BTA/…/….. **Construction of a satellite-imagery-based geomorphological maps of the Uyuni salt flat, Altiplano Basin, Bolivia.**

7-7-2014 **S.C. van der Graaf**

Abstract

The sedimentary systems terminating in the Salar de Uyuni, hold modern days analogues for the Permian and Triassic thin-bedded crevasse-splay deposits. In the future, these crevasse-splay deposits could be of importance for hydrocarbon production in Northwest Europe. In this study, the possibility of using satellite data to construct geomorphological maps is investigated. In order to do this, QGIS is used to determine the watershed and river paths based on ASTER GDEM satellite data, and to classify the land cover based on Landsat 7 ETM+ satellite data. The results generated by QGIS are compared to open source available datasets and high resolution Google Earth Pro images. The watershed and river paths determined in QGIS are accurate for high relief areas. In low relief areas, the results are not trustable. Here, the watershed and river path results seems to be controlled by the errors in the ASTER GDEM data. The resolution of the used Landsat 7 ETM+ is not large enough to distinguish various rock types accurately. A map that indicates the amount of clay within the basin could be created using the Landsat 7 ETM+ images though. Lithological classification using satellite images is a promising technique, with a lot of applications. However, to perform accurate lithological classification higher resolution satellite images and more detailed lithological information is necessary.

Content

Chapter 1. Introduction

The Northwest European gas province is a mature area in which the production of gas from conventional reservoirs is declining. Unconventional tough gas reservoirs in low-N/G stratigraphic intervals may constitute a secondary source of fossil energy to prolong the gas supply in the future. However, to date, production of these fine-grained, low-permeable reservoirs has been hampered by the economic risks related to the uncertainties in their size, shape, spatial distribution and reservoir properties.

Contemporary sedimentary systems terminating in the Uyuni salt flat, Altiplano Basin, Bolivia, hold excellent modern day analogues to Permian and Triassic thin-bedded crevasse-splay deposits in a semi-arid, low net-to-gross, lower coastal plain environment. Understanding the architecture of these systems helps to develop quantitative water- and sediment balance models, constraining the boundary conditions for the deposition of the before-mentioned sediment bodies.

This project aims to define the Salar de Uyuni watershed and to construct basin-scale geomorphological maps based on various satellite image data. In this project, watershed- and river path- maps are created in QGIS using digital elevation model. Furthermore lithological features are derived using the Landsat 7 satellite images. The results are compared to other open-source data, and validated using Google Earth Pro. The results of this project might help understanding the formation of possible reservoir rocks, that hold analogous to the reservoirs found in Northwest Europe.

The results will serve as a basis for the development of process-based physical models of thin-bedded fluvial deposition in a low-N/G floodplain environment. In combination with outcrop-analogue and subsurface-core studies, these models will allow building predictive static reservoir models of tough gas reservoirs. The ultimate goal of this study is to improve our knowledge of thin-bedded crevasse-splay deposits and to reduce the uncertainties in the development of tough gas reservoirs.

Chapter 2. Geological Setting

The Salar de Uyuni, the largest salt pan in the world, is located in the southwestern part of Bolivia (Figure 1). The salt lake is located in the Altiplano, part of the central Andes, and has an average elevation of around 3653 m above sea-level. The basin in which the lake is found is bordered by the Cordillera Oriental and the Cordillera Occidental mountain chains. This region of the Altiplano is dominated by volcanics of the Central Volcanic Zone [Tibaldi, 2008].This volcanic arc is formed by the subduction of the Nazca-plate under the South-American continent. Most of the volcanics in this area are andesitic of composition and Cenozoic of age.

The lake is a small remnant of a much larger lake that existed more than 10.000 years ago in the Pleistocene. The lake had water levels ranging up to 75 meters above its current level, and extended more than 100 km to the north of its current position. Today the main inflow comes from the Rio Grande de Lipez to the south of the lake. The salt in the lake is composed mainly of halite and gypsum, but also possesses economic quantities of lithium, potassium and boron minerals. [Orris, 1995]

Figure 1: The location of the Salar de Uyuni in Bolivia and South America.

Chapter 3. Data and Methods

There are multiple open source resources providing datasets regarding the hydrology and lithology of the Salar de Uyuni basin. In this chapter, the datasets and methods used in this study are described. In §3.1 the used datasets are covered. In §3.1.4. the data from Google Earth Pro is covered. In §3.1.2. the datasets provided by GeoBolivia are described. In §3.1.3. the datasets provided by the United States Geological Survey are described. §3.1.4. is about the used Aster GDEM data provided by NASA. In §3.2 the used methods are covered. In §3.2.1 is described how Google Earth Pro is used in this study. In §3.2.2 the used functions and tools in QGIS are described. Paragraph §3.2.3 describes how reliably land cover and rock types can be recognized using the Landsat ETM+ images.

§3.1.1. Google Earth Pro data

Google Earth Pro is the commercial version of Google Earth. Unlike the freeware version of Google Earth, Google Earth Pro is able to import GIS images and data. Google Earth Pro images are not only acquired from satellite images, but from superposition of satellite images and aerial photography. In most densely populated areas the resolution of Google Earth Pro lies within the decimeter scale. Unfortunately, the resolution is poorer in scarcely populated areas, because less aerial photographs are recorded. In the area of interest the resolution is approximately 1 meter. This resolution is more of less 900 times better than both the Landsat and Aster GDEM images, because these images consist of 30x30 meter resolution.

Google Earth Pro is equipped with a digital elevation model (DEM), which enables it to see height differences on the map and create profiles. The DEM included in Google Earth Pro is derived from the Shuttle Radar Topography Mission (STRM) elevation data, acquired by the National Aeronautics and Space Administration (NASA), that have been improved with supplementary DEM data to fill gaps in the SRTM data. [Google Inc, 2014]

§3.1.2. GeoBolivia datasets

GeoBolivia is a project established by the Bolivian government. The aim of this project is to store and distribute geographical data concerning Bolivia. On the website of the GeoBolivia project, geographical datasets can be downloaded. River network maps and watershed maps of different scales are used in this study. Also a geological map provided by GeoBolivia is used for better understanding of the geology of the area. The data from GeoBolivia is downloadable from *geo.gob.bo*.

§3.1.3. USGS datasets

The United States Geological Survey (USGS) is a scientific agency of the United States government. The USGS provides multiple open source datasets, that can be downloaded from the website. In this study, interpreted river network- and watershed maps, and also satellite images (Landsat ETM+) from the USGS are used. The river network- and watershed maps are downloaded from *hydrosheds.cr.usgs.go*, the Landsat ETM+ satellite images are downloaded from *earthexplorer.usgs.gov*.

§3.1.3.1. HydroSHEDS

The river network and watershed maps from the USGS are produced using HydroSHEDS. HydroSHEDS (Hydrological data and maps based on Shuttle Elevation Derivatives at multiple Scales) is a mapping product developed by the World Wildlife Fund (WWF) that provides hydrographic information based on SRTM elevation data acquired at 1 arc-seconds resolution (approximately 90 meters at the equator). Before inputting the data in HydroSHEDS, 3 arc-seconds data is generated by averaging the corresponding 3x3 kernels of the 1 arc-second data. The HydroSHEDS program calculates a conditioned elevation map, a drainage direction map and a flow accumulation map. Using these results, a watershed map and river network map are created. In this step the resolution is coarsened to reduce the calculation time. Therefore, the best available river network and drainage basins files have a resolution of 15 arc-seconds (approximately 500 meters at the equator) instead of 3 arc-seconds. [Lehner, 2013]

§3.1.3.2. Landsat 7 ETM+ data

Landsat 7 Enhanced Thematic Mapper Plus (ETM+) data is used to study hydrological and lithological features of the Salar de Uyuni Basin. Landsat 7 is the seventh satellite of the Landsat program. The Landsat program is the longest running project for acquisition of satellite images of the surface of the Earth and is operated by NASA and USGS. The main goal of the program is to refresh the global satellite images archive.

