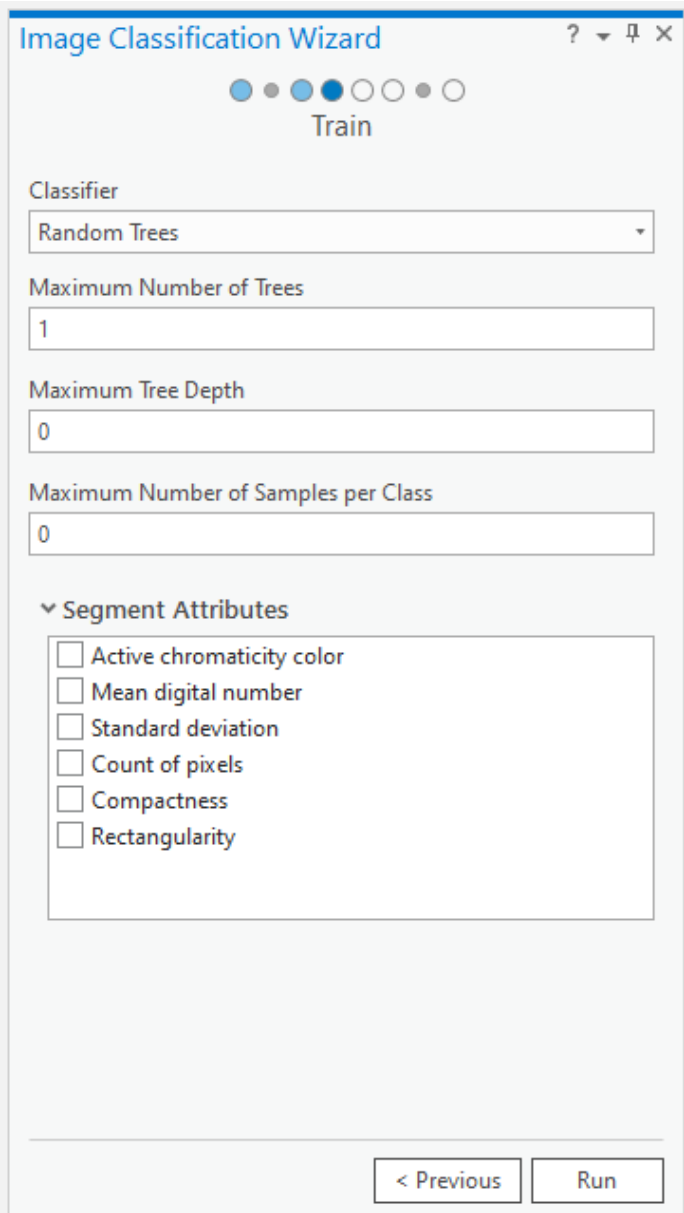
In this step, if an object-based classification was chosen and no segmented image was put as input in Step 1, a segmented image is generated. In this case, this step was not used as the necessary segmented images had previously been generated using the Segment Mean Shift command, as will be explained later in sub-chapter 4.1. The function of the various modifiable parameters, which are the same as those shown in the figure, will therefore be described in detail in the just mentioned sub-chapter.

Step 3: Training Samples Manager



This step is only active if a supervised classification is carried out and it is the one regarding the choice of the training samples to be used, showing the number of samples selected for each class and the corresponding percentage in area with respect to the total. In addition, if you have already previously selected training samples, it is here that you can modify them before starting the classification, for example by selecting new pixels or modifying the classes, incorporating one with another, adding new ones or subdividing them into subclasses.

Step 4: Train



The image shows a software window titled "Image Classification Wizard" with a standard Windows-style title bar (minimize, maximize, close buttons). Below the title bar is a progress indicator consisting of seven circles; the fourth circle from the left is filled with blue, and the word "Train" is centered below it. The main area of the window contains several input fields and a list of checkboxes. The "Classifier" is set to "Random Trees" in a dropdown menu. Below it are three text input fields: "Maximum Number of Trees" with the value "1", "Maximum Tree Depth" with the value "0", and "Maximum Number of Samples per Class" with the value "0". A section titled "Segment Attributes" with a downward arrow icon contains a list of six checkboxes, all of which are unchecked: "Active chromaticity color", "Mean digital number", "Standard deviation", "Count of pixels", "Compactness", and "Rectangularity". At the bottom right of the window are two buttons: "< Previous" and "Run".

Image Classification Wizard

Train

Classifier
Random Trees

Maximum Number of Trees
1

Maximum Tree Depth
0

Maximum Number of Samples per Class
0

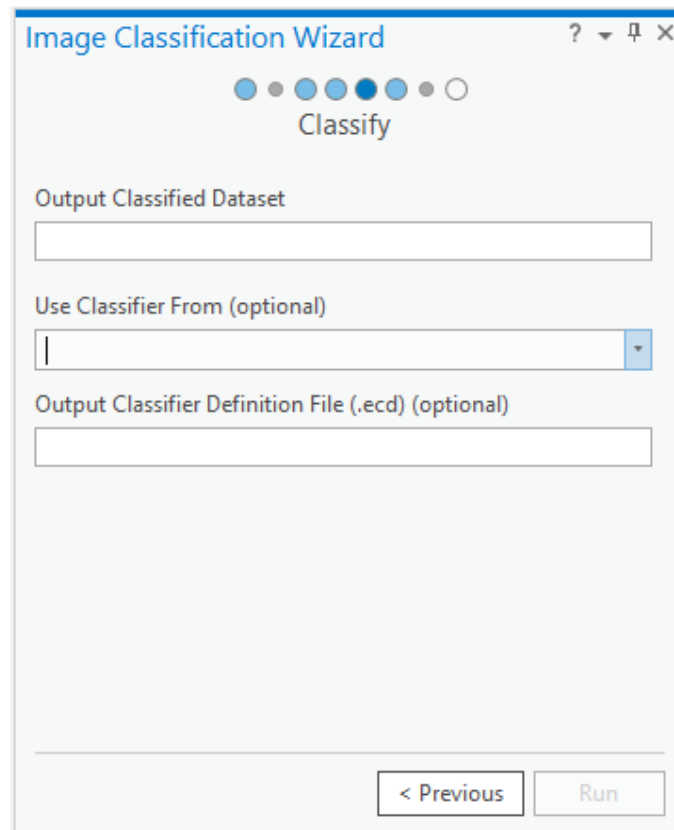
Segment Attributes

- ☐ Active chromaticity color
- ☐ Mean digital number
- ☐ Standard deviation
- ☐ Count of pixels
- ☐ Compactness
- ☐ Rectangularity

< Previous Run

This step involves selecting a specific classifier if the supervised mode has been selected in Step 1 (if the unsupervised mode has been selected the choice will be univocal since there is only one classifier of this type: the ISO cluster), and setting its parameters, as well as the segment attributes to be taken into account at the time of classification if it is object-based. The case of the Random Trees classifier is given as an example, but the different options will be explored in more detail in sub-chapters 4.1 and 4.2. This step generates a preview of how the classified image will be; it is not saved in the geodatabase and it is used just to check whether the result obtained is the desired one, otherwise it is possible to change the parameters to set and the choice of segment attributes. When you are satisfied with your choices, you can proceed to the next step.

Step 5: Classify



The image shows a software window titled "Image Classification Wizard" with standard window controls (help, maximize, close) in the top right. Below the title bar is a progress indicator consisting of seven circles; the fifth circle is filled with blue, and the word "Classify" is centered below it. The main area of the wizard contains three input fields: "Output Classified Dataset" (a text box), "Use Classifier From (optional)" (a dropdown menu with a blue arrow on the right), and "Output Classifier Definition File (.ecd) (optional)" (a text box). At the bottom right, there are two buttons: "< Previous" and "Run".

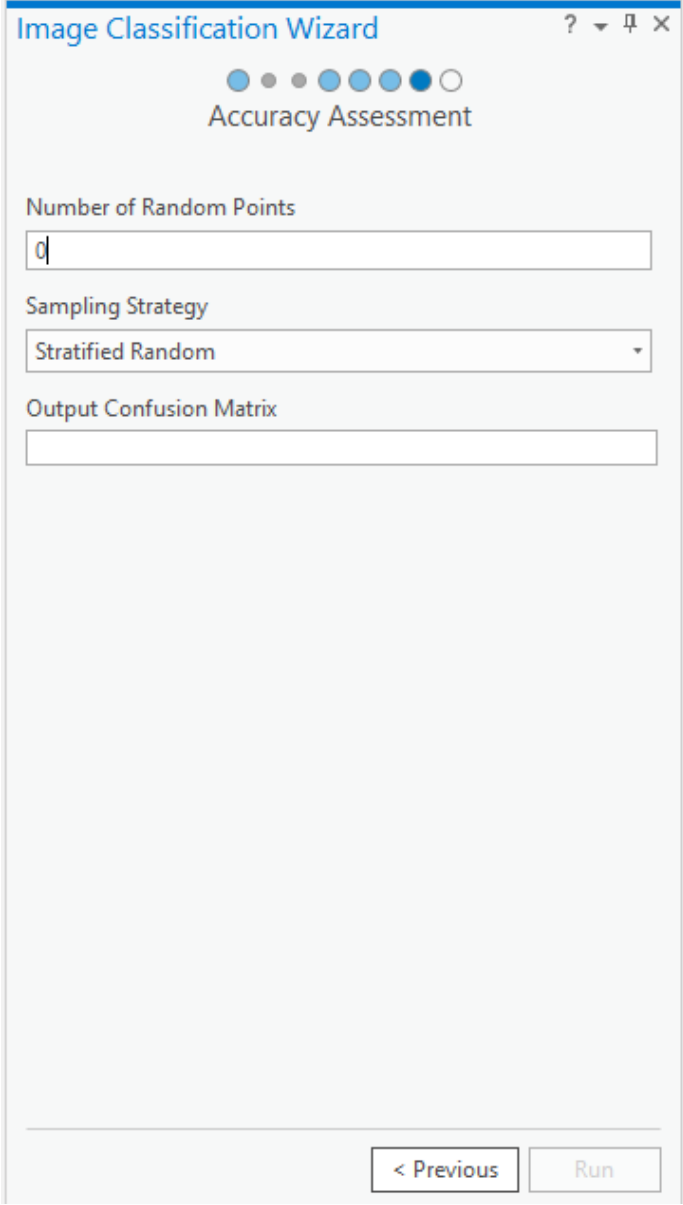
It is now that the classified image is generated; the two optional lines that can be seen in the image refer in the first case to the choice of which of the various previews previously generated you want to use (by default the last one which was created is considered), while the second allows you to save, in addition to the image, an .ecd file, which contains all the necessary information to operate an ESRI classification.

Step 6: Merge classes

Old Class	New Class
rock/ soil	rock/ soil
snow/ ice	snow/ ice
water	water

In this Step, if many classes have been generated in the previous one, it is possible to merge them (for example, if the vegetation was previously divided into many different classes, here it is possible to merge them into a single one). This tool was never really useful for this study as the classes generated were already few and very different the one from the other, so unmergeable.

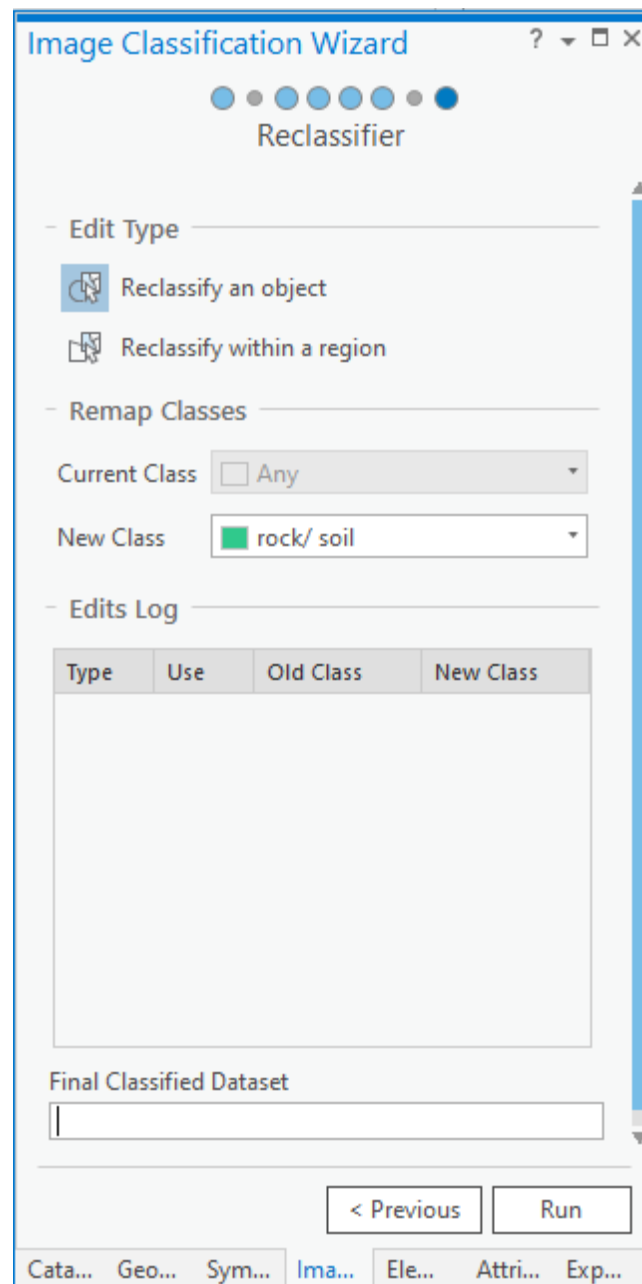
Step 7: Accuracy Assessment



The screenshot shows a software window titled "Image Classification Wizard" with a standard Windows-style title bar (minimize, maximize, close buttons). Below the title bar is a progress indicator consisting of seven circles; the sixth circle from the left is filled with blue, indicating the current step is "Accuracy Assessment". The main area of the dialog contains three labeled input fields: "Number of Random Points" with a text box containing the value "0"; "Sampling Strategy" with a dropdown menu currently set to "Stratified Random"; and "Output Confusion Matrix" with an empty text box. At the bottom right of the dialog, there are two buttons: "< Previous" and "Run".

If a Reference Dataset had been supplied in Step 1, this dialogue box is activated to generate the confusion matrix, setting the number of points to be used and the criteria for distributing them in the space. In this case, this step was actually replaced by the use of individual ArcMap commands, which will be described in more detail in sub-chapter 4.4, since at the time that the classified images were produced, ground truth was not yet available.

Step 8: Reclassify



This command is used, if small point areas have been misclassified, to assign them to the class they should belong to. It is possible to draw a polygon (Reclassify within a region), so the area it encloses will be reclassified, or select a segment (Reclassify an object) to reclassify.

It is a useful operation if there are small errors that can be corrected in this quick way instead of reclassifying the whole image, however, it has never been used, in order to see the behavior of the different classifiers and especially because, in the case of the SAR image classifications, the errors were too many to be corrected in this way.

4.1 SAR

First of all, the SAR image was subjected to preparation operations, for which the SNAP software was used. These were in particular: subset, calibration, filtering, georeferencing and geometric correction and finally the conversion of the band into decibels. The image was then segmented and classified, firstly in an unsupervised manner and then in a supervised one; both of them were object-based.

The reason why an object-based classification was chosen instead of a pixel-based one is that the latter only takes into account the individual characteristics of the pixel in order to assign it to a certain class, while the former groups neighboring pixels on the basis of their similarity, and since SAR images are affected by speckle, a pixel-based classification would have been too affected by this noise, and the final image would have presented a 'salt and pepper' effect, as Dewi (Dewi et al., 2019) states.

Seeing that in the image resulting from the unsupervised classification there was a strong noise due to the speckle effect, to overcome the problem, as proposed by Lee (Lee et al., 1989), the Lee filter was applied multiple times to the original image: this would reduce the noise and accentuate the peaks in the histogram so as to make segmentation by thresholds easier. Even though this segmentation was not the one used in this case, the idea was that applying the Lee filter multiple times would still help in segmenting the image and decrease the possibility of misclassifying individual segments.

What was done was then to repeat the same process of segmentation and unsupervised classification with three images to which the Lee filter, the Lee Sigma filter and the Refined Lee filter had been applied three times respectively. Then the one that gave the best result was chosen and classified in a supervised manner with the three classifiers available on ArcGIS: Maximum Likelihood, Random Trees and Support Vector Machine.

The steps followed are illustrated in the flowchart in Figure 4.1.1 and will now be explained in detail in sub-chapters 4.1.1, 4.1.2, 4.1.3 and 4.1.4.

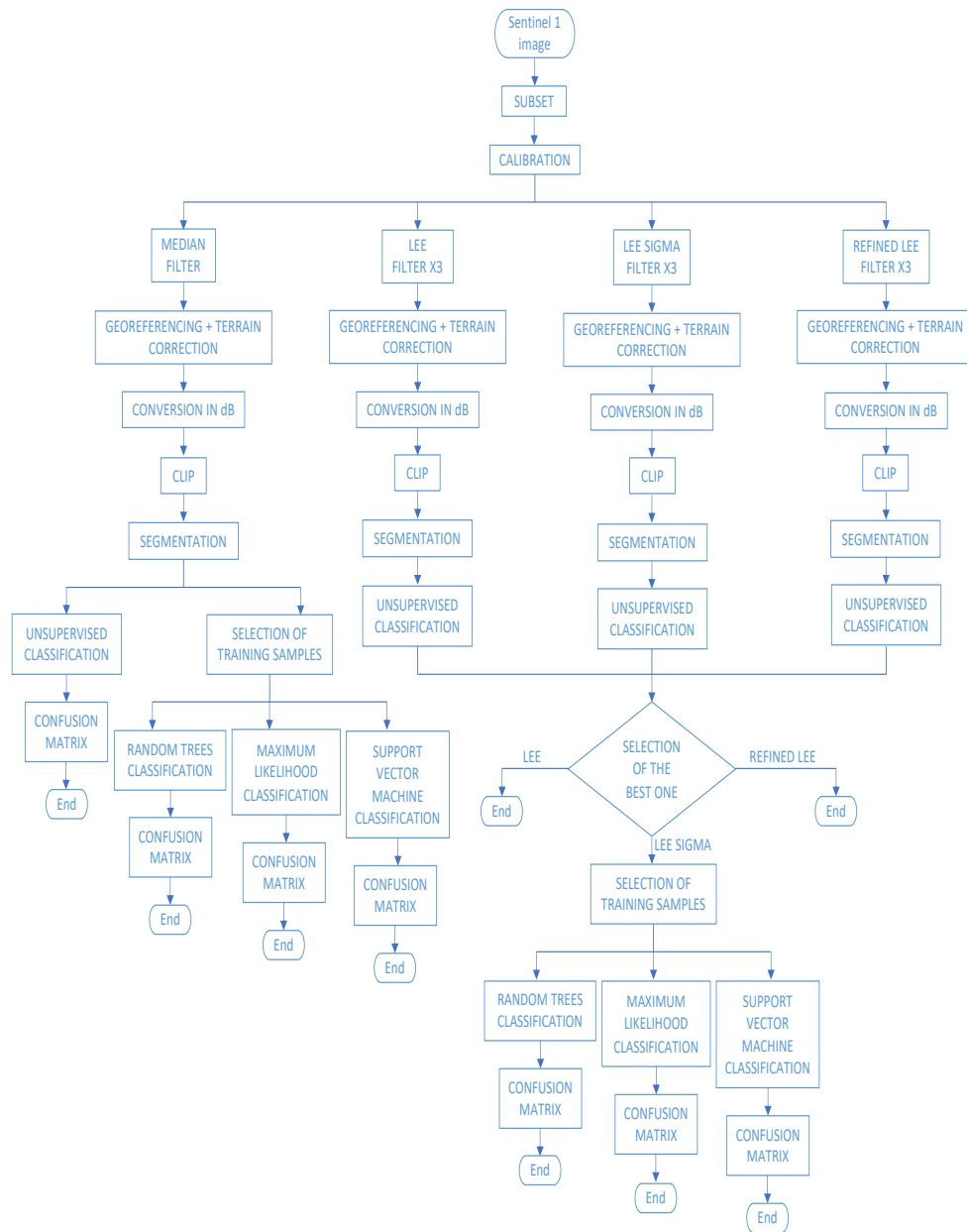


Figure 4.1.1: Flowchart of SAR image

4.1.1 PREPARATION OF THE IMAGE

The SAR image preparation operations prior to segmentation and subsequent classification were almost completely carried out using the software SNAP, with the exception of the last clip with the shapefile of the coastline, which instead took place in ArcGIS. They will now be illustrated, following the chronological order in which they occurred.

SUBSET

The subset operation is necessary because the downloaded image includes an area much larger than just the island of interest; it must also be done in order to lighten the subsequent operations.

To do this, the Subset command belonging to SNAP's Raster package was used; it allows you to enter the geographical or pixel coordinates of the crop you want to make, as well as the bands you want to take into account during the operation. The coordinates used are shown in Figure 4.1.1 and the bands considered were both amplitude and intensity. The other two windows that can be seen in the figure relate to other data included in the folder in which the image is contained: the metadata and the tie-point grids, images that indicate the trend of latitude, longitude, incidence and elevation angles and the slant range time. In both the windows it was set to consider everything.

A preview of the area that will be cropped is indicated with a blue rectangle on the original image in the command window.

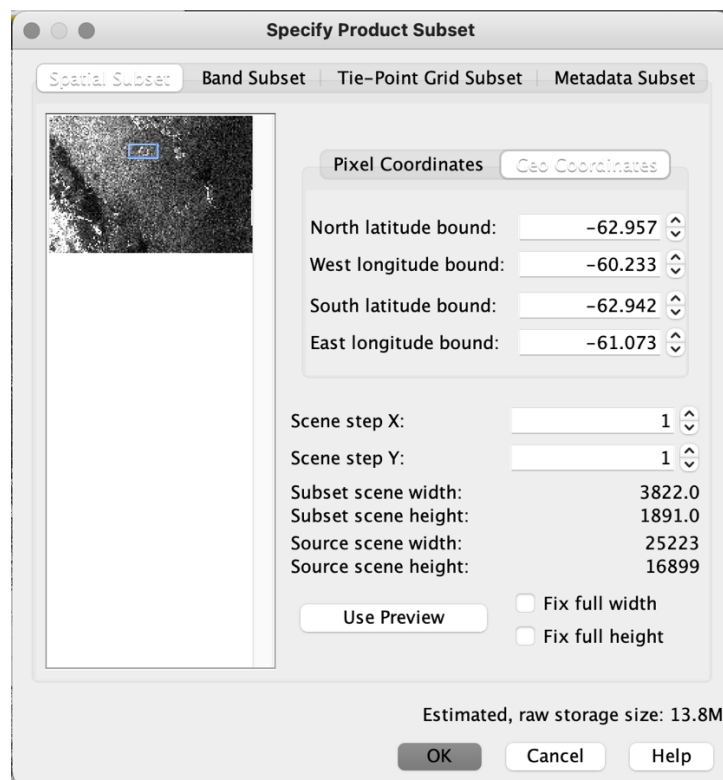


Figure 4.1.1.1: Subset

CALIBRATION

In order to normalize the intensity of the pixels it is necessary to calibrate the image; in order to do this the Calibrate command of the Radar package was used. This command allows you to choose which of the polarizations to use in case of a multipolarized image, but in this case there was only one polarization, the HH. It is then possible to choose whether to output a sigma 0, gamma 0 or beta 0 band, where the difference lies in how the values are normalized: in the case of sigma 0 they are normalized using the ground area, in the case of gamma 0 using the illuminated area projected in the look direction and instead with beta 0 they are modulated using the tangent of the local angle of incidence. The choice was sigma 0.

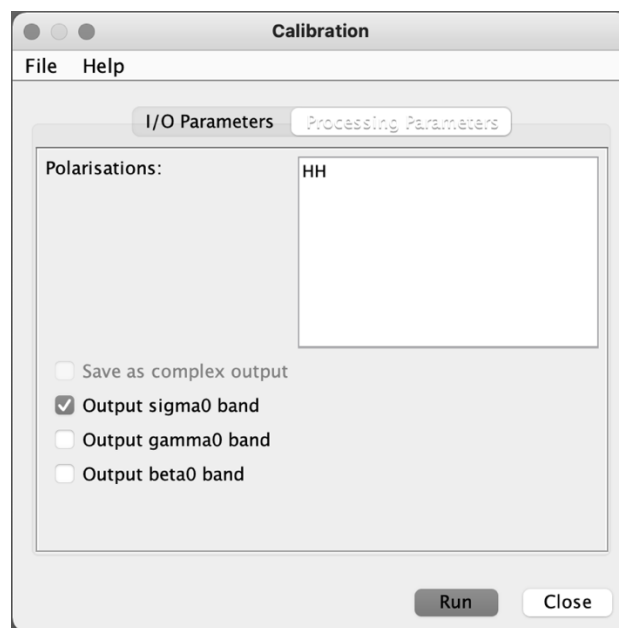


Figure 4.1.1.2: Calibration

FILTER

At this point, in order to reduce the speckle noise characteristic of this type of image, various filters were used: firstly, all of the filters proposed by SNAP were tried, and the best result seemed to be given by the Median filter, so it was applied, then, as suggested by Lee (Lee et al., 1989), we opted for a multiple application of the Lee filter, of which we also tried the two modified versions present in SNAP: the Lee Sigma and the Refined Lee.

The command to be used to do this is the Single Product Speckle Filter, from the Radar package (Figure 4.1.1.3), in which it is possible to choose between the available filters to be applied and then set the filter parameters such as the dimensions of the filtering window, which has always been set at 3x3, with the exception of the Refined Lee filter, for which it was not possible to define it, as it always operates with a 7x7 window.

In the case of the Lee Sigma filter, it was also possible to choose the sigma value, which was set at 0.9.

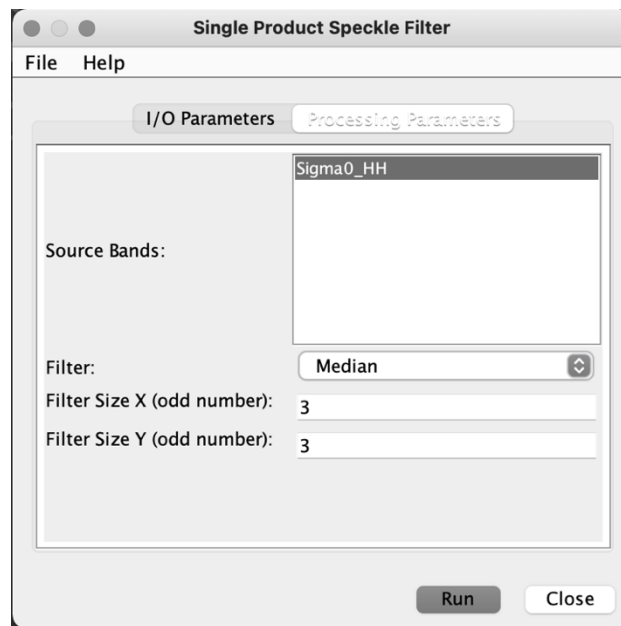


Figure 4.1.1.3: Single Product Speckle Filter

GEOREFERENCING AND GEOMETRIC CORRECTION

For the georeferencing and geometric correction of the image, the Range-Doppler Terrain Correction command of the Radar package was used, using the DTM obtained from the 2005 topographic map cited in chapter 3 and setting the UTM zone 20 South, WGS84 as the reference system. As far as the DTM resampling method is concerned, bilinear interpolation was chosen. In Figure 4.1.1.4 it is possible to see how the various parameters of the command have been set.

Source Bands: Sigma0_HH

Digital Elevation Model: External DEM

External DEM: iDecepcion/UCA_MDT_02m_IslaDecepcion.tif ...

External DEM No Data Value: 0.0

DEM Resampling Method: BILINEAR_INTERPOLATION

Image Resampling Method: BILINEAR_INTERPOLATION

Source GR Pixel Spacings (az x rg): 10.0(m) x 10.0(m)

Pixel Spacing (m): 10.0

Pixel Spacing (deg): 8.983152841195215E-5

Map Projection: UTM Zone 20, South / World Geodetic System 1984

☒ Mask out areas without elevation ☐ Output complex data

Output bands for:

☒ Selected source band ☐ DEM ☐ Latitude & Longitude

☐ Incidence angle from ellipsoid ☐ Local incidence angle ☐ Projected local incidence angle

☐ Apply radiometric normalization

☐ Save Sigma0 band Use projected local incidence angle from DEM

☐ Save Gamma0 band Use projected local incidence angle from DEM

☐ Save Beta0 band

Auxiliary File (ASAR only): Latest Auxiliary File

Figure 4.1.1.4: Range-Doppler Terrain Correction

CONVERSION OF THE BAND TO DECIBELS

The last operation is to convert the band to dB, which allows us to obtain values that are much easier to work with, and for this we used the special SNAP Linear to/from dB command (Figure 4.1.1.5). Figures 4.1.1.6 and 4.1.1.7 show the histograms of the band before and after the conversion, to show how the values change.

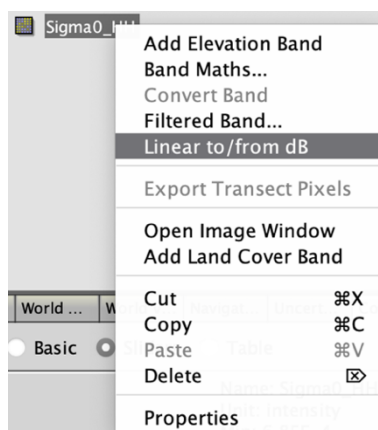


Figure 4.1.1.5: Linear to/ from dB

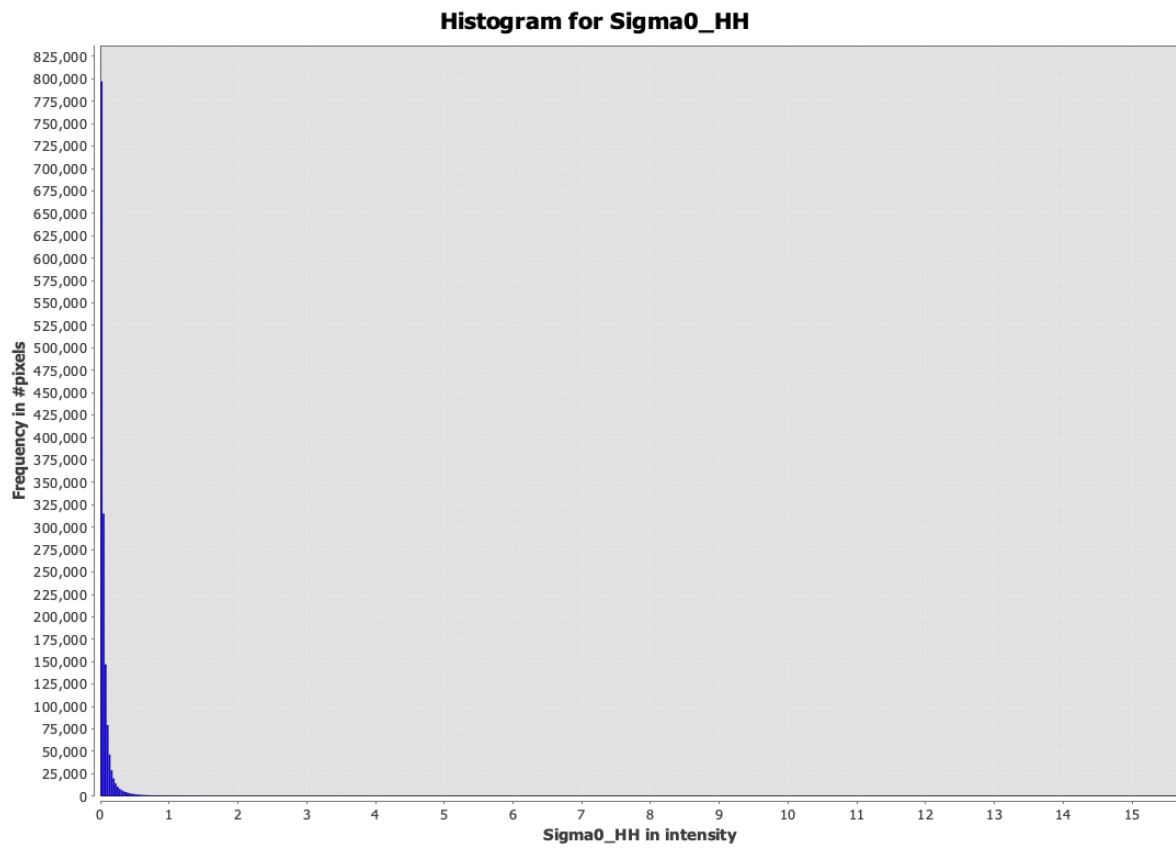


Figure 4.1.1.6: Histogram of the band before the conversion to dB

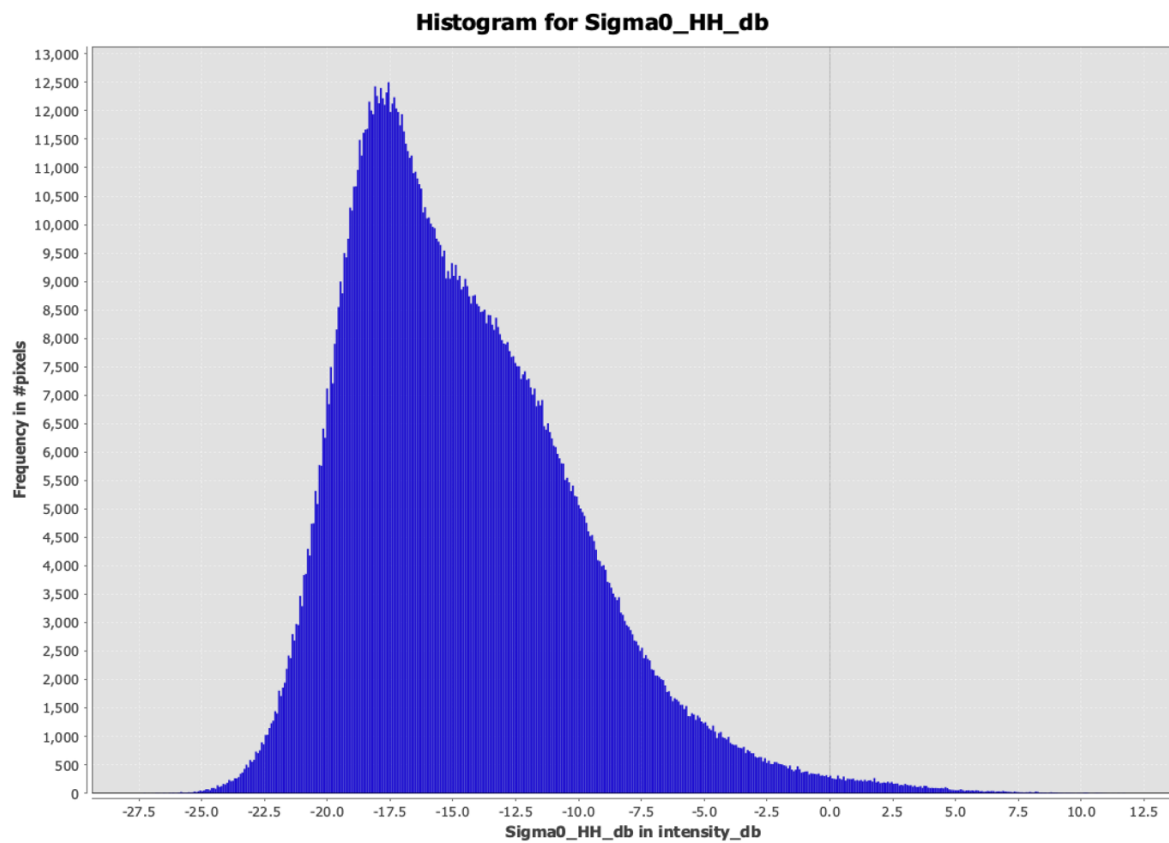


Figure 4.1.1.7: Histogram of the band after the conversion to dB

EXPORT IN GEOTIFF FORMAT

The last operation carried out in SNAP was to export the band in GeoTIFF format, so that the segmentation and classification work could be continued in ArcGIS.

CLIPPING

The preparation of the SAR image concludes with its clipping with the shapefile of the island's coastline, obtained from the 2005 topographic map, as already indicated above. The command used in this case was the Clip, which belongs to the Raster section of the Data Management Tools, shown in Figure 4.1.1.8. It is sufficient to place the image to be clipped as input, select an area by entering coordinates or insert a shapefile (as in this case) and, optionally, indicate which No Data value is to be placed on the final image, which has been set equal to -9999.

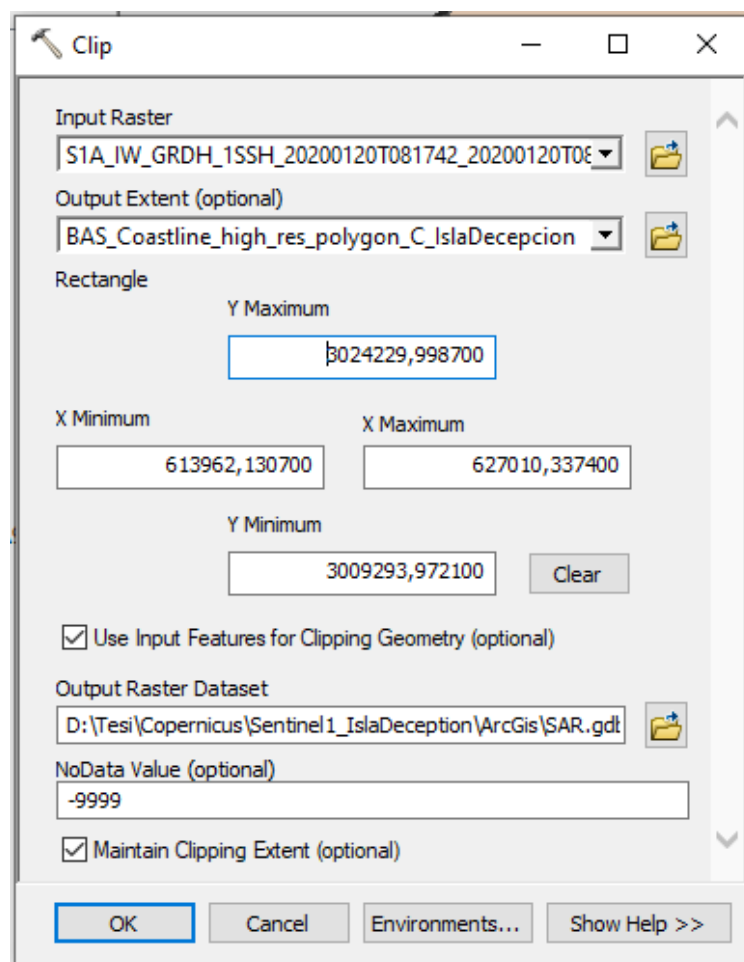


Figure 4.1.1.8 Clip

4.1.2 SEGMENTATION

In order to carry out an object-based classification, it is necessary to segment the image, an operation that can be done using the Segment Mean Shift command of the Spatial Analyst tools of ArcGIS. This command requires a raster image as input and allows you to change the spectral detail, the spatial detail, the minimum segment size in pixels and the bands to be used to segment the image (Figure 4.1.2.1). In particular, spectral detail indicates how much importance is given to the spectral differences between cells, and therefore a high spectral detail means placing pixels with slightly different spectral characteristics in different segments; while spatial detail indicates how important the proximity between the elements constituting the image is, so that larger values lead to a more fragmented image while smaller values lead to a more homogeneous image. Both spatial detail and spectral detail move in a range from 1 to 20.

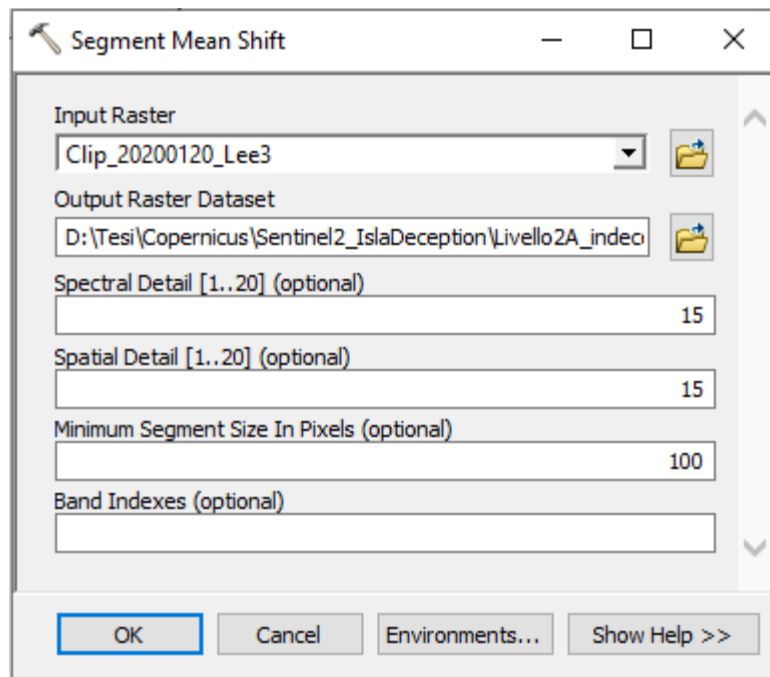
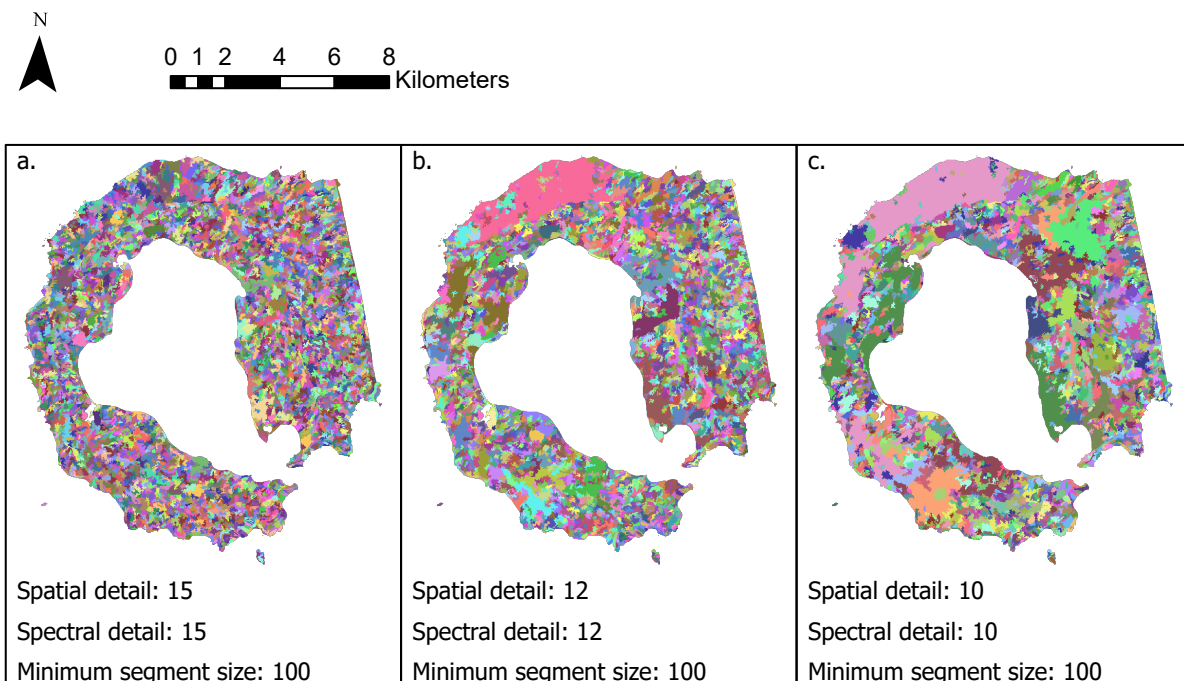


Figure 4.1.2.1: Segment Mean Shift

In this case, several tests were carried out by changing spatial detail and spectral detail, while the minimum segment size, which indicates the minimum size we want to give to a segment in terms of number of pixels, was considered appropriate to be 100 pixels, which corresponds to an area of 10000 m² (about 0.01% of the area of the island). Figures 4.1.2.2 a., b. and c. illustrate how the result varies with the variation of spatial detail and spectral detail in the image with the Median filter.

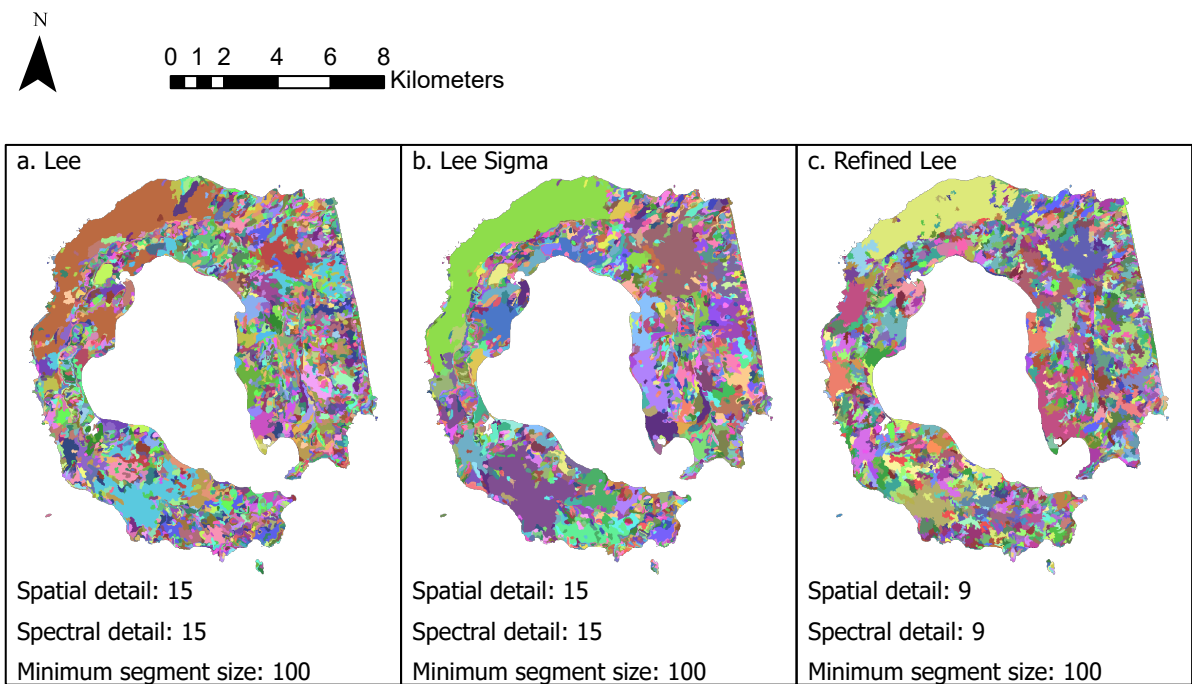


Figures 4.1.2.2: Effect of the change of spatial detail and spectral detail on the image with Median filter

In the end, the most suitable combination for this image turned out to be the one with spatial detail and spectral detail of 12, being able to segment the image sufficiently but at the same time to recognize homogeneous parts, such as the orange stripe located in the north, which is bare soil/ rock. The further reduction of spatial detail and spectral detail led to excessively large segments, as in the image on the right, within which more than one land cover type was included (undesired effect). On the contrary, an increase of spatial and spectral detail resulted in an overly fragmented segmentation, with excessively small segments, whose grouping into classes would be too difficult.

On the other hand, with regard to the images to which the Lee, Lee Sigma and Refined Lee filters were applied, the parameters used to obtain an optimal segmentation were different from those of the image with Median filter: spatial and spectral detail of 15 were chosen for the images to which the Lee and Lee Sigma filters were applied and spatial and spectral detail of 9 for the image to which Refined Lee was applied. The behavior was the expected one in the case of Lee and Lee Sigma images, because when the speckle is reduced, greater spatial and spectral detail are needed to obtain segments of the same size as those of an image affected by speckle, which leads to excessive fragmentation, while what was not expected was the behavior in the case of the Refined Lee image, where it was necessary to even reduce the spatial and spectral detail compared to those used in the original image.

The results of the three segmentations are shown in Figures 4.1.2.3 a., b. and c.



Figures 4.1.2.3: Lee, Lee Sigma and Refined Lee images after segmentation

4.1.3 UNSUPERVISED CLASSIFICATION

Having the segmented images, it was possible to proceed with unsupervised object-based classifications; to do this the Classification Wizard illustrated in Chapter 4 was used.

Referring to the steps indicated in this chapter, in the first step it was indicated that the classification was the one just mentioned, for the classification scheme the NLCD2011 (National Land Cover Database of 2011) scheme was set and the segmented images created in sub-section 4.1.2 were inserted, starting with the image to which the Median filter had been applied and then continuing with the other three. Using this setting, the steps that were turned on were 4, 5, 6 and 8.

In step 4, that is, the classifier training step, it was possible to set various parameters such as the number of classes into which the image was to be divided, the attributes of the segments to be considered in the grouping into classes, and various other parameters such as the number of iterations to be carried out or the maximum number of merges to be made in each iteration (Figure 4.1.3.1).

Image Classification Wizard

Train

Classifier
ISO Cluster

Maximum Number of Classes
3

Maximum Number of Iterations
20

Maximum Number of Cluster Merges per Iteration
5

Maximum Merge Distance
0,5

Minimum Samples Per Cluster
20

Skip Factor
10

▼ Segment Attributes

- ☐ Active chromaticity color
- ☐ Mean digital number
- ☐ Standard deviation
- ☐ Count of pixels
- ☐ Compactness

< Previous Run

Figure 4.1.3.1 ISO cluster training

What was done was to try out different classifications by changing the number of classes in which to group the segments and the attributes of the segments to be taken into account when grouping them into classes, while, regarding the other parameters, those proposed by default by ArcGIS were set, as shown in Figure 4.1.3.1.

As for the segment attributes, the possible one were [3]:

- active chromaticity color: refers to the RGB color of the segment, calculated on the original image;
- mean digital number: the average DN of the segment, calculated on the image in pixels;
- standard deviation: the standard deviation of the segment, calculated on the image in pixels;
- count of pixels: the number of pixels that are included in the segment;
- compactness: a number between 0 and 1 that indicates the degree of compactness or circularity of a segment, where 1 corresponds to a perfect circle;
- rectangularity: a value between 0 and 1 indicating the degree of rectangularity of the segment, where 1 corresponds to a perfect rectangle.

First of all, the application of the different attributes was tested, leaving the number of classes in which to classify the image high, in order to see how they behave, and what was seen for the image with the Median filter was that:

- the active chromaticity color and the mean digital number led to a classification of the image influenced by the morphology of the island and the illumination more than by the ground cover types;
- the standard deviation was the attribute that seemed to be less influenced by morphology and illumination and was able to vaguely identify the two glaciers present;
- count of pixels and compactness led to an excessive fragmentation of the island, in small and apparently insignificant areas;
- the rectangularity also presents the problem of excessive fragmentation and in addition it seems to not recognize the glaciers.

These results are shown in Figure 4.1.3.2.

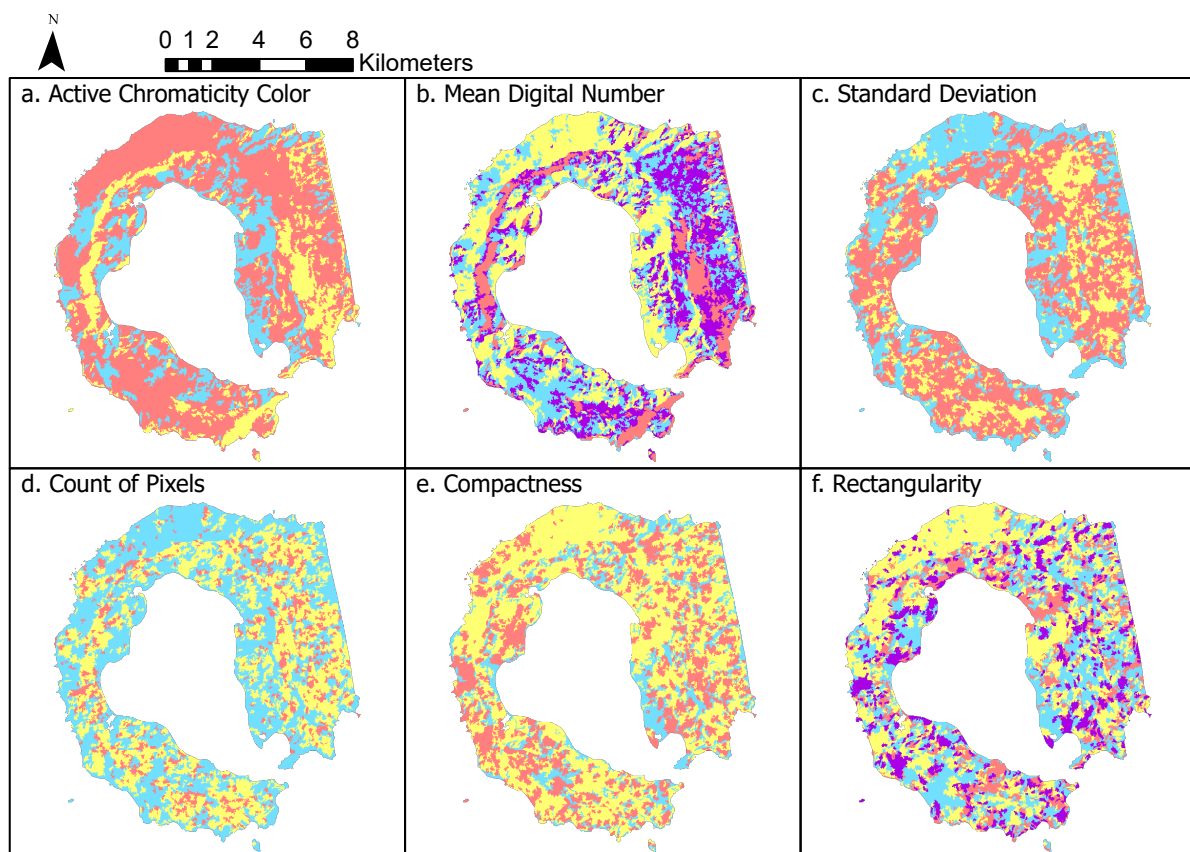
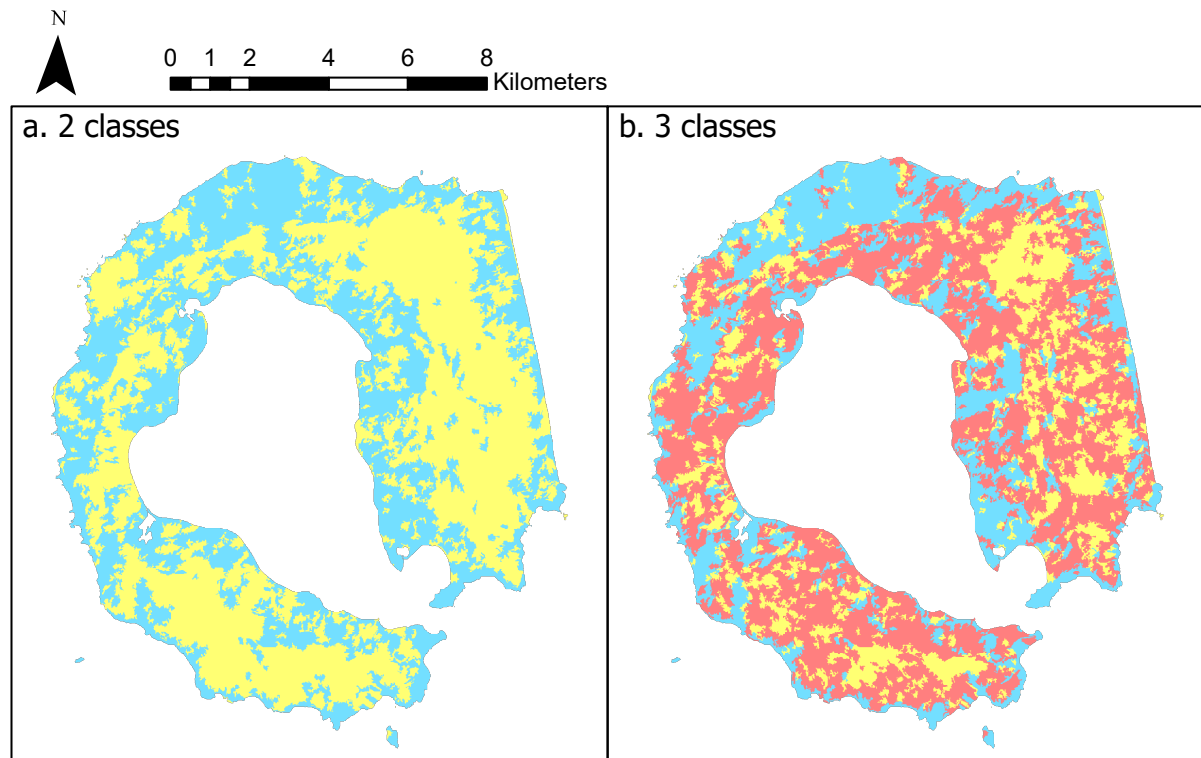


Figure 4.1.3.2: effect of the different segment attributes on the image with Median filter

It was stated that the best choice was to use only the standard deviation, as the other attributes were worsening the result by conditioning the image with radar illumination and island morphology or by causing fragmentation.

At this point, the number of classes was changed, considering a minimum of two to divide what was bare ground from what was covered by snow or ice. In fact, placing a higher number of classes showed that the maximum number in which the classifier could divide them was

three, and that the third class did not correspond to a new ground cover, water presumably, but only to noise, due to speckle (Figures 4.1.3.3). The image chosen was therefore the one consisting of only two classes.



Figures 4.1.3.3: division in 2 and 3 classes using the ISO cluster classifier and standard deviation as segment attribute

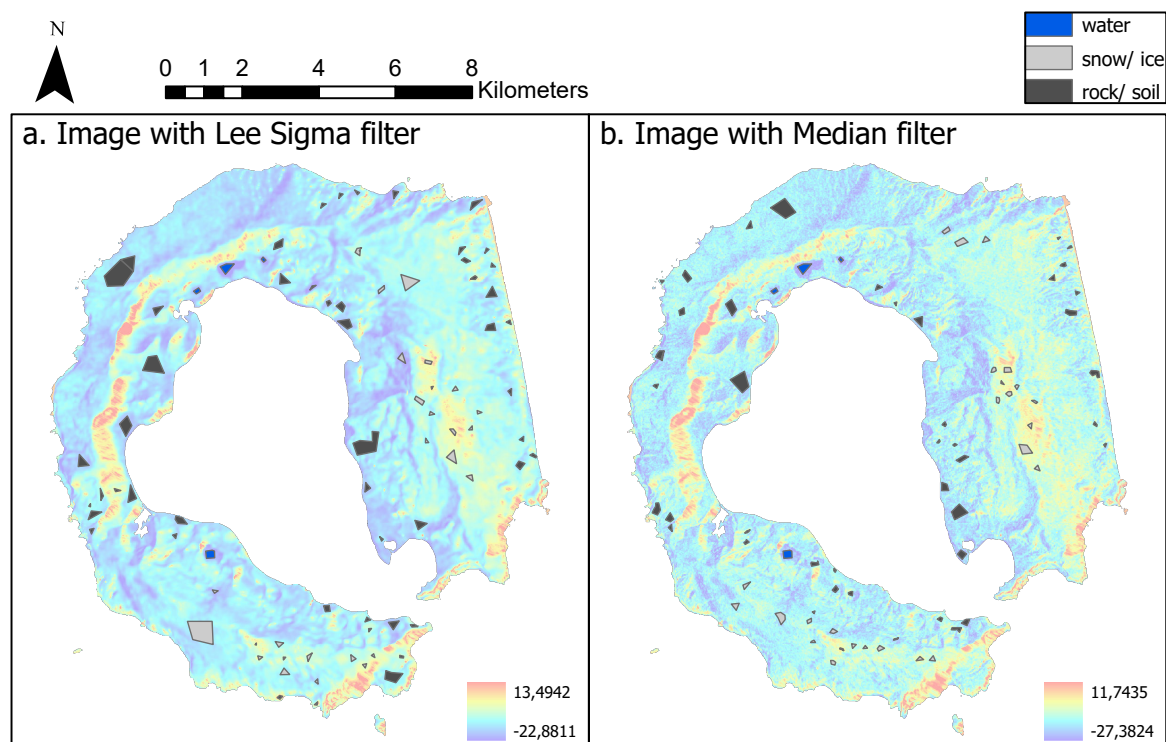
The classification was therefore practically finished: the step following this previous training one was carried out. It produced and saved the final images, while the two remaining steps were of no use, Step 6 because, as already mentioned in chapter 4, it serves to unite various classes into a single one when many are produced, but this is not the case, and Step 8 because, although the speckle generates a lot of point noise, which would therefore be correctable with this tool, there are too many errors to be corrected by hand.

The same process was then repeated for the remaining three images, repeating the search for the most suitable segment attributes in each case and the optimal number of classes. It turned out that the number of classes able to give the best result was always three, while, as far as the segment attributes were concerned, in the case of the image to which the Lee filter had been applied, the most suitable was the active chromaticity color, in the case of the image to which the Lee Sigma filter had been applied, the combination of active chromaticity color, mean digital number and standard deviation was the best choice, and in the case of the image to which the Refined Lee filter had been applied, the active chromaticity color and the mean digital number provided the best result.

The best result was obtained with the image to which the Lee Sigma filter had been applied, which was able to recognize the entire strip of rock/ soil located in the north of the island and to approximate the two glaciers, although with little precision. For this reason, it is with it (and with the image to which the Median filter was applied) that the supervised classifications will be carried out.

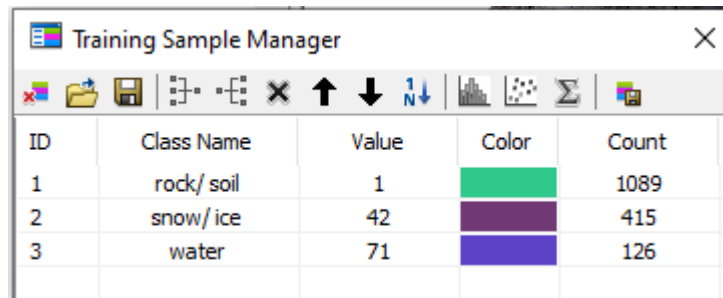
4.1.4 SUPERVISED CLASSIFICATION

To classify the image in a supervised manner, the necessary training samples must be generated before proceeding with the Classification Wizard in ArcGIS Pro. To do this, three separate polygonal shapefiles were created for each of the two segmented images (Median filtered image and Lee Sigma filtered image): snow/ice, water and soil/rock, and the polygons belonging to them were drawn using the segmented images as a base and the previous day's optical image as a reference in order to understand which ground cover corresponded to each zone, as it was impossible to tell by looking at the SAR image alone. In order to draw the various polygons, care was taken to distribute them more or less homogeneously over the area of the island, trying to include significant parts and not transitional parts between two covers, and always being careful that each polygon fell within a single segment and not across more than one, so as not to distort the results. In addition, an overall area greater than 1% was sampled for both images. The distribution of the selected areas is shown in Figure 4.1.4.1.



Figures 4.1.4.1: Training samples selected

At this point, once the shapefiles were ready, they were opened with the Training Sample Manager of the Image Classification toolbar of ArcMap (Figure 4.1.4.2), here it is possible to see each class created with its relative size and modify the value and color assigned if desired. From here the signature file necessary for supervised classification was created.



ID	Class Name	Value	Color	Count
1	rock/ soil	1	Green	1089
2	snow/ ice	42	Purple	415
3	water	71	Blue	126

Figure 4.1.4.2 Training Sample Manager

Having now all the necessary elements, the two images were classified in a supervised manner using the Classification Wizard of ArcGIS Pro, following the steps already illustrated in chapter 4.

The classification was set as supervised and object-based (for the reason already explained in chapter 4.1), the segmented image and the training samples were inserted as input, and the classification scheme was set as "generate from training samples", selecting the training samples created before. The steps that were activated in this way were 3, 4, 5, 6 and 8.

Step 3 is related to the Training Samples Manager and is used to modify the classes, adding new ones or eliminating them; however, since this operation has already been done previously, it has been left unchanged.

In the next step, the classifier is chosen between the three available for supervised classification: Maximum Likelihood, Random Trees and Support Vector Machine, and its parameters are set, as well as the segment attributes to be considered, since the classification is object based. For each classifier, several attempts were made with regard to the segment attributes, similarly to what was done for the unsupervised classification (see chapter 4.1.3), while the parameters set for each were as follows:

- Maximum Likelihood: no parameters could be set;
- Random Trees: three parameters could be changed:
 1. The maximum number of trees to be used, which is higher the more accurate the classification, but also the longer the processing time. This parameter was set equal to 500, as Belgiu recommends (Belgiu et al., 2016);
 2. The maximum depth of a tree, which corresponds to how many branches a tree is allowed to make, and was set equal to 3 as this is the number of variables involved;
 3. The maximum number of samples per class: it was set to consider all the training samples inserted;

- Support Vector Machine: the only customizable option was to change the number of training samples to be considered, and the value 0 was set, which is equivalent to considering them all.

Several temporary images were generated at this stage, and then the best ones were chosen to be produced in the next step, where the selection criteria were glacier recognition and minimization of misclassification caused by radar illumination and island topography. Furthermore, in several cases, with the Maximum Likelihood and Support Vector Machine classifiers, it was noted that the water class added a lot of noise to the image rather than actually recognizing the lakes present, so when this occurred, the classification process was repeated from the beginning, only considering the snow/ice and rock/soil training samples and then splitting the image into these two classes. Figure 4.1.4.3 shows an example of what was said: the image with Median filter classified in two (a.) and three (b.) classes using Maximum Likelihood classifier, considering Active Chromaticity Color and Standard Deviation as segment attributes, while to consult all the other cases in which this two-class classification was made because of the confusion generated by water's training samples, consult the appendix.

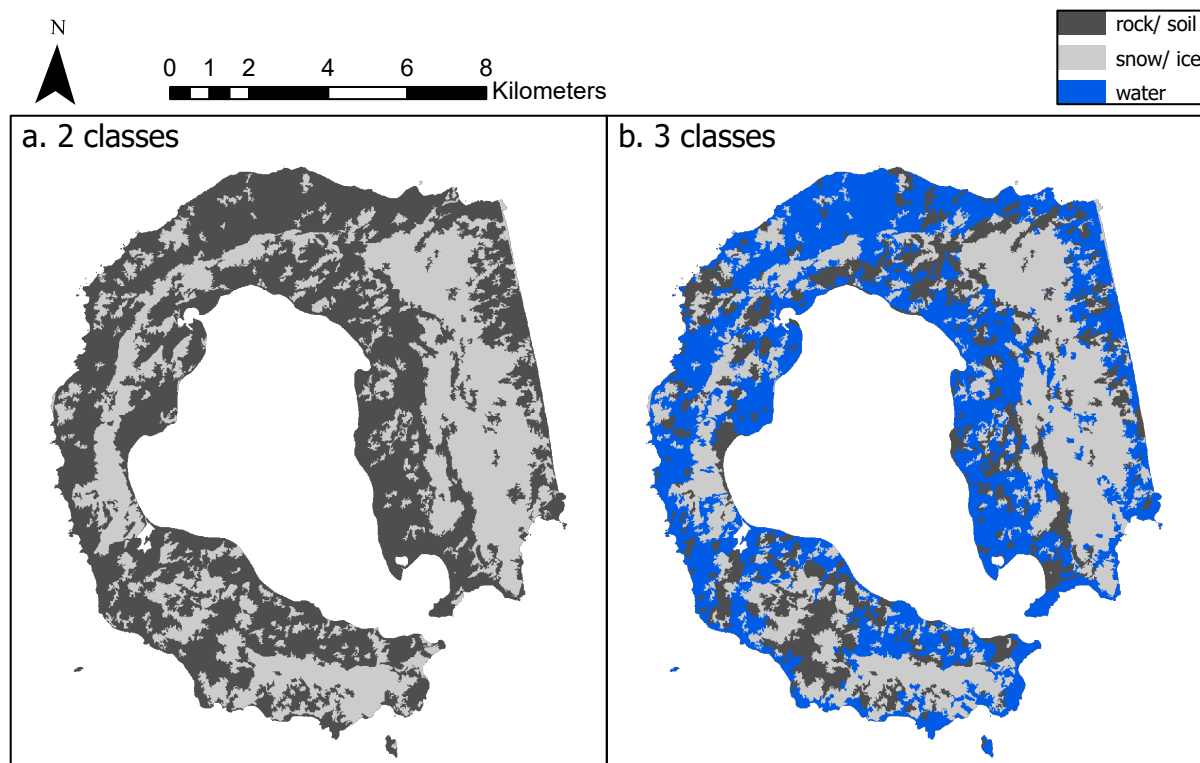


Figure 4.1.4.3: Example of bad classification caused by water training samples

Concerning the two remaining steps, the class merge step was not useful because, as already explained, the classes into which the image has been divided are few and very different from each other, so they cannot be merged. Similarly, the last step of reclassification, which is used

to correct small localized errors, was not carried out because the errors are largely punctual, but they are many and correcting them one by one would have taken too long. Repeating the process with the image to which the Lee Sigma filter had been applied, the results obtained were worse than with the image to which the Median filter had been applied. The appendix shows a comparison between the various classifications obtained from the image with the Median filter and the results obtained by considering the same segment attributes for the image with the Lee Sigma filter.

4.2 OPTICAL

The optical image required fewer operations prior to classification than the SAR image, as the data from Level 2A are already ready to use, presenting the Bottom-of-Atmosphere reflectance values, however area cropping and band resampling operations were carried out, so that the necessary indices could then be calculated with the ArcGIS Raster Calculator. This initial preparation phase was followed by the three classifications: unsupervised, supervised and index-based, where the unsupervised and supervised classifications were chosen to be pixel-based, because these classifications are more immediate than the object-based ones (used for the SAR image), there was no point noise such as speckle that made it necessary to also consider texture features, and the main ground covers present (soil/ rock, water, snow/ice and clouds) were well separated in the characteristics space, as will be seen later with the index-based classification.

These phases will be described in detail in the respective sub-chapters, similarly to what was done for the SAR image, while Figure 4.2.1 shows a flowchart illustrating the processes carried out.

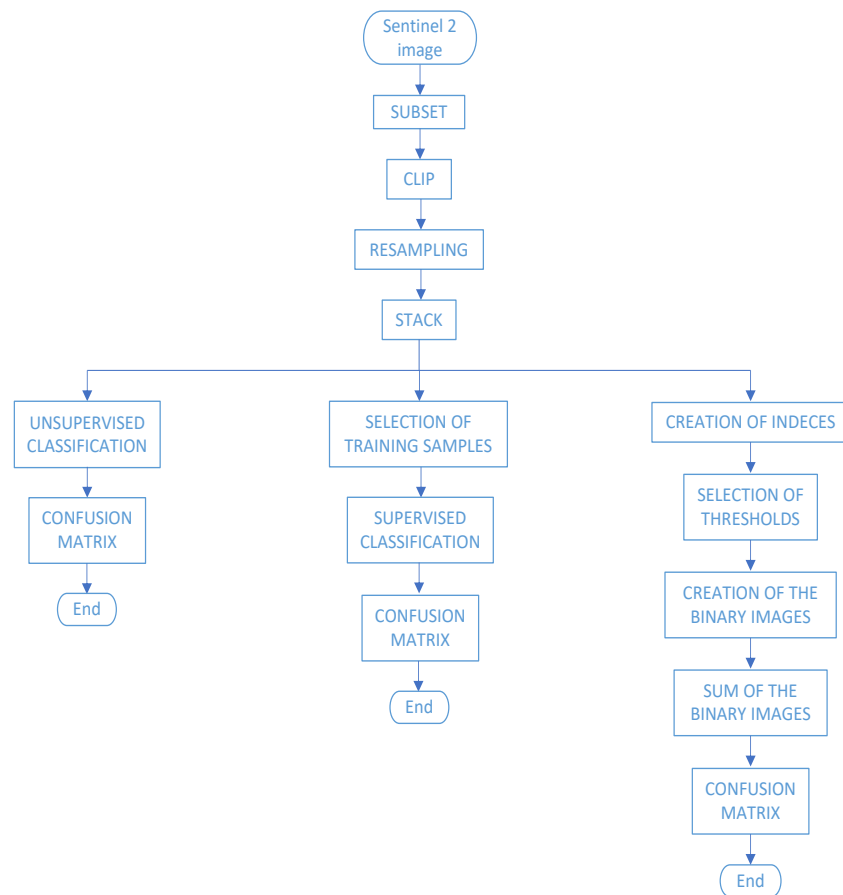


Figure 4.2.1: Flow chart of optical image

4.2.1 PREPARATION OF THE IMAGE

The operations needed to prepare the Sentinel 2 image were not as many as the ones needed for the Sentinel 1; they were carried out firstly using SNAP and then with ArcGIS.

The operations performed will now be explained in chronological order.

SUBSET

As for Sentinel 1 image, Sentinel 2 image also includes an area much larger than the one of interest, so the first operation was to crop it. The command used was Subset, of the Raster package, and the coordinates entered were the same as in the SAR image, reported in chapter 4.1.1. In this case, however, in the band subset window only one band was selected at a time and the operation was repeated for each band, because later, in the Export phase, it will not be possible to export the whole stack together in GeoTIFF format, instead it will be necessary to export band by band and then reunite them in a stack in ArcGIS. Figure 4.2.1.1 shows how the band subset window appears.



Figure 4.2.1.1 Band subset

EXPORT

After cropping the bands, we exported them in GeoTIFF format, one by one as mentioned in the subset explanation.

RESAMPLING

In order to be able to carry out operations between the image bands and be able to calculate the indices for the hierarchical classification, it was necessary that the bands had cells of the same size; therefore a resampling was carried out, bringing the lower resolution bands to a resolution of 10x10. The command to do this is Resample, from the Raster package of the Data Management Tools, where the band to be resampled is entered as input, the size of the

cells to be obtained as output (10x10) is imposed together with the interpolation method to be used to assign the values to the new cells; in particular, there are four available in ArcGIS: nearest, bilinear, cubic and majority. Of these, nearest and majority are the most suitable methods if one is resampling an image of discrete values, since they do not average the values present, but assign the value of the nearest cell in the case of nearest and the most frequent value in a 4x4 filtering cell in the case of majority; while for images with continuous values one uses bilinear, which assigns to the pixel the weighted average over the distances of the four nearest cells, and cubic, which determines the value to be assigned by performing a cubic convolution on the 16 nearest cells. The choice was therefore between bilinear and cubic, but since cubic can give values outside the range of those in the image and requires a long processing time, bilinear was selected.

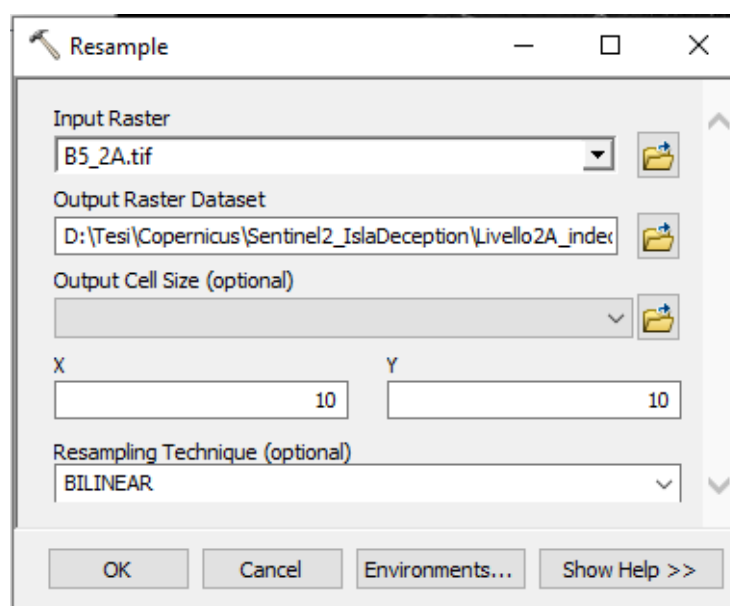


Figure 4.2.1.2 Resample

STACK

At this point, all of the bands except for 1, 9, and 10, as they are not of interest for the study, are grouped into a stack using the Composite Bands command from the Raster section of the Data Management Tools, where you simply select all the bands you want to stack and they are then grouped together. The command interface is shown in Figure 4.2.1.3.

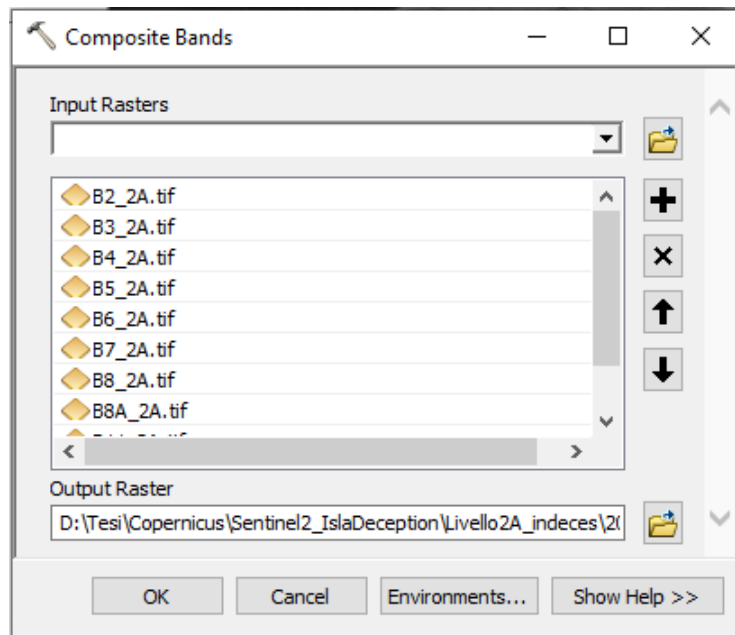


Figure 4.2.1.3 Composite bands

CLIP

Finally, the image was clipped with the shapefile of the island's coastline, in the same way explained in chapter 4.1.1.

4.2.2 UNSUPERVISED CLASSIFICATION

The first of the classifications performed with the optical image was, as with the SAR one, the unsupervised one. Again, the Classification Wizard of ArcGIS Pro was used and the classification was pixel based, as mentioned in chapter 4.2. Bands 1, 9 and 10 of the image were not used, as they were not of interest for the purposes of the study, as already mentioned in chapter 3.

Following the thread indicated in chapter 4, in the first screen of the Classification Wizard, the classification we set to be unsupervised and pixel based; this resulted in the activation of steps 4, 5, 6 and 8. The NLCD2011 was chosen as the classification scheme, as had already been done for the SAR image (see chapter 4.1.3). The next step is therefore that of training the classifier and, as for the SAR image, the parameters proposed as default by ArcGIS were considered (Figure 4.1.3.1), but the number of classes was changed. In this case, moreover, obviously the segment attributes to be considered did not appear, since the classification is pixel based.

By setting the number of classes equal to 4 (water, clouds, snow, rock), it was noticed that the classification was not really well done and for this reason different number of classes were tried (Figure 4.2.2.1). In particular, what happened was that with smaller numbers of classes there was a tendency to assign to the same class elements that in reality belonged to different ground covers, while with an excessive increase in the number of classes the image was

confusing and obviously what in reality was a single class was subdivided into more subclasses, not necessarily in a correct manner, in the sense that a subclass was not necessarily homogeneous and could include elements coming from different ground covers. The best option seemed to be the division into 7 classes, which did not lead to excessive division and thus confusion, and at the same time had fairly homogeneous subclasses.

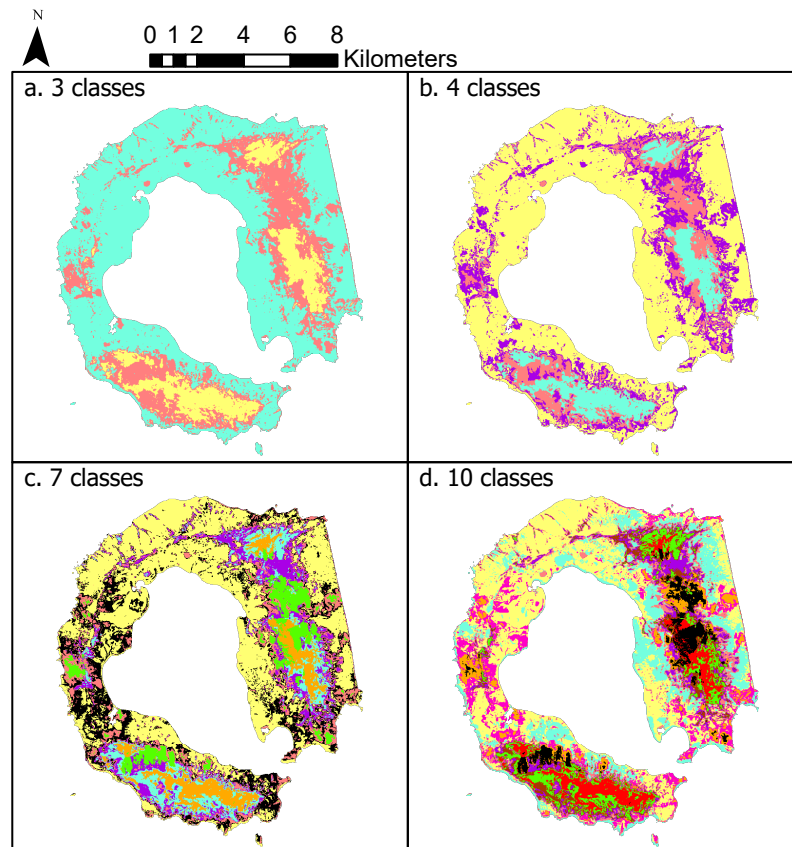


Figure 4.2.2.1 Comparison of division into different numbers of classes

Once the image with 7 classes was chosen, it was set as the one to be used in the next step of creating the actual classified image (these were just temporary versions that were not saved in memory).

It is now time to move on to the next step of class assignments and what it does is, having the list of generated classes and those of NLCD2011, assign to each of the classes considered the ground cover to which it corresponds (see Figure 4.2.2). In this case, the covers listed were not those present in the image, except in the case of water, so the other three of interest to us (rock, snow and clouds) were added. Figure 4.2.2.3 shows the command for adding a new class where what you do is assign the class a color and a value, as well as a brief description. Now having all the necessary classes, we proceeded to assign to each classified class the corresponding ground cover to which it belongs (those just generated) and in doing so we combined the 7 classes into 3 final classes (snow, clouds and rock, because water was not well recognized and belonged to the snow class). At the end of the process of class assignment just described, the image was complete, as the last step of correcting point errors did not take

place, even if it would have been possible, because with SAR images it was never used and therefore comparing manually corrected images with others left in the raw state would not have been fair.

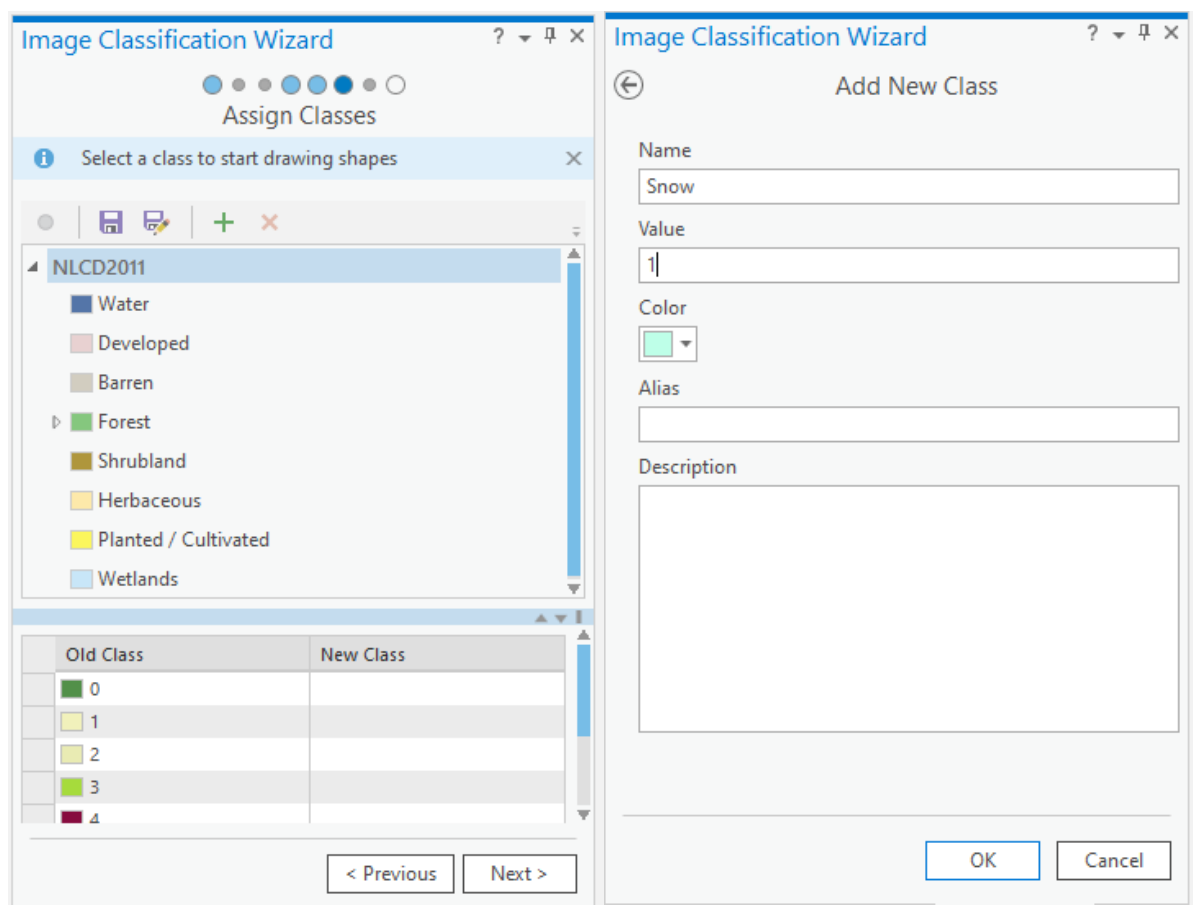


Figure 4.2.2.2. Assign Classes and Figure 4.2.2.3: Add New Class

4.2.3 SUPERVISED CLASSIFICATION

The next classification was supervised and pixel based and was carried out using the Classification Wizard in ArcGIS. The procedure was very similar to that of the supervised classification of the SAR image (chapter 4.1.4), with the difference that, since in this case the classification was object based, it did not require a segmented image.

The first step was to select the necessary training samples for the four macro ground covers present: rock/ soil, snow/ ice, water and clouds. The procedure for the generation of the training samples and the signature file was similar to that carried out for the SAR image, so please refer to chapter 4.1.4 for more details. In this case, however, cloud cover was present, so four classes were generated instead of three: rock/ soil, water, snow/ ice and clouds. As with the SAR image, training samples were selected so that they were evenly distributed over the island, and an area greater than 1% of the island's surface was sampled for reliability. In

addition, while the difference between snow and rock was evident and it was easy to select the corresponding training samples, the same cannot be said for the difference between snow and clouds, so the NDSI index, known to render well the contrast between snow and clouds, was used to identify the clouds. The lakes were easily recognizable, but reference was made to the 2005 topographic map to be sure. Figure 4.2.3.1 shows the training samples selected for the optical image.

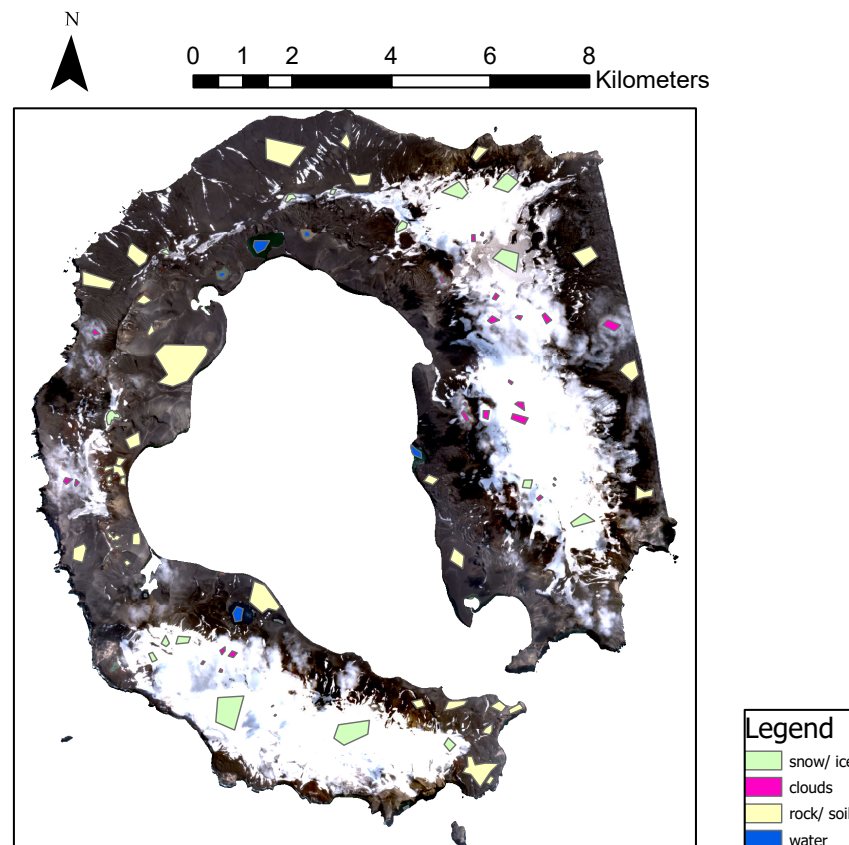


Figure 4.2.3.1: training samples selected for the Sentinel 2 image

At this point, it was possible to proceed with the supervised classification, which was carried out excluding bands 1, 9 and 10, since they were not interesting for the purposes of the study, as already mentioned in chapter 3. The steps followed are the same as those indicated in the supervised classification of the SAR image, with the only difference that this classification is pixel based, therefore when the different classifiers are used, the parameters that can be set are only those relative to them and there's no section dedicated to segment attributes, since in this case there is no segmented image.