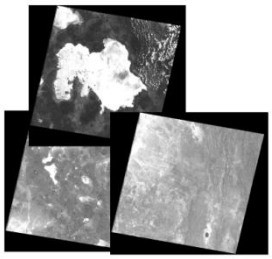

Figure 2: The raw data (band 1) scenes in QGIS.

Landsat 7 ETM+ images consist of eight spectral bands with a spatial resolution of 30 meters for bands 1 to 7. The resolution for band 8 is 15 meters. The approximate scene size is 170 km north-south by 183 km east-west per scene. Landsat scenes are multispectral photographs of a certain region acquired at a certain time. The used scenes can be seen in Figure 2. The name, wavelength and resolution of each band are given in Table 1.

Band	*Name*	*Wavelength (µm)*	*Resolution (m)*
1	Blue	0.45-0.52	30
2	Green	0.52-0.60	30
3	Red	0.63-0.69	30
4	Near Infrared	0.77-0.90	30
5	Shortwave IR-1	1.55-1.75	30
6	Thermal IR	10.40-12.50	30
7	Shortwave IR-2	2.09-2.35	30
8	Panchromatic	0.52-0.90	15

Table 1: Band designations for Landsat 7 ETM+ satellites. [Sabine, 1999]

The Landsat 7 ETM+ images scenes used are acquired between September 2002 and March 2003. The images are available for free on the USGS website. The images have been processed to Standard Terrain Correction Level 1 (L1T). This correction provides radiometric and geometric accuracy by incorporating ground control points, and also uses a digital elevation model (DEM) to provide topographical accuracy.

The geometric accuracy of Landsat 7 ETM+ data is at least 250 meters. However, definitive ephemeris is used for the Level 1 correction of the Landsat 7 ETM+ images, which increases the geometric accuracy to within 30-50 meters. Ephemeris data shows the position and velocity of the satellite at the time the images were taken. Definitive ephemeris data is measured in one-minute intervals, and thus allows more accurate geometrical correction for the Landsat 7 ETM+ images. [USGS, 2014]

§3.1.4. ASTER GDEM data

The Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) on the National Aeronautic and Space Administration (NASA) spacecraft Terra is capable of collecting in-track stereo using near infrared cameras. The stereo pairs have been used to produce global digital elevation models (GDEM) with a scene-size of approximately 60x60km. The data has a spatial resolution of 30x30 meters.
In this study, ASTER GDEMis used to determine the river paths and watershed in QGIS. The ASTER GDEM images used are downloaded from *reverb.echo.nasa.gov*. The newest version available, version 2, is used in this study. In total, eleven ASTER GDEM scenes are used in this study.

The ASTER GDEM version 2 has a horizontal resolution of 2.4 arc seconds, this is equal to 72 meters. Validation research has found that the ASTER GDEM has a mean vertical error of -0,20 meter, and an additional error based on the type of land cover. The vertical elevation error varies from -2.3 meter to 5 meter. Both the horizontal and vertical resolution are similar to those of the SRTM data used in HydroSHEDS. The average difference between the ASTER GDEM version 2 and the SRTM data ranges between -5 and -9 meters. [Meyer, 2011]

§3.2.1. Google Earth Pro

The watershed of the whole basin and the sub-basins are determined using Google Earth Pro. To get a better understanding of the watershed and the separate river systems within the basin, this was done before the proceedings in QGIS.

Around the Salar de Uyuni there are multiple high relief areas, such as mountain ranges and volcanos. The main borders of the watershed can be visually interpreted in high relief areas. In order to clarify the location of the border, the DEM height is exaggerated three times.

At the west and southwest side of the Salar de Uyuni lake there are many small (salt) lakes. For some of these lakes it is clear whether these would overflow into the Salar the Uyuni or not. For the remaining lakes, height profiles are created. Multiple lines are drawn where the lakes appears to overflow into the basin. For all these lines profiles are created. Since the area around the Salar de Uyuni is very dry, small height differences (a few meters) are enough to exclude these lakes and salt flats from the basin watershed.
Several profiles are created in flat areas, for example the large flat northeast of Salar de Uyuni, separating it from the smaller Salar de Coipara. Since most of the river paths are invisible or absent here, the border of the watershed follows the highest points in the profiles.

The total basin is divided into sub-basins for separate river systems. These sub-basins are visually interpreted by looking at the river paths and the location of mountains and volcanos. Because the relief is low close to the Salar de Uyuni, the borders between the separate systems are sometimes undistinguishable. In these areas straight lines were drawn towards the Salar de Uyuni.

The available and calculated watershed datasets are compared to the watershed and sub-basins created in Google Earth Pro afterwards. Also the available and calculated river paths are compared to Google Earth Pro images. In Google Earth Pro, all river paths greater than approximately 1 meter are visible. In this study, the results from Google Earth Pro are considered ‘ground truth’, because these images have by far the highest resolution and thus the highest confidence level.

§3.2.2. QGIS

QGIS is a free and open source geographic information system (GIS), that has applications to visualize, edit, and analyze geographical data. In this study, the newest version, QGIS 2.2 Valmiera, is used. This version can be downloaded from *www.qgis.org*. Other open source GIS packages, such as GRASS GIS, are integrated in QGIS and thus functions created by these packages can be used with QGIS. Together with optionally installed packages, the tools and functions in QGIS are similar to those in the licensed version of ArcGIS.

Besides standard viewing, editing and analyzing tools, a few advanced functions and plugins are used in this study. For determination of hydrological features using the DEM data, two main functions are used: *r.watershed* and *r.terraflow*. For an atmospheric correction on the Landsat ETM+ data, the *i.atcorr* function is used. These functions are all created by GRASS GIS and included in QGIS. For image classification, the *semi-automatic classification plugin* is used.

§3.2.2.1. r.terraflow

The flow paths are calculated in QGIS using the *r.terraflow* function created by GRASS. *r.terraflow* computes the flow direction raster taking the DEM raster as input. The flow directions are computed using either the SFD (single flow direction or D8) model or the MFD (multiple flow direction) model. Both methods inspect a 3x3 window around every cell. The SFD method assigns the flow direction to the one steepest neighbor, while the MFD method assigns the flow direction to all downslope neighbours (Figure 3).

Figure 3: Example of flow routing for SFD (left) and MFD (right). The numbers indicate the height for the different pixels, the red arrows show the routing of the flow to downslope neighbours for both methods.

Both the SFD and the MFD method cannot compute flow directions for cells in flat areas, where cells have the same height as all their neighbours or do not have downslope neighbours. The function *r.terraflow* has a different way of handling flat areas that spill out (plateaus) and flat areas that don't spill out (sinks). On plateaus, the flow tends to go towards the edges of the plateau, the spill cells. On sinks, the function assigns the flow by flooding the terrain until all the sinks are filled and then assigning the flow direction (Figure 4).

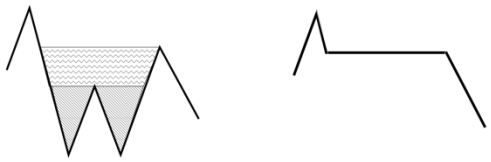

Figure 4: Filling sinks until overflow to form a filled terrain.