In the first step it was set to carry out a supervised pixel based classification and to use the scheme generated by the training samples, and the generated training samples were inserted as input. Step 3 of the Training Samples Manager was superfluous in this case, because the training samples were already ready as they had been prepared previously, and there were no modifications to be made.

In the classifier training step, all three supervised classifiers present in ArcGIS were used: Maximum Likelihood, Random Trees and Support Vector Machine, for which the parameters set were the followings:

- Maximum Likelihood: there are no parameters to set;
- Random Trees:
 1. Maximum number of trees to be used: 500, as suggested by Belgiu (Belgiu et al., 2016);
 2. Maximum depth of a tree: 4, which is the number of variables involved;
 3. Maximum number of samples per class: all selected training samples were considered;
- Support Vector Machine: the only parameter to be set was the number of training samples to be taken into account, and here too it was decided not to exclude any.

After this training phase, the final classified images were created. In fact, the two remaining steps were not useful, that of class merge because the classes present were few and very different from each other, and that of reclassification because the SAR classified images have never been corrected, and correcting the optics would make the comparison between the two unequal.

4.2.4 INDEX CLASSIFICATION

Another method to classify an optical image is to use a series of indices aimed at separating the different classes of terrain, exploiting the presence of two distinct and easily separable distributions in the index histogram, as already studied by Keshri (Keshri et al., 2009) and Shukla (Shukla et al., 2016). In this particular case, the covers present were: water, bare soil/rock, ice mixed with debris, ice and snow, and in particular the indices taken into consideration were: NDWI, MNDWI, MuWIR, NDSI, NDGI, NDSII; for more information on them please refer to chapter 1.2.

In the case of the first three, it was assessed which of them would best achieve the separation of the water features present, i.e. the lakes located in the northern and southern part of the island, near the coast facing Port Foster, the NDSI was used to separate what was snow/ice cover from what was not, and to extract the clouds that obscured part of the island, the NDGI to separate what was pure ice and snow from what was mixed with debris and finally the NDSII to separate ice from snow. The process followed is shown in the flow chart in Figure 4.2.4.1.

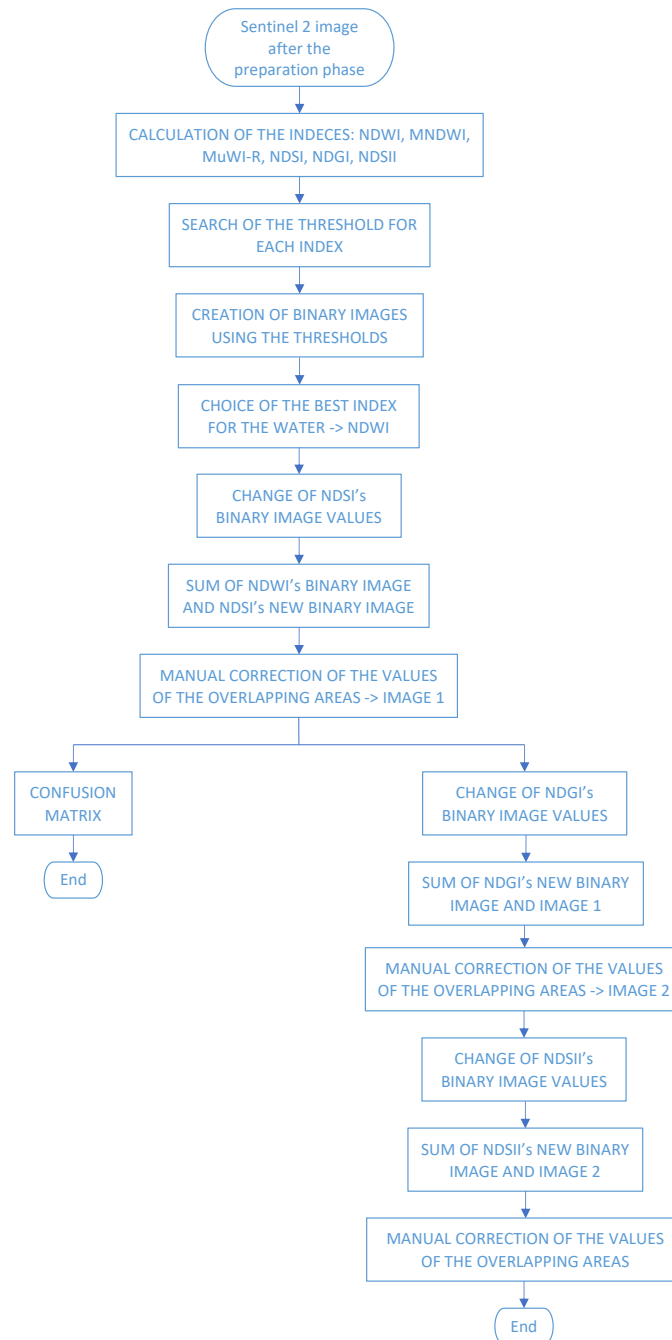


Figure 4.2.4.1 Flow chart of the process

The first step necessary for the purposes of this classification was therefore that of calculating the different indices to be used, for which the formulas reported in chapter 1.2 were used and implemented using the Raster Calculator of the Spatial Analyst Tools package, that

consents to carry out mathematical operations between different rasters, using the rasters of the individual bands that had been resampled, as explained in chapter 4.2.1. Figure 4.2.4.1 shows as an example the expression used to calculate the NDWI index and the indices obtained as a result of what has just been explained are shown in the Appendix.

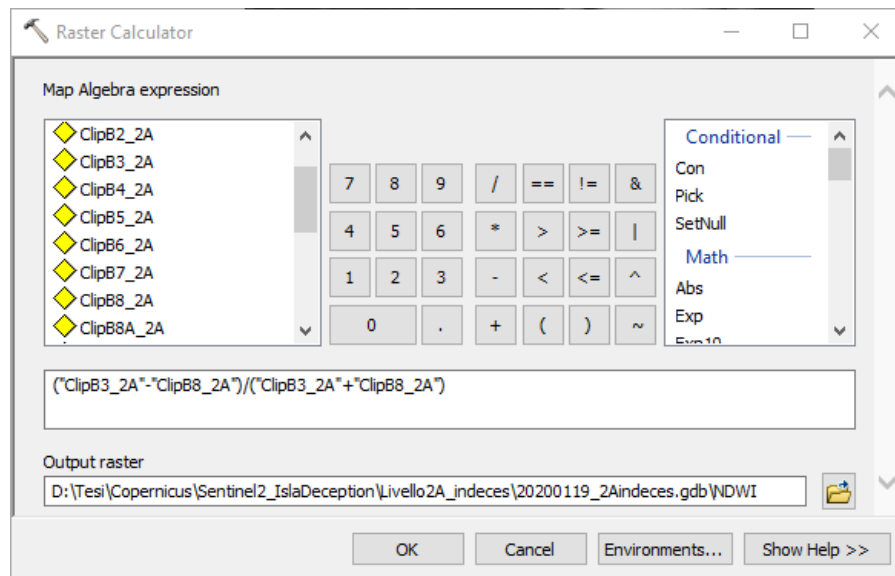


Figure 4.2.4.2 Calculation of NDWI with Raster Calculator

For each index, the threshold value able to divide its histogram into two distinct distributions, which would represent different land covers, was searched. To do this, the program ENVI was used, and in particular its Interactive Stretching function, which allows the visualization of the histogram of the image, and Density Slice (Figure 4.2.4.3), with which it is possible to color different intervals of the histogram: in this way it was verified that the threshold value previously chosen by looking at the histogram was actually able to separate the land cover classes as desired. In Figure 4.2.4.3 an example of the command Density Slice is shown, in which the range set will be colored as red in the image; the command also allows you to add more ranges and color them with different colors.

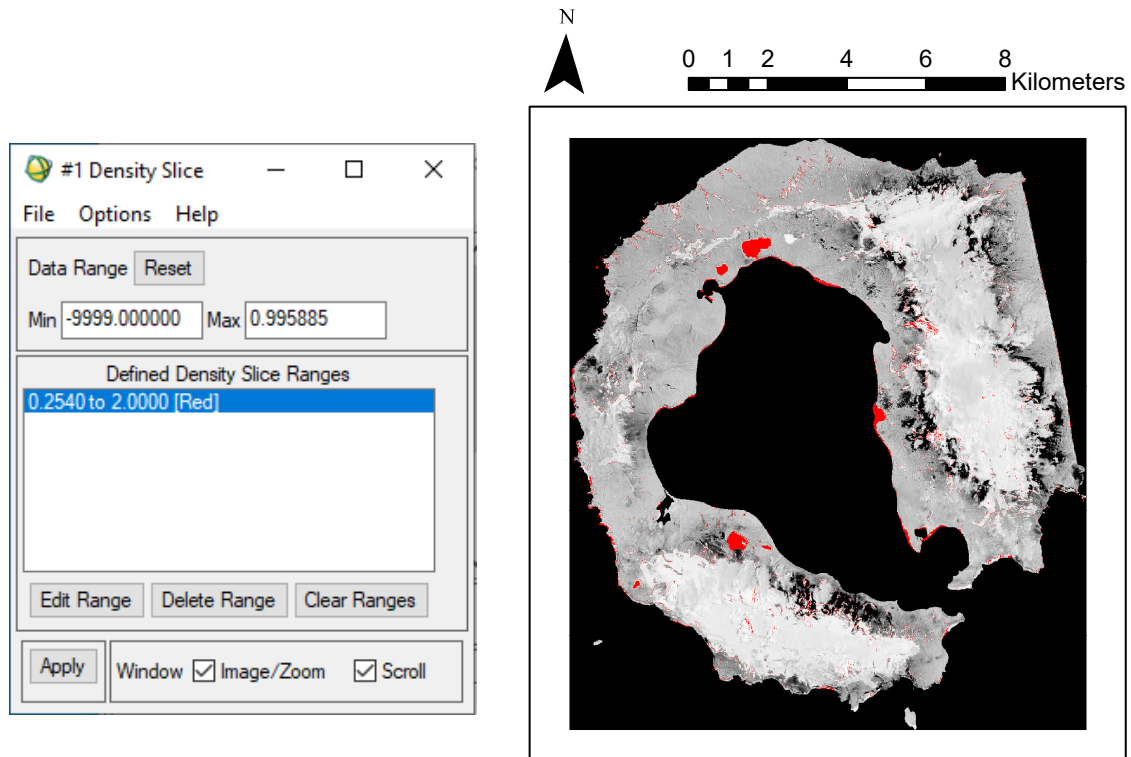


Figure 4.2.4.3 Density Slice command on the left and effect produced on MNDWI image on the right (highlight of red areas)

The values identified for each index are given in Table 4.2.4.1.

Index		
NDWI	$x \geq 0,175$: water	$x < 0,175$: other
MNDWI	$x \geq 0,254$: water	$x < 0,254$: other
MuWI-R	$x \geq 1,378$: water	$x < 1,378$: other
NDSI	$x \geq 0,113$: snow and ice, also mixed with debris	$x < 0,113$: other
NDSI clouds	$0,05 \leq x \leq 0,6$: clouds	$x < 0,05$, $x > 0,6$: other
NDGI	$-0,038 \leq x \leq 0,126$: pure snow and ice	$x < -0,038$, $x > 0,126$: other
NDSII	$x \geq 0,026$: pure snow	$x < 0,026$: other

Table 4.2.4.1: thresholds

Once these threshold values had been found, ArcGIS software was used again, and in particular the Raster Calculator, to create binary images, i.e. composed only of 0 and 1. This was done by imposing conditions based on the limits found in ENVI: for example, by imposing "MuWIR>1.378" in the Raster Calculator, an image was obtained where what is water has a value of 1 and what is not has a value of 0.

With regard to the limits identified, in some cases the choice was clear from the beginning while in others more than one value seemed suitable; for example in the case of the NDSI the

values 0.113 and 0.599 were identified, the former ensures that all the snow cover is highlighted, including the stripes near the outer perimeter of the island, but also includes clouds, while the latter omits these details but does not include clouds. The former was the option chosen because it would also mask out the clouds, thus eliminating these misclassified areas, and under ideal conditions, in the absence of clouds distorting the result, the former would be the most suitable boundary.

Figure 4.2.4.4 shows the histogram of the NDSI index with the lines in correspondence of the identified limits, while Figure 4.2.4.5 shows the effect of the choice of the two different limits on the selection of land cover.

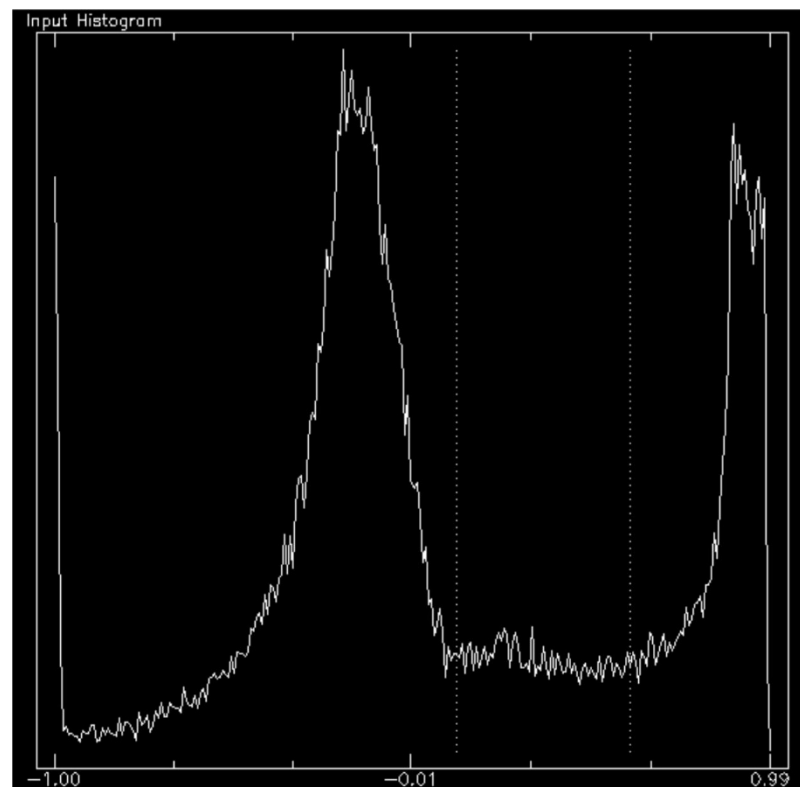


Figure 4.2.4.4: Histogram of NDSI with lines in correspondence of the identified thresholds (0,113 e 0,599)

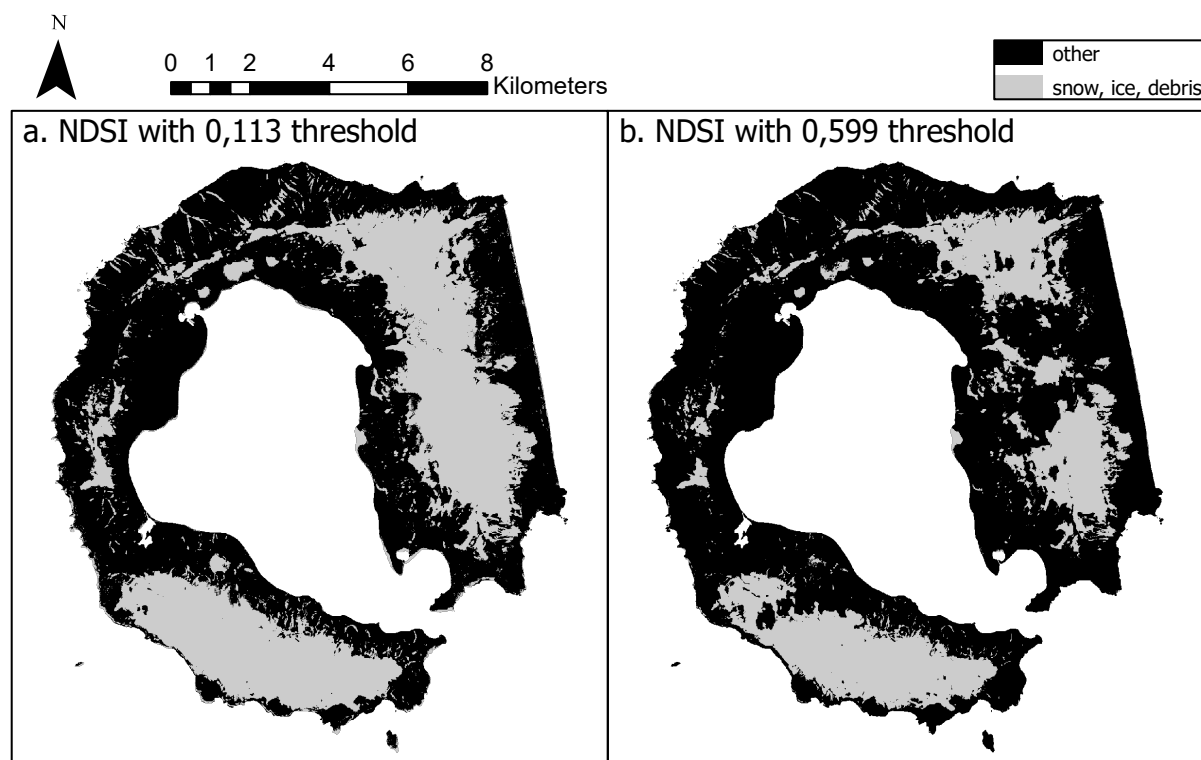


Figure 4.2.4.5: Comparison of the images generated by selecting the two different limits mentioned for the NDSI

A similar situation also occurred with the NDSII, whereby a limit of -0.029 selected more areas of snow but highlighted some clouds, while one of 0.026 left out some strips of snow but didn't include clouds. In reality, as can be seen from the images in the appendix, the snow areas added in the first case were more peripheral, and since this index only classifies snow, without including ice or debris, it was deemed appropriate to choose the value 0.026, which considered a more circumscribed and localized coverage, especially in the central part of the two glaciers. Unfortunately, in this case it is not possible to be very precise using only the naked eye, as it is impossible to distinguish pure snow from snow mixed with ice, and no other means are available to help with this.

As far as the identification of the most suitable index for water separation is concerned, it turned out that none of the three indices were able to identify all the lakes; in the case of the NDWI and MNDWI, a lake to the north was left out, and in the case of MuWI-R, a lake to the south. Furthermore, as can be seen from the appendix, in the case of the MuWI-R index, an area around the snow cover is also classified as water, this could be due to the fact that there is melting snow at the base of the mountain. The same can be seen in the ESA classification, in the "quality_scene_classification" band of the image (Figure 4.3.1 in chapter 4.3).

In order to find a solution to this different selection of lakes, we tried to combine the results of NDWI and MuWI-R, and MNDWI and MuWI-R, so as to highlight all the lakes, and in order

to do this, the steps taken were two: a first sum of the two rasters, with which we will obtain a third that will be composed of three values: 0 for the rock, 1 for the areas classified as water by one of the two indices and 2 for the areas classified as water by both. The second step was then to substitute, with the Raster Calculator, the number 2 with the number 1, using the conditional expression and writing: Con("NDWI_MuWIR" == 2, 1, "NDWI_MuWIR").

In order to choose which case was the best, the five confusion matrices (for NDWI, MNDWI, MuWI-R, NDWI-MuWI-R union, MNDWI-MuWI-R union) were created, using as ground truth the one generated in the water ground truth generation sub-section of chapter 4.3 (Figure 4.3.7). The process of generating the confusion matrices will be explained in detail in Chapter 4.4 and the confusion matrices are shown in tables 4.2.4.2, 4.2.4.3, 4.2.4.4, 4.2.4.5, 4.2.4.6.

NDWI_confmatrix							
	OBJECTID *	ClassValue	C 0	C 1	Total	U Accuracy	Kappa
▶	1	C_0	9836	7	9843	0,999289	0
	2	C_1	84	56	140	0,4	0
	3	Total	9920	63	9983	0	0
	4	P_Accuracy	0,991532	0,888889	0	0,990885	0
	5	Kappa	0	0	0	0	0,547788

Table 4.2.4.2: NDWI confusion matrix

MNDWI_confmatrix							
	OBJECTID *	ClassValue	C 0	C 1	Total	U Accuracy	Kappa
▶	1	C_0	9788	10	9798	0,998979	0
	2	C_1	143	49	192	0,255208	0
	3	Total	9931	59	9990	0	0
	4	P_Accuracy	0,985601	0,830508	0	0,984685	0
	5	Kappa	0	0	0	0	0,38488

Table 4.2.4.3: MNDWI confusion matrix

MuWIR_confmatrix							
	OBJECTID *	ClassValue	C 0	C 1	Total	U Accuracy	Kappa
▶	1	C_0	9678	15	9693	0,998452	0
	2	C_1	245	43	288	0,149306	0
	3	Total	9923	58	9981	0	0
	4	P_Accuracy	0,97531	0,741379	0	0,973951	0
	5	Kappa	0	0	0	0	0,241215

Table 4.2.4.4: MuWI-R confusion matrix

NDWI_MuWIR_confmatrix							
	OBJECTID *	ClassValue	C 0	C 1	Total	U Accuracy	Kappa
▶	1	C_0	9638	3	9641	0,999689	0
	2	C_1	296	53	349	0,151862	0
	3	Total	9934	56	9990	0	0
	4	P_Accuracy	0,970203	0,946429	0	0,97007	0
	5	Kappa	0	0	0	0	0,254526

Table 4.2.4.5: NDWI + MuWI-R confusion matrix

MuWIR_MNDWI_confmatrix							
	OBJECTID *	ClassValue	C 0	C 1	Total	U Accuracy	Kappa
▶	1	C_0	9522	0	9522	1	0
	2	C_1	403	57	460	0,123913	0
	3	Total	9925	57	9982	0	0
	4	P_Accuracy	0,959395	1	0	0,959627	0
	5	Kappa	0	0	0	0	0,212501

Table 4.2.4.6: MNDWI + MuWI-R confusion matrix

In the case of the NDGI, instead of considering values greater than a certain limit, values within a certain range were selected, in order to exactly delimit the second peak, this way areas that were not snow and ice but had reflectance values located in the later part of the histogram were not considered.

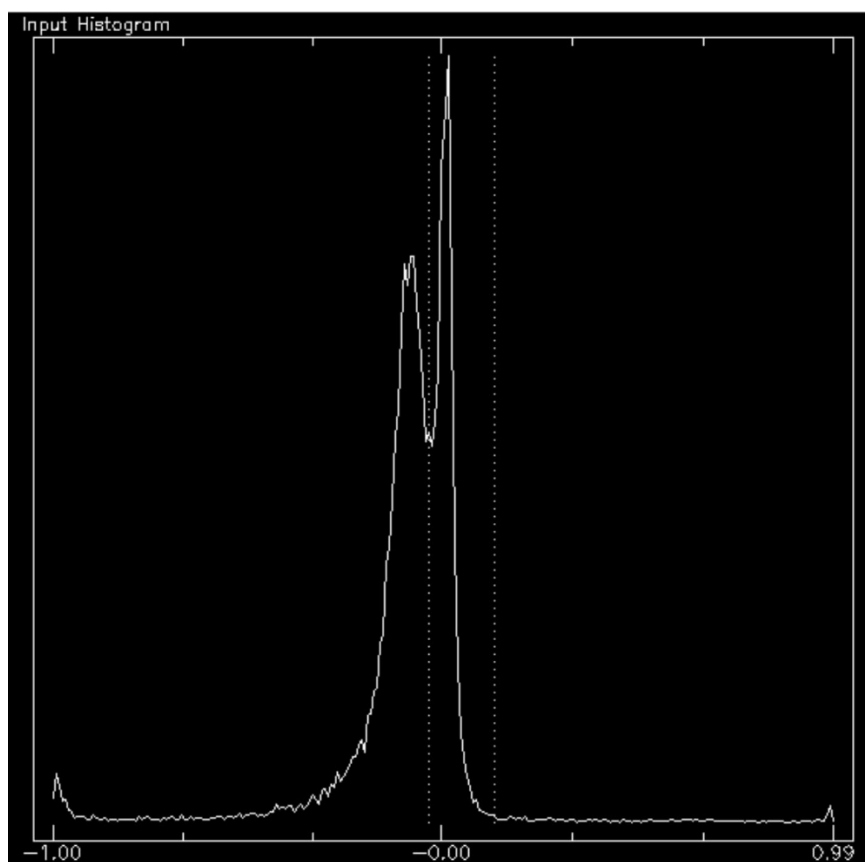


Figure 4.2.4.6: NDGI histogram. The interval considered is the one inside the two dotted lines

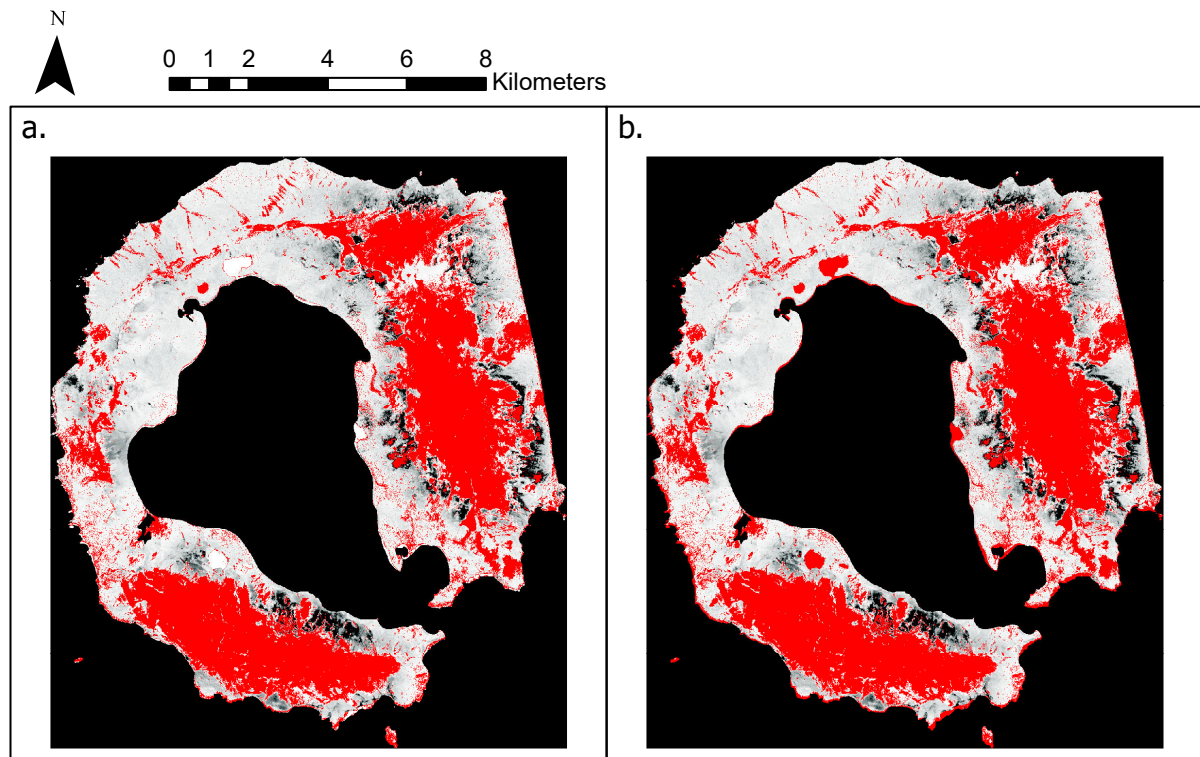


Figure 4.2.4.7: Difference between highlighted regions (red) if you consider only the indicated interval (a.) and those highlighted if we consider the whole part of the histogram to the right of the limit between the two peaks (b.)

For the identification of the clouds, the same steps were taken as for the other ground covers, i.e. to use an index (in this case the NDSI) and look for what range of values would separate the desired class, but then another step was added, in which the areas of the main clouds were drawn over, in particular the one located in the center of Mount Pond and the strip in the north-west part of Mount Kirkwood, in order to create a polygonal shapefile that could then be used as a mask in the generation of the confusion matrices of the optical images, since the parts covered by clouds are the most prone to misclassification, in order to highlight a possible improvement in the results by removing this critical part. In order to create the mask, the ArcGIS editor was used and a new empty polygonal shapefile was generated, then the main clouds that could be seen in the binary raster just created by imposing the limit identified with ENVI were manually drawn over. Figure 4.2.4.8 shows the mentioned raster and the cloud mask.

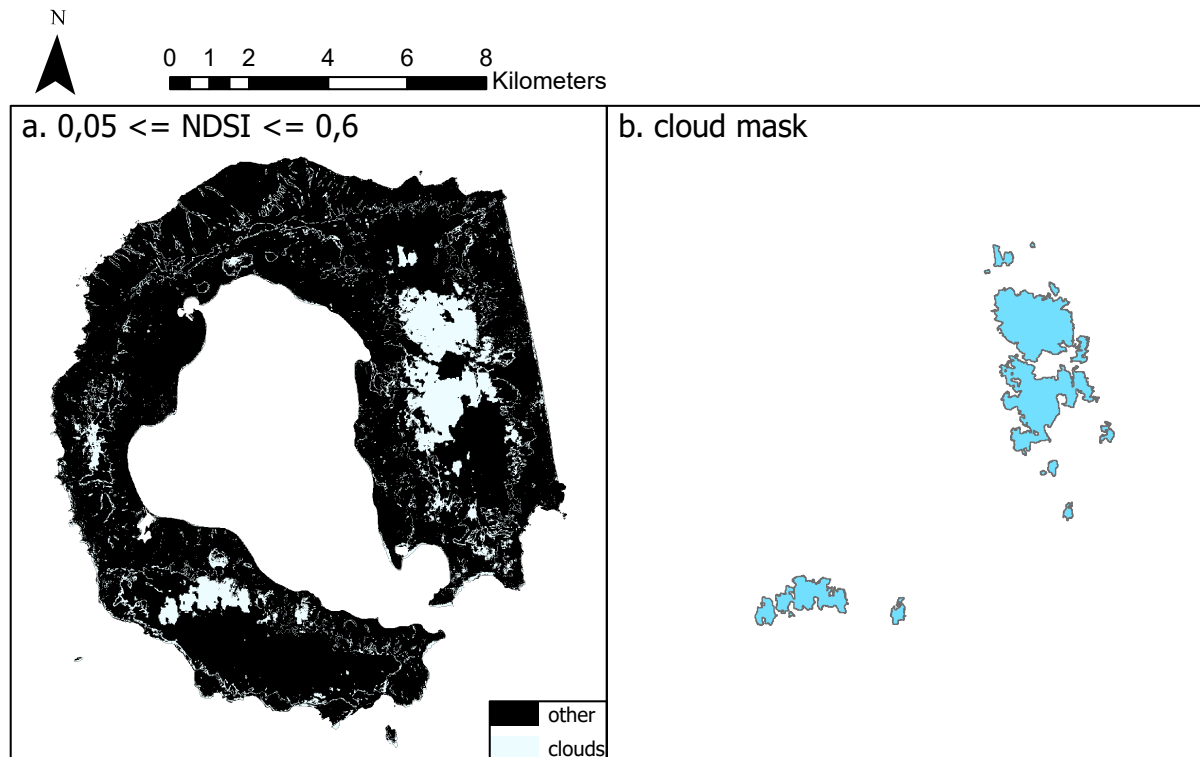


Figure 4.2.4.8: Binary image obtained from the NDSI, using the extremes of the indicated interval, and polygonal cloud mask obtained by drawing the clouds of the image on the left

In order to produce the final image, as far as the index for classifying water is concerned, it was found that NDWI was the best. In fact, by analyzing the confusion matrices shown in Figures 4.2.4.2, 4.2.4.3, 4.2.4.4., 4.2.4.5, 4.2.4.6, it can be seen that it is the index with the greatest accuracy in separating water, with a k of 0.547. In particular, it can be seen that in the two cases of combining two indices, i.e. MuWIR + MNDWI and MuWIR + NDWI, the k was significantly reduced compared to the use of single NDWI and MNDWI, contrary to what was expected. The reason for this is that the MuWIR also classified the perimeter part of the two snow covers as water, probably because it was in the process of melting, while the other indices, as well as ground truth, only classified the open water bodies as water. Unfortunately, not for all the indices present a ground truth, able to quantify their goodness, was found: in particular, no ground truth was found that was able to separate pure snow from pure ice, and as regards the separation of snow and ice mixed with debris from rock, these are usually considered in a single class, since the color of snow and ice mixed with debris is no longer white but brownish, and therefore similar to that of rock. The final images proposed are therefore two: one that presents the distinction between rock/ soil, snow/ ice and water, thus using the results provided by NDWI and NDSI, and a complete one, whose classes are: water, rock, snow and ice mixed with debris, pure ice, pure snow. This is because for the former it is possible to assess its goodness through the ground truth available, while for the latter it is

not, but it could still be a starting point for research, especially as regards ground truths capable of separating snow from ice and snow and ice mixed with debris from rock.

The two rasters to be merged to make the first image are both binary and one has values of 1 for water and 0 for the rest and the other has values of 1 for snow and 0 for the rest, so before summing them it is necessary to change these values, otherwise the water and snow classes would be merged. With the Raster Calculator, using the conditional expression given in formula 4.2.4.1, the value 1 was substituted with the value 2 in the NDSI binary raster, so that the final image summed from the two rasters will have the rock with a value of 0, the water with a value of 1 and the snow with a value of 2, analogously to the ground truth (see chapter 4.3).

`Con("NDSI_thr"==1, 2, "NDSI_thr")` (4.2.4.1)

Where:

- NDSI_thr is the binary raster obtained by applying the threshold identified with ENVI to the NDSI raster
- 1 is the value of snow, which will be replaced with 2
- 2 is the value that will be assigned to the snow

At this point, the raster obtained was summed to the binary raster for water; this resulted in a raster with 4 values instead of the expected 3, due to the classification of water as snow in the case of the NDSI index. Using an expression similar to 4.2.4.1, the value 3 was then set to 1, thus including it in the water class. Please refer to the Appendix for the final image obtained.

The image just described was the first (image 1 in the flow chart), whose goodness can be verified using the ground truth described in chapter 4.3, while for the second, the results of the NDGI and NDSII indices were added following a procedure similar to that just described, again using the Raster Calculator, and the substantial steps were the followings:

- the value 1 was substituted with the value 10 in the binary image of the NDGI;
- this last image was added to the image obtained by summing the NDWI and NDSI results described above. At this point, it was possible to note the overlapping zones of the areas selected by the NDGI, which took on values of 10, 11 or 12. This means, on the basis of the knowledge gained from the previously used indices, that the NDGI classified an area of rock as snow-ice in the first case, in the second an area of water cover, and the last is the correct case, in which pure parts of the snow-ice selected previously by the NDSI are classified as snow-ice;
- on the basis of the considerations made in the previous point, using the conditional function in the Raster Calculator, the value 10 was replaced with the value 0, 11 with 1 and 12 with 3, creating the new snow-ice class. The image resulting from this step (image 2 in the flow chart)

will therefore have value 0 for rock, 1 for water, 2 for snow and ice also mixed with debris, 3 for pure snow and ice;

-an analogous procedure was now followed with the NDSII, whereby the value 1 was substituted with the value 10 in the binary image relative to the index and this new image was summed to the one obtained from the previous step;

-the image obtained has, in addition to the values 0, 2 and 3, 10, 11, 12 and 13. This means that, in the case of 11, the NDSII classified as snow watery coverings, selected by the NDWI, in that of 12 it classified as pure snow what, according to the NDSI, was a mixture of snow, ice and debris, and in that of 10 the same thing happened but with rock. As for 13, this is the correct situation, where the index classified as snow cover part of what had been classified as pure snow and ice by the NDGI.

Based on these considerations, the number 10 was replaced with 0, 11 with 1, 12 with 2 and 13 with 4, thus forming the new class of pure snow. The final image obtained is therefore composed of the following classes: 0 rock, 1 water, 2 snow and ice mixed with debris, 3 pure snow and ice, 4 snow cover.

To consult the two final images obtained from this classification, as well as the binary images generated by applying the limits to the different indices, and their histograms, please refer to the Appendix.

4.3 PREPARATION OF THE GROUND TRUTH

With regard to the ground truth, it was not possible to use one of the classifications provided by the space agencies already mentioned because, considering the classification provided by ESA for the image of 19.01.2020, it was not usable as it was incomplete, and the same applies to those of the two nearest dates not affected by cloud cover (20.12.2019 and 08.02.2020). Speaking of the classification provided by the USGS on 09.02.2020 (the closest date usable and without cloud cover), the bodies of water were not classified, even though they were present and not covered or disappeared for other reasons (in fact, they are not always present, but in this case it was possible to identify them in both the 19.01 and 09.02 images), and the rock was incorrectly classified as water. The solution conceived was therefore to combine two different sources: the Pixel Quality Assessment band of the 09.02.2020 Landsat 8 image for the snow cover, a dgn of the 2005 topographic map for deriving the lakes, and the shapefile of the coastline, already applied to clip the analyzed images as explained in chapter 4, to be used as a base, considering that the rock/ soil will be by exclusion all the remaining part of the island not covered by snow or water bodies.

For more precise information about the details of the Landsat image and the cartography which was the source of the dgn and the coastline, please refer to chapter 3.

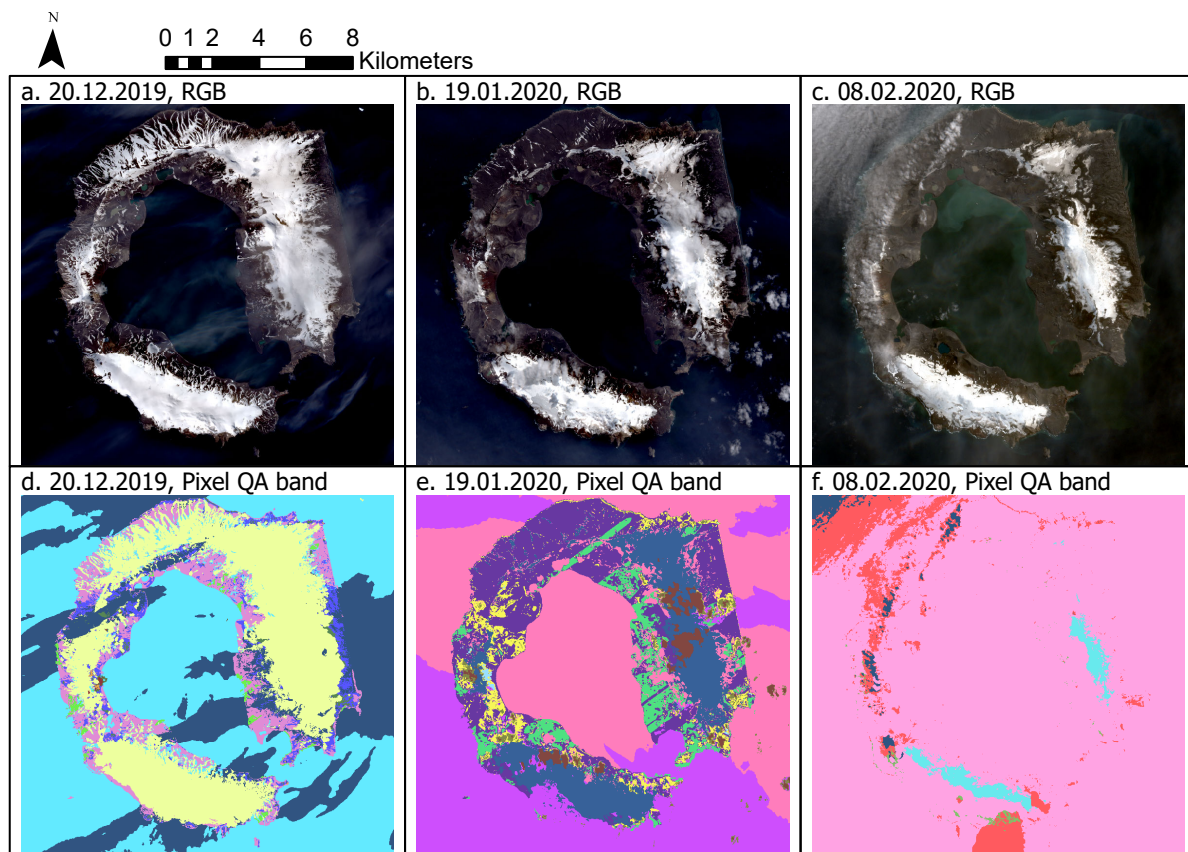


Figure 4.3.1: RGB bands combination of the images acquired on the dates 20.12.2019 (a.), 19.01.2020 (b.) and 08.02.2020 (c.) and ESA-supplied classifications of the area of interest on the same dates

Preparation of the ground truth of snow

First of all, it is necessary to point out that it should be kept in mind that the snow cover is very variable and, if you do not have an image of the same day at the same time, it will be different. In particular, analyzing the weather data acquired during the Antarctic campaign of the Spanish Land Army (Campaña Antártica del Ejército de Tierra, 2019-2020) already reported in sub-chapter 2.4, it emerges that in the period 19.01.2020-09.02.2020 the temperature was always above 1 °C, temperature occurring only four times, while the vast majority of days the minimum temperature was 2 °C, with sporadic maximums of 5 °C (five days) and lows of 3 °C (three days) on the hottest days. In addition, there was no further snowfall throughout the period, but the days were roughly divided into half rainy and half sunny. This led to the melting of some of the snow present, so that the image of 09.02 shows a more limited snow cover than that of 19.01, as can be seen in Figure 4.3.3.

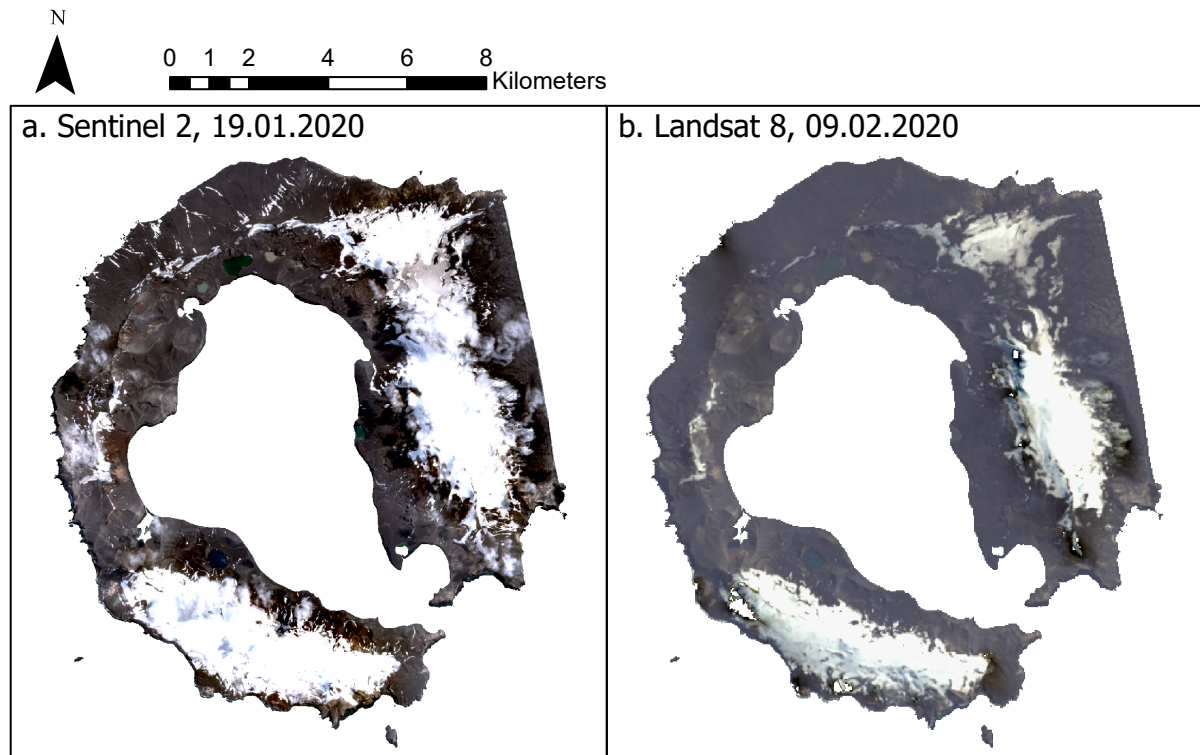


Figure 4.3.3: comparison between the Sentinel 2 image of 19.01.2020 (a.) and the Landsat 8 image of 09.02.2020 (b.), the difference in snow cover can be seen

For the preparation of the ground truth of the snow, first of all the Landsat image was opened in ArcGIS: unlike the Sentinel images that required a first conversion of the bands in GeoTIFF format using the SNAP software, this one was already provided in the format just mentioned. It was then cropped using the shapefile of the coastline and bands 1-7 (those contained in the folder minus the Quality Assessment bands) were merged into a stack, using the commands already mentioned in Chapter 4 for the Sentinel 2 image.