Routing the water using the created flow direction raster, *r.terraflow* calculates the flow accumulation raster. The flow accumulation raster represents all the river paths in the area. Within the catchment area of the basin, the rivers paths are calculated using both the SFD and MFD method. Because the whole Aster GDEM dataset was too large to process at once, the dataset had to be divided into approximately 20 subparts for each method. Unfortunately, some errors were created on the edges of the subparts during this process. An example of such error can be seen in Figure 6. To minimise these errors, the subparts were kept as large as possible. [Toma, 2001]

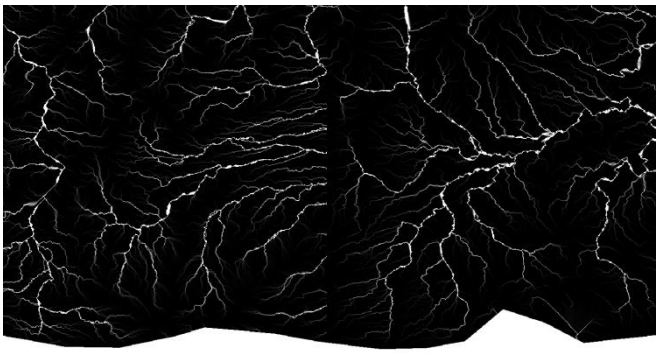

Figure 5: An example of an error in the SFD results caused by subdivision of the dataset.

§3.2.2.2. r.watershed

The *r.watershed* function makes use of the A* (or A^T) least-cost search algorithm to find the best stream paths from the DEM-data. In this algorithm, all possible flow-paths are inspected and the 'cost' of traveling along this path is determined. The path with the lowest cost to travel is chosen as the ideal flow-path. The least-cost search algorithm is similar to the SFD algorithm in *r.terraflow*. The flow path map created is used to generate the flow accumulation and watershed map. [Ehlschlaeger, 1989]

In the function, a threshold value for the watershed area has to be entered. This value is the amount of cells that is the minimum catchment size. Since the Aster GDEM data has a resolution of 30 x 30 meters, each cell

includes an area of 900 m². So, for example, if one wants to create catchments of at least 10 km², a threshold value of 10 km²/ 900 m² has to be entered in the function. In this study, a threshold value of 50 million pixels is used to calculate the watershed of the whole basin is. In order to calculate the subbasins within the basin, threshold values of 1, 2, 3 and 4 million are entered.

In order to create more accurate watersheds, a raster depressions map has been created. This map indicates actual depressions or sinkholes in the area that are large enough to store water runoff from a storm event. In Google Earth Pro, polygons of the recognizable lakes around the basin were created. These polygons were exported and merged into one vector layer. This vector layer had to be converted to a raster map, using the *v.to.rast.value* function in GRASS. This function converts a vector layer into a raster layer with a value of 1. Vector data is represented in points, lines and polygons, while raster data is represented as a grid of individual pixels. The resulting raster map is the input for the *r.watershed* function. The watershed map that takes real depressions into account is calculated for the same threshold values as mentioned above.

§3.2.2.3. i.atcorr

Since the Landsat ETM+ satellite images are only processed to a level 1 correction, an atmospheric correction needs to be done in order to perform accurate image classification. This is done for all Landsat 7 ETM+ images using the GRASS *i.atcorr* function in QGIS. *i.atcorr* outputs an estimate of the true surface reflectance by removing effects of atmospheric gases, aerosols and thin cirrus clouds. The function performs an algorithm based on the specific 6S parameters, that take into account the geometrical conditions, date and time, longitude, latitude, atmospheric model, aerosols model, visibility, elevation sensor height and sensor band. The 6S parameter choices can be found in the GRASS *i.atcorr* manual. The 6S parameter choices used are derived from literature values and the Landsat metadata files, or estimated using the function manual provided by GRASS. An example of the corrected data is given in Figure 6.

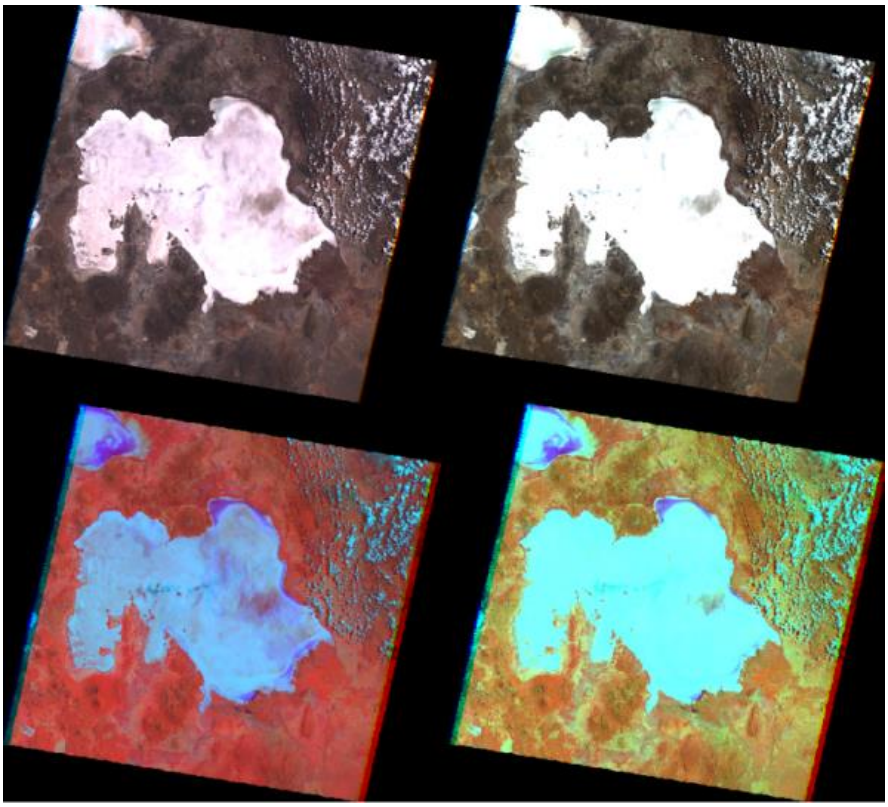

Figure 6: Landsat ETM+ band combinations before [left] and after [right] the atmospheric correction. The upper two images show the stack created by combining bands 3,2 and 1 (Natural Color image), and the lower two images show the stack created by combining bands 6, 4 and 2. The atmospheric correction enhances the contrast between different land cover types.

§3.2.2.4. Semi-Automatic Classification Plugin

The *Semi-Automatic Classification Plugin* allows for supervised classification of remote sensing images. Using this plugin, multiple training areas can be created. The spectral signatures of these training areas are automatically calculated and displayed in a spectral signature plot. These plots show the reflectance of the training area for each band entered. On the horizontal axis the band number is given, on the vertical axis the reflectance. Usually, the reflectance values are scaled from 0 to 1. Here a value of 0 represents truly black areas, where all light is absorbed, and a value of 1 truly white areas, where all light is reflected.

A composite of all Landsat bands is created in QGIS using the virtual raster creation tool. This tools merged all bands in one file. Now, in the properties menu of the Landsat composite, different band combinations can be created easily. This facilitates the recognition of different landcovers.

In the plugin training files (ROIs, 'region of interest') of recognizable rock types can be created (Figure 7). In the Landsat composite, multiple areas where this specific rocktype is found need to be selected. The plugin automatically calculates the mean and standard deviation of the reflectance of each band for a certain rocktype included in the Landsat composite. This is plotted in the spectral reflectance plot. Which rock type is situated where is derived from both the the geological map from the GeoBolivia website (geo.gob.bo) and the specific standard landsat responses, as described in §3.2.3.

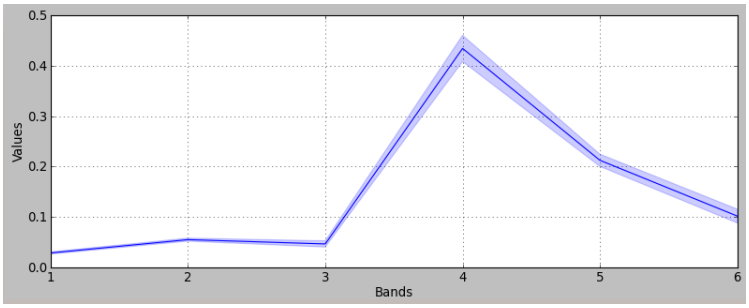

Figure 7: An example of the one spectrum signature for a certain ROI. The blue line is the mean response of each band, the surrounding light blue area indicates the standard deviation.

After the ROI files are created, the classification can be done. The plugin examines every pixel seperately, and assigns a certain ROI to it. There are three available classification algorithms in this plugin; these are based on maximum likelihood, minimum distance or spectral angle mapping. The maximum likelihood algorithm assigns the ROI with the highest probability to each pixel. The minimum distance algorithm assigns the ROI with the minimum distance to the mean value to the pixel. The spectral angle mapping method calculates the angles between the reflectance of each pixel and the reference training spectra (Figure 8). If the angles for a certain ROI lie within the entered threshold around the reference training spectra, this ROI is assigned to the pixel. [Congedo, 2014]

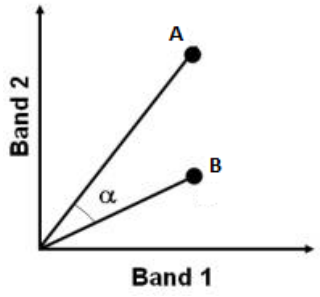

Figure 8: An example of the spectral angle between band 1 and 2. Suppose A is the plotted reflectance for band 1 and 2 for the ROI, and B the reflectance for the examined pixel, than α is the spectral angle between A and B.

§3.2.3. Landsat Imagery Interpretation

Photons are absorbed in minerals by several processes. The variety of absorption and their wavelength dependence makes it possible to derive information about the chemistry of a mineral from its reflected or emitted light. The reflectance or absorbance can be measured using sensors that are sensitive to a certain wavelength of light.

Each mineral has a particular spectrum, depending on its components. On the USGS digital spectral library *speclab.cr.usgs.gov,* the spectra of many minerals can be found. In Figure 10 and 11 the reflectance of halite, sand, quartz, and clay minerals, downloaded from the USGS digital spectral library, are shown.

Since Landsat 7 does not measure the spectrum throughout. The average values for the reflectance for all minerals for each Landsat band is calculated using Excel. The thick lines in the images represent the average reflectance as measured by Landsat 7.

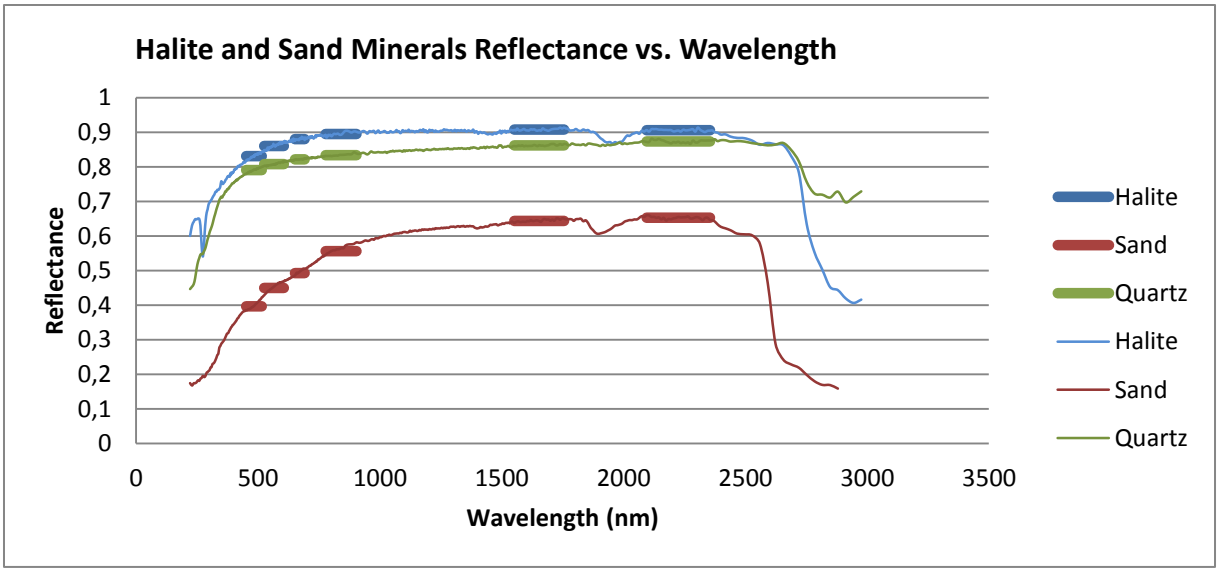

Figure 9: The reflectance of halite, sand and quartz. The thin lines are the spectra of these halite, sand and quartz, the thick lines show the average reflectance in the available Landsat bands.

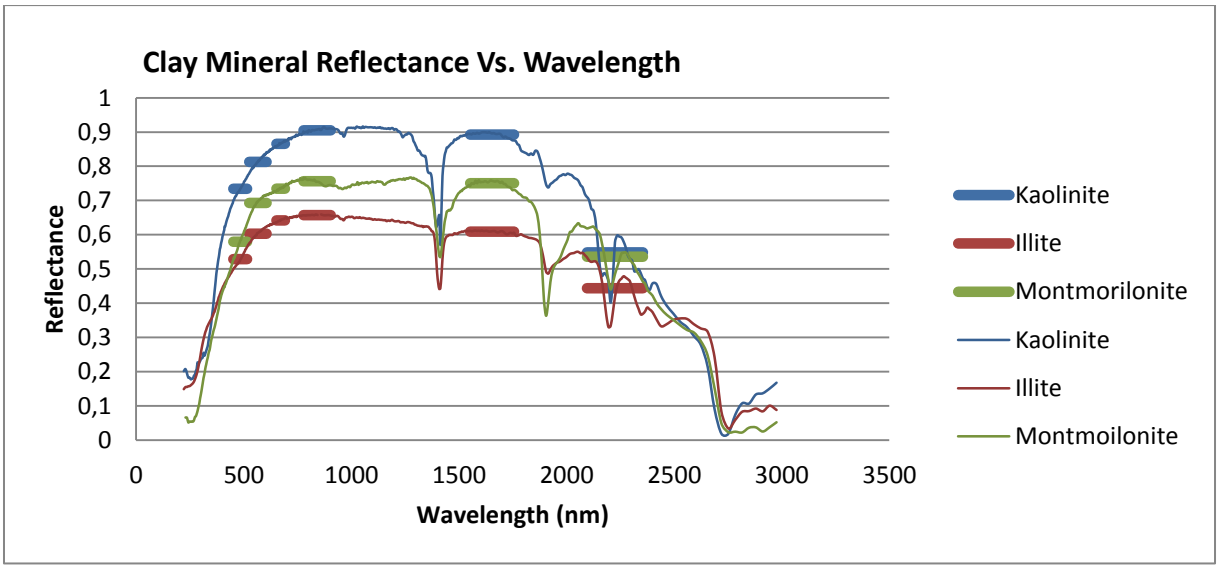

Figure 10: The reflectance of clay minerals. The thin lines are the spectra of these clay minerals, the thick lines show the average reflectance in the available Landsat bands.