The band of interest, that of Pixel Quality Assessment, is presented in a different way from the classifications provided with the Sentinel 2 images: the classes here are in fact not explained with a descriptive field in the attribute table, but instead only 16-bit numbers are shown. In particular, each bit or pair of bits is associated with a land cover class, and the value assigned to it depends on the level of confidence with which the land cover analyzed can be said to correspond to that land cover class (Figure 4.3.4).

Bit	Flag Description	Values
0	Fill	0 for image data 1 for fill data
1	Dilated Cloud	0 for cloud is not dilated or no cloud 1 for cloud dilation
2	Cirrus	0 for Cirrus Confidence: no confidence level set or Low Confidence 1 for high confidence cirrus
3	Cloud	0 for cloud confidence is not high 1 for high confidence cloud
4	Cloud Shadow	0 for Cloud Shadow Confidence is not high 1 for high confidence cloud shadow
5	Snow	0 for Snow/Ice Confidence is not high 1 for high confidence snow cover
6	Clear	0 if Cloud or Dilated Cloud bits are set 1 if Cloud and Dilated Cloud bits are not set
7	Water	0 for land or cloud 1 for water
8-9	Cloud Confidence	00 for no confidence level set 01 Low confidence 10 Medium confidence 11 High confidence
10-11	Cloud Shadow Confidence	00 for no confidence level set 01 Low confidence 10 Reserved 11 High confidence
12-13	Snow/Ice Confidence	00 for no confidence level set 01 Low confidence 10 Reserved 11 High confidence
14-15	Cirrus Confidence	00 for no confidence level set 01 Low confidence 10 Reserved 11 High confidence

Figure 4.3.4: bit values of the Landsat 8 Pixel QA band (source: USGS)

To decode these values it was necessary to use a Toolbox compatible with ArcGIS and downloadable from the USGS website ([25]). It comprised two tools: the "Decode QA" and the "Extract QA Bands", which work in the following way: the first adds a descriptive field to the attribute table of the Pixel QA band and it fills in interpreting the bits of the band (it is optionally possible to choose not to select the bits with low confidence, so as to include in the descriptive field only the classes that are most likely to represent the ground covers), while the second extracts the various classes of the band individually and forms a new image for each class. The tool used in this case was Decode QA, so it was sufficient to insert as input the Pixel QA band of the image (original, unclipped); the result is shown in Figure 4.3.5.

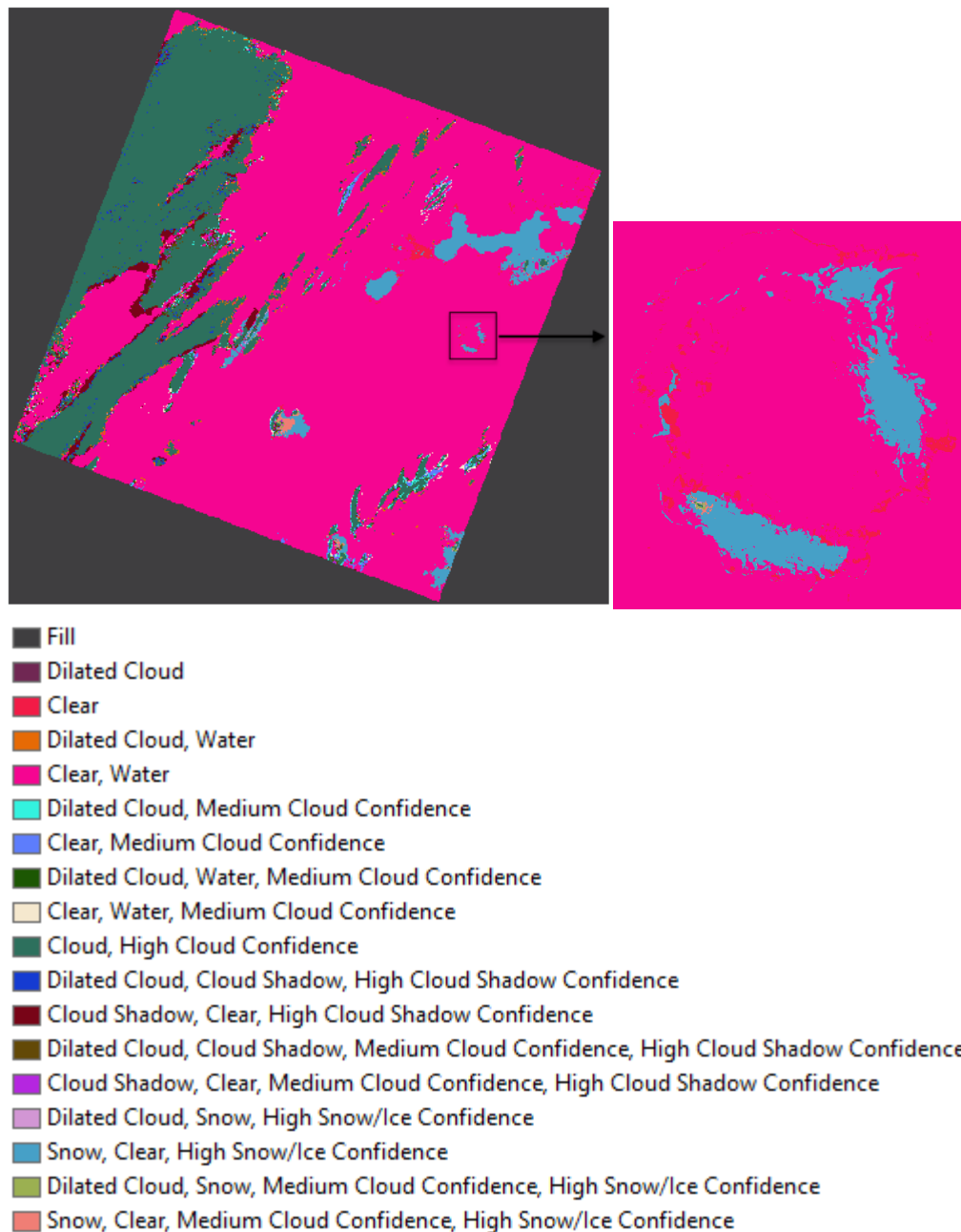


Figure 4.3.5 Descriptions added by the Decode QA command

The Pixel QA band does not present a completely good classification, the soil is classified as water and there are also anomalous areas, in the north-west part of the snow cover on the Mount Kirkwood glacier and in a small area of the snow cover above the Mount Pond glacier, which, using the visible bands, we can see that are detected as No Data, while in the automatic classification they are associated to different covers without a true match. However, the snow cover is detected well and will be extracted and then combined with the lakes, obtained from another source, as explained later, to form the final ground truth.

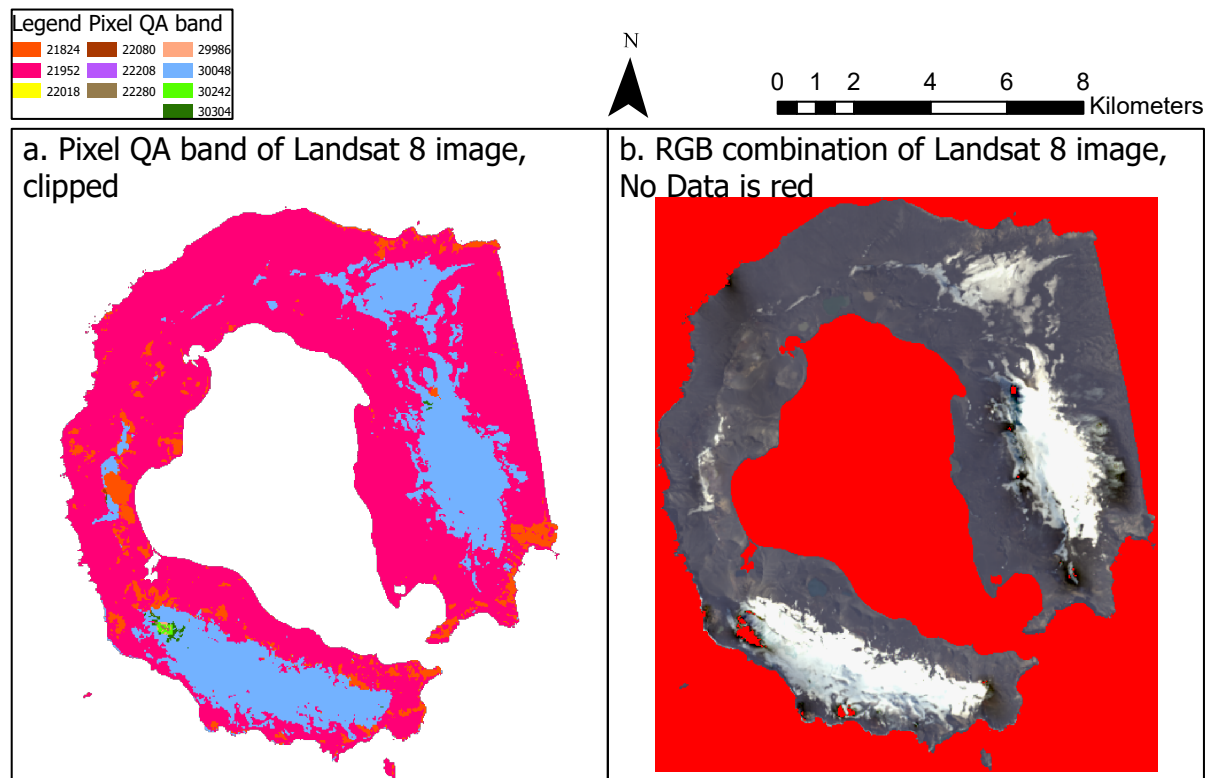


Figure 4.3.6: Pixel Quality Assessment Band of the Landsat 8 image (a.) and true color display of the image (b.), where the No Data has been colored red

Since it is evident that in both cases the anomalous zones belong to the snow cover, to generate the ground truth of the snow, the large yellow zone with a value of 30048 was considered, incorporating the small anomalous zones, using the Raster Calculator command of the Spatial Analyst Tools package of ArcGIS Pro. Specifically, the conditional operator was used, imposing the following expression:

`Con("Clip_LC08_QA_PIXEL", 1, 0, "Value >= 22018")`

Where:

- Clip_LC08_QA_PIXEL is the Pixel Quality Assessment band of the image, clipped with the shapefile of the island's coastline
- 1 is the value to assign to pixels for which the condition is verified
- 0 is the value to assign to pixels for which the condition is not verified
- Value >= 22018 is the condition set

The condition Value >= 22018 was imposed, with Value being the field of the band indicating the various values assigned to the different ground covers, since, as can be seen from the legend in Figure 4.3.6, all the values of the anomalous zones and the snow value were greater than or equal to this number, while the rock and the clouds had the two lowest assigned numbers. It was therefore easy to group the snow cover in this way. The raster obtained will therefore take on values of 1 where snow is present and 0 where rock is present and it is shown in Figure 4.3.7.

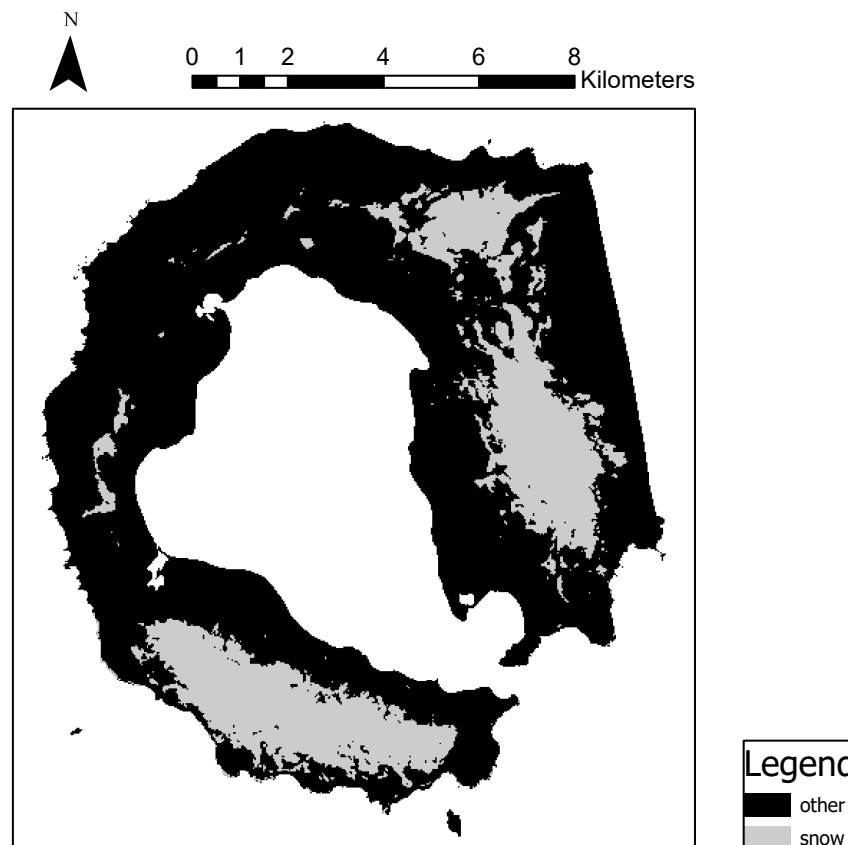


Figure 4.3.7: ground truth of the snow

Preparation of the ground truth of water

The dgn used to derive the lakes present on the island is composed of various layers, mainly relating to the delineation of glaciers and lakes present, as well as areas subject to temporary flooding, watercourses, toponymy and contour lines and geodetic stations. Of these, the polygonal layer that included the lakes was considered (Figure 4.3.8) and modified with the ArcGIS editor, selecting and deleting anything that was not of interest, i.e. anything that was not the lakes. At this point, what was obtained was a shapefile consisting only of the lakes, which therefore needed to have the shape of the island as a base. For this reason, this shapefile and the coastline shapefile were merged with the Union command of ArcGIS Geoprocessing and then the resulting shapefile was converted into a Raster using the Polygon to Raster command, using the cell center as the assignment type and a cell size of 10x10, being that of the Sentinel 2 bands after resampling.

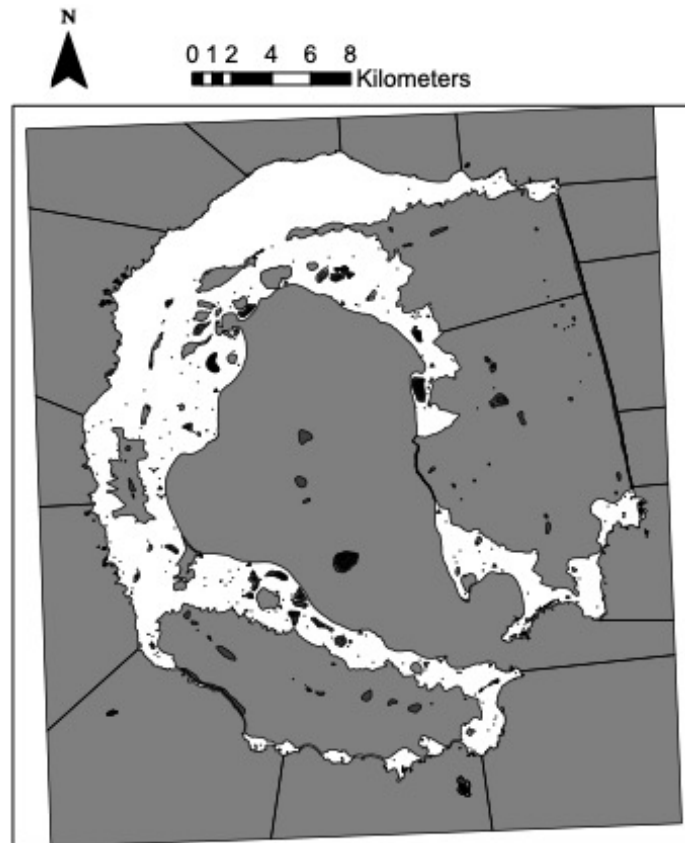


Figure 4.3.8: dgn from which the lakes were extracted

At this point, in order to evaluate the goodness of the indices used to classify the water in chapter 4.2.4 and thus choose which one was better, it was necessary to have a ground truth that had the same values for the corresponding classes; we therefore set 1 as the value for the water pixels and 0 for the rest, using the Raster Calculator conditional, in the same way as we did for the indices and explained in chapter 4.2.4.

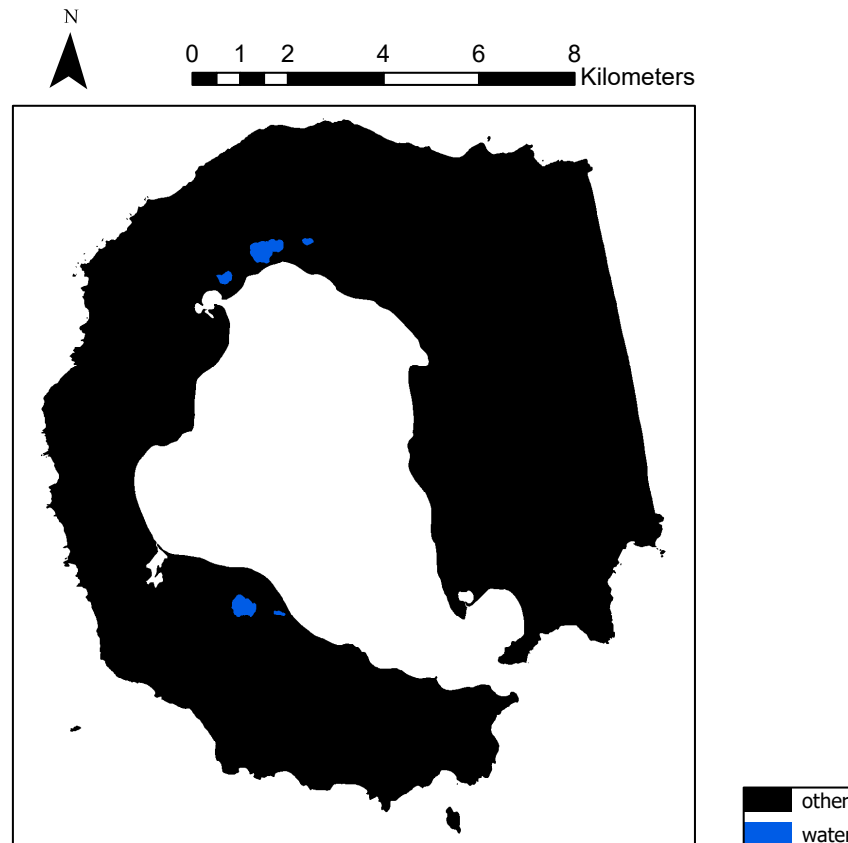
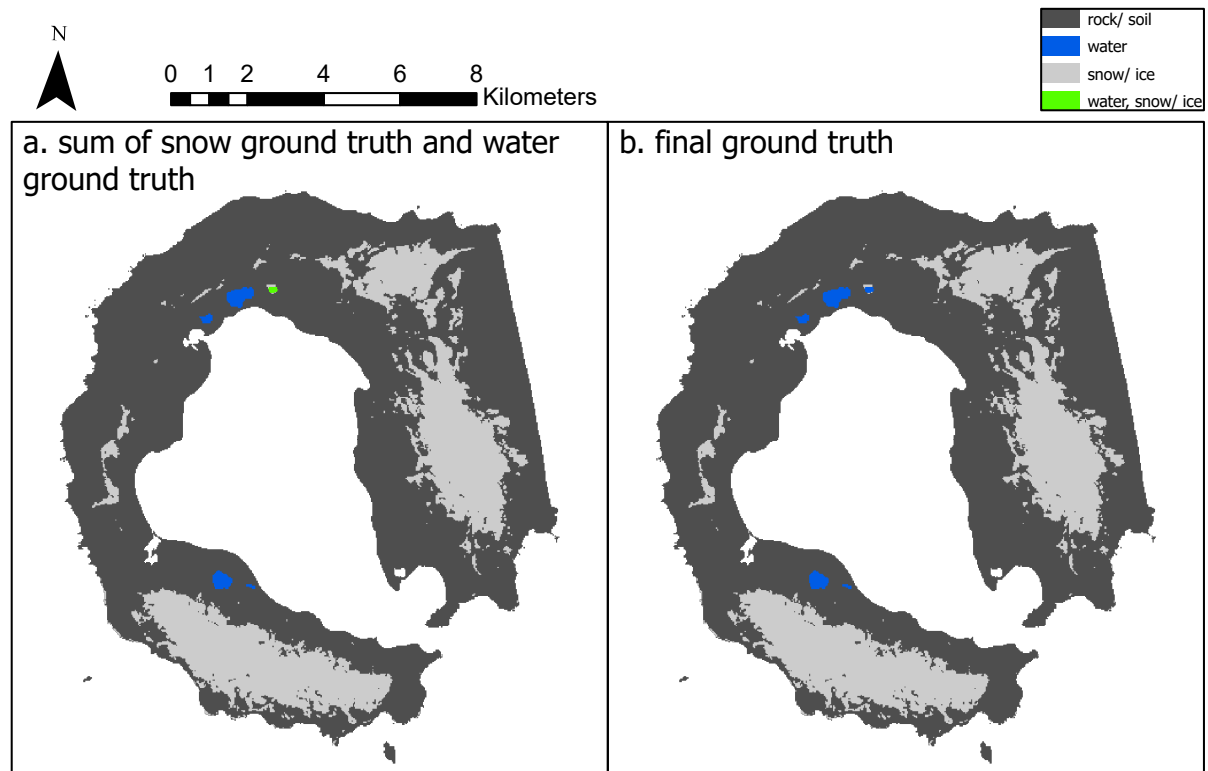


Figure 4.3.9: ground truth of water

Final ground truth

To obtain the global ground truth of the island, the two rasters resulting from the two processes were merged, as follows. Since both rasters are composed of 1 and 0, summing them would have meant merging the water class with the snow class (since they do not overlap and are both equal to 1); therefore, the value 2 was substituted for snow in the first raster, again using the conditional expression with the Raster Calculator, as in chapter 4.2.4, and then summed with the water ground truth, again using the Raster Calculator. Analyzing the summed raster, it was noted that one of the lakes to the north had been classified as snow in the Landsat image, so its value was not 0 in the raster I generated but 2, and adding it to the ground truth of the water gave it a value of 3 (Figure 4.3.10). This is evidently an error on the part of the USGS classification, since consulting the 2005 topographic map and analyzing the image in real colors it is evident that it is a lake. Therefore, using the Raster Calculator, with the same conditional function used in chapter 4.2.4, the value 3 was replaced by the value 1, thus incorporating that lake into the water class. The resulting image is the final ground truth and is shown in Figure 4.3.11.



Figures 4.3.10: sum of the two ground truths (a.), the error of the lake in the north (green) can be seen, and final ground truth (b.)

4.4 REALIZATION OF THE CONFUSION MATRICES

The tools used to measure the goodness of the classifications are the confusion matrices. In order to create them in ArcGIS it is necessary to carry out a number of steps, which will now be described. First of all, in order to compare each land cover of a classified image with its corresponding one belonging to the ground truth, it is necessary for these to have the same values; the first thing to be done was therefore to change all the values of the classified raster, random and assigned by the program during the classification phase, and replace them with the same values of the ground truth, that are: 0 for rock, 1 for water and 2 for snow-ice; to do this the Raster Calculator of the Spatial Analyst package was used once again, with the conditional expression, using a formula similar to 4.2.4.1, in several subsequent steps. At this point, it is necessary to generate ground control points, in which the values of the two images (classified image and ground truth) will be compared. There's a proper command to do this, the Create Accuracy Assessment Points of the Spatial Analyst Tools of ArcGIS (Figure 4.4.1), in which you put as input the classified image, then, from a drop-down list you select that it is the classified image that you are inserting and not the ground truth, you set the

number of points that you want to generate (10000 to cover an area equal to 1% of the island) and the way to distribute them. There are three possible distributions:

- Stratified Random, whereby the distribution of points generated is proportional to the area of each class;
- Equalized Stratified Random, whereby the points are distributed equally among the classes;
- Random, whereby the points are distributed randomly across the image area.

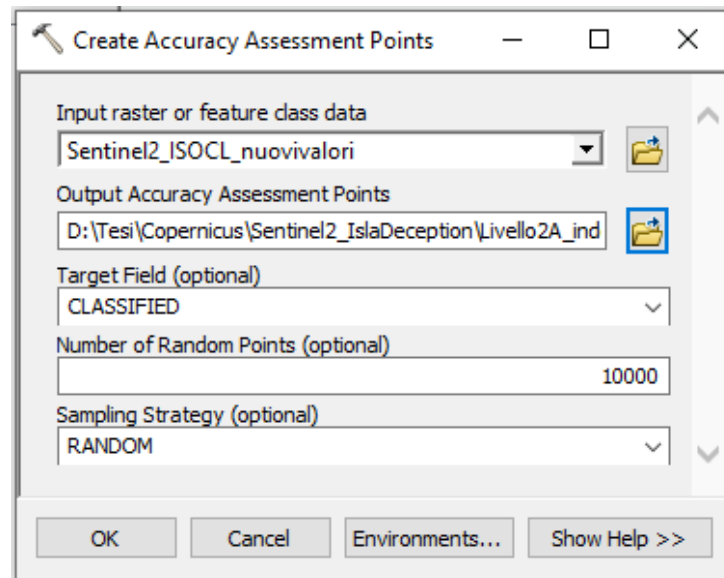


Figure 4.4.1 Create Accuracy Assessment Points

The third option was chosen, because the areas of the different classes were very different from each other (for example, the area of the water class was much smaller than those of rock and snow-ice), so the option Equalized Stratified Random did not make sense, and at the same time we did not want to create concentrations of points in some areas rather than others, as would happen with the option Stratified Random, the Random method seemed therefore the most suitable. The output is a point shapefile of 10000 points in whose attribute table there are two columns: one relative to the classified image (called Classified) and one relative to the ground truth (called GrndTruth). Having set as input that the image was the classified one, the fields filled will be those of the Classified column, while the other will be filled by "-1"; if instead it had been set that the input image was the ground truth, the opposite would have happened. Figure 4.4.2 shows an example of the attribute table of the shapefile generated by the command.

OBJECTID *	SHAPE *	Classified	GrndTruth
1	Point	0	-1
2	Point	0	-1
3	Point	0	-1
4	Point	0	-1
5	Point	0	-1
6	Point	0	-1
7	Point	0	-1
8	Point	0	-1
9	Point	0	-1
10	Point	0	-1

Figure 4.4.2 Example of the result of the Create Assessment Points command

The second command to use is Update Accuracy Assessment Points (Figure 4.4.3), from the same package, in which the inputs to be inserted are instead the ground truth raster and the shapefile of the points just generated; then you select from a drop-down menu equal to the first one that in this case it is the ground truth that is being inserted and not the classified image, and the result will be that the shapefile described above will now have the -1 mentioned replaced by the values assumed by the ground truth in those points.

Figure 4.4.3: Update Accuracy Assessment Points

The last command to use is Compute Confusion Matrix (Figure 4.4.4), which, starting from the point shapefile and comparing the two fields Classified and GrndTruth, generates the confusion matrix.

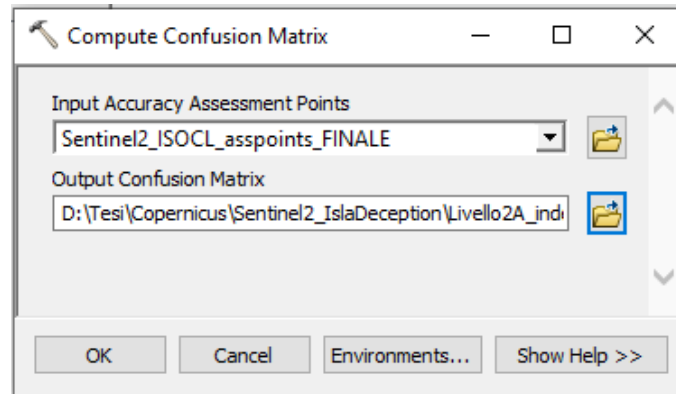


Figure 4.4.4: Compute Confusion Matrix

For the supervised and unsupervised classifications of the optical image, this process was repeated twice: the first time following what has just been described, and the second time using shapefiles clipped with the cloud mask generated in chapter 4.2.4, in order to check whether by eliminating the most of the part covered by clouds, which is what normally leads to error, the goodness of the classification was better. In particular, in order to use this shapefile as a mask, it was merged with that of the coastline using the Union command of ArcGIS Geoprocessing, shown in Figure 4.4.5, which requires the shapefiles you want to merge as input and also allows you to optionally select which attributes you want to be transferred to the merge shapefile (in this case it was imposed to consider all of them, otherwise it would have been possible to exclude the FID, the numbering field, or to consider only the FID, neglecting all the other attributes), the minimum distance between the coordinates ("XY Tolerance", it has not been set) and if you want to eliminate any gaps which can be present with filling polygons, by not checking the box "Gaps Allowed" (in this case it has been set to allow the gaps).

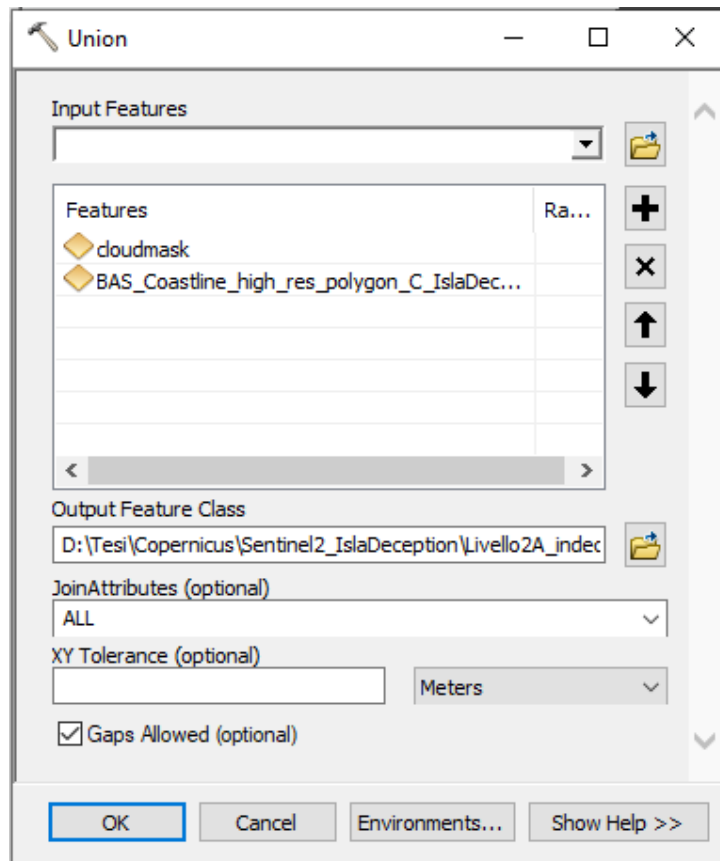


Figure 4.4.5: Union

This way a single shapefile was obtained, which was composed of different polygons, representing the clouds and the complementary part of the island not covered by the clouds. It was then edited with the Editor, in order to eliminate the polygons of the clouds, and thus obtaining a shapefile of the surface of the island with holes in correspondence of the clouds, shown in Figure 4.4.6.

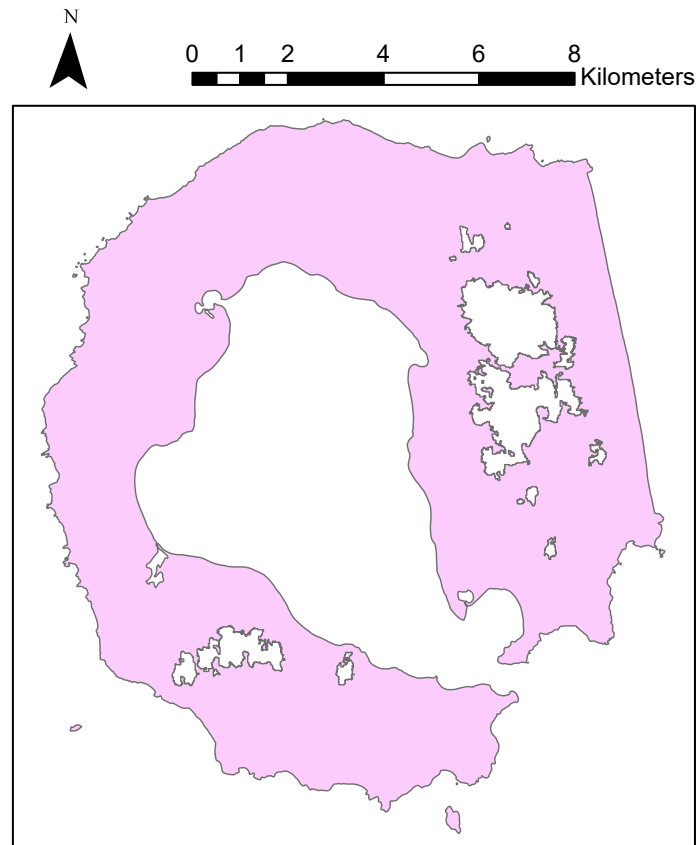


Figure 4.4.6: Shapefile of the island with holes in correspondence of the clouds

Using this shapefile, the classifications obtained from the optical image and ground truth were clipped and the operations to create the control points and confusion matrices were carried out again, following the procedure described above; using this trick, no control points were generated in the areas affected by cloud cover, and it is assumed that the confusion matrices will therefore report better k.

5. RESULTS AND DISCUSSION

The results obtained from the different classifications operated, i.e. the classified images and the confusion matrices, both for the SAR image and for the optical one, are illustrated below. In particular, because in the case of the SAR image several classifications were produced with different combinations of segment attributes, the visually best one was chosen for each classifier in the case of the image with the Median filter, and the confusion matrix was created; instead, for the image with the application of the Lee Sigma filter, the effect of the classification produced using the same segment attributes was then reported.

In the confusion matrices shown, the columns relate to the ground truth classes, while the rows relate to the classified image, and the classes will have the following values: 0 rock/ soil, 1 water, 2 snow/ ice, 4 clouds in the case of the optical image classifications and illumination in the case of the SAR image (the latter is not a type of ground cover, but was left as a third error class in cases where reducing the number of classes to two would lead to a deterioration in the quality of the classified image, badly incorporating the effect of illumination). The column related to class 4 for ground truth will always report only 0, as there were no clouds or effects due to illumination in it, since it was derived from a topographic map and an optical image, where this disturbance is not present.

The U_Accuracy column relates to User's Accuracy, while the P_Accuracy column relates to Producer's Accuracy (see chapter 1.2); on the basis of these measurements and k , the goodness of the classification will be quantified.

As already mentioned, the original confusion matrices and those generated after the application of the cloud mask will be available for each classification operated with the optical image.

To see all of the other images which were produced, please refer to the appendix.

5.1 SAR

5.1.1 UNSUPERVISED CLASSIFICATION

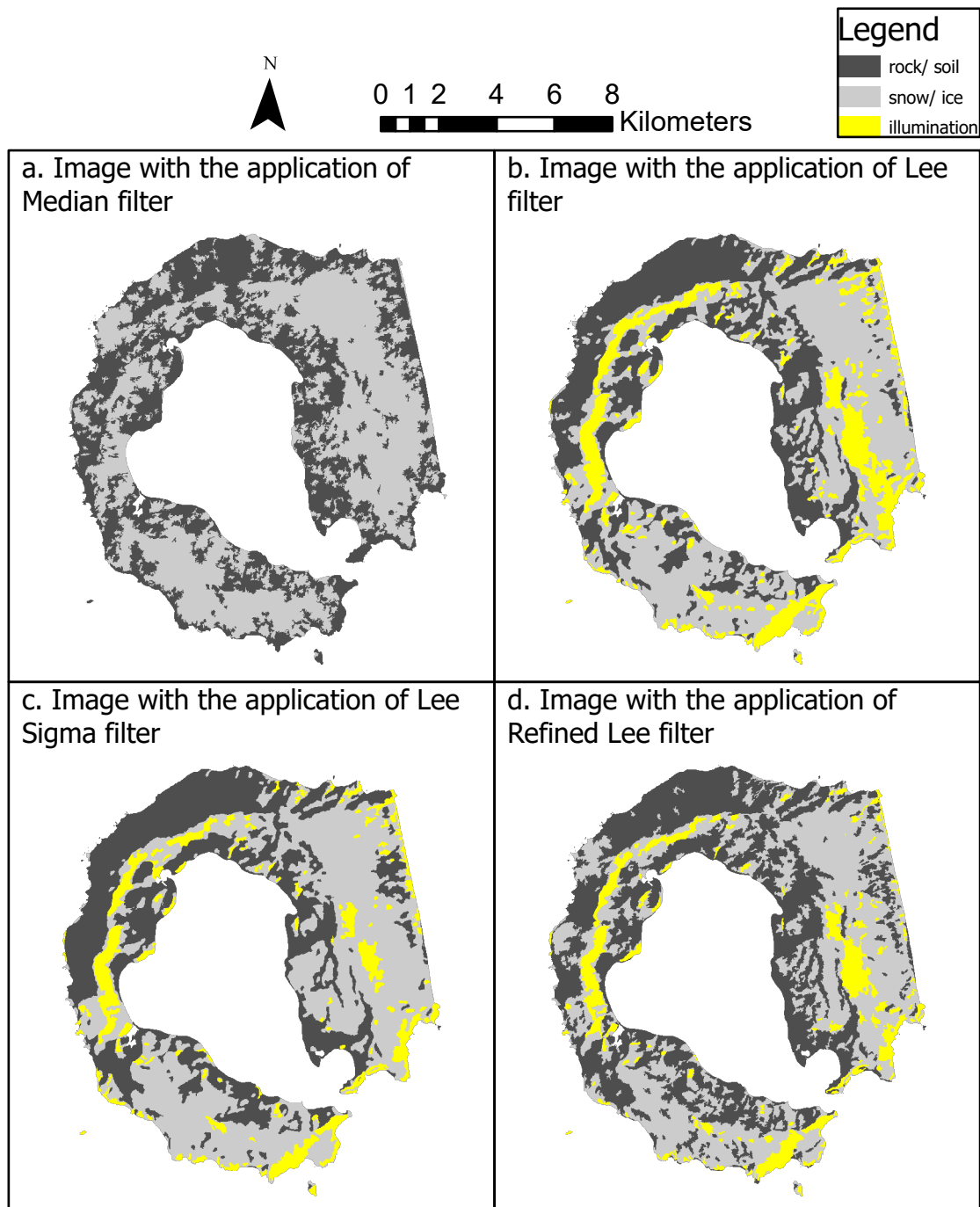


Figure 5.1.1.1 Images resulting from ISO cluster classifications

Sentinel1_med_ISOCL_confmatrix								
OBJECTID *	ClassValue	C 0	C 1	C 2	Total	U Accuracy	Kappa	
1	C_0	3975	48	344	4367	0,910236	0	
2	C_1	0	0	0	0	0	0	
3	C_2	3612	11	1984	5607	0,353843	0	
4	Total	7587	59	2328	9974	0	0	
5	P_Accuracy	0,523922	0	0,852234	0	0,597453	0	
6	Kappa	0	0	0	0	0	0,248607	

Table 5.1.1.1: Confusion matrix of the image with Median filter classified with ISO cluster classifier

Sentinel1_Lee_ISOCL_confmatrix								
OBJECTID *	ClassValue	C 0	C 1	C 2	C 4	Total	U Accuracy	Kappa
1	C_0	3308	45	372	0	3725	0,888054	0
2	C_1	0	0	0	0	0	0	0
3	C_2	3255	5	1396	0	4656	0,299828	0
4	C_4	1109	2	482	0	1593	0	0
5	Total	7672	52	2250	0	9974	0	0
6	P_Accuracy	0,431178	0	0,620444	0	0	0,471626	0
7	Kappa	0	0	0	0	0	0	0,130133

Table 5.1.1.2: Confusion matrix of the image with Lee filter classified with ISO cluster classifier

Sentinel1_LeeSig_ISOCL_confmatrix								
OBJECTID *	ClassValue	C 0	C 1	C 2	C 4	Total	U Accuracy	Kappa
1	C_0	3504	64	313	0	3881	0,90286	0
2	C_1	0	0	0	0	0	0	0
3	C_2	3187	0	1745	0	4932	0,353812	0
4	C_4	909	1	242	0	1152	0	0
5	Total	7600	65	2300	0	9965	0	0
6	P_Accuracy	0,461053	0	0,758696	0	0	0,526744	0
7	Kappa	0	0	0	0	0	0	0,196146

Table 5.1.1.3: Confusion matrix of the image with Lee Sigma filter classified with ISO cluster classifier

Sentinel1_RefLee_ISOCL_confmatrix								
OBJECTID *	ClassValue	C 0	C 1	C 2	C 4	Total	U Accuracy	Kappa
1	C_0	3783	45	481	0	4309	0,87793	0
2	C_1	0	0	0	0	0	0	0
3	C_2	3234	3	1388	0	4625	0,300108	0
4	C_4	676	0	352	0	1028	0	0
5	Total	7693	48	2221	0	9962	0	0
6	P_Accuracy	0,491746	0	0,624944	0	0	0,519072	0
7	Kappa	0	0	0	0	0	0	0,14497

Table 5.1.1.4: Confusion matrix of the image with Refined Lee filter classified with ISO cluster classifier

In the unsupervised classification of the SAR image, water was never detected: when the classifier was forced to consider three classes instead of two, the third class did not result from the distinction of water, but from the illumination of the satellite, which is why the row associated to the class water (1) is full of 0. Column 4 is always empty because there is no

illumination class in the ground truth. In order to know which segment attributes were considered in each case, please refer to chapter 4.1.3.

The classification of the image to which the Median filter had been applied produced an image in which there was a recognition of glaciers, even if approximate, but the rock areas were not "clean", but affected by several areas erroneously classified as snow-ice. In the case of the three images to which the Lee, Lee Sigma and Refined Lee filters had been applied, these localized errors decreased, but the effect of illumination, represented by the yellow class, was strong.

Of the three, the image in which the point errors were most reduced was Lee Sigma, which in fact had the highest k , while the image in which their influence was still most present was the one with the Refined Lee filter applied. However, the lowest k was reported by the image with the Lee filter applied, although it was only slightly different from that of Refined Lee (1.5%). The k -values were always very low, for all images, but the highest occurred in the image with only the Median filter applied. This seems to be due to the fact that in this image the Producer's Accuracy of the snow-ice class is substantially higher than in all the other images (23% higher than the images with Lee filter and Refined Lee and 10% higher than the image with Lee Sigma), while the other Accuracies do not vary much, there's a maximum difference of 5% for the User's Accuracy of snow and of 3.3% in that of rock. There is more variability in the case of the Producer's Accuracy of the rock: the image with Median filter has an accuracy 9% higher than the Lee filter image, 6% higher than the Lee Sigma filter image and 3% higher than the Refined Lee filter image.

What essentially happened is that the multiple application of the Lee, Lee Sigma and Refined Lee filters reduced the speckle effect, and consequently the localized misclassification errors of the rock/ soil, but also resulted in a loss of accuracy of the delineation of the snow-ice cover. The contrasts that were maintained and highlighted after the application of the filters were not those between the different ground covers, but those of morphology and illumination, with a more or less marked effect depending on the filter; this led to a reduction in the goodness of classification. The non-detection of water, on the other hand, is supposed to be due to the small size of this cover compared to the others, so that when a third class was considered, the (wrong) class attributed to the effect of illumination was larger than that of water, which was present in small, localized areas as a kind of noise.

The reason for a low Producer's Accuracy for rock is that several parts of it were placed in the water class instead, its User's Accuracy is high because almost always what was classified as rock was actually rock. As far as the snow-ice class is concerned, its User's Accuracy is always very low and this is because it often contains points of rock, while its Producer's Accuracy is better, although not very high, because several parts of this cover were able to be identified, although not with precision. The estimation was not good for either class but what can be said is that the rock was underestimated while the snow-ice was overestimated. For example, the eastern part of the snow cover of the Pond and Goddard mountains almost always extends to the coast, whereas in reality it does not; what does extend to the coast is the glacier below. In fact, comparing Figure 5.1.1 with the

topographic map (Figure 3.10) it appears that the pattern of what was classified as snow-ice in the various unsupervised classifications of the SAR image is similar to that of the glaciers shown on the map rather than to the snow cover present at the time the image was taken. What may have happened is that the glaciers were recognized and not the debris or snow cover on them.