A problem with any kind of reflectance remote sensing is that some wavelengths are not available due to absorption of the elements in the Earth its atmosphere. There are corrections available to reduce the effects of the atmosphere. In this study QGIS modules are used to perform this correction (see section §3.2.3). However, as can be seen in Figure 11, for some wavelengths the atmospheric transmittance is zero. Here the absorptions of these elements in the atmosphere causes gaps at wavelengths that cannot be corrected. For example, the presence of water in the atmosphere makes it unable to measure the reflectance near both 1,4 µm and 1,9 µm. These areas are very significant in distinguishing various minerals, because many minerals contain amounts of H_2O- or OH groups and thus show differences in reflectance here. [Clark, 1999]

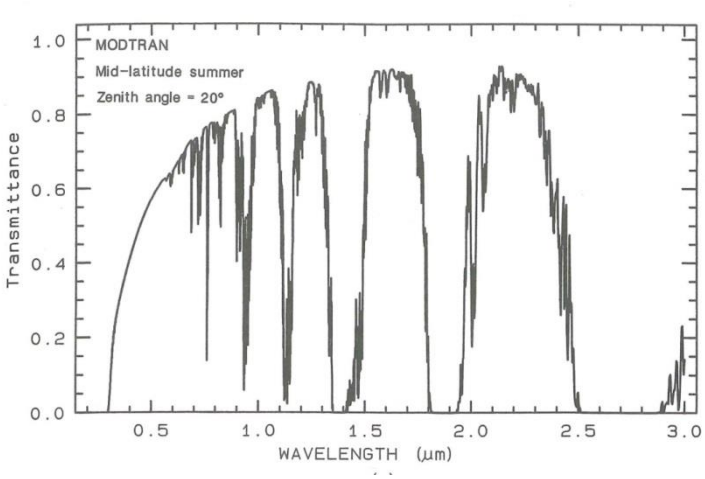

Figure 11: The atmospheric transmittance against wavelength. [Clark, 1999]

The majority of the minerals is identified by their absorption peaks. The spectral resolution of the Landsat data used is insufficient to allow such identification of individual minerals. Such identification is only possible using imaging spectrometer aircraft instrument, which have a spectral resolution of 0.1 µm or better. For example, the different clay minerals show characteristic absorption peaks around 1300 nm, 1800 nm and 2200 nm. As can be seen, only the peaks around 2200 nm are included in the Landsat 7 data. The Landsat bands do not cover the whole spectrum of the minerals. Furthermore, the wavelength range of the bands is too large to distinguish individual peaks.

However, certain mineral groups can be distinguished using satellite images using the overall reflectance at certain wavelength regions. Figure 12 shows the typical wavelengths at which certain mineral groups are distinguishable because of their characteristic reflectance or absorbance peaks here. [Lang, 1999]

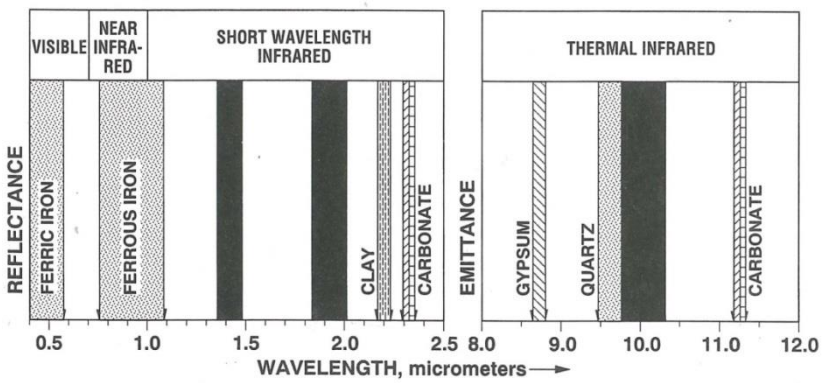

Figure 12: Visible to infrared wavelength positions of spectral bands that are characteristic for certain minerals. The wavelengths shown in black are not available due to atmospheric absorption. The areas near 1,4 µm and 1,9 µm are black due to atmospheric water absorption, the areas near 10 µm due to atmospheric ozone absorption. [Lang, 1999]

Not all these wavelength regions are covered in Landsat 7. However, the wavelength region used for distinguishing clay is available. Since clay minerals have an absorbance peak around 2.2 µm (Figure 10), the reflectance measured for band 7 (2.09-2.35 µm) is lower than the reflectance for band 5 (1.55-1.75 µm). For non-clay minerals the reflectance of band 7 is approximately equal to the reflectance of band 5, as can be seen for halite, quartz and sand in Figure 9. The difference between the reflectance of band 5 and band 7 (Figure 13) is thus an indication for the amount of clay. In Figure 13 an example is shown, areas that are darker in band 7 than band 5 indicate large amounts of clay. A map of the indicated amounts of clay can be created in QGIS by subtracting band 7 from band 5.

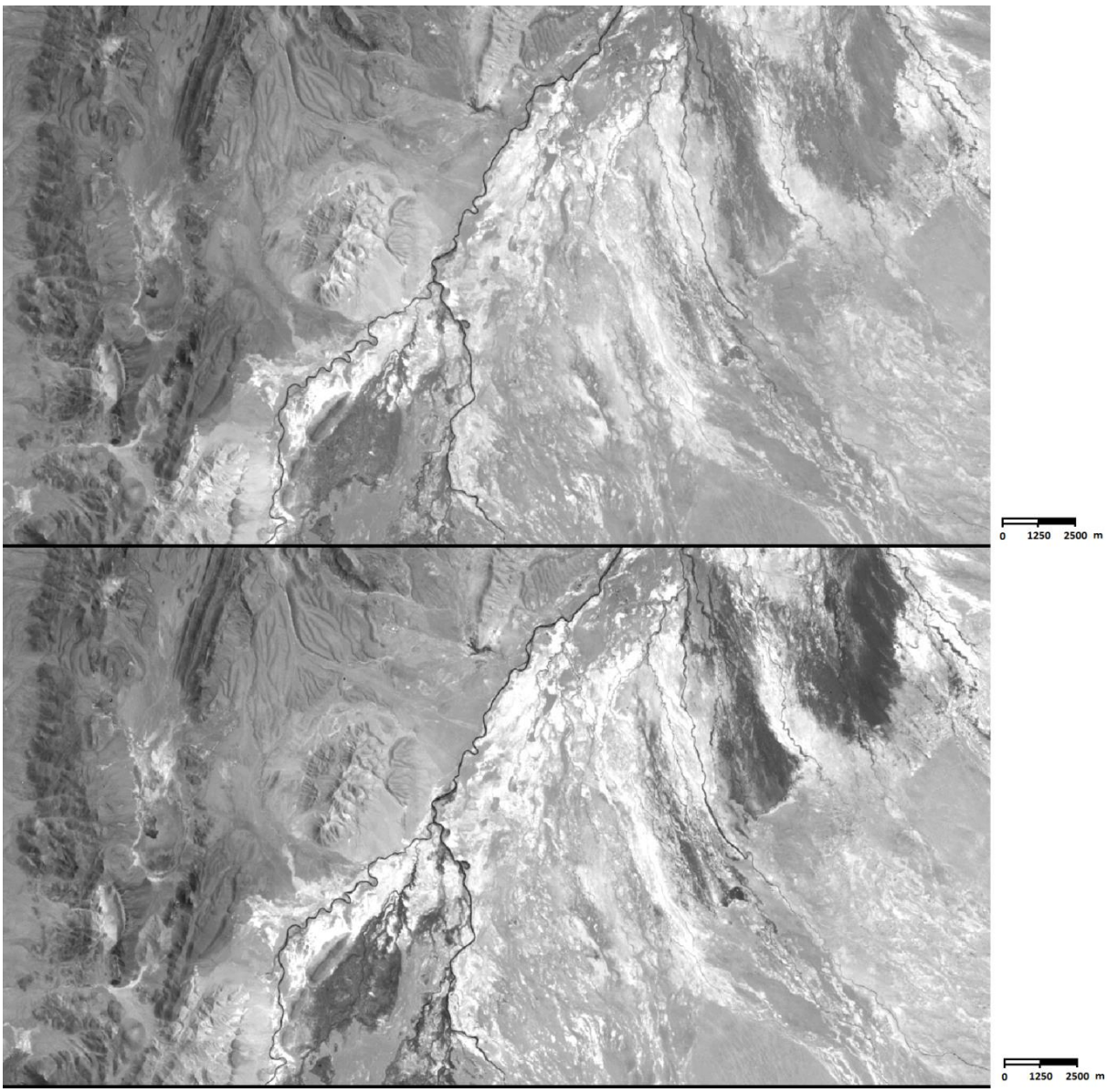

Figure 13: Landsat ETM+ band 5 [upper picture] and band 7 [bottom picture]. Areas that have higher absorbance (black) in band 7 than in band 5 indicate large amounts of clay. Areas that have approximately the same reflectance indicate areas with large amounts of sand or volcanic rocks.