5.1.2 SUPERVISED CLASSIFICATION

MAXIMUM LIKELIHOOD

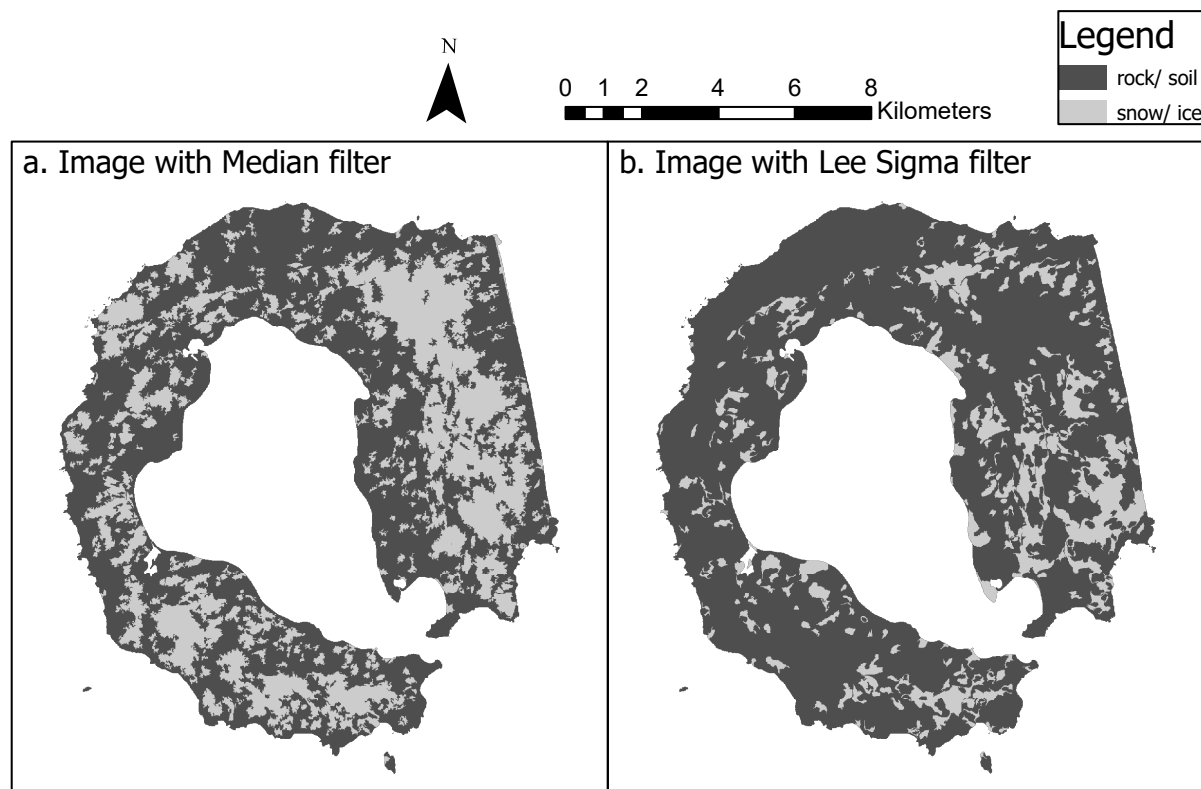


Figure 5.1.2.1 Images with Median (a.) and Lee Sigma (b.) filters, classified with Maximum Likelihood classifier

Sentinel1_Med_ML_2class_color_stddev_compact_confmatrix								
	OBJECTID *	ClassValue	C 0	C 1	C 2	Total	U Accuracy	Kappa
▶	1	C_0	5433	41	842	6316	0,860196	0
	2	C_1	0	0	0	0	0	0
	3	C_2	2270	15	1372	3657	0,375171	0
	4	Total	7703	56	2214	9973	0	0
	5	P_Accuracy	0,70531	0	0,619693	0	0,682342	0
	6	Kappa	0	0	0	0	0	0,26029

Table 5.1.2.1 Confusion matrix of the image with Median filter classified with Maximum Likelihood classifier

Sentinel1_LeeSig_ML_2class_color_stddev_compact_confmatrix								
	OBJECTID *	ClassValue	C 0	C 1	C 2	Total	U Accuracy	Kappa
▶	1	C_0	6389	42	1772	8203	0,778861	0
	2	C_1	0	0	0	0	0	0
	3	C_2	1174	15	581	1770	0,328249	0
	4	Total	7563	57	2353	9973	0	0
	5	P_Accuracy	0,844771	0	0,246919	0	0,698887	0
	6	Kappa	0	0	0	0	0	0,09946

Table 5.1.2.2 Confusion matrix of the image with Lee Sigma filter classified with Maximum Likelihood classifier

The image chosen as the best in the case of this classifier was the one resulting from the consideration of the segment attributes Active Chromaticity Color, Standard Deviation and Compactness and from the use of the rock/ soil and snow/ ice training samples, leaving out the water ones. The reason for the latter is that the addition of the water samples did not lead to the recognition of the lakes, instead it caused a major error in the classification (see appendix). In the original image, the lakes are actually distinguishable because they are darker than the surrounding rock, so the fact that in the Maximum Likelihood classification these training samples produced such a large error may mean that the problem lies in the classifier's grouping algorithm, which is not suitable for classifying this image: in fact, using Random Trees and the Support Vector Machine, it will be seen that water produced minimal errors of a completely different magnitude from these ones.

Classification using the Maximum Likelihood classifier produced slightly better results in terms of k for the image with Median filter, but much worse in the case of the one with Lee Sigma filter. In the first case, the extension of the rock areas erroneously classified as snow-ice has been reduced, but not their spatial distribution, as they continue to cover the whole island; in the second case, instead, there is no more recognition of glaciers nor the influence of illumination, the image produced is inconsistent with the one analyzed, the only positive note seems to be the concentration of the areas classified as snow-ice in the part of the island affected by Mount Pond. The User's and Producer's Accuracy of class 2 in the case of the image with Lee Sigma filter reflect what has just been said: both are very low because it was not possible to identify the snow cover present, and the points belonging to the snow class are not always the real ones: several areas of rock have been classified as snow. As a consequence of the small extent of snow there is the large extent of rock, which led to higher User's and Producer's Accuracy for this class. As for the Median filtered image, the Producer's Accuracy is much higher than that obtained with the unsupervised classification, the reason for this is that in this case, as mentioned above, the rock areas misclassified as snow-ice have reduced in areal extent, although not in spatial distribution, so the number of points belonging to the rock class has increased. What has instead decreased is the Producer's Accuracy of the snow and the reason of this is that the areas classified as snow have been greatly reduced and are affected by a pronounced point noise: if we compare this image with the one classified with ISO cluster (with the same filter), we can see that the latter presents much more homogeneous glaciers. The other accuracies remained more or less the same.

RANDOM TREES

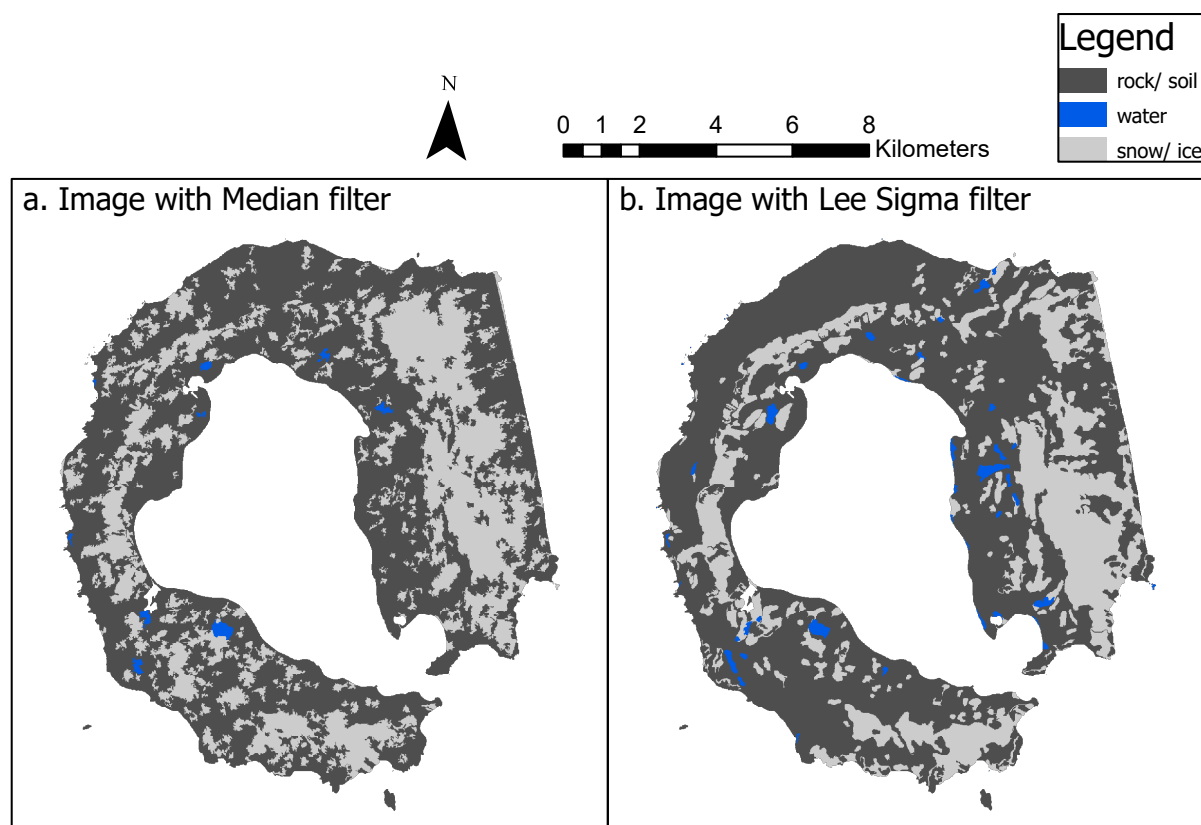


Figure 5.1.2.2 Images with Median (a.) and Lee Sigma (b.) filters, classified with Random Trees classifier

Sentinel1_Med_RT_3class_color_stdev_confmatrix							
OBJECTID *	ClassValue	C 0	C 1	C 2	Total	U Accuracy	Kappa
1	C_0	5524	34	928	6486	0,851681	0
2	C_1	29	29	5	63	0,460317	0
3	C_2	2084	6	1327	3417	0,388352	0
4	Total	7637	69	2260	9966	0	0
5	P_Accuracy	0,723321	0,42029	0,587168	0	0,690347	0
6	Kappa	0	0	0	0	0	0,268795

Table 5.1.2.3 Confusion matrix of the image with Median filter classified with Random Trees classifier

Sentinel1_LeeSig_RT_3class_color_stdev_confmatrix							
OBJECTID *	ClassValue	C 0	C 1	C 2	Total	U Accuracy	Kappa
1	C_0	5762	17	1248	7027	0,81998	0
2	C_1	112	16	28	156	0,102564	0
3	C_2	1775	17	993	2785	0,356553	0
4	Total	7649	50	2269	9968	0	0
5	P_Accuracy	0,753301	0,32	0,437638	0	0,679274	0
6	Kappa	0	0	0	0	0	0,188798

Table 5.1.2.4 Confusion matrix of the image with Lee Sigma filter classified with Random Trees classifier

The image chosen in this case was the one resulting from the choice of the segment attributes Active Chromaticity Color and Standard Deviation and using the training samples of all three classes, in this case in fact the training samples relative to the water class did not cause an excessive error. In the case of the image with Median filter, the k that occurred is similar to the one already obtained for the ISO cluster classification and the Maximum Likelihood one, while in the case of the image with the Lee Sigma filter, it is similar to that which occurred with the ISO cluster classification. In the first image, the snow cover in the eastern part of the island was identified, but not that in the southern part. Once again it is evident the influence of the speckle, as it leads to punctual errors in the classification of the soil/ rock as snow/ ice; this effect is reduced in the second image thanks to the multiple application of the Lee Sigma filter, however it is evident how in this case the illumination of the satellite has led to the incorrect classification of a part of rock/ soil as snow, in the North-West part of the island. In the second image too the snow cover on Mount Pond was partly recognized, but not the one on Mount Kirkwood, only a small area in the south. In both images there was a recognition of one lake in the south and one in the north, and in the second image the other two in the north were also recognized: one correctly classified as water and the other incorrectly as snow. In addition, in the latter a part of the sea that had been included by clipping with the coastline, located to the east, facing Port Foster, was recognized but classified as snow. This greater sensitivity to water in the second image, however, also led to more point errors than in the first image, in fact it can be seen that the User's Accuracy relative to class 1 is much greater in the first image. The failure to recognize all the lakes has led to low Producer's Accuracy. The User's and Producer's Accuracy for the rock/ soil were good and the one of the snow was again quite low, due to the failed delineation of the snow cover, as mentioned above. In the first image the accuracies were higher because the snow cover on Mount Goddard was detected, while in the second image it was not. The k obtained in the case of the image with Median filter was almost the same as the one obtained with Maximum Likelihood classifier, while the k of the image with Lee Sigma filter was double the one obtained in the case of Maximum Likelihood classifier.

SUPPORT VECTOR MACHINE

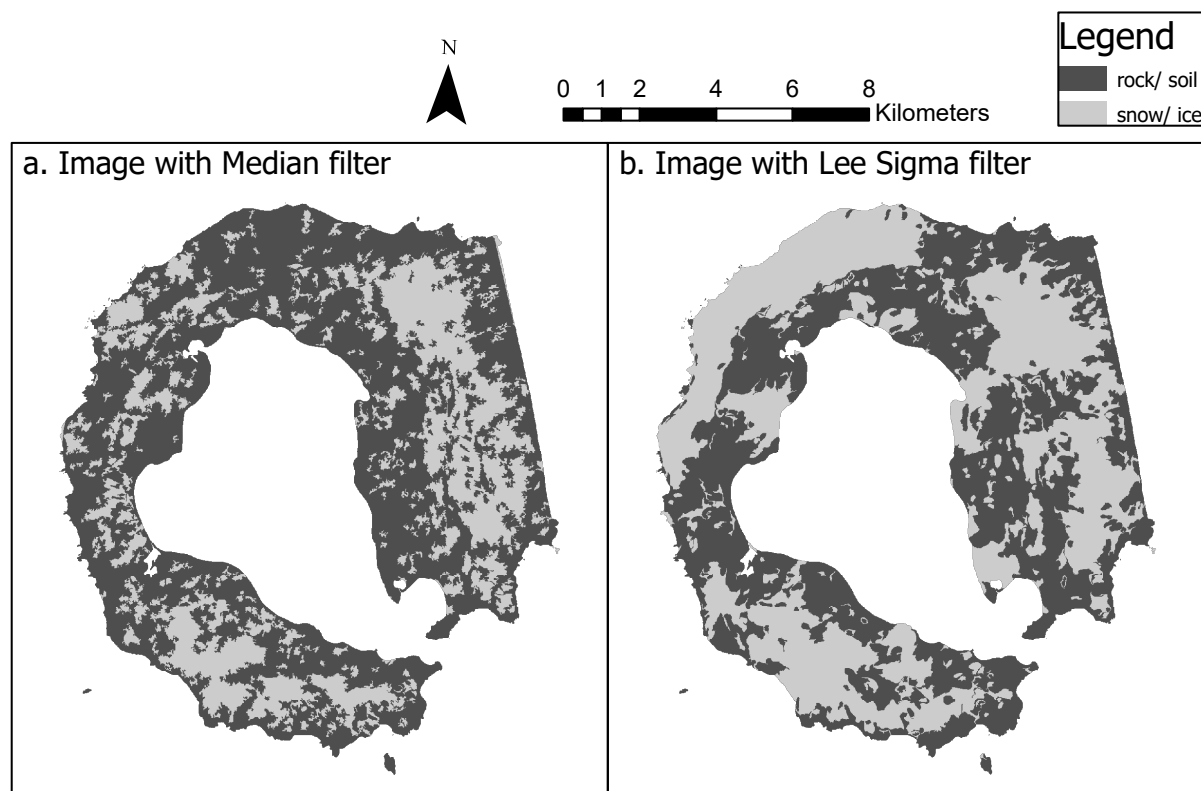


Figure 5.1.2.3 Images with Median (a.) and Lee Sigma (b.) filters, classified with Support Vector Machine classifier

Sentinel1_Med_SVM_3class_stdev_count_confmatrix							
OBJECTID *	ClassValue	C 0	C 1	C 2	Total	U Accuracy	Kappa
1	C_0	5647	51	882	6580	0,858207	0
2	C_1	0	0	0	0	0	0
3	C_2	1940	2	1451	3393	0,427645	0
4	Total	7587	53	2333	9973	0	0
5	P_Accuracy	0,744299	0	0,621946	0	0,711722	0
6	Kappa	0	0	0	0	0	0,311131

Table 5.1.2.5 Confusion matrix of the image with Median filter classified with Support Vector Machine classifier

Sentinel1_LeeSig_SVM_2class_stdev_count_confmatrix							
OBJECTID *	ClassValue	C 0	C 1	C 2	Total	U Accuracy	Kappa
1	C_0	4260	48	794	5102	0,834967	0
2	C_1	0	0	0	0	0	0
3	C_2	3385	1	1489	4875	0,305436	0
4	Total	7645	49	2283	9977	0	0
5	P_Accuracy	0,557227	0	0,652212	0	0,576225	0
6	Kappa	0	0	0	0	0	0,146204

Table 5.1.2.6 Confusion matrix of the image with Lee Sigma filter classified with Support Vector Machine classifier

The image chosen in this case was the one resulting from the use of the segment attributes Standard Deviation and Count of Pixels and the consideration of all training samples. In the case of the image to which Lee Sigma filter was applied, the use of the training samples of water led to a misclassification, causing all the island to be classified as rock/ soil, and the same happened using as segment attributes only Standard Deviation, as can be seen from the appendix. For this reason, we report the image obtained from the consideration of only the snow/ice and rock/soil training samples, so it is divided into two classes.

In the case of the image with the Median filter applied, the k obtained was the highest, errors caused by the speckle still occurred but did not prevent the recognition of the two snow-ice macrozones. The accuracies were quite high for rock/ soil and lower for snow-ice, in line with previous classifications, while the row in the table relating to class 1 is null because water was not classified, despite the inclusion of its training samples. This may have occurred because of the small extent of the watery areas, which were not recognised as a separate class but were incorporated into another, confused with the rock/soil. The same confusion between soil and water occurred with the Lee Sigma filter image, where the Kendall Terrace soil strip was misclassified, causing an overestimation of the snow cover, as opposed to the image obtained with Maximum Likelihood classifier. This resulted in a lowering of the User's Accuracy of the snow/ice class and the Producer's Accuracy of the soil class. The row corresponding to the water class is obviously empty because, as anticipated, the image resulting from the use of only the training samples of snow and soil is reported.

5.2 OPTICAL

5.2.1 UNSUPERVISED CLASSIFICATION

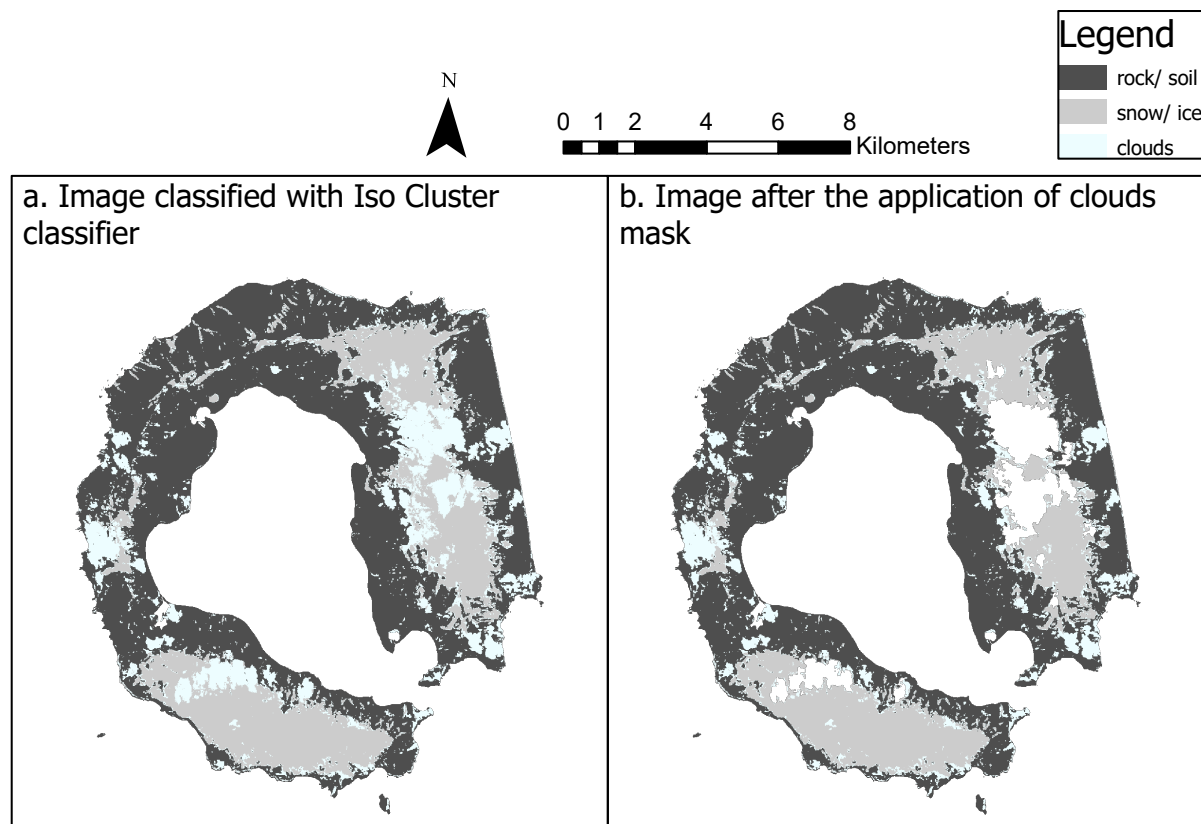


Figure 5.2.1.1 Image classified with ISO cluster classifier, with and without subsequent application of clouds mask

Sentinel2_ISOCL_confmatrix								
OBJECTID *	ClassValue	C 0	C 1	C 2	C 4	Total	U Accuracy	Kappa
1	C_0	5706	54	29	0	5789	0,985662	0
2	C_1	0	0	0	0	0	0	0
3	C_2	608	9	1842	0	2459	0,749085	0
4	C_4	1280	5	428	0	1713	0	0
5	Total	7594	68	2299	0	9961	0	0
6	P_Accuracy	0,751383	0	0,801218	0	0	0,757755	0
7	Kappa	0	0	0	0	0	0	0,51547

Table 5.2.1.1 Confusion matrix of the image classified with ISO cluster classifier

Sentinel2_ISOCL_MASKED_confmatrix								
OBJECTID *	ClassValue	C 0	C 1	C 2	C 4	Total	U Accuracy	Kappa
1	C_0	6126	48	31	0	6205	0,987268	0
2	C_1	0	0	0	0	0	0	0
3	C_2	629	3	1855	0	2487	0,745879	0
4	C_4	1128	7	137	0	1272	0	0
5	Total	7883	58	2023	0	9964	0	0
6	P_Accuracy	0,777115	0	0,916955	0	0	0,800984	0
7	Kappa	0	0	0	0	0	0	0,564175

Table 5.2.1.2 Confusion matrix after the application of clouds mask

This classification can be considered relatively good: the two snow-ice covers located on the glaciers were well delineated, and the clouds were identified, although there may have been an overestimation (however, it must be considered that the cloud mask was generated by the user from an index and is therefore not an absolute truth; the clouds may have been over- or underestimated). However, water was not classified; only two of the lakes present, located in the northern part of the island, were identified, but one was classified as snow/ ice and the other as clouds. This can also be seen from the confusion matrix, where the row for class C1 is filled with 0.

The use of the cloud mask results in a slight increase in k and in Producer's Accuracy for class C2 (by 10%). This was the expected result, since the clouds mainly cover the part related to snow, and therefore parts of the island affected by snow will be classified as clouds when covered by them: this leads to a decrease in Producer's Accuracy because the snow class will lack points that belong to it. By removing these zones with the mask, the accuracy will automatically increase.

The reason why the water class has not been created in the classified image is due to the fact that the areas of water have a small extension compared to the rest of the coverage, in addition, the lakes that have been identified have been classified one as snow and the other as clouds, this is due to the fact that it's always the same substance but in different physical states for the three ground covers. Probably, if the extent of the water had been of the same order as that of the snow, the two classes would have been divided, but what is supposed to have happened is that the lakes were incorporated in one case into the snow class and in the other into the cloud class in the various iterations because of their small area.

5.2.2 SUPERVISED CLASSIFICATION

MAXIMUM LIKELIHOOD

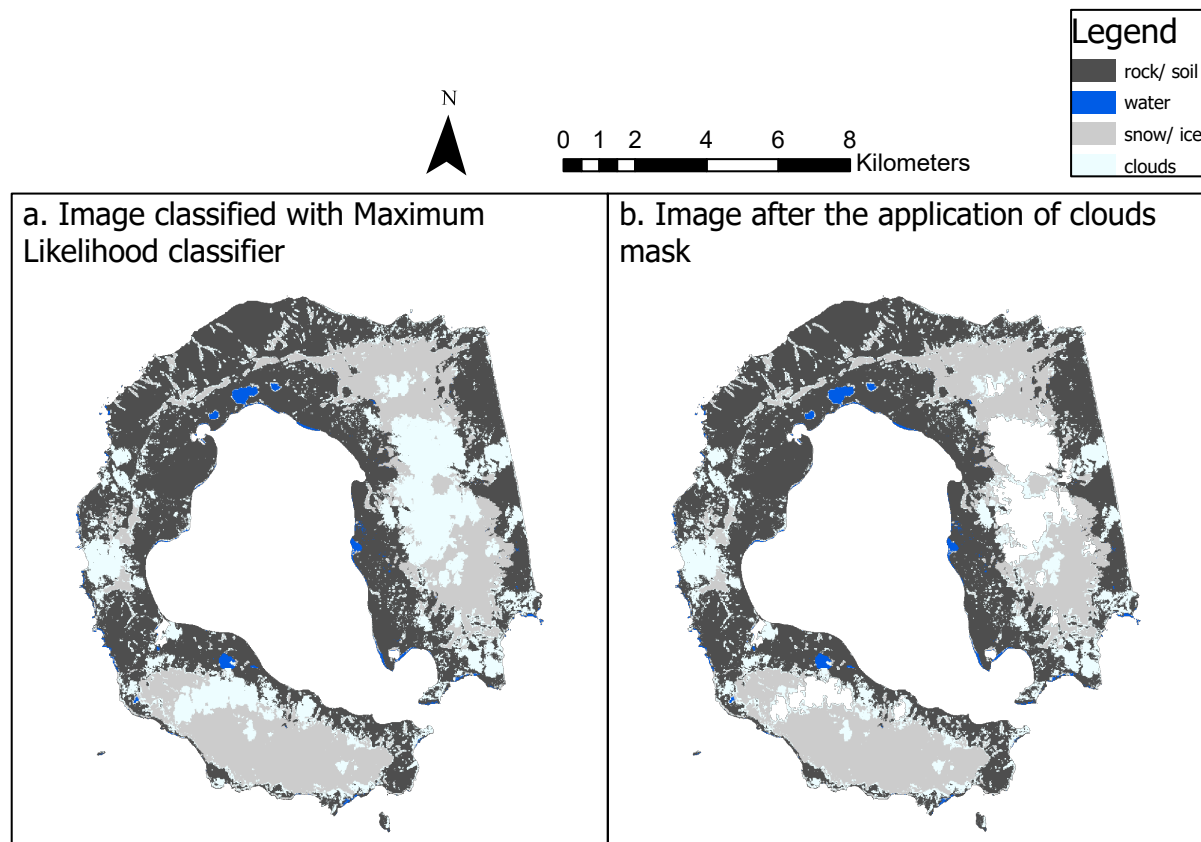


Figure 5.2.2.1 Image classified with Maximum Likelihood classifier, with and without subsequent application of clouds mask

Sentinel2_ML_confmatrix									
	OBJECTID *	ClassValue	C 0	C 1	C 2	C 4	Total	U Accuracy	Kappa
▶	1	C_0	5023	4	6	0	5033	0,998013	0
	2	C_1	65	45	1	0	111	0,405405	0
	3	C_2	864	2	1570	0	2436	0,644499	0
	4	C_4	1684	4	696	0	2384	0	0
	5	Total	7636	55	2273	0	9964	0	0
	6	P_Accuracy	0,657805	0,818182	0,690717	0	0	0,666198	0
	7	Kappa	0	0	0	0	0	0	0,400785

Table 5.2.2.1 Confusion matrix of the image classified with Maximum Likelihood classifier

Sentinel2_ML_MASKED_confmatrix									
	OBJECTID *	ClassValue	C 0	C 1	C 2	C 4	Total	U Accuracy	Kappa
▶	1	C_0	5411	3	4	0	5418	0,998708	0
	2	C_1	58	41	1	0	100	0,41	0
	3	C_2	967	6	1706	0	2679	0,636805	0
	4	C_4	1481	1	296	0	1778	0	0
	5	Total	7917	51	2007	0	9975	0	0
	6	P_Accuracy	0,683466	0,803922	0,850025	0	0	0,717594	0
	7	Kappa	0	0	0	0	0	0	0,451442

Table 5.2.2.2 Confusion matrix after the application of clouds mask

The k in the Maximum Likelihood classification was quite low and, analyzing the User's Accuracy and Producer's Accuracy, this seems to be mainly caused by commission errors in the water class, plus other commission errors in the snow class and omission in the rock and snow classes. Applying the cloud mask, there is an increase in Producer's Accuracy of the snow class while the increases in the others accuracies can be considered almost negligible. This resulted in a slight increase in k .

Analyzing the image, it can be seen that there has been an overestimation of the cloud cover, with the presence of various punctual errors in the rock/ soil class which have led to the lowering of its Producer's Accuracy; the classification of snow/ ice can be considered to be roughly good and the same can be said for water.

The increase in Producer's Accuracy of the snow/ ice class when applying the cloud mask was expected, because the removal of the parts covered by clouds that actually belong to snow reduces the percentage of points falling mistakenly classified as snow. Its User's Accuracy, however, is not as high and this is due to the fact that the snow cover of the ground truth is less than that of the analyzed image.

An unexpected result is instead that of the User's Accuracy value of the water class, which is very low. Comparing the classified images with the ground truth (figure 4.3.10) it can be seen that in the latter only the lakes present on the island are considered as water, while in the former several strips adjacent to the coast can be seen as belonging to this class. This is probably due to the fact that the coastline is not univocal but varies with the tides, and what has happened is that by clipping the image with the shapefile of the coastline, a small strip of sea was also included, and that is not present in the ground truth.

RANDOM TREES

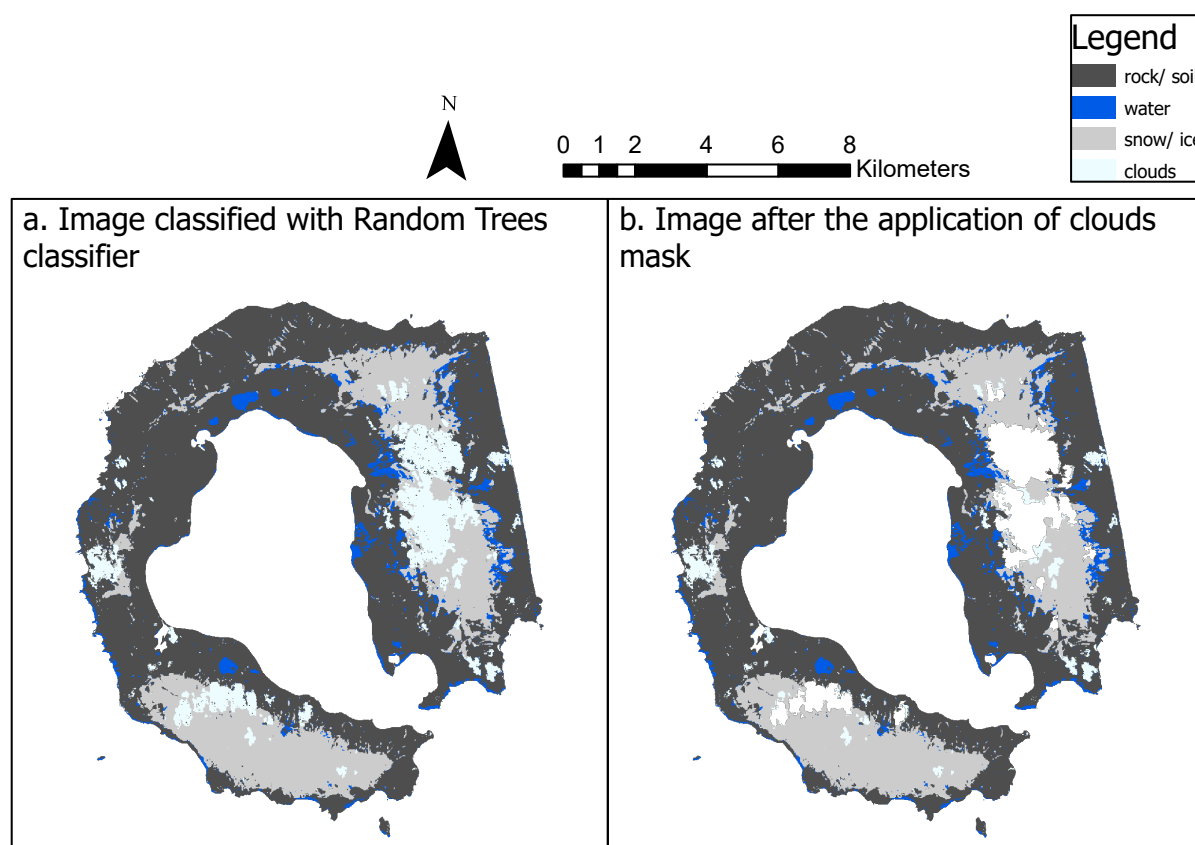


Figure 5.2.2.2 Image classified with Random Trees classifier, with and without subsequent application of clouds mask

Sentinel2_RT_confmatrix									
OBJECTID *	ClassValue	C 0	C 1	C 2	C 4	Total	U Accuracy	Kappa	
1	C_0	6126	3	43	0	6172	0,992547	0	
2	C_1	419	61	26	0	506	0,120553	0	
3	C_2	653	0	1711	0	2364	0,723773	0	
4	C_4	443	0	475	0	918	0	0	
5	Total	7641	64	2255	0	9960	0	0	
6	P_Accuracy	0,801728	0,953125	0,758758	0	0	0,792972	0	
7	Kappa	0	0	0	0	0	0	0,560018	

Table 5.2.2.3 Confusion matrix of the image classified with Random Trees classifier

Sentinel2_RT_MASKED_confmatrix									
OBJECTID *	ClassValue	C 0	C 1	C 2	C 4	Total	U Accuracy	Kappa	
1	C_0	6451	7	37	0	6495	0,993226	0	
2	C_1	475	66	25	0	566	0,116608	0	
3	C_2	662	0	1867	0	2529	0,738236	0	
4	C_4	249	0	126	0	375	0	0	
5	Total	7837	73	2055	0	9965	0	0	
6	P_Accuracy	0,823147	0,90411	0,908516	0	0	0,841345	0	
7	Kappa	0	0	0	0	0	0	0,634983	

Table 5.2.2.4 Confusion matrix after the application of clouds mask

The classification carried out using the Random Trees classifier gave very high User's Accuracy and Producer's Accuracy values, and the k could therefore have been appreciable, but what led to its lowering was the overestimation of the water class, as can be seen from the images. It can be seen that for this reason the User's Accuracy of the water class was very low.

As always, the Producer's Accuracy of the snow class increased when the cloud mask was applied, by about 15%. Its User's Accuracy was the lowest after that of the water class, and, as mentioned above, the reason is that the snow cover of the ground truth is smaller than that of the analyzed image.

The snow and cloud cover estimations were good, as was the rock estimation.

The reason for assigning the contours of the snow cover to the water class can be attributed to the fact that in these areas the snow present is wetter than in the center, located in higher areas of the mountain, its partial melting may therefore have led to its assignment to the water class by the classifier.

SUPPORT VECTOR MACHINE

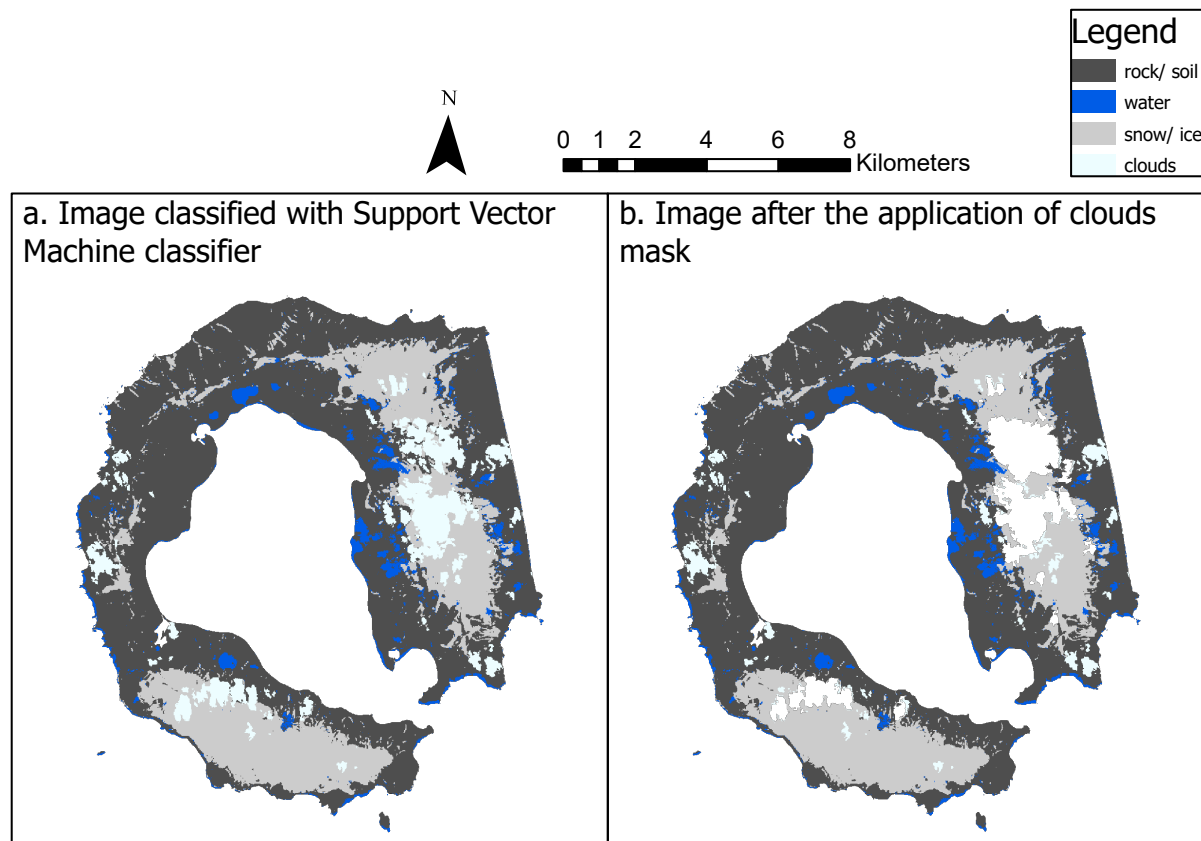


Figure 5.2.2.3 Image classified with Support Vector Machine classifier, with and without subsequent application of clouds mask

Sentinel2_SVM_confmatrix									
	OBJECTID *	ClassValue	C 0	C 1	C 2	C 4	Total	U Accuracy	Kappa
▶	1	C_0	6128	4	45	0	6177	0,992067	0
	2	C_1	361	47	34	0	442	0,106335	0
	3	C_2	791	0	1741	0	2532	0,687599	0
	4	C_4	433	0	385	0	818	0	0
	5	Total	7713	51	2205	0	9969	0	0
	6	P_Accuracy	0,794503	0,921569	0,789569	0	0	0,794062	0
	7	Kappa	0	0	0	0	0	0	0,556354

Table 5.2.2.5 Confusion matrix of the image classified with Support Vector Machine classifier

Sentinel2_SVM_MASKED_confmatrix									
	OBJECTID *	ClassValue	C 0	C 1	C 2	C 4	Total	U Accuracy	Kappa
▶	1	C_0	6429	6	40	0	6475	0,992896	0
	2	C_1	401	58	24	0	483	0,120083	0
	3	C_2	834	0	1859	0	2693	0,690308	0
	4	C_4	226	0	87	0	313	0	0
	5	Total	7890	64	2010	0	9964	0	0
	6	P_Accuracy	0,814829	0,90625	0,924876	0	0	0,837615	0
	7	Kappa	0	0	0	0	0	0	0,62288

Table 5.2.2.6 Confusion matrix after the application of clouds mask

The situation that occurred in this case was very similar to that of Random Trees: there was an overestimation of the aqueous coverage which led to a very low User's Accuracy for its class. The other accuracies were on average high, with the lowest being 0.69, relative to the snow class, as it happened with the other classifications, due to the fact that the snow cover of the analyzed image is more extensive than the one of ground truth. Again, using the cloud mask, the Producer's Accuracy of the snow class increases, more precisely by 13.5 %.

In this case, it seems that the overestimation of the water class occurred not only in the perimeter areas of the snow cover, but also in zones where it was confounded with small clouds. The reason for this can be attributed to the fact that in the end it is the same substance in different physical forms, as already explained in chapter 5.2.1: in the unsupervised classification one lake was classified as snow cover and another as clouds; here the opposite has happened, so snow and clouds have been classified as water, but it is still the same problem of confusing these classes.

5.2.3 INDEX CLASSIFICATION

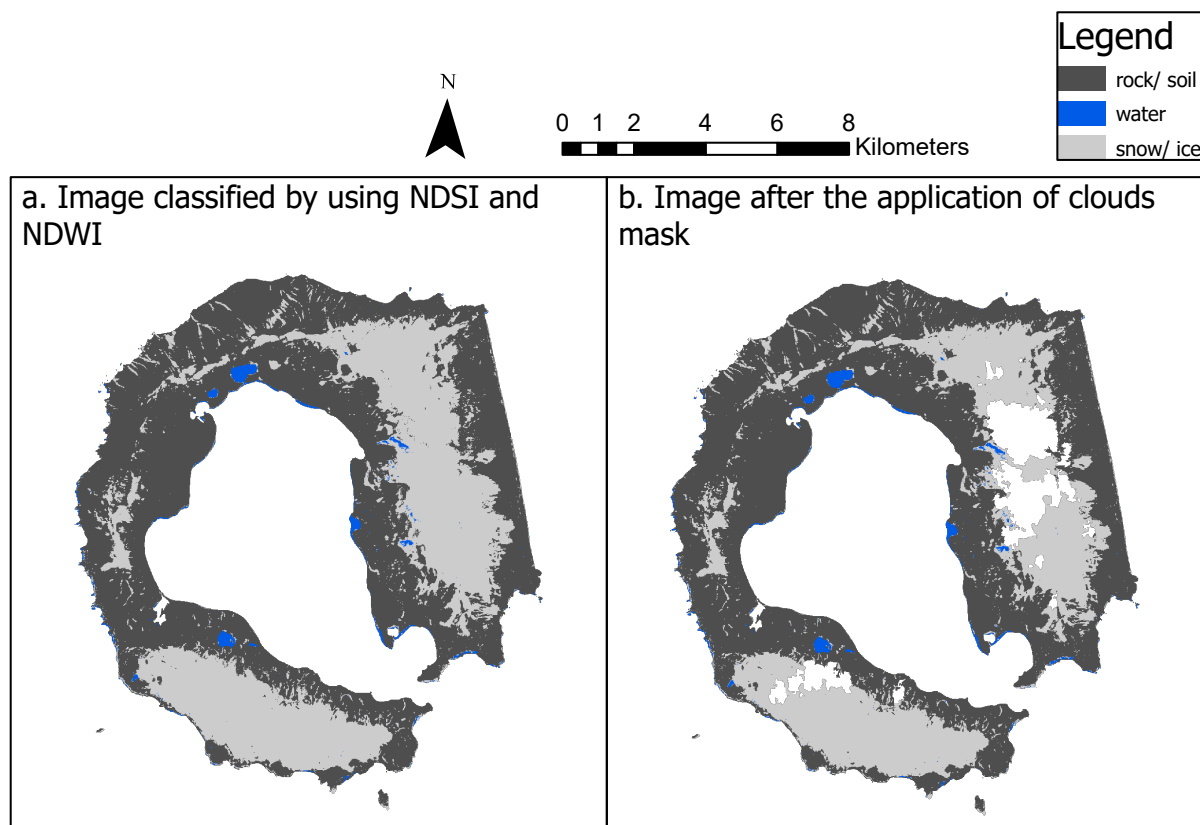


Figure 5.2.3.1 Image classified with NDSI and NDWI (a.) and effect of the subsequent application of clouds mask (b.)

classificataindici_confmatrix							
OBJECTID *	ClassValue	C 0	C 1	C 2	Total	U Accuracy	Kappa
1	C_0	6096	4	11	6111	0,997545	0
2	C_1	61	50	7	118	0,423729	0
3	C_2	1497	4	2236	3737	0,598341	0
4	Total	7654	58	2254	9966	0	0
5	P_Accuracy	0,796446	0,862069	0,992014	0	0,84106	0
6	Kappa	0	0	0	0	0	0,64218

Table 5.2.3.1 Confusion matrix of the image classified with NDSI and NDWI

classificataindici_MASKED_confmatrix							
OBJECTID *	ClassValue	C 0	C 1	C 2	Total	U Accuracy	Kappa
1	C_0	6519	1	18	6538	0,997094	0
2	C_1	84	52	13	149	0,348993	0
3	C_2	1273	4	2006	3283	0,611027	0
4	Total	7876	57	2037	9970	0	0
5	P_Accuracy	0,827704	0,912281	0,984782	0	0,860281	0
6	Kappa	0	0	0	0	0	0,663003

Table 5.2.3.2 Confusion matrix after the application of clouds mask

This classification was the one that resulted in the highest k, so the misclassified parts were minimal: mainly the lake to the north-east which was classified as snow cover and some small

strips of snow cover were classified as water. The application of the cloud mask did not bring about an appreciable improvement in this case as with the other classifications, and this is attributable to the fact that with the use of the indices the influence of clouds was diminished by using bands that did not detect them. Again, the lowest User's Accuracy occurred for the water class and the snow class, and this is attributable to the same reasons already mentioned: the difference between the snow cover of the treated image and that of ground truth, and the water along the coastline of the treated image, not present in ground truth. With regard to the lake that has been classified as snow-ice, in reality it is a phenomenon that has always occurred in all the binary images used to separate the snow cover (snow from ice, pure snow and ice from that mixed with debris, and that what had snow and ice from that what was bare ground), used as a step prior to the creation of this image and the other complete image that is proposed (Figure 5.2.3.2). This is supposed to occur because actually it is always water but in different forms, so even if an index is designed to divide snow from other ground covers it is possible to commit error in the presence of water. Another error that occurred was that of the NDWI's classification of some snow strips as belonging to the water class; it is assumed that, since they are perimeter strips, they are wet and melting snow, so their high liquid water content led to their inclusion in this class.