Chapter 4. Results

In this chapter all results are presented. In the first part of the chapter, §4.1., the different datasets and results for the watershed map are presented. In §4.2. the different river path datasets and results are presented. In the last paragraph, §4.3., the results of the Landsat image interpretation are presented.

§4.1 Watershed definition results

In this paragraph the watershed definition results are discussed. A major part of the watershed results are added to APPENDIX A. In §4.1.1. the watershed data from GeoBolivia are discussed, in §4.1.2. the watershed data from WWF are discussed, in §4.1.3. the watershed results from the *r.watershed* function in QGIS are discussed.

§4.1.1 GeoBolivia watersheds

Figure A1 shows the 'true' watershed, the division of the sub-basins and the outline of the Salar de Uyuni, as created in Google Earth Pro. The total drainage area of the Salar de Uyuni is 1405 km². The area of the Salar de Uyuni itself is approximately 610 km². In total eleven sub-basins can be distinguished. The total area of these sub-basins varies from approximately 50 km² (the smallest sub basin in the western side of the basin) to 600 km² (the drainage area of the Rio Grande de Lipez).

Figure A2 to A5 show the watersheds as provided by GeoBolivia and WWF. Figure A2 shows the main watershed of GeoBolivia in blue. The GeoBolivia watershed follows the mountainous region on the eastern site of the basin well. Only a few deviations with the 'true' Google Earth Pro watershed can be seen. On the southern and western site of the region, the main GeoBolivia watershed incorrectly includes some (salt)lakes that do not actually flow into the Salar de Uyuni. Furthermore, it can be seen that the GeoBolivia watershed is cut off at the borders of Bolivia.

With a few exceptions, it can be seen that the main GeoBolivia watershed follows the summits in high relief areas. The GeoBolivia watershed seems to follow the highest points in low relief areas. This can be seen on the profiles created in Google Earth Pro.

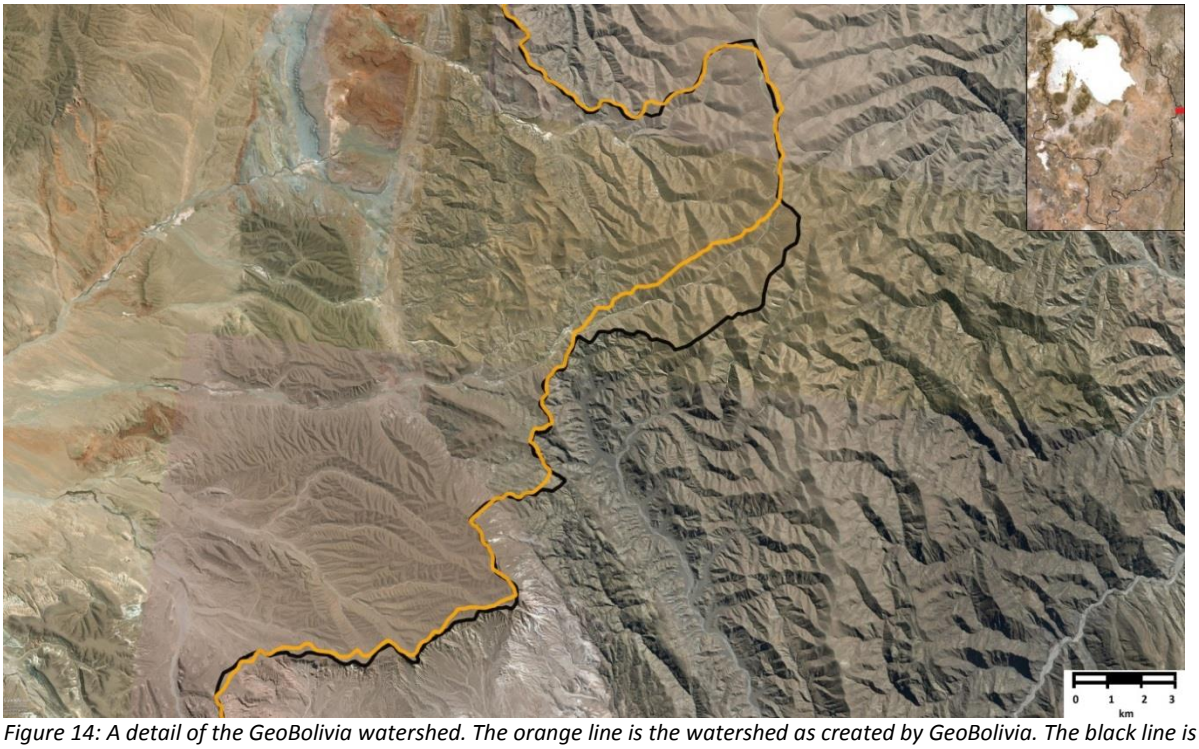

Figure 14: A detail of the GeoBolivia watershed. The orange line is the watershed as created by GeoBolivia. The black line is the 'true' watershed created in Google Earth Pro. The GeoBolivia watershed follows the mountains well, but occasionally incorrectly includes or excludes parts of rivers from the drainage area.

In figure A3 the watershed subdivision is shown. Here it can bee seen that most of the (salt)lakes that do not flow into the Salar de Uyuni have their own drainage area, not continuing into the Salar de Uyuni. Globally, the sub-basins that do flow into the Salar de Uyuni match the 'true' Google Earth Pro sub-basins. However, in the lower relief areas the borders are assigned incorrectly. This can be seen east from the Salar de Uyuni. For example, one of the borders crosses the delta at the southmost side of the Salar de Uyuni. Within the Salar de Uyuni the subdivision of the drainage basin is continued, which results in artifacts.

In figure A4 further subdivision of the drainage basin by GeoBolivia is shown. Since the river systems are supposed to flow towards the Salar de Uyuni, one would expect the further division to be enlongated shapes from the sides of the main watershed towards the Salar de Uyuni. This is not the case.

§4.1.2 WWF watershed

In figure A5 the main watershed of WWF is shown. On the global scale, the 'true' watershed and the WWF watershed seems to differ slightly. On a small scale it becomes clear that the resolution of the WWF watershed is in fact very low (Figure 15). The borders of the WWF watershed are blocky and do not follow the summits accurately. On average the error compared to the 'true' watershed lies within zero to two kilometers. The WWF has not subdivided the watershed.