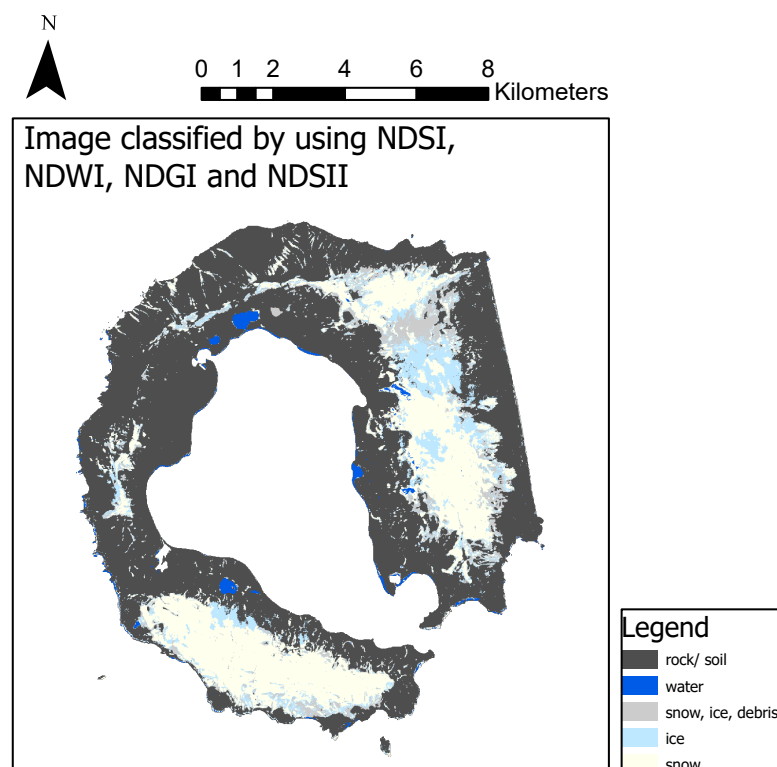


Figure 5.2.3.2 Image classified with NDSI, NDWI, NDGI, NDSII

In the image resulting from the application of the four indices, it can be seen that two new classes have been added to the previous one, which overlap the snow/ ice class, but cover a smaller area; this is what was expected (the pure snow and ice will be a part of the total snow/

ice class, in which the debris is present, and again, the snow will be a part of the snow/ ice total class). For the newly added classes, however, there is no ground truth and it is therefore not possible to quantify the goodness of the classification, but a qualitative analysis of the image produced can be made. The addition of the result of the last index concerning the separation of snow from ice seems to be affected by the clouds, as it leaves some areas uncovered in correspondence with the location of these, where the coverings classified by the two previous indices emerge. Apart from this, the trend seems to be coherent, the areas gradually left uncovered by these last two indices are mainly those around the perimeter of the snow covers, which are supposed to be those most mixed with debris, as they are at the foot of the mountains and more humid, and that present in the part between Mount Pond and Mount Goddard, this is therefore also an area in depression, in which the snow will consequently be more humid and mixed with debris as in the perimeter areas of the mountains. It can be seen that in the image used as ground truth, taken after the one analyzed, this part has been affected by melting of the snow present, which confirms what has just been said.

6. CONCLUSIONS

The k values reached in the confusion matrices of the Sentinel 2 image classifications were not those desirable with an optical image, and this is due to the fact that a ground truth of the same date was not available, the consequence of which is a much reduced snow cover in the ground truth compared to the image analyzed; this has led to a User's Accuracy for this class never exceeding 75%. This implies that what the algorithm detects as classification errors, being snow points in the classified image and rock/ soil points in the ground truth, are in fact only a consequence of reduced snow cover and not a misclassification.

In addition, the image was subject to partial cloud cover, which not only did not allow for a full comparison of the optical image classifications with SAR ones in the covered areas, but also led to errors in the covered areas, lowering the quality of the optical image classifications: when applying the cloud mask, there has always been an increase in the Producer's Accuracy of the snow-ice class of 10-15%.

By using SAR images, the cloud problem doesn't exist, but the classification is strongly influenced by satellite illumination and island morphology, effects that are amplified when speckle noise is reduced by using filters. Classifications have never achieved acceptable results: in the case of the image with a single application of the Median filter, punctual errors caused by the speckle were very evident in the classified images, it must however be recognized that the snow cover on the Pond and Goddard Mountains was almost always recognized; in the case of the image with multiple application of the Lee Sigma filter, the correct recognition of the ground covers was almost never verified, except in some cases of detection of part of the snow cover, it seems that the multiple application of a speckle reduction filter also led to the reduction of the contrasts between the different ground covers, thus preventing their recognition. In any case, the effect of illumination and the morphology of the island has always been strong and influential in SAR images.

The k of the classification based on the use of indices was the highest, and this is because it was possible to avoid the influence of clouds by using certain bands; this is therefore a road that can be explored in the field of optical classification, if we want to overcome the problem of cloud cover.

In conclusion:

- single polarimetry SAR backscattering images are not suitable for this type of classification in an area of this type, where strong contrasts between covers are not present and geometries are not simple;
- the classifications made with the optical image would have been better had it not been for the influence of the cloud cover;
- the ground truth used was not optimal due to the different snow cover, and therefore led to unreal errors when computing the confusion matrices, but the classifications already provided by NASA and ESA were unusable;

- the classification based on indices was the best as it seems to have led to a partial circumvention of the cloud problem, so more studies could be carried out in this field.

7. BIBLIOGRAPHY

Baraldo A., Rapalini A.E., Böhnel H., Mena M., “Paleomagnetic study of Deception Island, South Shetland Islands, Antarctica”, *Geophysical Journal International*, Vol. 153, pp. 333-343, 2003

Belgiu M., Dragut L., “Random forest in remote sensing: A review of applications and future directions”, *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 114, pp. 24–31, 2016

Breiman L., “Random Forests”, 2001

Campaña Antártica del Ejército de Tierra, Oficina de Comunicación, “Diario de Operaciones, XXXIII Campaña Antártica del Ejército de Tierra”, 2019-2020

Dalziel I.W.D., “Tectonic evolution of a fore-arc terrane, southern Scotia Ridge, Antarctica”, *Geological Society of America Special Paper*, Vol. 200, 1984

Dewi F., Lailia A., Jatmiko R.H., “Object based classification for land cover using Sentinel-1A in Yogyakarta”, 2019

ESA, “Sentinel-2 User Handbook”, 2015

ESA, Thales Alenia Space France, “Sentinel-1 Product Definition”, 2016

ESA, Thales Alenia Space France, “Sentinel-2 Products Specification Document”, 2018

Fukunaga K., Hostetler L.D., “The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition”, *IEEE Transactions on Information Theory*, Vol. IT-21, No. 1, pp. 32-40, 1975

Gao B.C., “NDWI-A Normalized Difference Water Index for Remote Sensing of Vegetation Liquid Water From Space”, *Remote Sensing of Environment*, Vol. 58, pp. 257-266, 1996

Geyer A., Álvarez-Valero A. M., Gisbert G., Aulinas M., Hernández-Barreña D., Lobo A., Marti J., “Deciphering the evolution of Deception Island’s magmatic system”, www.nature.com/scientificreports/, 2018

Geyer A., Marti J., "The new worldwide collapse caldera database (CCDB): A tool for studying and understanding caldera processes", *Journal of Volcanology and Geothermal Research*, Vol. 175, pp. 334-354, 2008

Hald A., On the History of Maximum Likelihood in Relation to Inverse Probability and Least Squares, *Statistical Science*, Vol. 14, No. 2, pp. 214-222, 1999

Hall D.K., Riggs G., "Normalized-Difference Snow Index", 2010

Henriet J.P., Meissner R., Miller H., The Grape Team, "Active margin processes along the Antarctic Peninsula", *Tectonophysics*, Vol. 201, pp. 229-253, 1992

Keshri A.K., Shukla A., Gupta R.P., "ASTER ratio indices for supraglacial terrain mapping", *International Journal of Remote Sensing*, Vol. 30, No. 2, pp. 519–524, 2009

Kupidura P., "Comparison of Filters Dedicated to Speckle Suppression in SAR Images", *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, Vol. XLI-B7, pp. 269-276, 2016

Lee J.S., "A Simple Speckle Smoothing Algorithm for Synthetic Aperture Radar Images", *IEEE Transactions on Systems, Man, and Cybernetics*, Vol. SMC - 13, No. 1, pp. 85-89, 1983

Lee J.S., "Digital image enhancement and noise filtering by use of local statistics", *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. PAMI-2, No. 2, pp. 165-168, 1980

Lee J.S., "Refined Filtering of Image Noise Using Local Statistics", 1981

Lee J.S., Jurkevich I., "Segmentation of SAR Images", *IEEE Transactions on Geoscience and Remote Sensing*, Vol. 27, pp. 674-608, 1989

Liu K., Shi W., Zhang H., "A fuzzy topology-based maximum likelihood classification", *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 66, pp. 103–114, 2011

Luzón F., Almendros J., García-Jerez A., "Shallow structure of Deception Island, Antarctica, from correlations of ambient seismic noise on a set of dense seismic arrays", *Geophysical Journal International*, Vol. 185, pp. 737-748, 2011

Martí J., Geyer A., Aguirre-Díaz G., "Origin and evolution of the Deception Island caldera (South Shetland Islands, Antarctica)", *Bulletin of Volcanology*, Vol. 75(6), 2013

McFeeters S.K., "The Use of the Normalized Difference Water Index (NDWI) in the Delineation of Open Water Feature", *International Journal of Remote Sensing*, Vol. 17, pp. 1425-1432, 1996

Moreton S.G., "Quaternary tephrochronology of the Scotia Sea and Bellingshausen Sea, Antarctica", PhD Thesis, Cheltenham and Gloucester College of Higher Education, UK, 1999

Mountrakis G., Im J., Ogole C., "Support vector machines in remote sensing: a review", *ISPRS Journal of Photogrammetry and Remote Sensing*, Vol. 66, pp. 247–259, 2011

Oliva-Urcia B., Gil-Peña I., Maestro A., López-Martínez J., Galindo-Zaldívar J., Soto R., Gil-Imaz A., Rey J., Pueyo O., "Paleomagnetism from Deception Island (South Shetlands archipelago, Antarctica), new insights into the interpretation of the volcanic evolution using a geomagnetic model", *International Journal of Earth Sciences*, Vol. 105, pp. 1353-1370, 2016

Palais J.M., Kirchner S., Delmas R., "Identification and correlation of volcanic eruption horizons in a 1,000-year ice-core record from the South Pole", *Antarctic Journal of the United States*, Vol. 24, pp. 101-104, 1989

Shukla A., Ali I., "A hierarchical knowledge-based classification for glacier terrain mapping: a case study from Kolahoi Glacier, Kashmir Himalaya", *Annals of Glaciology*, Vol. 57(71), 2016

Shultz C.H., "Eruption at Deception Island, Antarctica, August 1970", *GSA Bulletin September*, Vol. 83(9), pp. 2837-2842, 1972

Smellie J.L., "Graham Land and South Shetland Islands", *Antarctic Research Series*, Vol. 48, pp. 303-312, 1990

Smellie J.L., "Lithostratigraphy and volcanic evolution of Deception Island, South Shetland Islands", *Antarctic Science*, Vol. 13(2), pp. 188-209, 2001

Smellie J.L., López-Martínez J., and others, "Geology and geomorphology of Deception Island", 2002, with accompanying maps. BAS GEOMAP Series, Sheets 6-A and 6-B, 1:25000. Cambridge, British Antarctic Survey

Stehman S.V., "Selecting and Interpreting Measures of Thematic Classification Accuracy", *Remote Sensing Environment*, Vol. 62, pp.77-89, 1997

USGS, "Landsat 8 Collection 2 (C2) Level 2 Science Product (L2SP) Guide", 2020

Valencio A., Mendía E., Vilas J., “Paleomagnetism and K-Ar age of Mesozoic and Cenozoic igneous rocks from Antarctica”, *Earth and Planetary Science Letters*, Vol. 45, pp. 61-68, 1979

Vapnik V., “Estimation of Dependences Based on Empirical Data”, 1979 (in Russian) (English translation: Springer Verlag, New York, 1982)

Wang Z., Liu J., Li J., Zhang D.D., “Multi-Spectral Water Index (MuWI): A Native 10-m Multi Spectral Water Index for Accurate Water Mapping on Sentinel-2”, *Remote Sensing*, 2018

Xu H., “Modification of Normalized Difference Water Index (NDWI) to Enhance Open Water Features in Remotely Sensed Imagery”, *International Journal of Remote Sensing*, 2006

WEBSITES

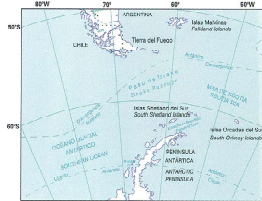
1. ArcGIS Desktop: Compute Confusion Matrix:
<https://desktop.arcgis.com/en/arcmap/latest/tools/spatial-analyst-toolbox/compute-confusion-matrix.htm>
2. ArcGIS Desktop: <https://desktop.arcgis.com/en/>
3. ArcGIS Pro: Compute Segment Attributes: <https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/compute-segment-attributes.htm>
4. ArcGIS Pro: How Iso Cluster works: <https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/how-iso-cluster-works.htm>
5. ArcGIS Pro: How Maximum Likelihood Classification works:
<https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/how-maximum-likelihood-classification-works.htm>
6. ArcGIS Pro: <https://www.esri.com/en-us/arcgis/products/arcgis-pro/resources>
7. ArcGIS Pro: Understanding segmentation and classification:
<https://pro.arcgis.com/en/pro-app/latest/tool-reference/spatial-analyst/understanding-segmentation-and-classification.htm>
8. Australian Antarctic Data Centre: Composite Gazetteer of Antarctica:
https://data.aad.gov.au/aadc/gaz/scar/display_name.cfm?gaz_id=108460
9. Copernicus Open Access Hub: <https://scihub.copernicus.eu>

10. Deception Island Website: climate: <http://www.deceptionisland.aq/climate.php>
11. Deception Island Website: <http://www.deceptionisland.aq>
12. ESA: Sentinel 1 Instrument Payload: <https://sentinel.esa.int/web/sentinel/missions/sentinel-1/instrument-payload>
13. ESA: Sentinel 1: <https://sentinel.esa.int/web/sentinel/missions/sentinel-1>
14. ESA: Sentinel 2 Level-2 Products: <https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/processing-levels/level-2>
15. ESA: Sentinel 2 Spatial Resolution: <https://dragon3.esa.int/web/sentinel/user-guides/sentinel-2-msi/resolutions/spatial>
16. ESA: Sentinel 2: <https://sentinel.esa.int/web/sentinel/missions/sentinel-2>
17. ESA: SNAP: <https://step.esa.int/main/toolboxes/snap/>
18. L3HARRIS: ENVI: <https://www.l3harrisgeospatial.com/Software-Technology/ENVI>
19. Multi-Resolution Land Characteristics (MRLC) Consortium: <https://www.mrlc.gov>
20. NASA: Landsat 8 Overview: <https://landsat.gsfc.nasa.gov/landsat-8/landsat-8-overview>
21. NASA: Landsat 8: <https://landsat.gsfc.nasa.gov/landsat-8>
22. USGS: Band Designations for the Landsat Satellites: https://www.usgs.gov/faqs/what-are-band-designations-landsat-satellites?qt-news_science_products=0#qt-news_science_products
23. USGS: Earth Explorer: <https://earthexplorer.usgs.gov>
24. USGS: Landsat 8: https://www.usgs.gov/core-science-systems/nli/landsat/landsat-8?qt-science_support_page_related_con=0#qt-science_support_page_related_con
25. USGS: Landsat Quality Assessment ArcGIS Toolbox: <https://www.usgs.gov/core-science-systems/nli/landsat/landsat-quality-assessment-arcgis-toolbox>

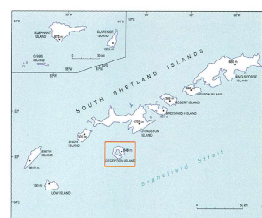
8. APPENDIX

8.1 GEOLOGIC MAP

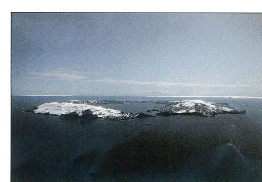
The geological map shown was compiled by Smellie and López-Martínez in 2000 and published by the British Antarctic Survey, together with the geomorphological map shown in 8.2 and an accompanying text, the book 'Geology and Geomorphology of Deception Island' (Smellie et al., 2002), for the BAS GEOMAP Series project.



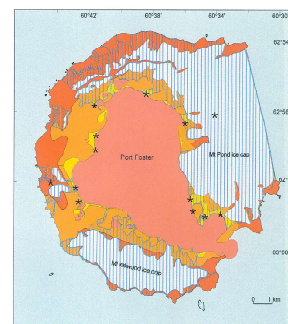
Location map



Sketch map of the South Shetland Islands, showing the main altitude each island and the location of Kingman Island.



Aerial view from the south

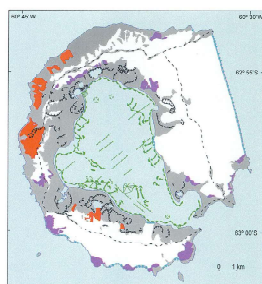


VOLCANIC HAZARDS (based on eruptions observed in 1942, 1907, 1903 and 1970)

Tonsamic, turbulent wave, ash fall, possible explosion, under stress	Site of highly explosive eruptions with ash and fall, possible lava flows, cinder fall
 Low-lying shoreline areas at risk from tonsamic	 Site of mildly explosive eruptions with lava flows and cinder fall, ash fall
 Area affected by mudflows (the eruptions under low aspect)	 Ash fall and possible tsunami affecting sea and small islands outside Christmas Island

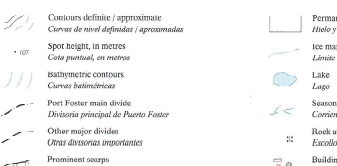
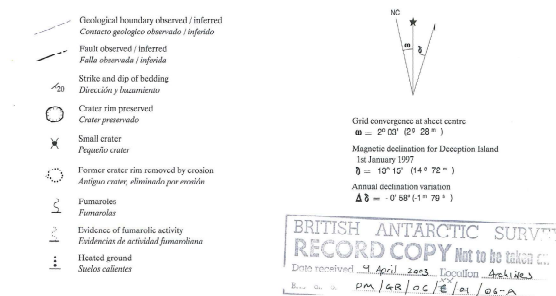
A **HAZARD ZONES**, including tsunami and not just volcanic and temperatures may vary

Volcanic hazards map (based on eruptions observed in 1842, 1967, 1969 and 1970). Prepared by J. L. Smellie, British Antarctic Survey



Geomorphology and marine ecology simplified from RAS GEOMAP 6-B

Lithostratigraphy *Litoestratigrafía* (All geological units are Quaternary / *Todas las unidades geológicas son cuaternarias*)

[illegible]Topographic symbols *Símbolos topográficos*Geological symbols *Símbolos geológicos*

8.2 GEOMORPHOLOGIC MAP

The geomorphological map shown was prepared by Smellie and López-Martínez in 2000 and published by the British Antarctic Survey, accompanying the geological map shown in 8.1 and the book 'Geology and Geomorphology of Deception Island' (Smellie et al., 2002), for the BAS GEOMAP Series project.

Compiled by J. López-Martínez¹, E. Serrano², J. Rey³ and J.L. Smellie⁴¹Universidad Autónoma de Madrid, ²Universidad de Cantabria, ³Instituto Español de Oceanografía, ⁴British Antarctic Survey, under Project ANT/91-1207, ANT/94-0606 and ANT/98-0225 of the Programa Español de Investigación en la Antártida. Participation in the field work by C. Criado, J.J. Durán, I.M. González-Casado and P. Nicolás is acknowledged.Topography based on Servicio Geográfico del Ejército-Universidad Autónoma de Madrid, 1994. *Deception Island map*, 1:25 000. Spanish Antarctic Cartography, Madrid, Servicio Geográfico del Ejército, with corrections, mainly to glaciers and streams, by the authors. The extent and type of glacier surfaces have been mapped by interpreting aerial photographs taken in 1979 by the Royal Navy, UK.

LEGEND LEYENDA

TOPOGRAPHY

TOPOGRAFÍA

- Contours on rock
- Curvas de nivel en roca
- Spot height, in metres
- Cota puntual, en metros
- Survey point
- Punto geodésico
- Port Foster main divide on rock
- Divisoria principal de Puerto Foster en roca
- Other major divides on rock
- Otras divisorias importantes en roca
- Crest or local divide on rock
- Cresta o divisoria local en roca
- Scarp
- Escarpes en roca
- Buildings / Storage tanks
- Edificios / Depósitos

ANIMAL COLONIES AND NESTS

COLONIAS Y NIDOS DE ANIMALES

- Penguin colony
- Plagüeros
- Weddell seals
- Focas de Weddell
- Seal Colony
- Colonia de lobos marinos
- Petrel nests
- Nidos de petreles
- Seagull nests
- Nidos de gaviotas, chamanes, etc.

GEOLOGICAL STRUCTURE AND STRUCTURAL LANDFORMS

ESTRUCTURA GEOLÓGICA Y FORMAS ESTRUCTURALES

- Strike and dip obfolding
- Dirección y buzamiento
- Fault with morphological expression
- Falla con significación morfológica
- Fumaroles
- Fumaroles
- Evidence of fumarolic activity
- Evidencias de actividad fumaroliana
- Heated ground
- Suelos calientes
- Tuff cone or cinder cone
- Cone de cenizas o cenizas
- Crater rim preserved / retreated / removed by erosion
- Cenit conservado / erosionado
- Main lava flow with scoriaceous surface
- Principal flujo de lava con superficie escoriada
- Volcanic bombs
- Bombas volcánicas
- Roundish volcanic blocks
- Bloques volcánicos redondeados

GLACIAL LANDFORMS AND DEPOSITS

FORMAS Y DEPÓSITOS GLACIARES

- Roches moutonnées, abraded and plucked surfaces
- Rocas aborregadas, pulimentadas o estradas
- Till, glacial deposit
- Til, depósito glacial
- Moraine ridge
- Cresta Morénica
- Ice-cored moraine ridge
- Cresta morénica con núcleo de hielo
- Fluvio-glacial cone
- Cone fluvio-glacial
- Lake, jokulhup deposits
- Depósitos de lake
- Present glaciers
- Glaciares actuales
- Glacier with debris-free surface
- Glaciar con superficie sin derrubios
- Glacier with debris / pyroclastic cover
- Glaciar con recubrimiento de derrubios / piroclásticos
- Ice margin
- Límite del hielo
- Contours on ice
- Curvas de nivel sobre hielo
- Port Foster main divide on ice
- Divisoria principal de Puerto Foster en hielo
- Other major divides on ice
- Otras divisorias importantes en hielo
- Crest or local divide on ice
- Cresta o divisoria local sobre glaciar
- Main scarps covered by ice
- Escarpes mayores cubiertos de hielo
- Scarp in ice
- Escarpes en hielo
- Crevasses
- Crías
- Ice cliff
- Escarpes de hielo
- Supraglacial depression
- Depresión supraglaciar
- Ice mounds
- Montes de hielo

PERIGLACIAL AND NIVAL LANDFORMS AND DEPOSITS

FORMAS Y DEPÓSITOS PERIGLACIARES Y NIVALES

- Subglacial lobes
- Lóbulos de subglaciación
- Kettle holes
- Hoyos de fusión
- Flat-floored valley
- Valle de fondo plano
- Even slope with pyroclasts
- Ladera regular con piroclastos
- Debris slope
- Derrubios de ladera
- Debris cone
- Ceno de derrubios
- Block field
- Acumulación de bloques

FLUVIAL AND LACUSTRINE LANDFORMS AND DEPOSITS

FORMAS Y DEPÓSITOS FLUVIALES E HIDROGRÁFICA

- Lake
- Lago
- Seasonally flooded depression
- Depresión con encharcamiento temporal
- Seasonal streams
- Corrientes estacionales
- V-shaped valley
- Valle encajado en V
- Alluvial fan
- Cono de deposición

AEOLIAN LANDFORMS

MODELO EÓLICO

- Wind-eroded surface and faceted rocks
- Superficie con delabración y cornos facetedos

MARINE LANDFORMS AND DEPOSITS

FORMAS Y DEPÓSITOS MARINOS

- Emergent landforms and deposits
- Formas y depósitos emergidos
- Rock awash
- Escallos descubiertos

Beach

Playa

- Cliff
- Acantilado

Submerged features and deposits

Formas y depósitos sumergidos

- Bathymetric contours
- Curvas batimétricas
- Coastal shelf edge
- Escarpes de plataforma litoral
- Cliff
- Acantilado

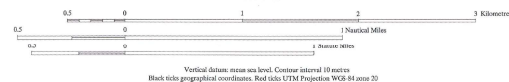
- Slope
- Talud
- Outer extent of subhorizontal deep sector
- Límite del sector profundo subhorizontal
- Scarp scar
- Cicatriz de despegue
- Channel
- Canal
- Linear runnels and grooves
- Surcos y marcas lineales
- Undulating sea-floor
- Fondo ondulado
- Normal fault (tectonically identified)
- Falla normal (identificada tectónicamente)
- Volcanic cone
- Cono volcánico
- Dome
- Domo
- Fluvial or fluvio-glacial fan
- Abanico fluvial o fluvio-glacial
- Progradation wedges
- Conos de progradación
- Depositional lobe
- Lóbulo deposicional
- Till, glacial deposit
- Til, depósito glacial

Sea-floor deposits

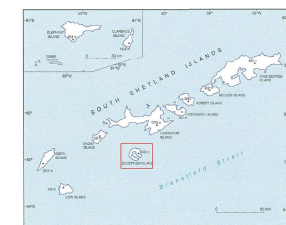
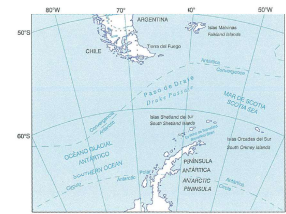
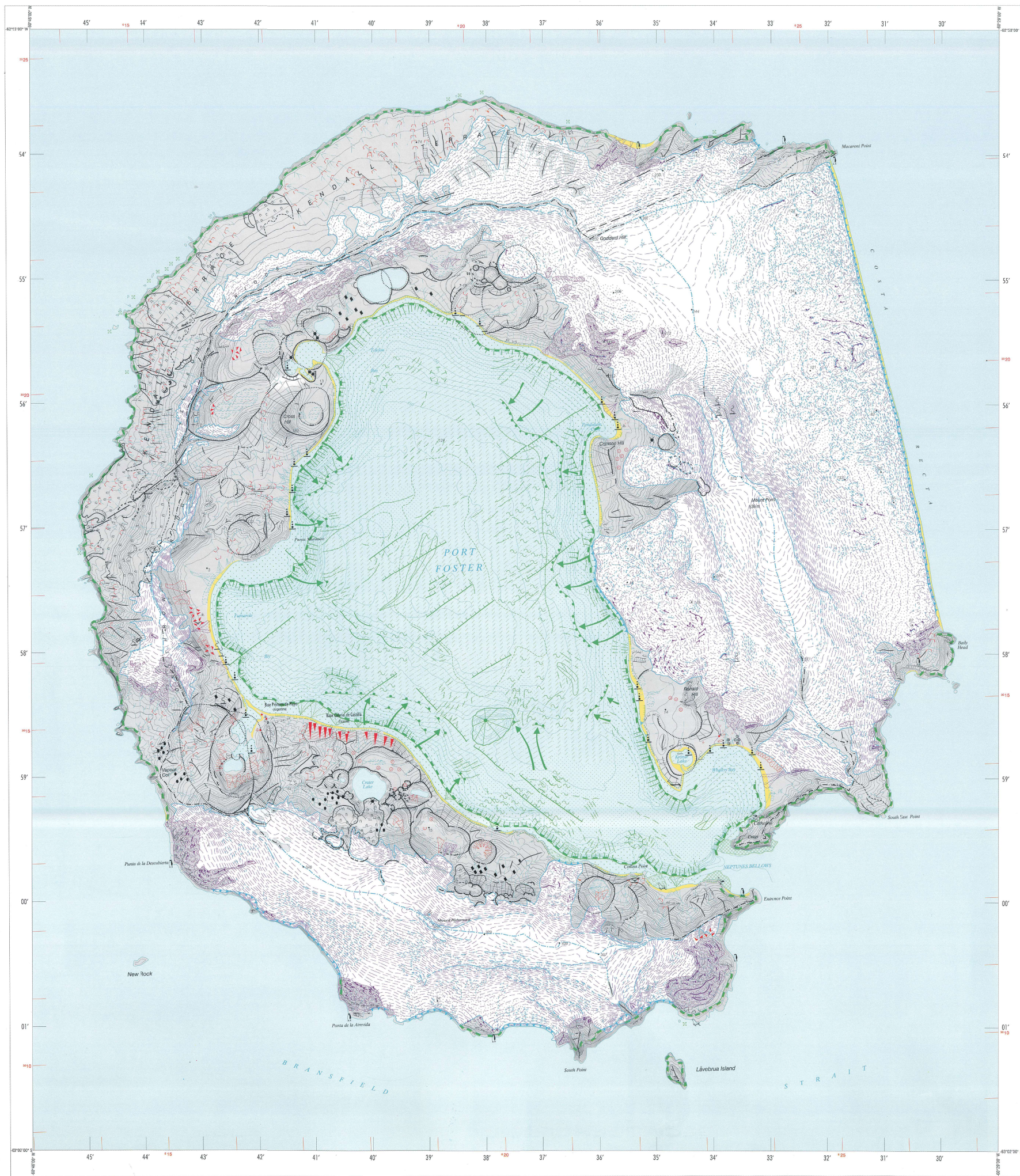
Depósitos de fondo marino

- Rock
- Roca
- Gravel (> 4 mm)
- Grava (> 4 mm)
- Sand (4 to 0.063 mm)
- Arena (4 a 0.063 mm)
- Muddy sand
- Arena fangosa
- Sandy mud
- Fango arenoso

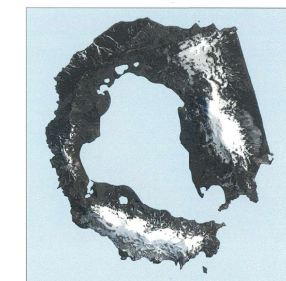
SCALE 1:25 000

Recommended reference for maps:
LÓPEZ-MARTÍNEZ, J. and others, 2000. *Geomorphological map of Deception Island*. BAS GEOMAP Series, Sheet 6-B, 1:25 000, with supplementary text. Cambridge, British Antarctic SurveyPublished by the BRITISH ANTARCTIC SURVEY, High Cross, Madingley Road, Cambridge CB3 0ET, England, as a part of the BAS GEOMAP Series. It is accompanied by a geological map (Sheet 6-A) and a supplementary text.
Map printed by CENTRO GEOGRÁFICO DEL EJÉRCITO, Darío Ganzo 8, 28024 Madrid, Spain.
© for the base map SERVICIO GEOGRÁFICO DEL EJÉRCITO, DP. M-9246-1994.

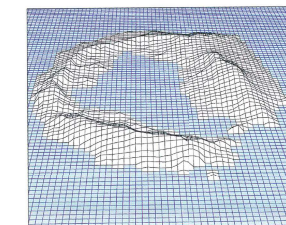
Digitization and preparation for multicoloured reproduction co-ordinated by Gral. R. Tur, Col. V. Martínez Valero, Tost. F. Hernández Chifuentes, Cte. M. Abril and Spjo J. M. Ortega López, Centro Geográfico del Ejército, Programa Español de Investigación en la Antártida, Project ANT 93 0961. Advice on map design by J.W. Thomson, British Antarctic Survey

Grid convergence at sheet centre
= 2° 03' (2° 28' N)
Magnetic declination for Deception Island
1st January 1997
= 13° 15' (14° 72' N)
Annual declination variation
ΔD = -0° 58' (1° 70' N)

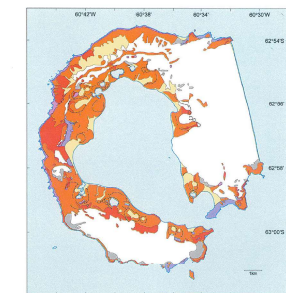
Sketch map of the South Shetland Islands, showing the main altitude of each island and the location of Deception Island



Landsat satellite image, combining TM 4, TM 3 and TM 2. Provided by the Spanish Antarctic Program, prepared by J. Calvet



Digital elevation model created by Centro Geográfico del Ejército, Spain

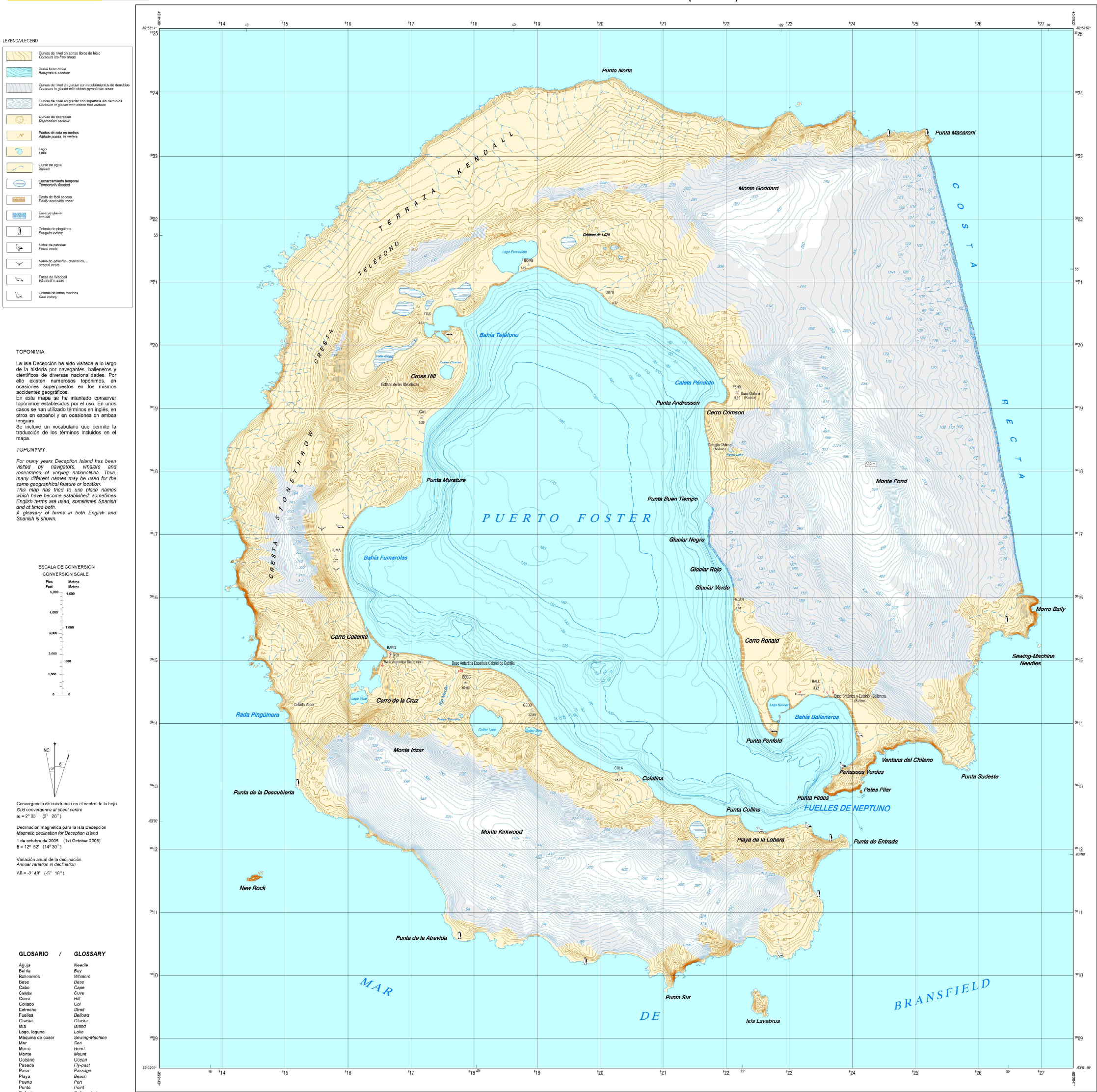


Geological map generalized from BAS GEOMAP 6-A by J.L. Smellie, British Antarctic Survey

BRITISH ANTARCTIC SURVEY
RECORD COPY Not to be taken away
Date received 9 April 2003 Location, Argentina
BAS Cat. No. PM/98/307/6/24/6-B

8.3 TOPOGRAPHIC MAP

The topographic map shown is the second edition of a map drawn up by the Centro Geográfico del Ejército Español in collaboration with the Universidad de Cádiz and dates from 2005.



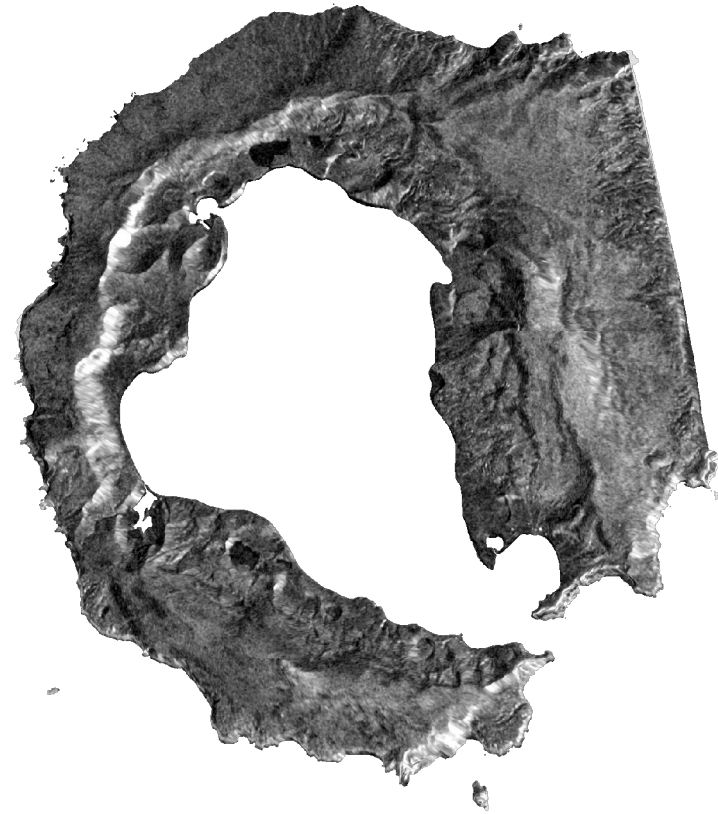
8.4 IMAGES USED

The three satellite images used are reported below, in particular, the RGB combinations of their bands are shown, furthermore, the ground truth produced according to the steps explained in detail in chapter 4.3 using as sources the Pixel QA band of the Landsat image and the topographic map of 2005 are also reported.

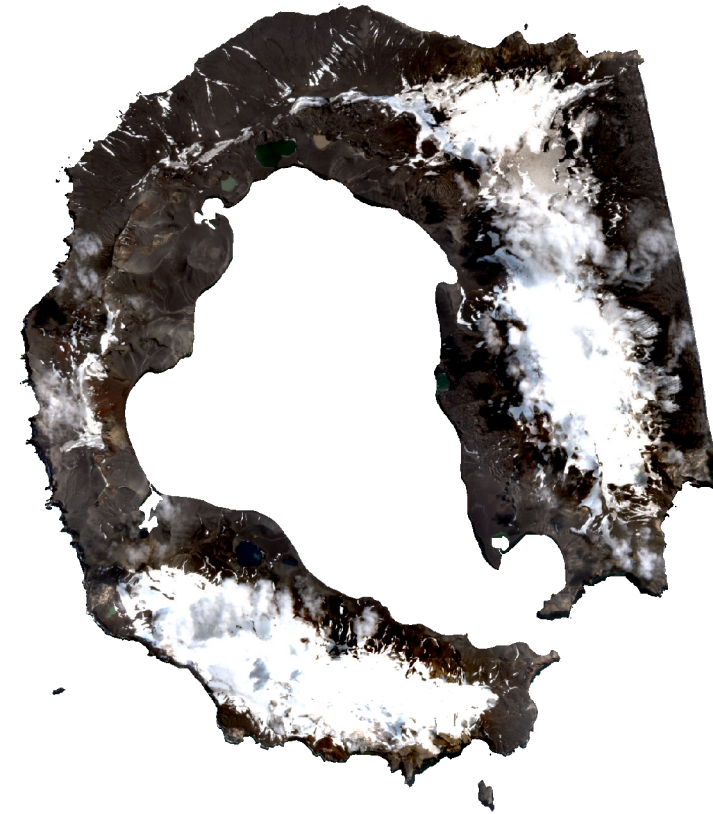


0 1 2 4 6 8 Kilometers

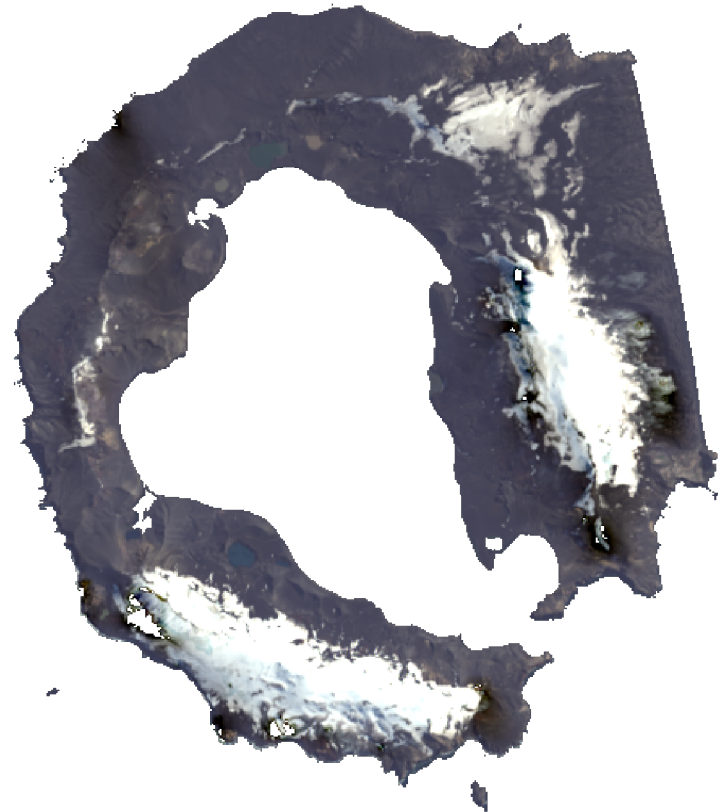
a. Sentinel 1 image



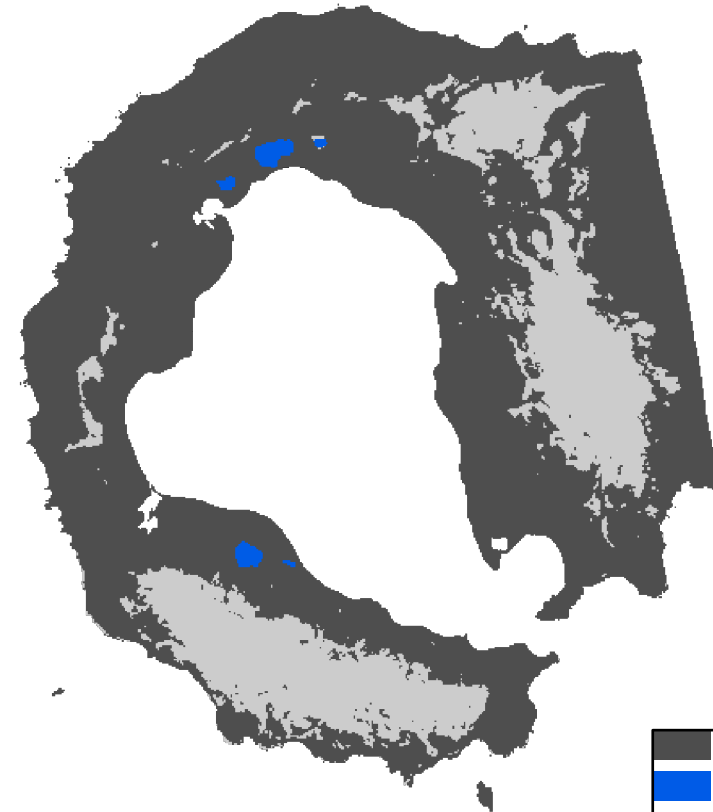
b. Sentinel 2 image



c. Landsat 8 image



d. Ground truth



rock/ soil
water
snow/ ice

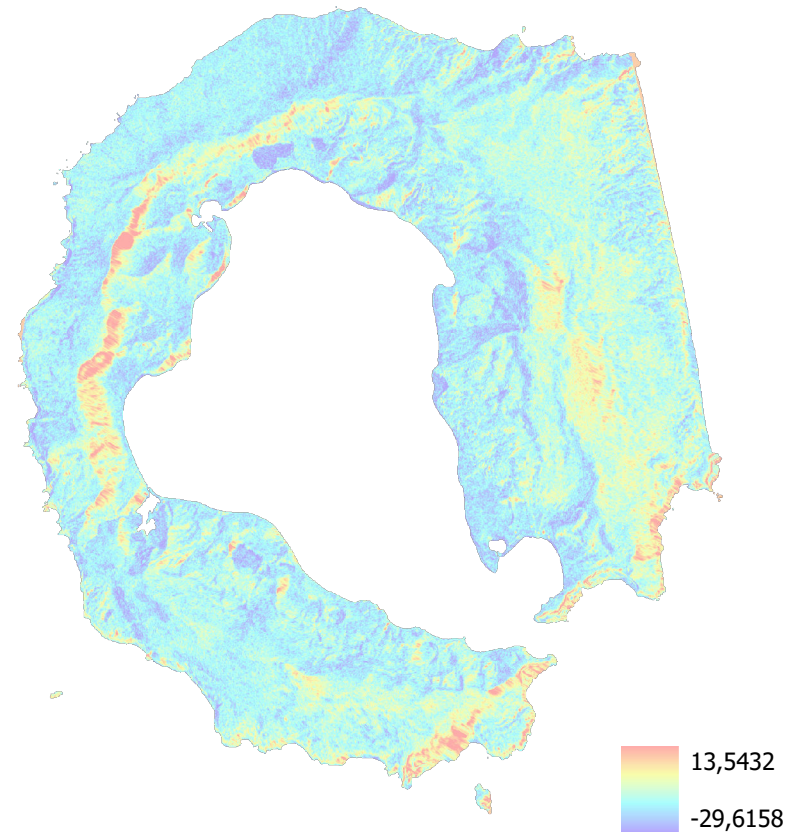
8.5 SAR

In this section we will first report the effect of applying the different filters on the image, then the four segmented images used for the classifications and finally the classifications, first unsupervised and then supervised; in particular, in the case of the supervised ones, we will report several images that were produced but not discussed in chapter 5 because they were worse than those chosen and discussed in that chapter.