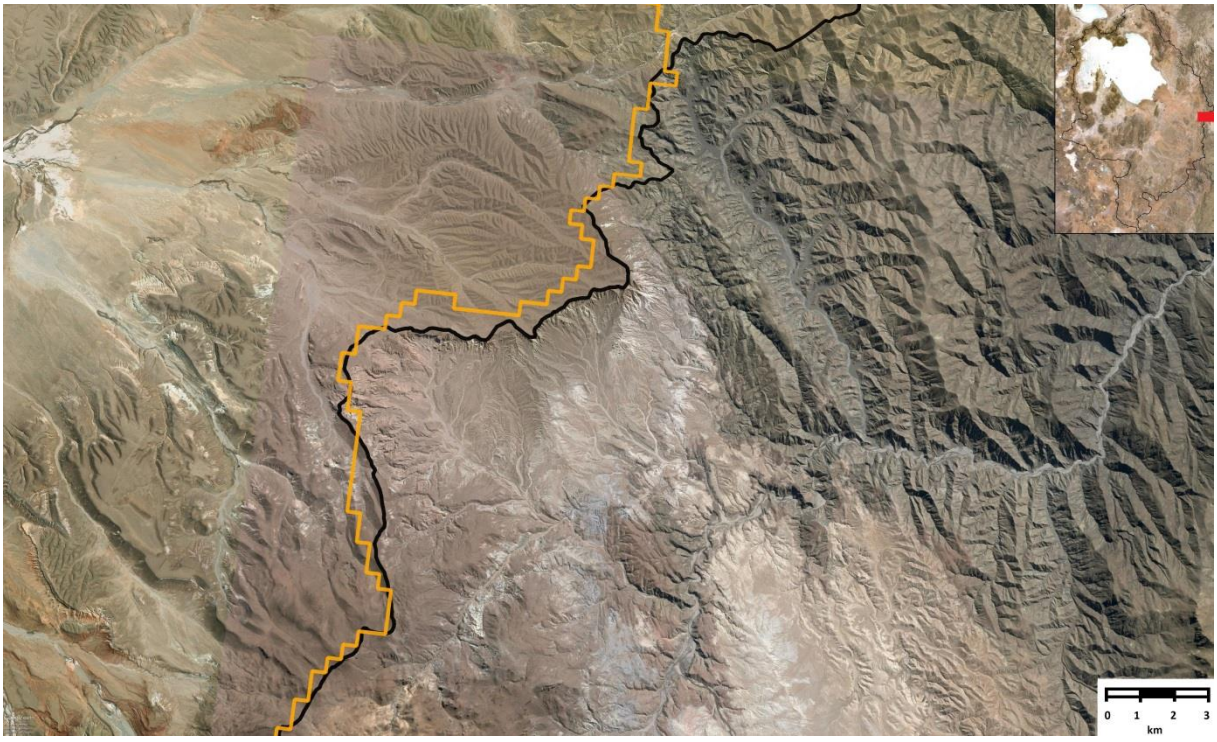

Figure 15: A detail of the WWF watershed. The orange line is the watershed as created by WWF. The black line is the 'true' watershed created in Google Earth Pro. The WWF watershed does follow the summits, but not very accurately.

§4.1.2 QGIS r.watershed results

The results of the *r.watershed* function are also compared with the 'true' watershed and sub-basins determined in Google Earth Pro. In figure A6 the outline of the 'true' main watershed , sub-basins and the outline of the salt lake imported in QGIS are shown. In all the results from the *r.watershed* function the 'true' main watershed and the outline of the salt lake are given to verify the results.

Figure A7 and A8 show the watershed calculated by *r.watershed* for increasing threshold size. In figure A7 the resulting drainage basins are given for threshold values of 1 million and 2 million. In figure A8 the resulting drainage basins for threshold values of 3 million and 4 million are given. The subdivision of the basin within the salt-lake is excluded from the *r.watershed* results, since subdivision within the Salar de Uyuni itself is unnecessary for the purpose of this research.

Figure A9 and A10 show the resulting watershed when inputting the location of real depression map. The depressions entered are shown in the map in light blue. In figure A9 the resulting drainage basins are given for threshold values of 1 million and 2 million. In figure A10 the resulting drainage basins are given for threshold values of 3 million and 4 million.

The outline of the watershed calculated using the locations of real depressions map matches to the 'true' outline of the watershed determined in Google Earth Pro. The calculated watershed seems to give a more accurate representation of the outline of the watershed in high-relief areas (Figure 16). However, in low-relief areas, the calculated watershed seems to incorrectly include or exclude areas from the watershed (Figure 17).

Inputting the real depressions map seems only to effects the outline of the main watershed, further subdivision of the basin remains unchanged. The subdivision created by *r.watershed* deviates from the subdivision of the watershed determined in Google Earth Pro. Increasing the threshold value does not seem to give more accurate results. Some of the 'true' sub basins were visible for threshold values of 1 and 2 million, but were merged together for threshold values of 3 and 4 million.

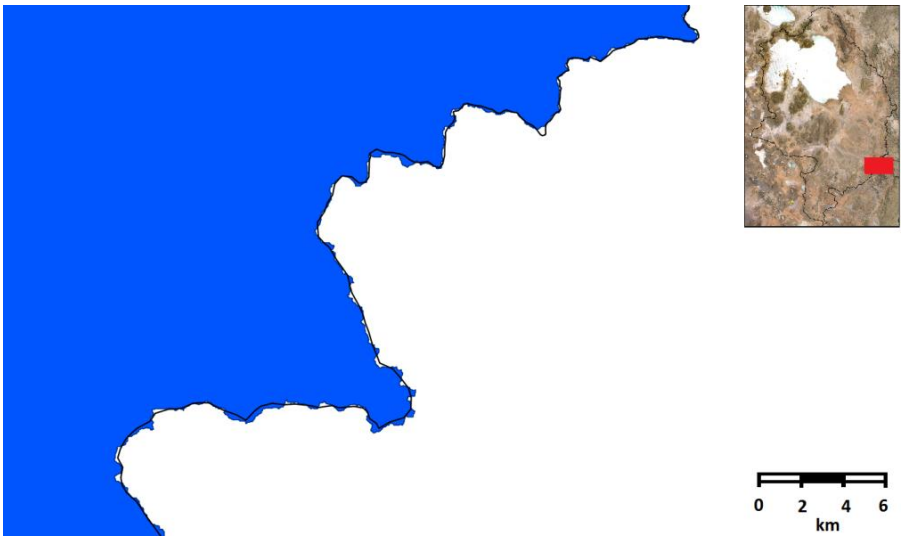

Figure 16: A part of the watershed as calculated by r.watershed (blue) in a high relief area. The black line indicated the 'true' watershed as defined using Google Earth Pro. The watershed calculated by r.watershed seems to follow the summits more accurately.

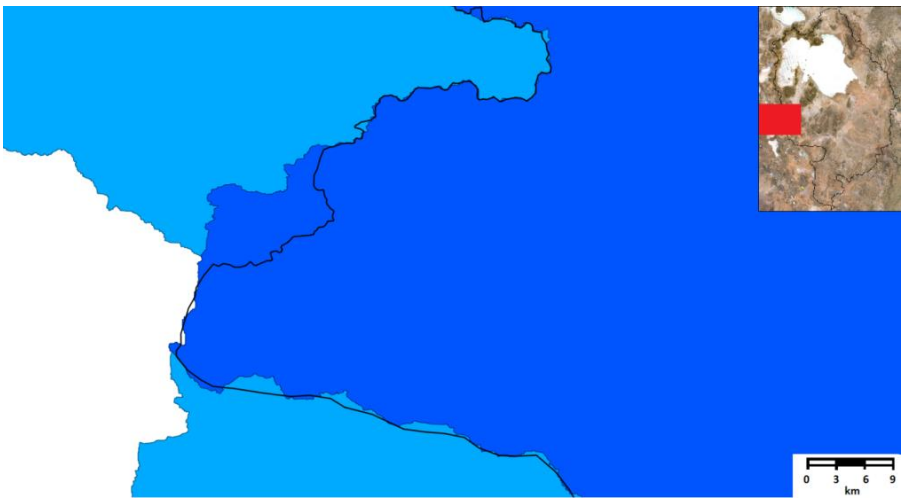

Figure 17: A part of the watersheds calculated by r.watershed in QGIS. The light blue area is the resulting watershed without inputting the location of real depressions map. The darker blue area is the resulting watershed, giving the location of real depressions map as input. The black line indicates the 'true' watershed as defined using Google Earth Pro.