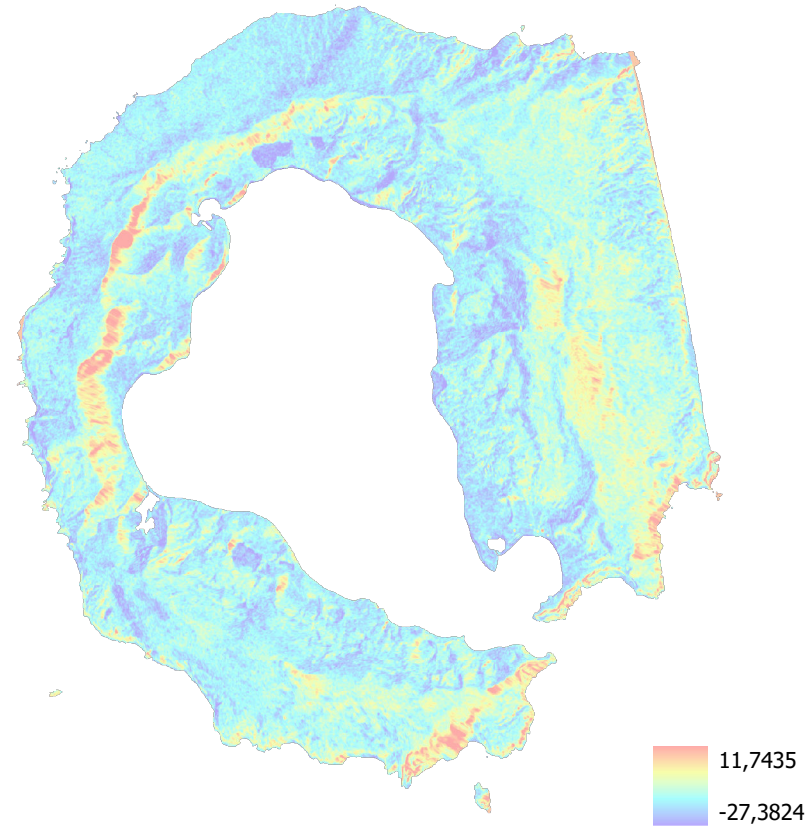
8.5.1 EFFECT OF THE FILTERS

A comparison is made between the processed but unfiltered image (therefore still cropped, calibrated, georeferenced and subjected to geometric correction) and the application of the different filters used for the study; in particular, in the case of the Lee, Lee Sigma and Refined Lee filters, the effect of the triple application of each filter is reported.

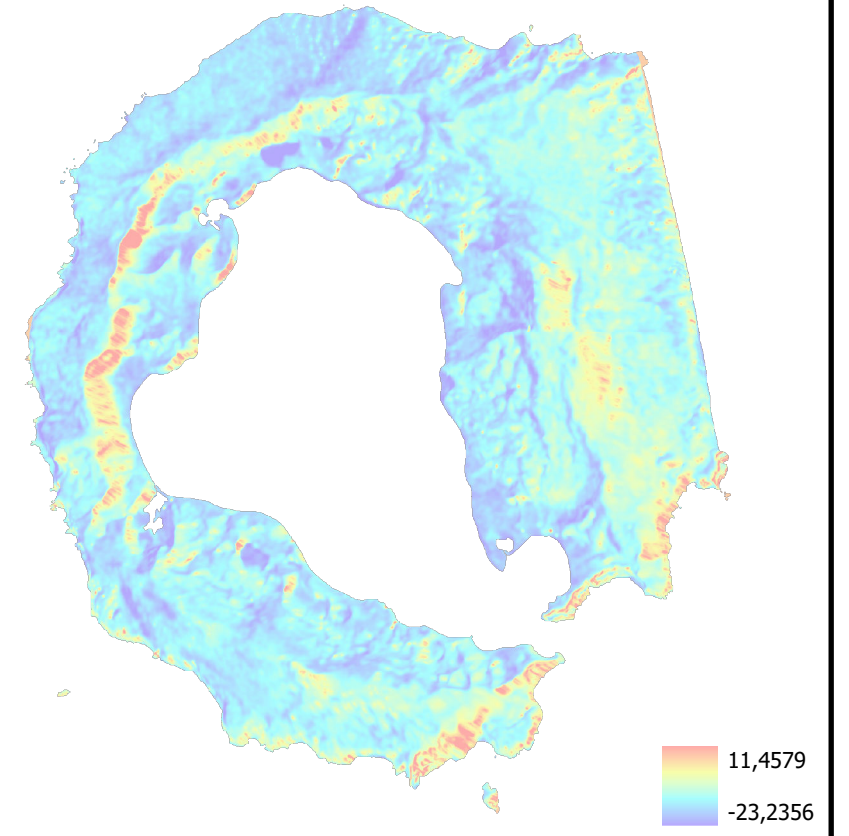
a. Image without filters



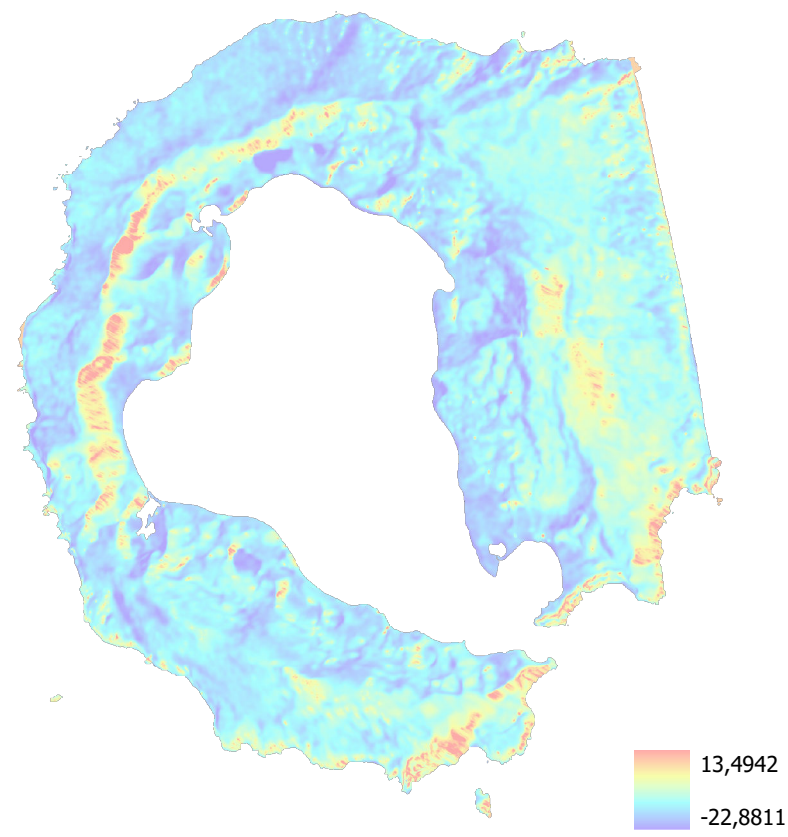
b. Image with an application of Median filter



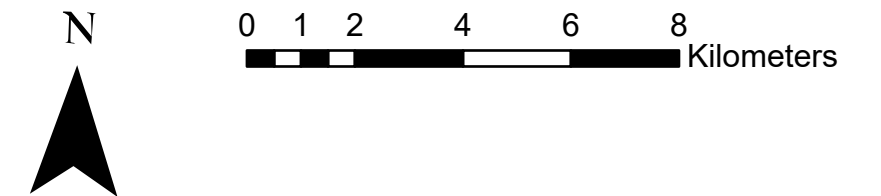
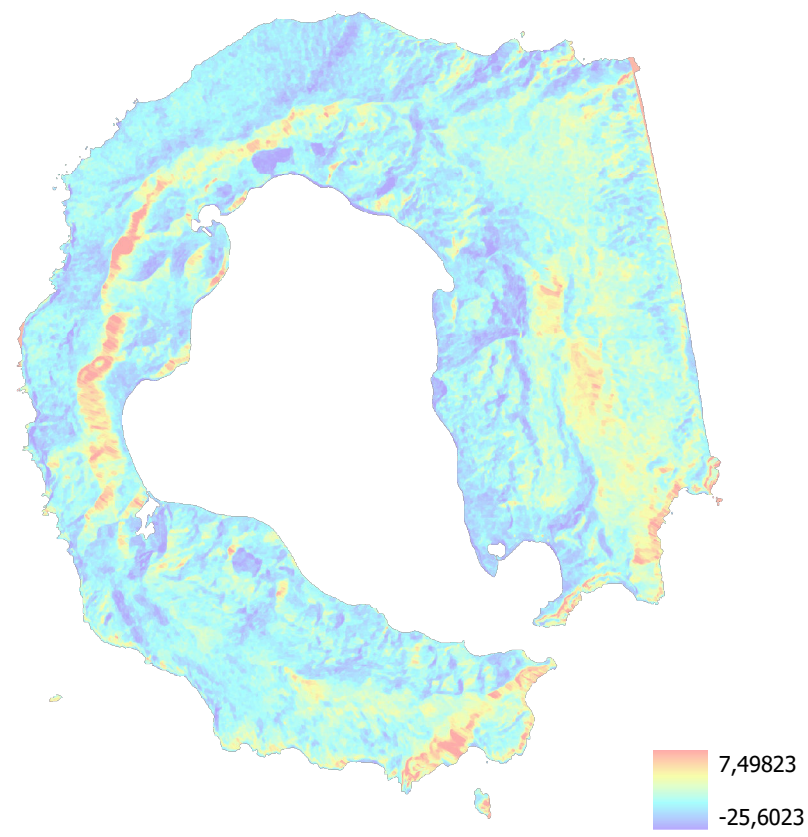
c. Image with three applications of Lee filter



d. Image with three applications of Lee Sigma filter



e. Image with three applications of Refined Lee filter



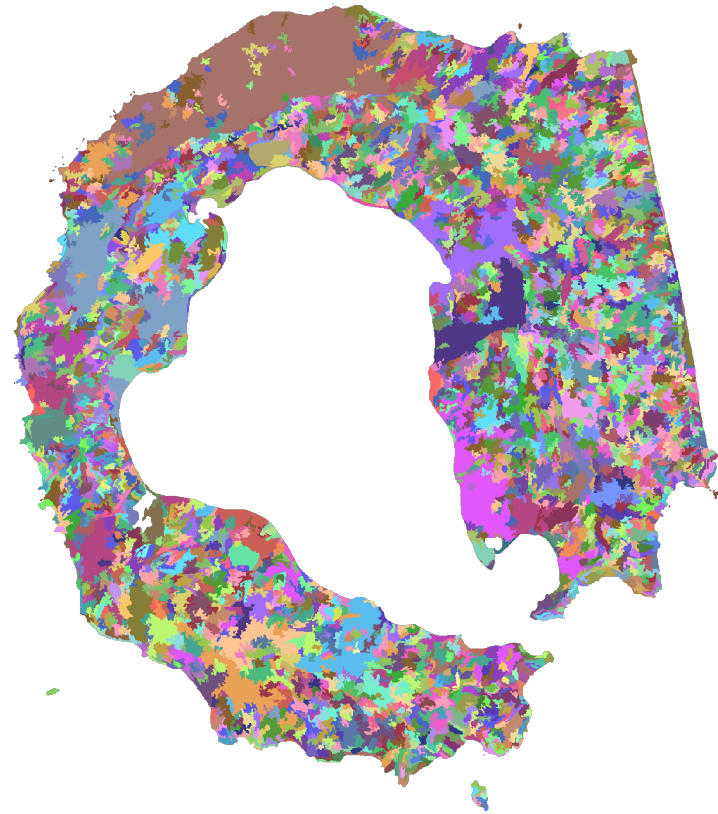
8.5.2 SEGMENTED IMAGES

The four segmented images that were then used for unsupervised classification are shown, one for each filtered image.



0 1 2 4 6 8 Kilometers

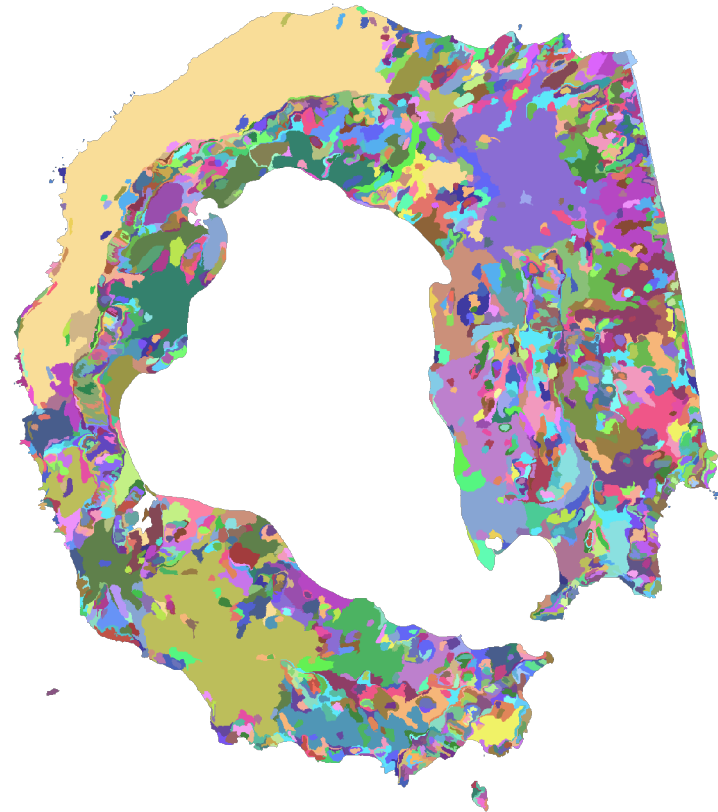
a. Segmentation of the image with Median filter



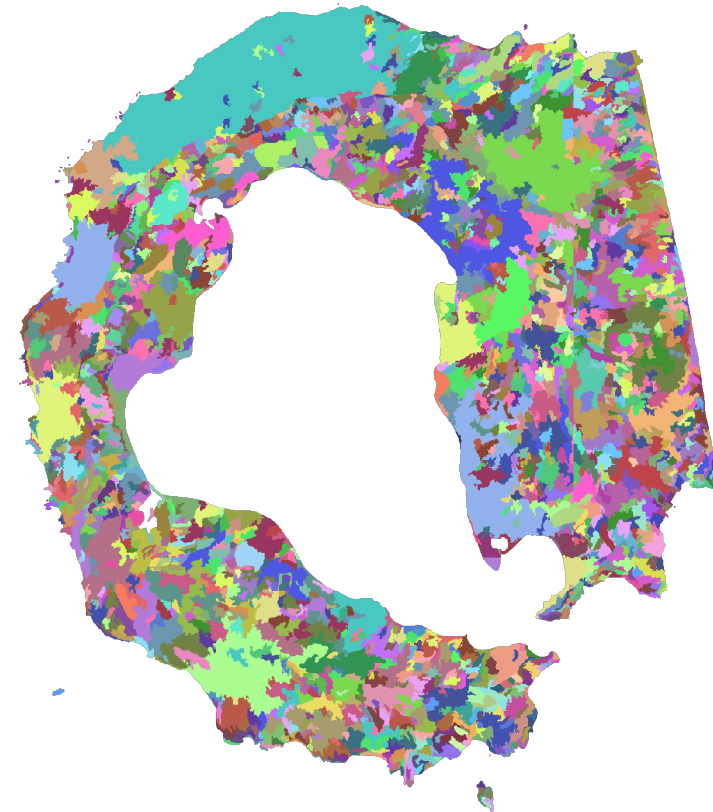
b. Segmentation of the image with Lee filter



c. Segmentation of the image with Lee Sigma filter

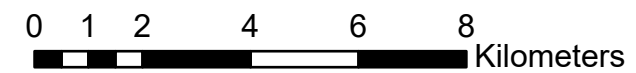


d. Segmentation of the image with Refined Lee filter

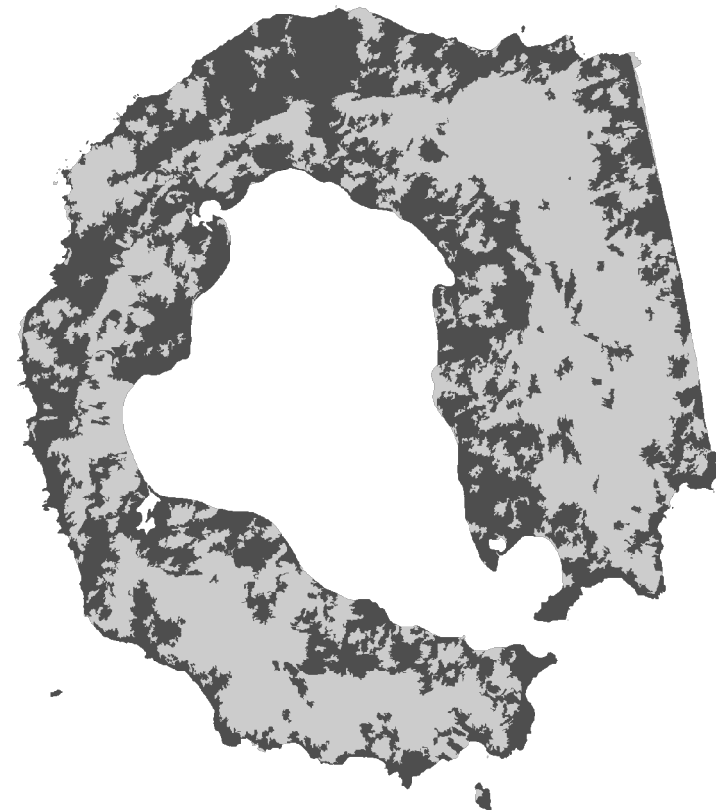


8.5.3 UNSUPERVISED CLASSIFICATION

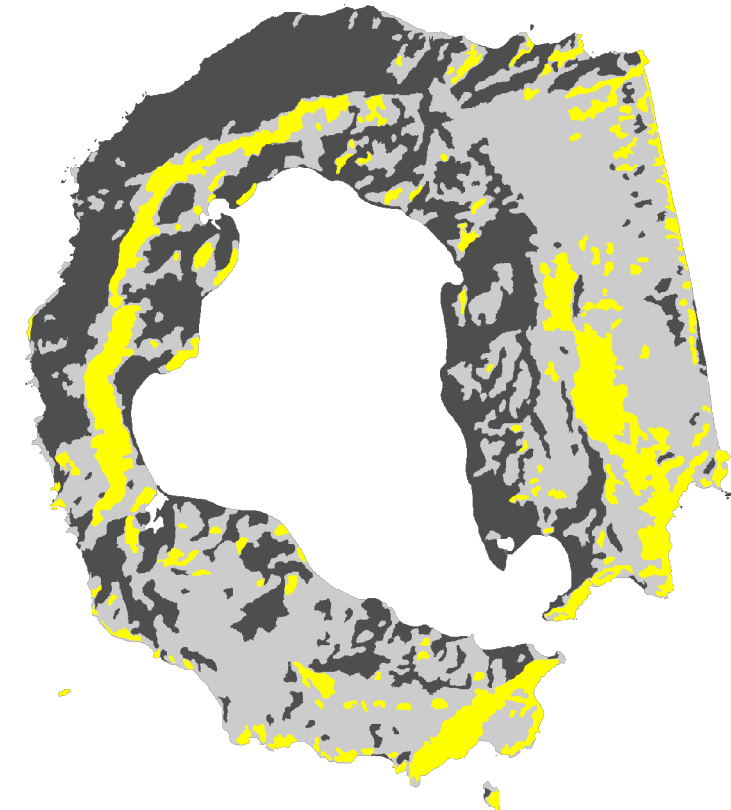
The four images obtained from the unsupervised classification of those in 8.4.2 are shown.



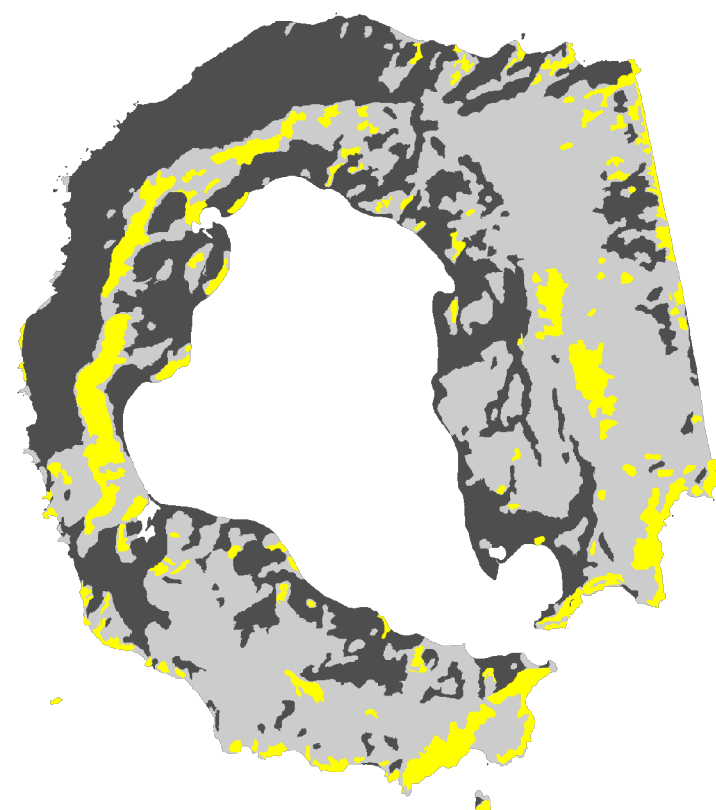
a. Unsupervised classification of the image with Median filter



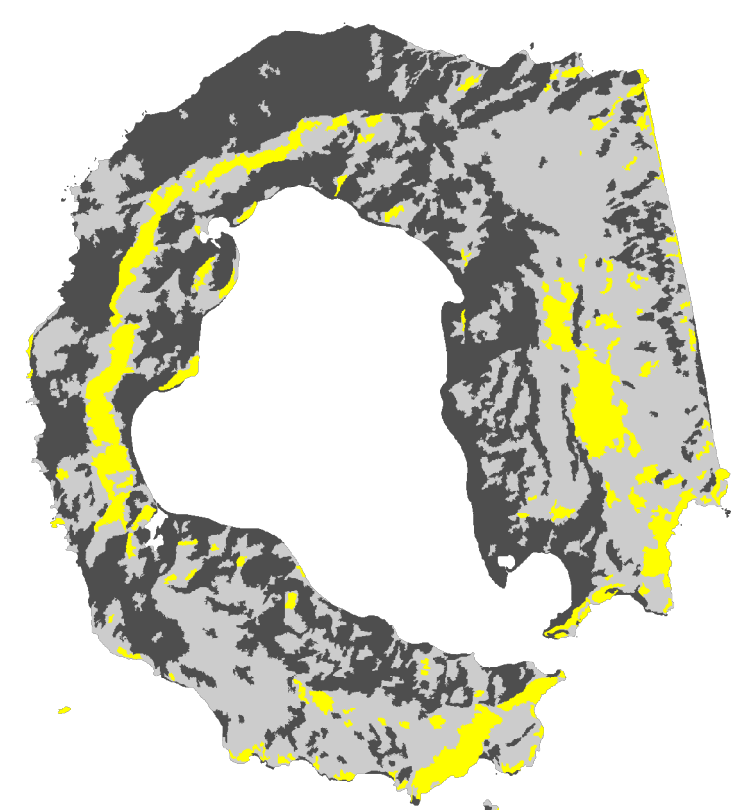
b. Unsupervised classification of the image with Lee filter



c. Unsupervised classification of the image with Lee Sigma filter



d. Unsupervised classification of the image with Refined Lee filter



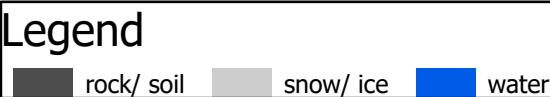
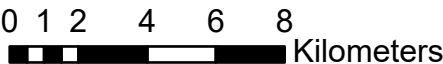
8.5.4 SUPERVISED CLASSIFICATION

The different images obtained with the supervised classification of the image with Median filter, and produced by changing the segment attributes and the number of classes, are shown: when the use of the training samples of the water class led to an excessive error in the classified image the number of classes was reduced from three to two, excluding the training samples of this last class; subsequently we report the result of applying the same combinations to the image with Lee Sigma filter.

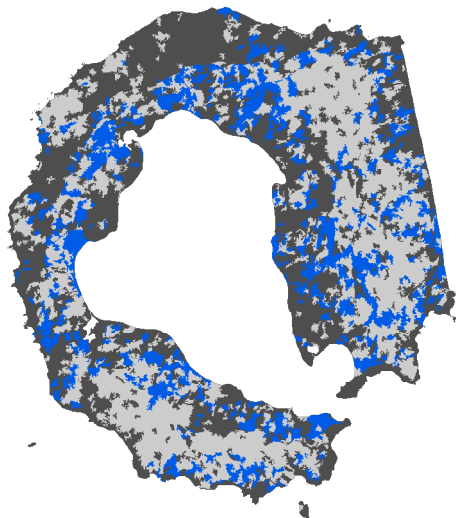
The images will be divided by classifier: Maximum Likelihood, Random Trees, Support Vector Machine, and in the cells of the tables the segment attributes used in the specific case will be indicated.

MAXIMUM LIKELIHOOD

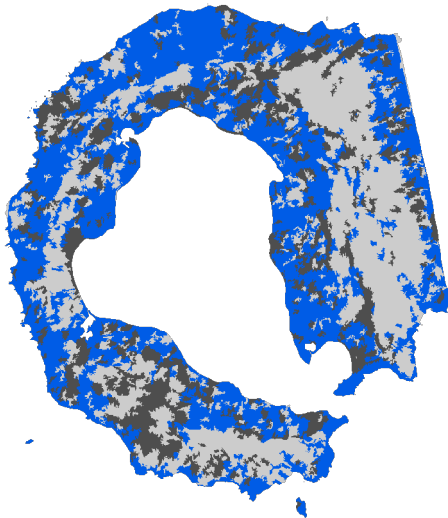
3 CLASSES, Image with Median filter



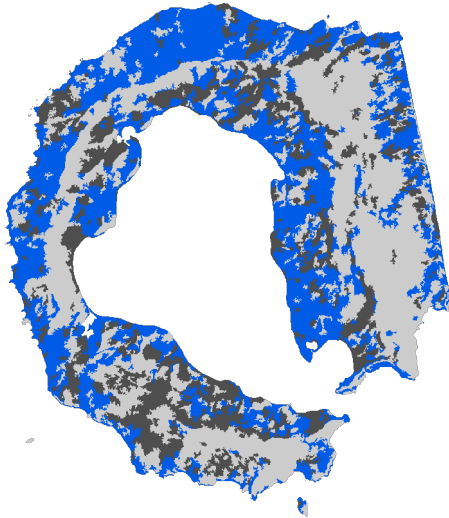
a. Standard Deviation



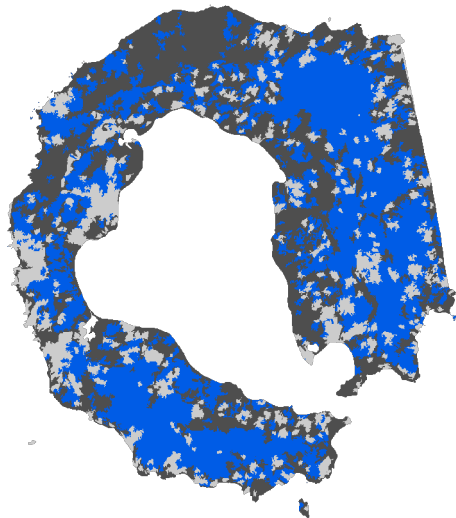
b. Active Chromaticity Color + Standard Deviation



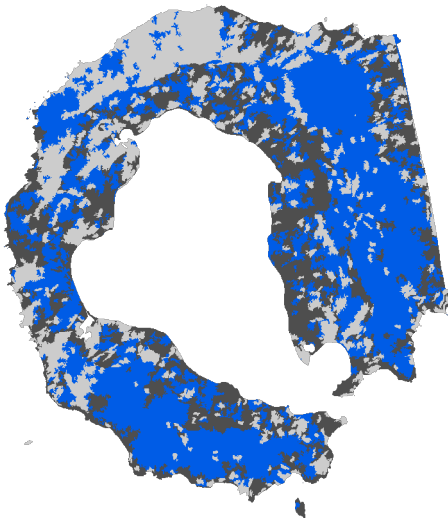
c. Mean Digital Number + Standard Deviation



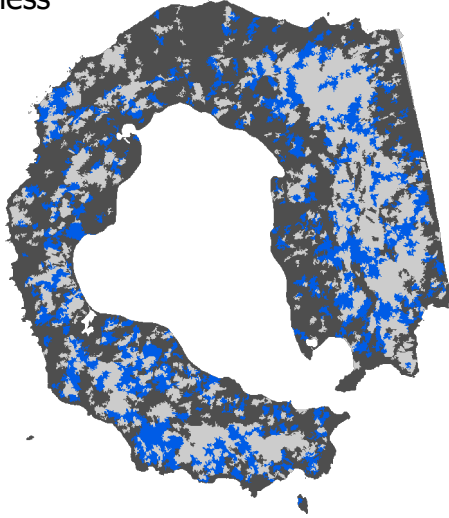
d. Standard Deviation + Compactness



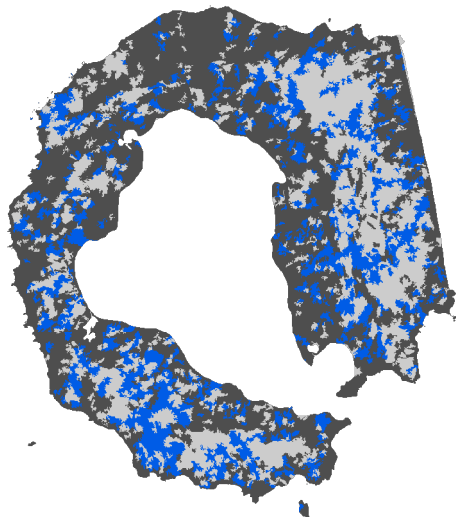
e. Standard Deviation + Rectangularity



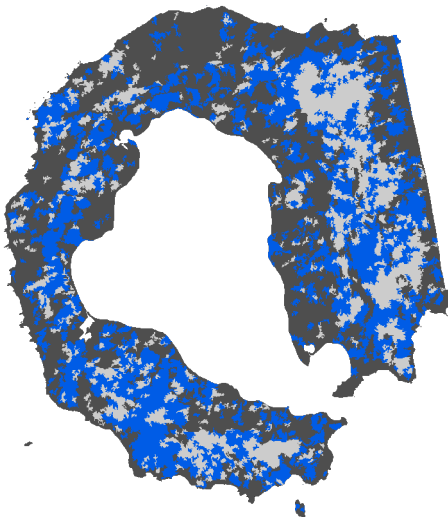
f. Mean Digital Number + Standard Deviation + Compactness



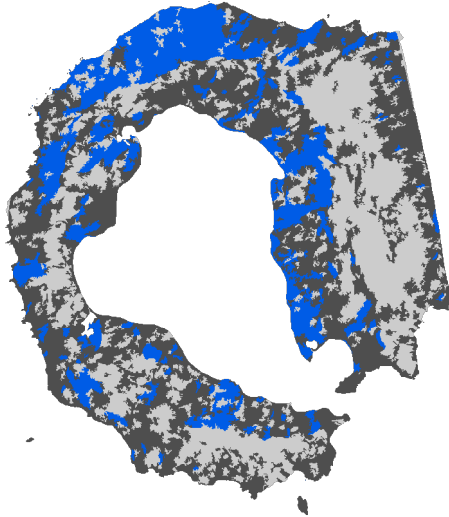
g. Active Chromaticity Color + Standard Deviation + Compactness



h. Active Chromaticity Color + Mean Digital Number + Standard Deviation + Compactness

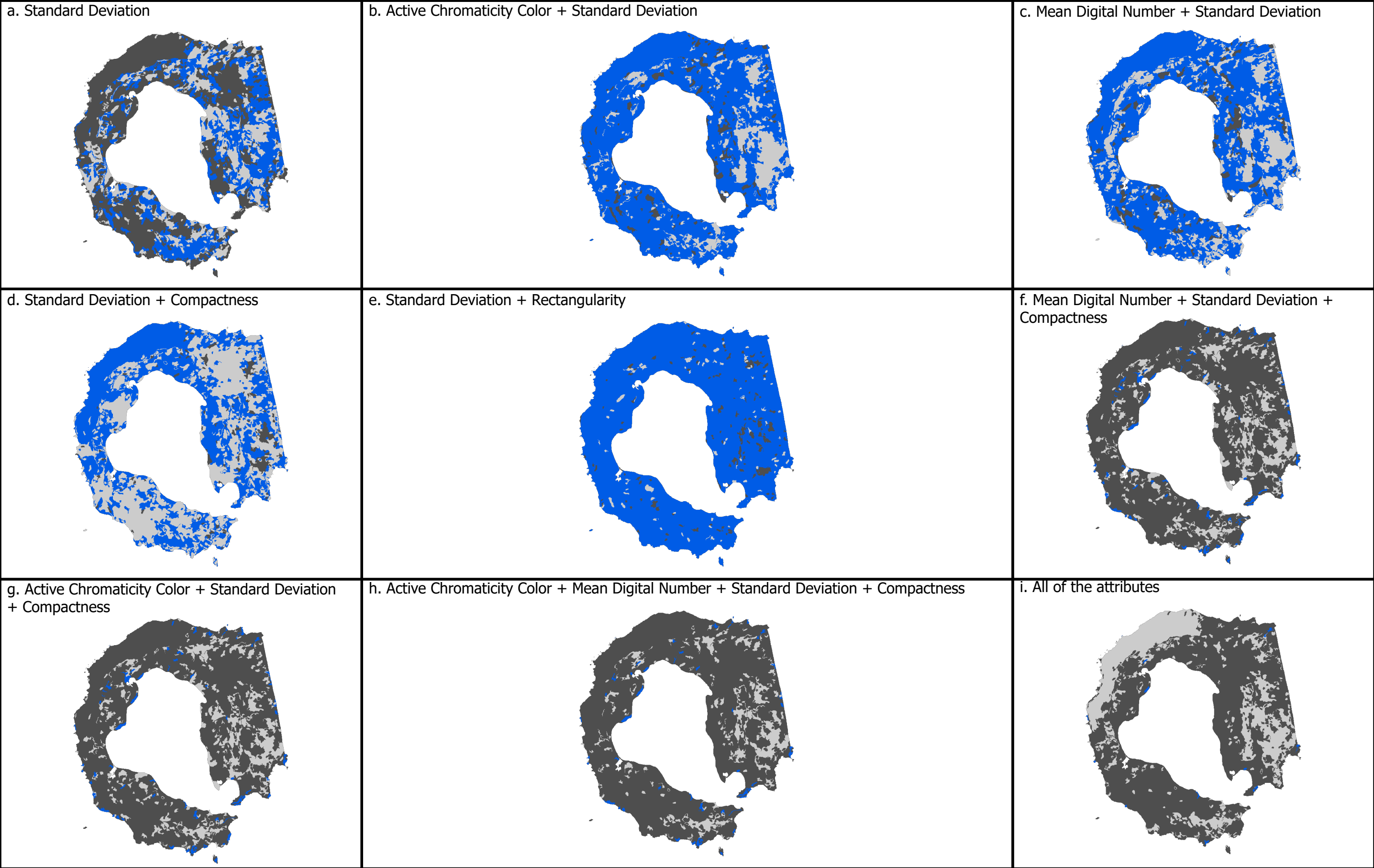
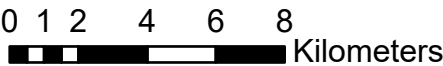


i. All of the attributes



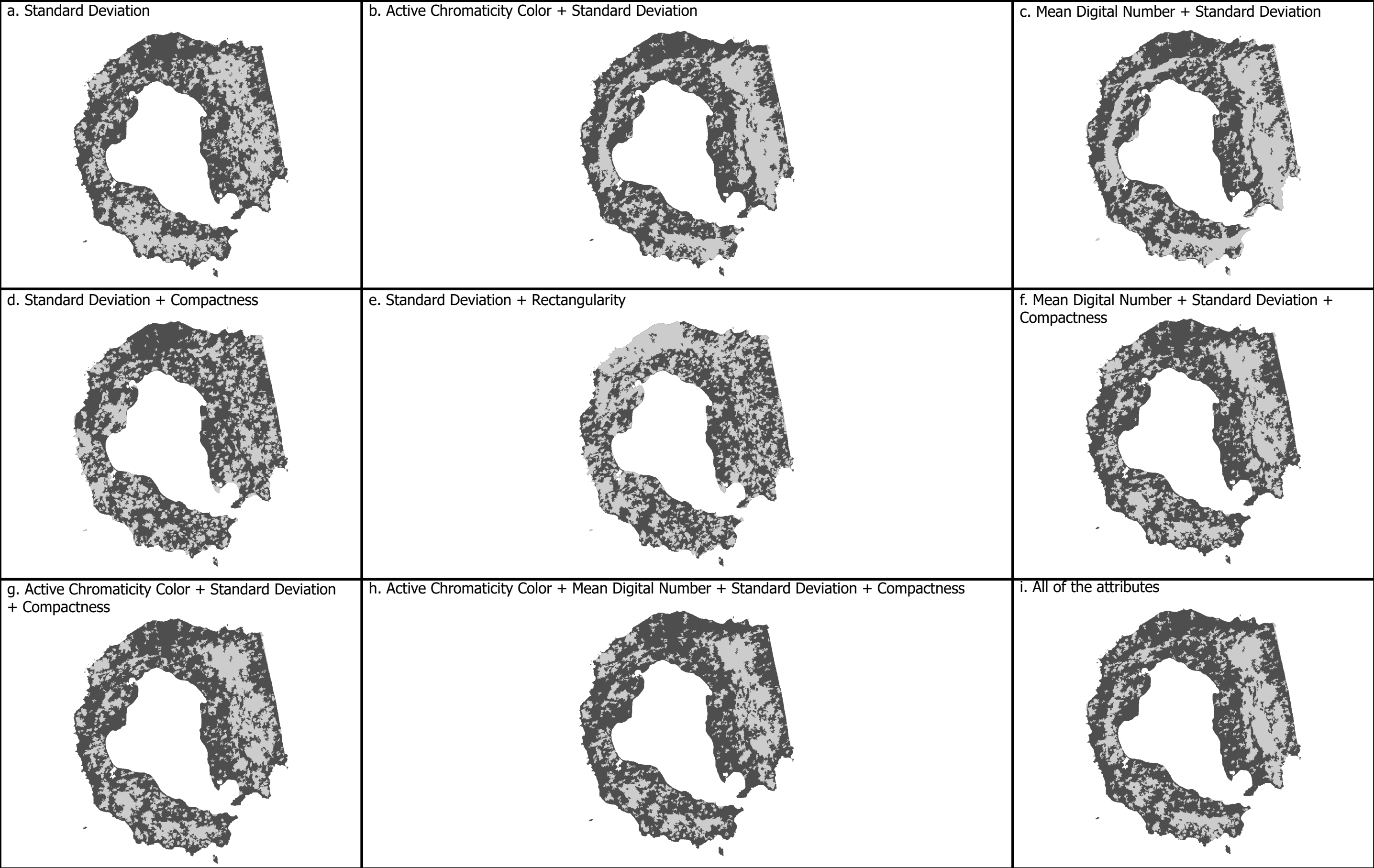
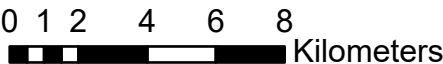
MAXIMUM LIKELIHOOD

3 CLASSES, Image with Lee Sigma filter



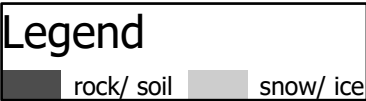
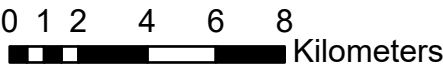
MAXIMUM LIKELIHOOD

2 CLASSES, Image with Median filter

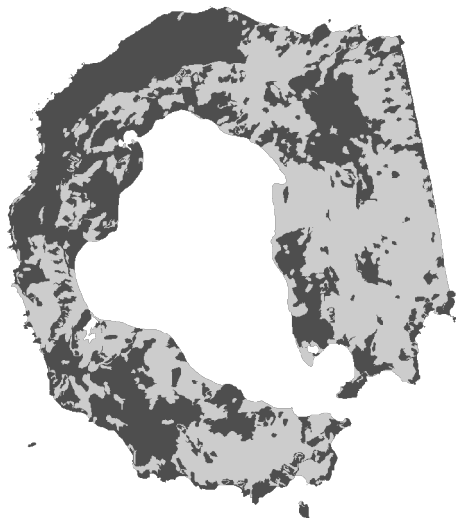


MAXIMUM LIKELIHOOD

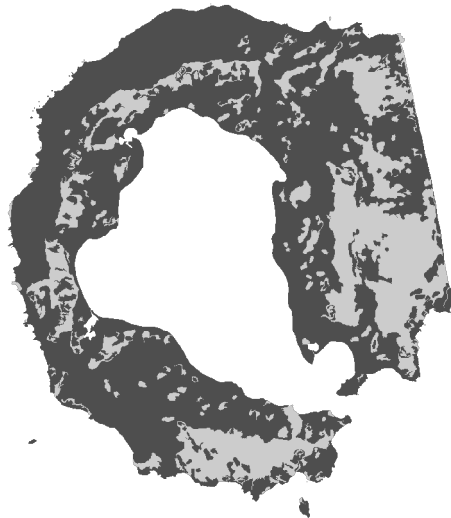
2 CLASSES, Image with Lee Sigma filter



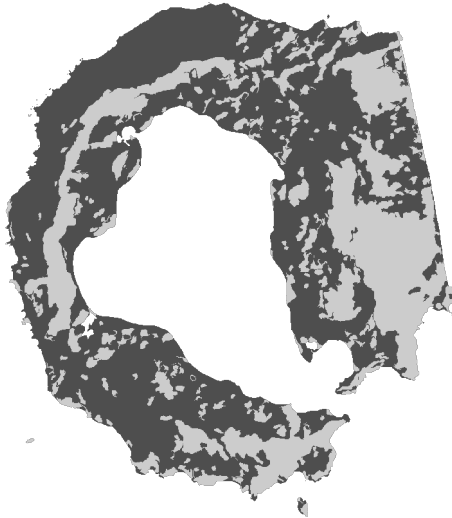
a. Standard Deviation



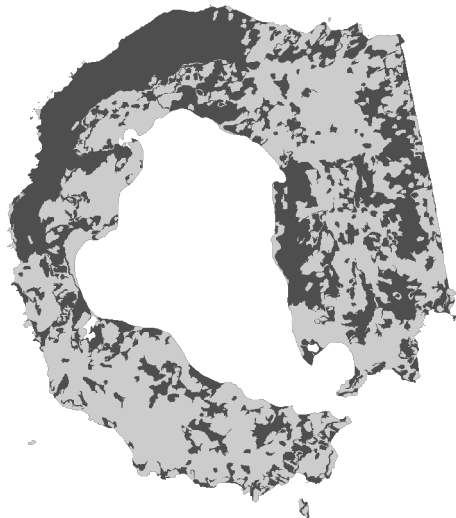
b. Active Chromaticity Color + Standard Deviation



c. Mean Digital Number + Standard Deviation



d. Standard Deviation + Compactness



e. Standard Deviation + Rectangularity



f. Mean Digital Number + Standard Deviation + Compactness



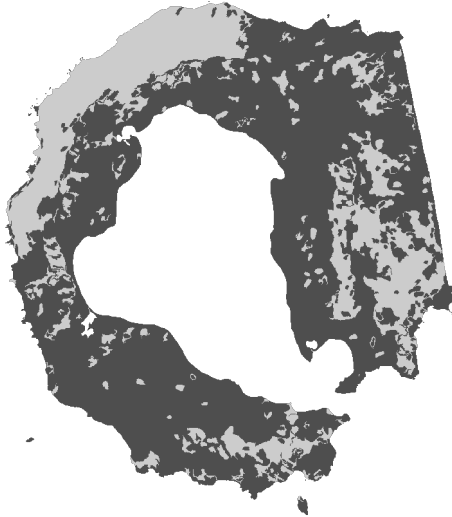
g. Active Chromaticity Color + Standard Deviation + Compactness



h. Active Chromaticity Color + Mean Digital Number + Standard Deviation + Compactness

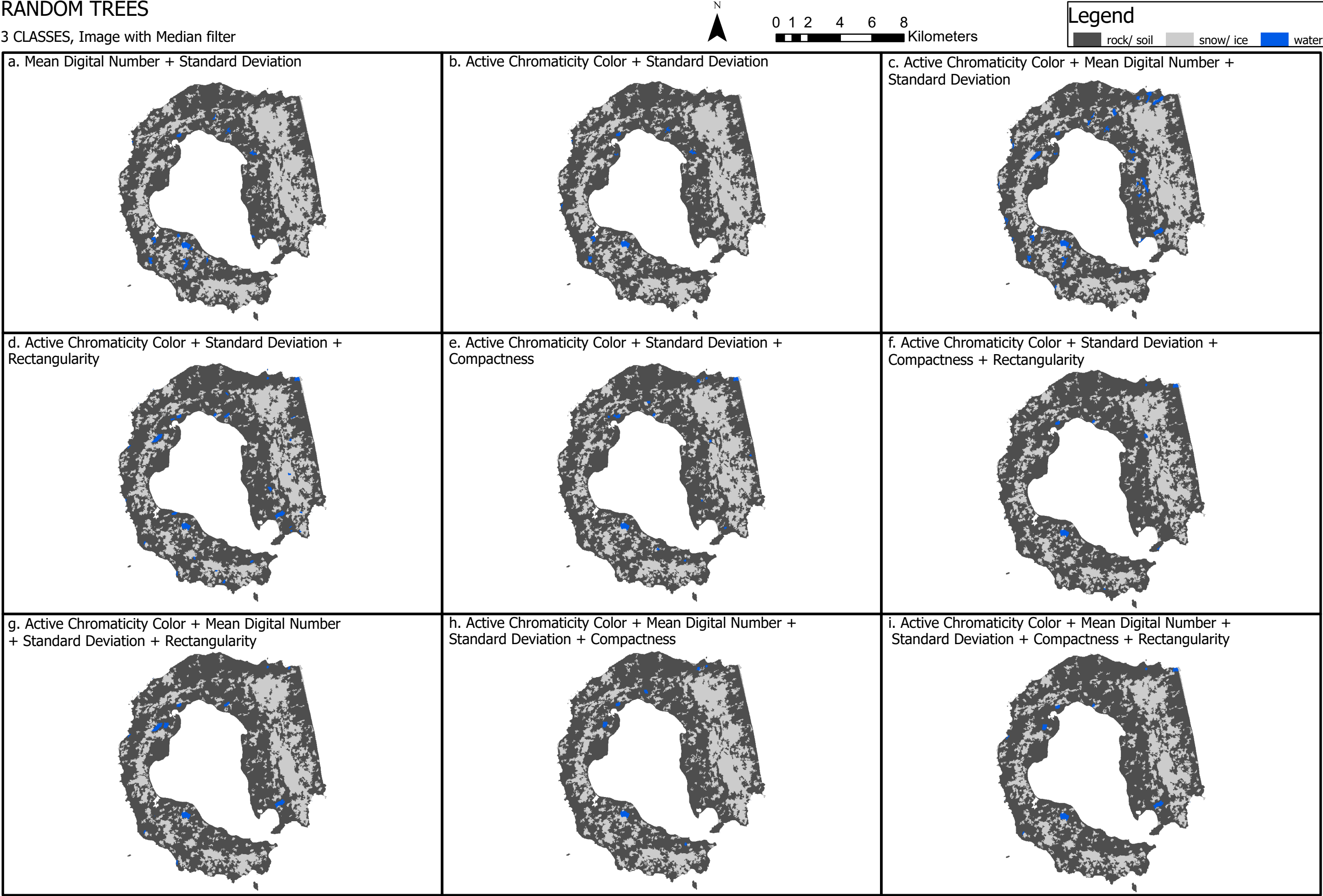


i. All of the attributes



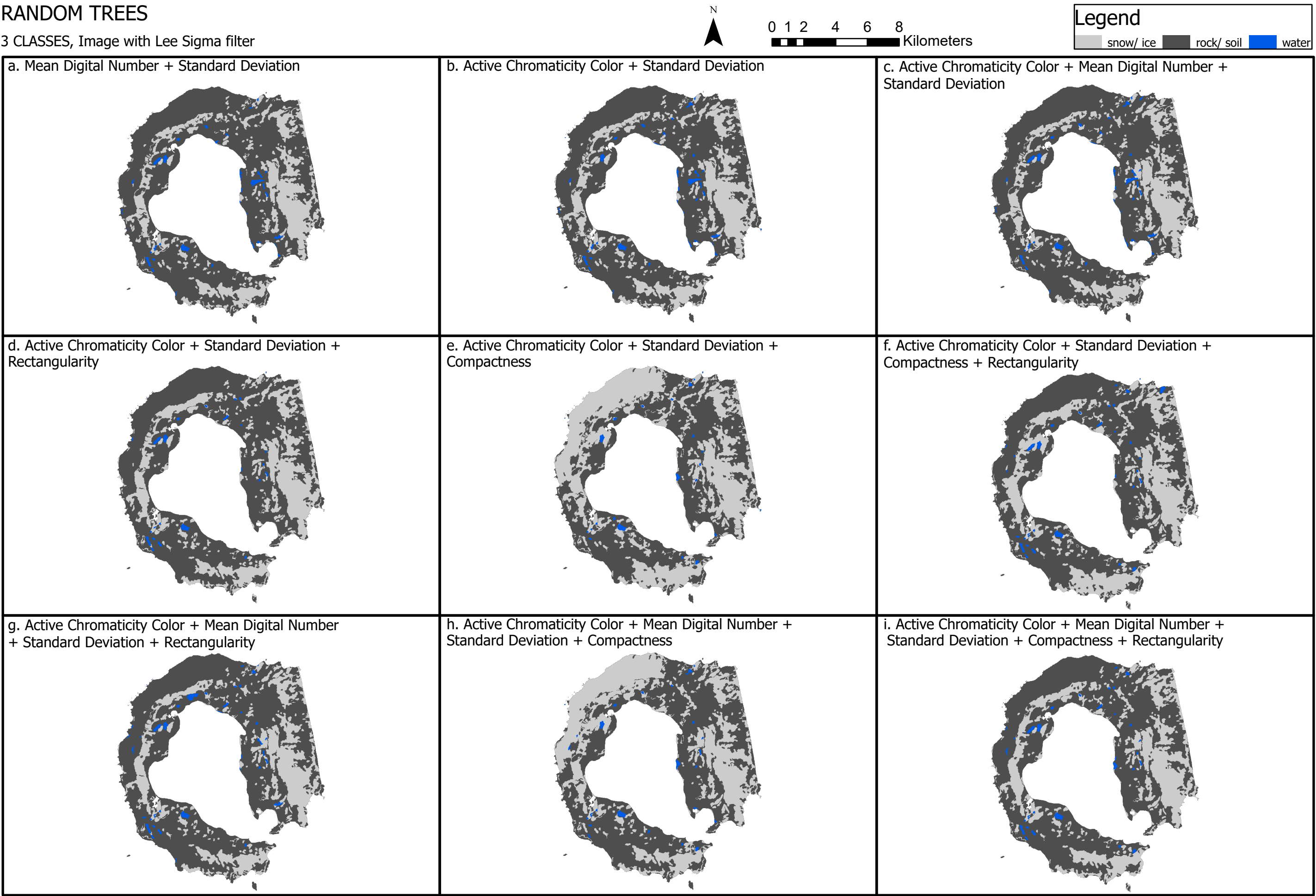
RANDOM TREES

3 CLASSES, Image with Median filter



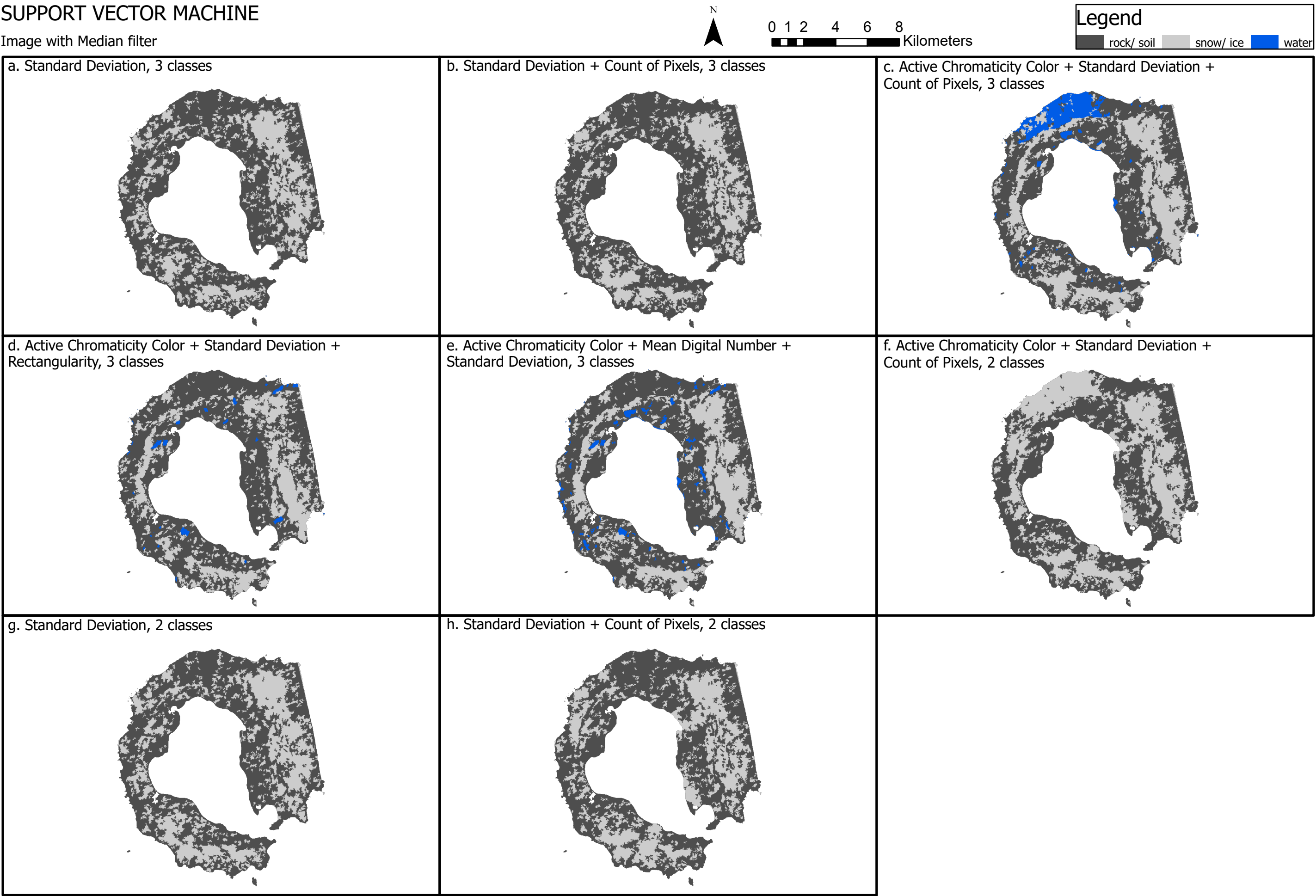
RANDOM TREES

3 CLASSES, Image with Lee Sigma filter



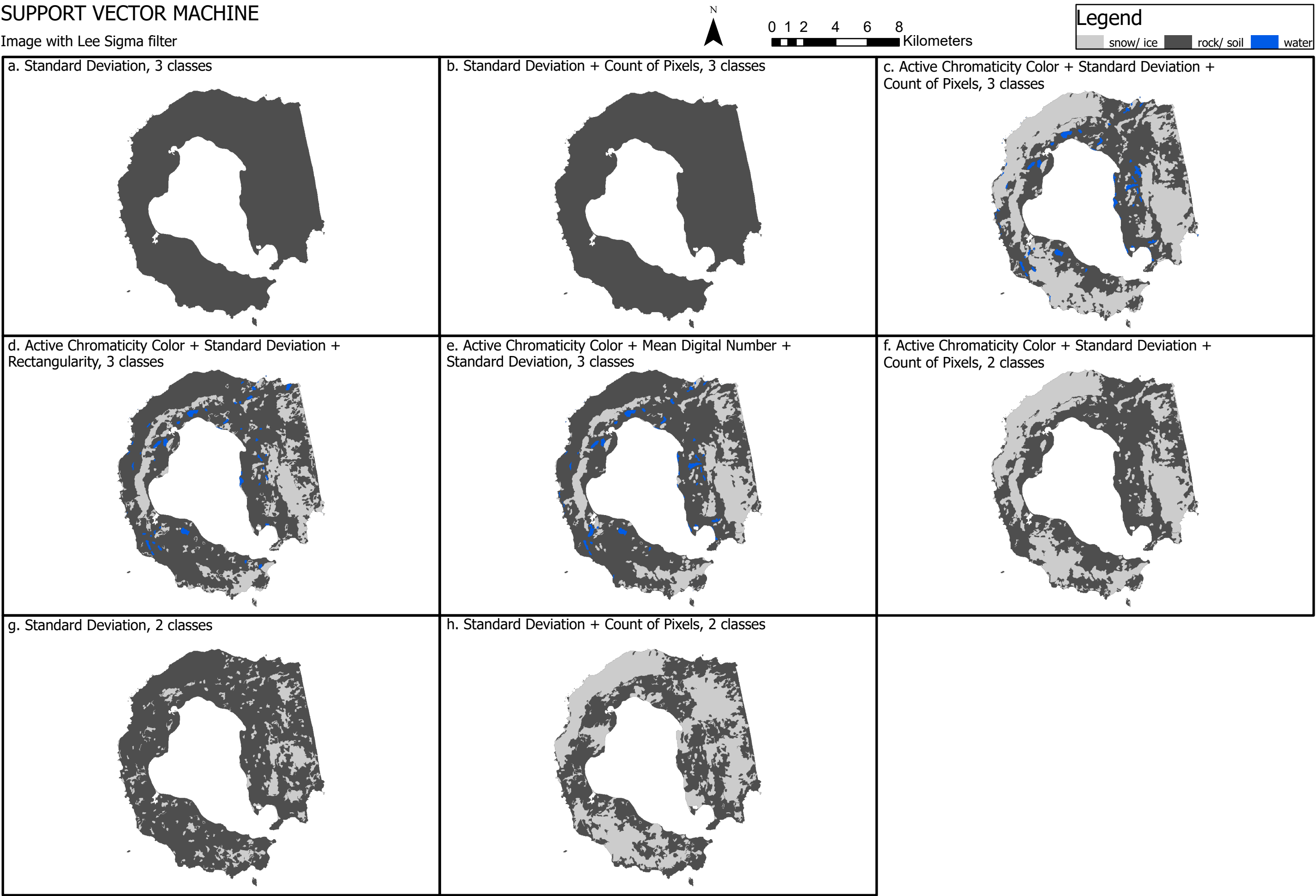
SUPPORT VECTOR MACHINE

Image with Median filter



SUPPORT VECTOR MACHINE

Image with Lee Sigma filter



8.6 OPTICAL

This section is divided into the three types of classification to which the optical image has been subjected. In addition, the index-based classification section shows not only the final classified images, but also the index images, the binary images obtained by applying limits to each index and the index histograms. The effect of applying the cloud mask is also shown each time.

8.6.1 UNSUPERVISED CLASSIFICATION

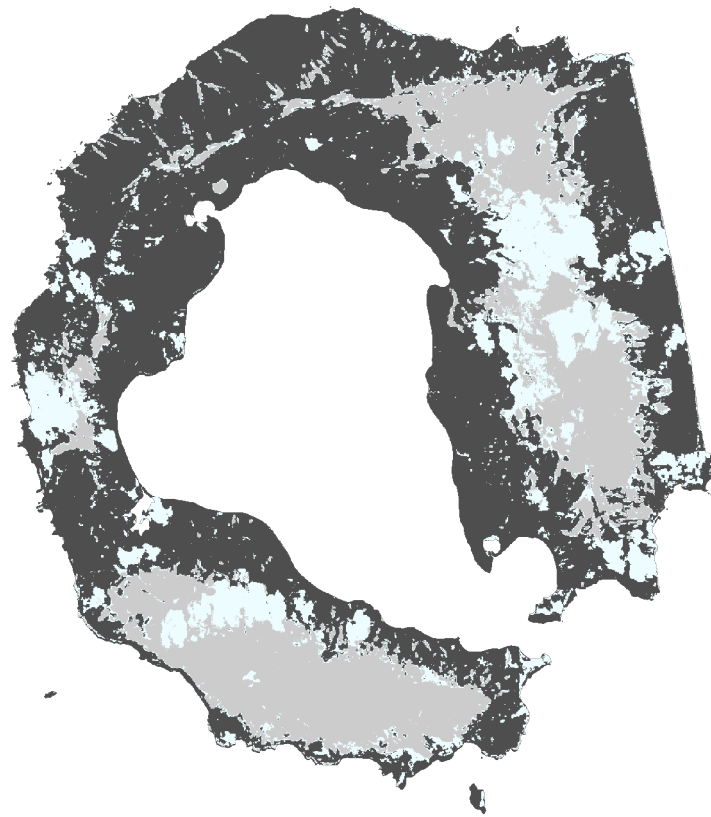
The image obtained with the unsupervised classification and the effect of applying the cloud mask are shown.



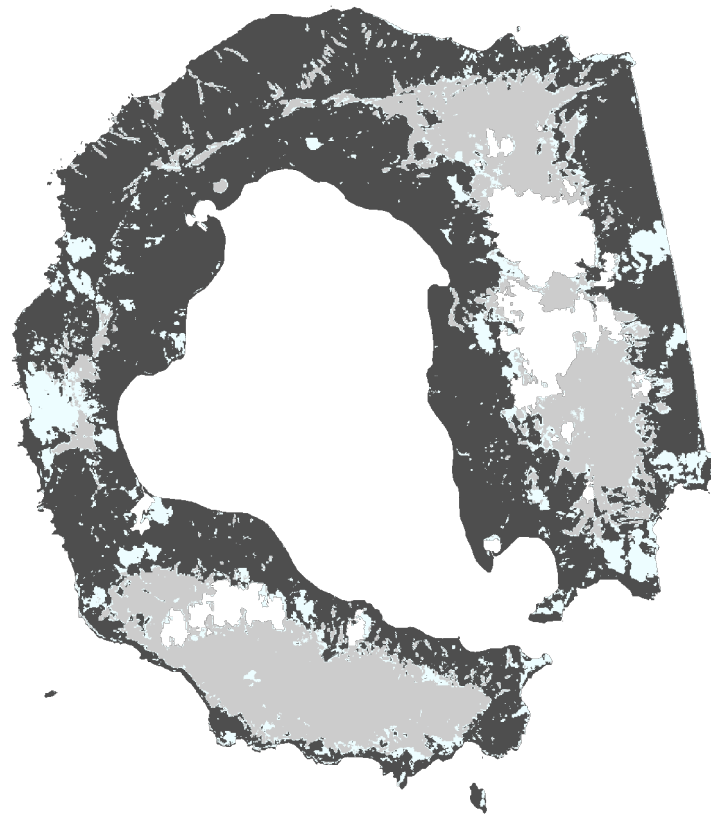
0 1 2 4 6 8 Kilometers

Legend
rock/soil snow/ice clouds

a. Classified image

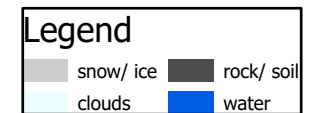
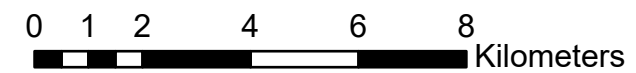


b. Application of cloud mask

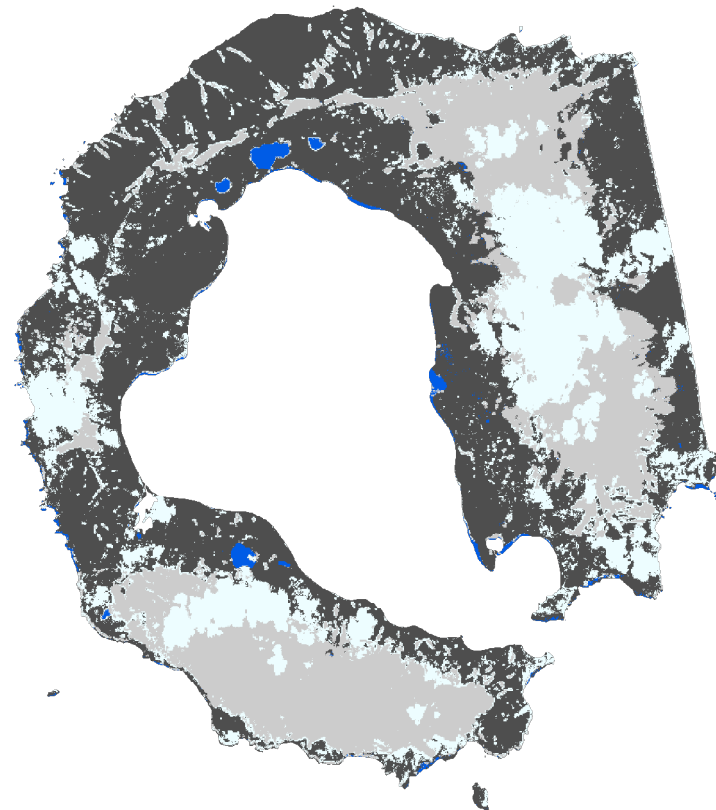


8.6.2 SUPERVISED CLASSIFICATION

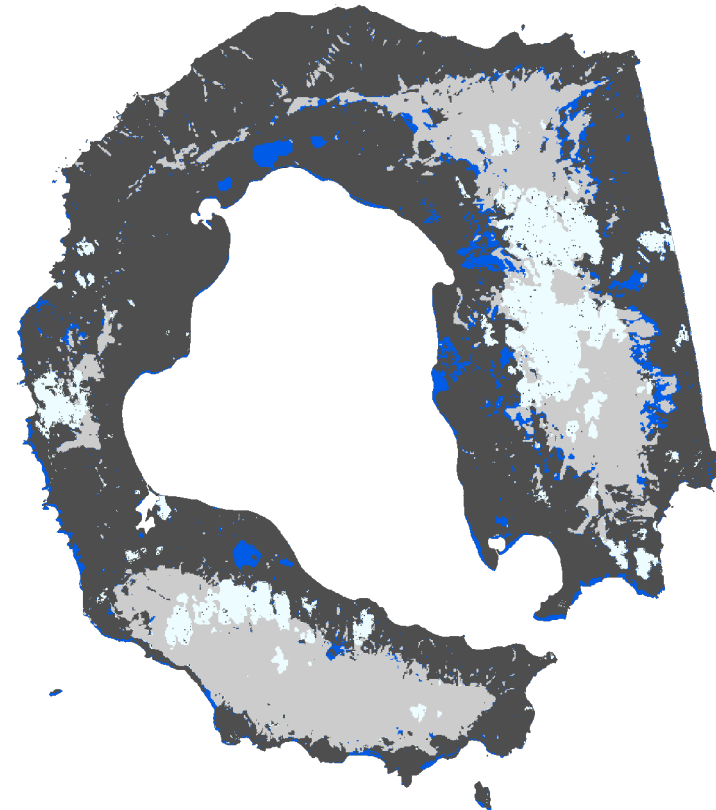
The images obtained with the three classifiers Maximum Likelihood, Random Trees and Support Vector Machine and then the application of the cloud mask to them are shown.



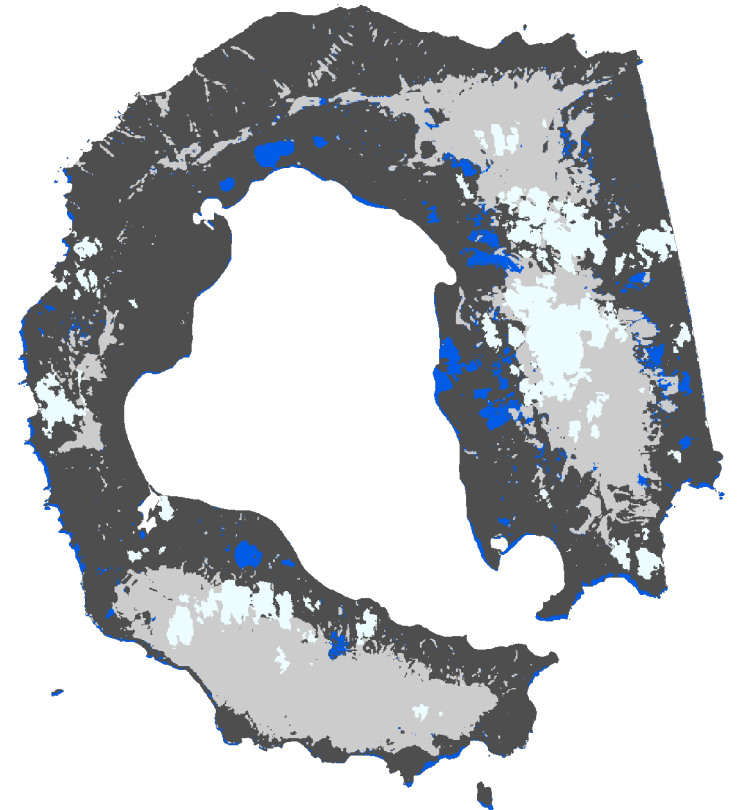
a. Maximum Likelihood



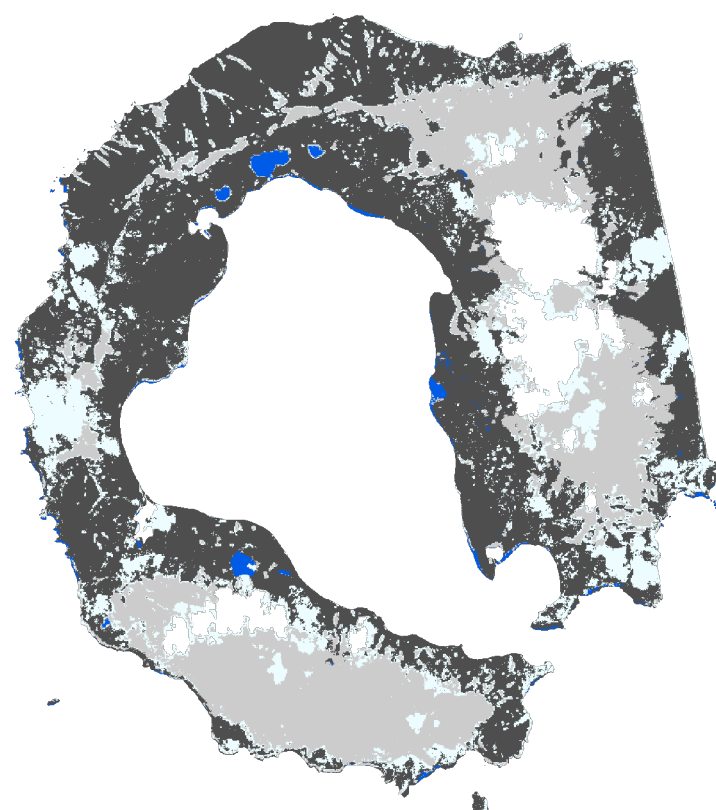
b. Random Trees



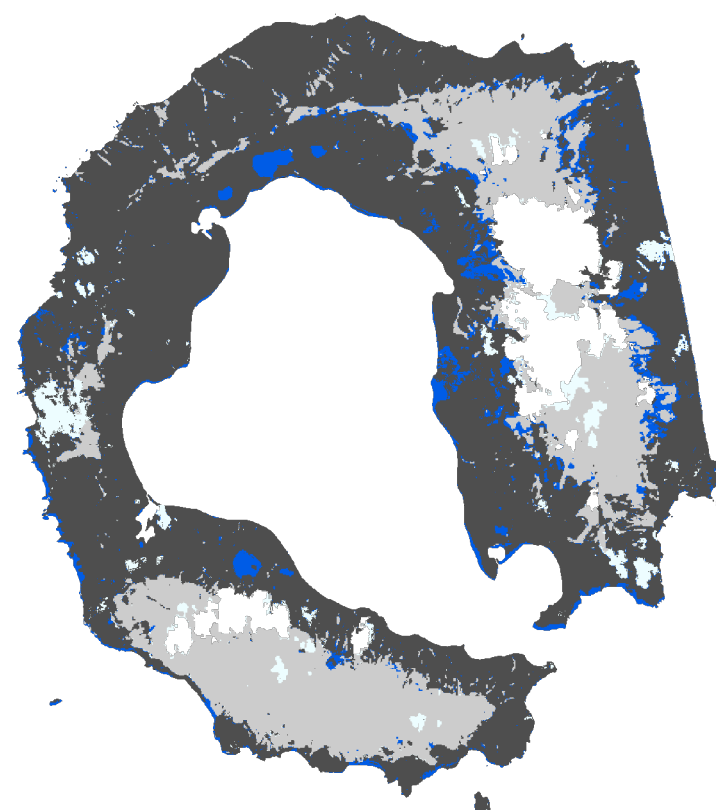
c. Support Vector Machine



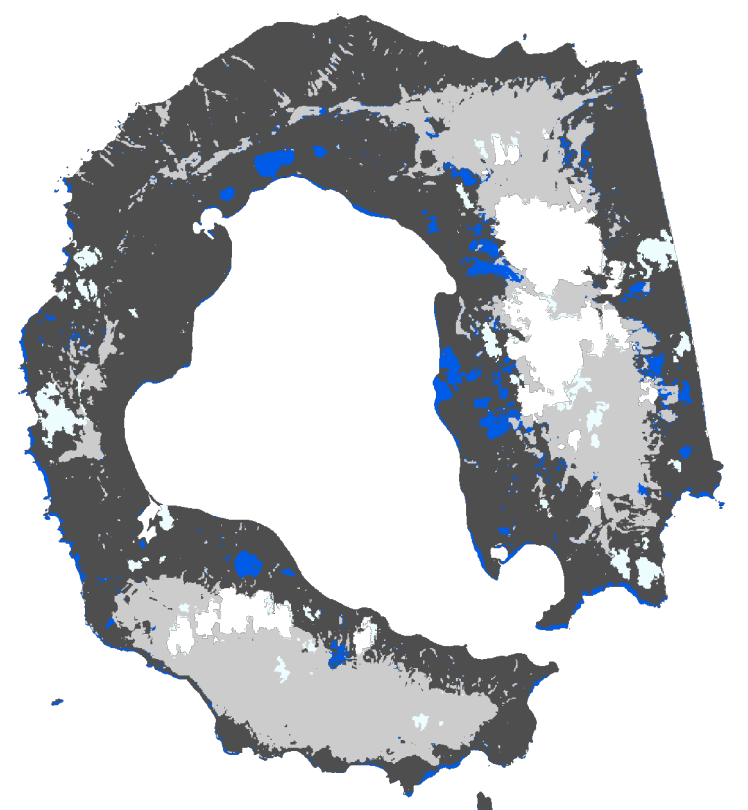
d. Maximum Likelihood + Cloud mask



e. Random Trees + Cloud mask



f. Support Vector Machine + Cloud mask

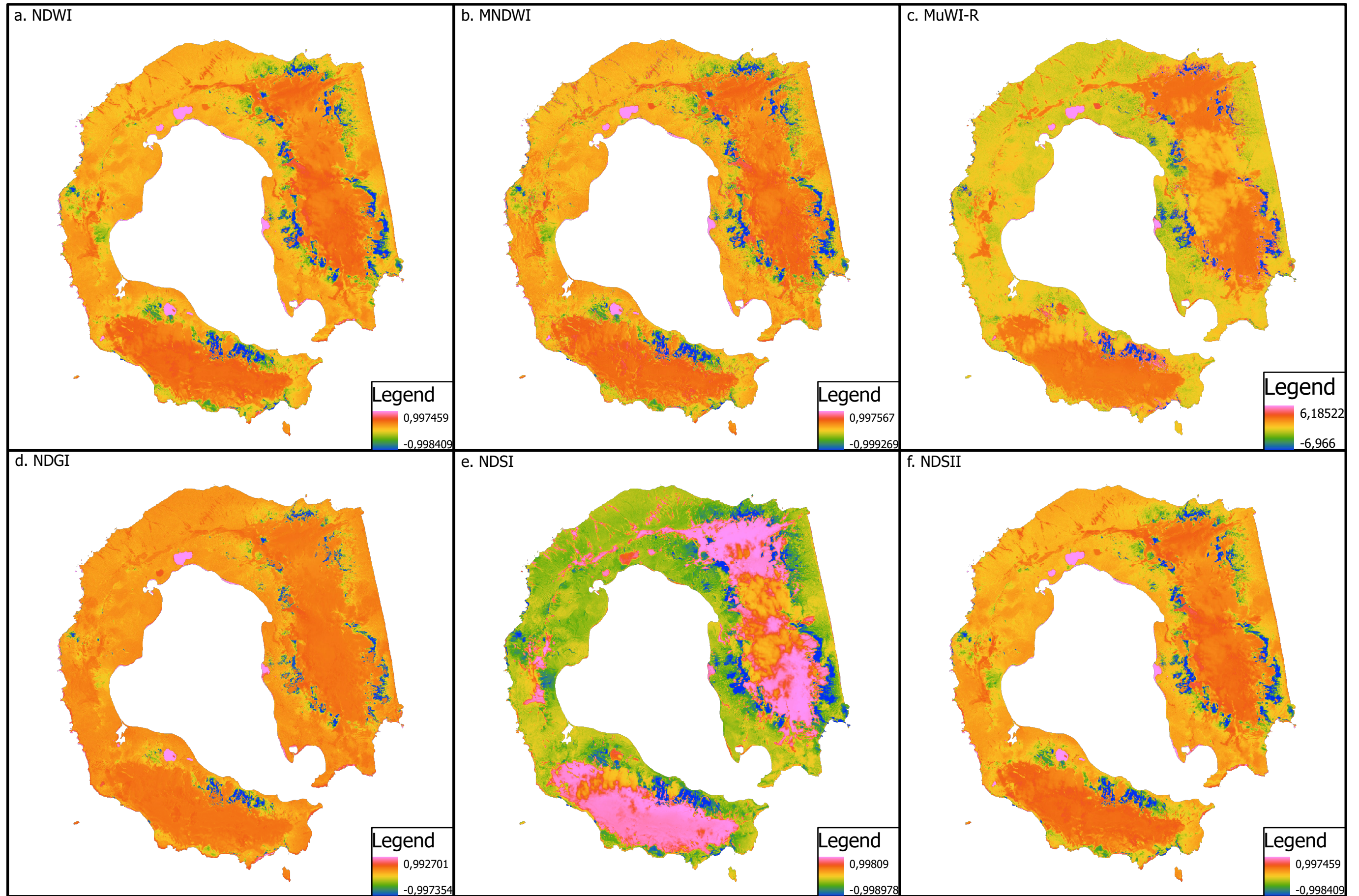


8.6.3 INDEX CLASSIFICATION

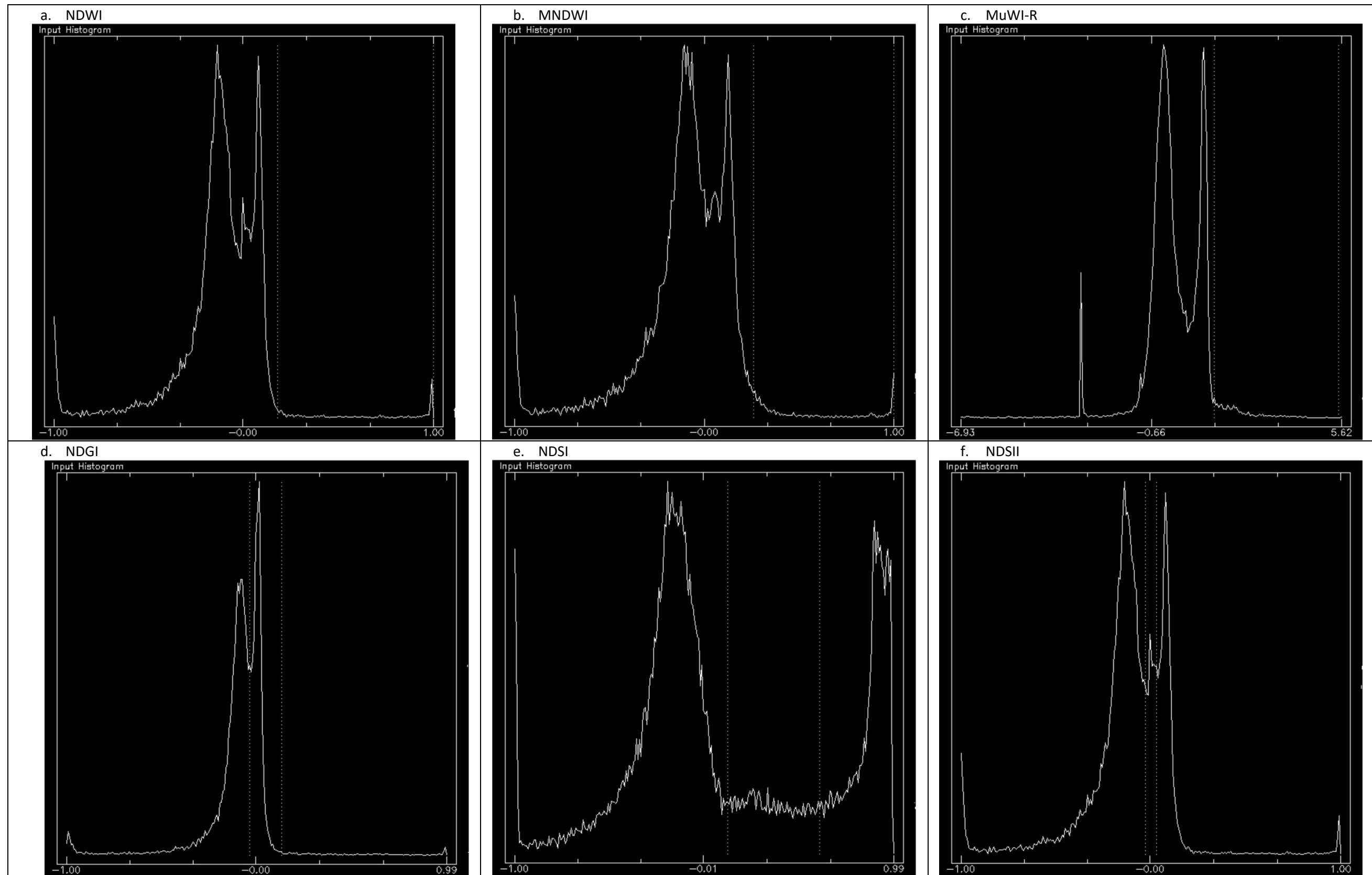
As anticipated in 8.5, the images of the indices, their histograms, the binary images obtained from the application of the limits identified for each and finally the final classified images, followed by the effect of the application of the cloud mask, are reported in this order.



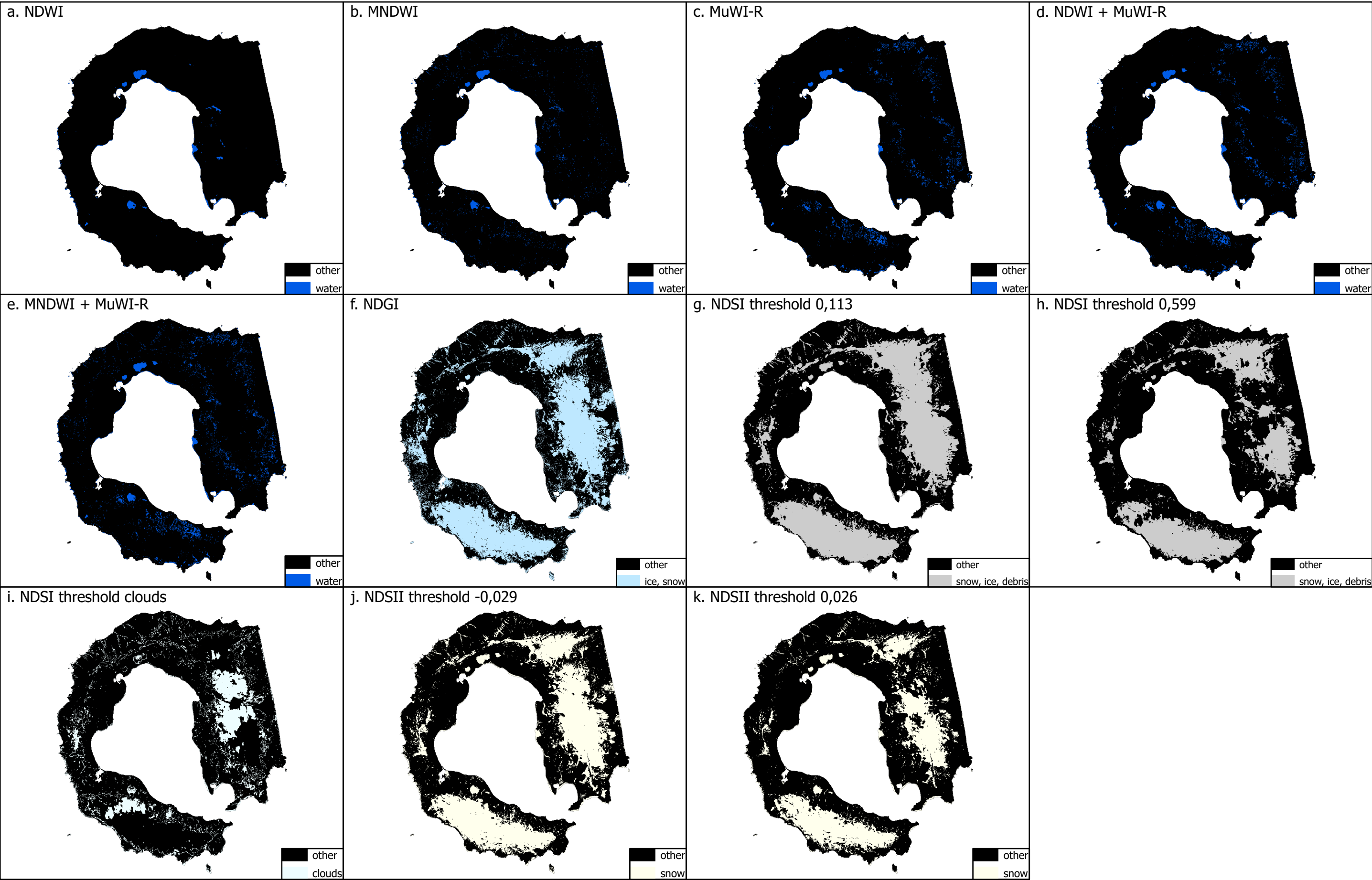
0 1 2 4 6 8 Kilometers



HISTOGRAMS



BINARY IMAGES



CLASSIFIED IMAGES