§4.2. River path definition results

In this paragraph the river path definition results are discussed. A major part of the river path results are added to APPENDIX B. In §4.2.1. the relief map of the area is discussed. In §4.2.2. the river path data from GeoBolivia and WWF are discussed, in §4.2.3. the results from the *r.terraflow* function in QGIS are discussed and in §4.2.4. the river responses of the Landsat 7 ETM+ images are discussed.

§4.2.1. Relief map

Figure B1 shows the relief map of the area around and within the basin. The relief map of the area around the basin is created using intervals of 100 meter. The relief map within the basin is created using intervals of 20 meters. The relief map is used to compare the river path results for high and low relief areas.

§4.2.2. GeoBolivia and WWF river path results

Figure B2 shows the watershed and the river path of both GeoBolivia and WWF. The dark blue lines indicate the rivers, and lighter blue lines indicate the watersheds. In Figures B3 and B4 the river path datasets are compared using Google Earth Pro. In Figure B3 the datasets are compared in an area with high relief. In Figure B4 the datasets are compared in a low-relief area, the delta of the Rio Grande de Lipez river.

The main river systems can be recognized in both the GeoBolivia and WWF dataset. However, Figure B3 shows that the GeoBolivia river paths incorrectly crosses the subdivision of the watershed. The river path calculated by HydroSHEDS incorrectly continue into the salt lake, too. Figure B3 shows that the main river path are included in both datasets. However, smaller river systems are not taken into account. Both datasets include incorrect river paths in low relief areas, as can be seen in Figure B4.

§4.2.3. QGIS r.terraflow results

The results of r.terraflow are compared to Google Earth Pro in Figure B5, B6 and B7. Figure B5 shows a comparison of the SFD and the MFD results in an area with high relief. Figure B6 shows a comparison of the SFD and the MFD results in an area with low relief, and Figure B7 shows a comparison of the SFD and MFD results of a small delta system at the western side of the Salar de Uyuni.

In high relief areas, the SFD and MFD river paths correspond with the river paths seen in Google Earth Pro. Since here most of the small rivers are absent in the MFD results, the SFD results seems to give a better representation of the river paths. In low relief areas, both the SFD and MFD results do not correspond the river paths seen in Google Earth Pro. However, in these areas MFD seems to give a better representation of the river paths.

§4.2.4. Landsat 7 ETM+ river path results

Figure B8 shows the responses of each Landsat band of two parallel river systems. In band 1, band 5, band 6 and 7 the rivers show a high absorbance, and are thus dark-colored. In band 3 and 4, some of the river path show a high reflectance, and are thus white-colored. Other river paths show a high absorbance, and are thus dark-colored. Small river systems cannot be recognized in the Landsat 7 ETM+ images.

§4.3. Landsat interpretation results

In this paragraph the Landsat interpretation results are discussed. In §4.3.1. the result of the difference map between band 7 and 5 is covered. In §4.3.2 the results from the *semi-automatic classification plugin* in QGIS are covered. A major part of the result can be found in APPENDIX C.

§4.3.1. Difference map results

Figure C1 shows the difference map between band 7 and band 5. In this map the more white areas indicate high amounts of clay. At the edges of the salt lake, and along river paths large amounts of clay are indicated. Some artifact can be seen in Figure C1. The black areas within the salt lake and at the edges of the basin are missing areas or areas containing errors. Figure C2 shows details of the difference map created in QGIS. The

uppermost figure shows a part of the drainage area of the Rio Grande de Lipez, the bottommost figure shows the difference map delta of the Rio Grande de Lipez.

§4.3.2. Semi-automatic classification plugin results

Figure C3 shows the spectral signatures of the ROI as calculated by the *semi-automatic classification plugin*. Note that band 6 is split up in two, so the band numbers 6 and 7 both represent a part of the thermal band. Although the salt does show resemblance for visible bands (band 1, band 2 and band 3), the calculated spectral signature differs from the spectral signature as expected from the USGS digital spectral library values. This also applies to the other spectral signatures calculated by the plugin. Furthermore, the calculated spectral signatures for slate and sandstone are nearly identical, while these two rocks can be distinguished in the true color image in QGIS.

Figure C4 shows the results of the *semi-automatic classification plugin* for the sandstone and slate ROI's, using the maximum likelihood algorithm with a threshold value of 99. Although the clay and salt seems to be excluded, large portions of the remaining land cover are incorrectly classified as slate or sandstone. For lower threshold values, even more area is incorrectly assigned to be slate or sandstone. Furthermore, artifacts can be seen on the edge between different Landsat scenes. Here the overall reflectance has changes a bit due to difference in acquisition time. This could not be corrected for in QGIS.

Figure C5 shows the results of the semi-automatic classification plugin for the clay ROI using the maximum likelihood algorithm with a threshold value of 50.. The clay ROI was selected at the edge of the lake, and likely contains amounts of salts Comparing the results to the difference map results §4.3.1., these are indeed indicated clay amounts.

For the same clay ROI, the minimum distance and spectral angle settings in the semi-automatic classification plugin are used, too. However, the results are incorrect. When using the spectral angle algorithm, the plugin did not seem to work. The plugin was unable to perform a classification and the output was an empty image for all threshold values entered. When using the minimum distance algorithm, the plugin incorrectly classified the salt lake as clay for any threshold value entered. An example of the results using the minimum distance algorithm is shown in Figure 18.

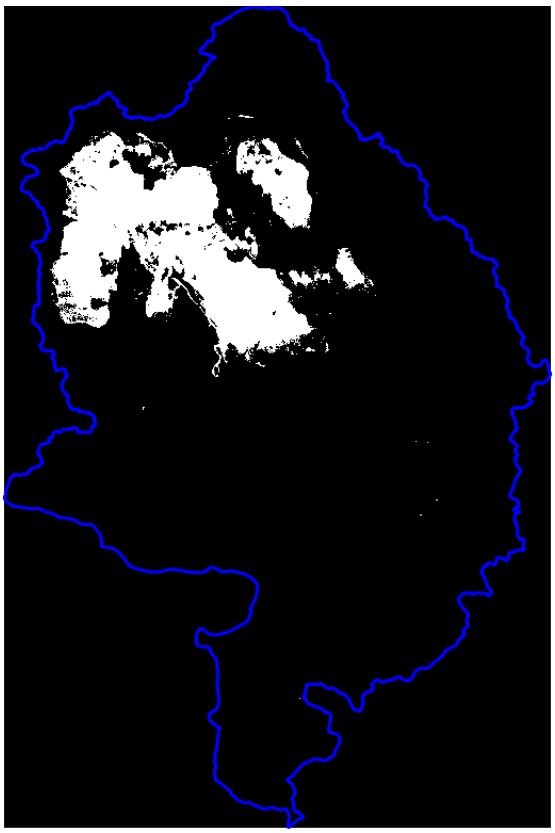

Figure 18: An example of the results using the minimum distance algorithm in the semi-automatic classification plugin.

Chapter 5. Discussion

The algorithms in QGIS clearly produce incorrect results in low relief areas, even in high relief areas the results show some discrepancies. Since the Aster GDEM data contain errors of approximately 10 meters in this area, the algorithms will probably still calculate partly erroneous watersheds and river paths.

As mentioned in chapter 3, using *r.terraflow*, the computers used were unable to process the whole Aster GDEM dataset at once. Since the data had to be subdivided into multiple tiles, parts of rivers will not be taken into account. This might result in a difference in calculated flow accumulation near the borders of the tiles.

The Landsat ETM+ satellite images used in this study are acquired in 2002 until 2003. Therefore, the river paths and land cover interpretation derived from these images are time equivalent to the other datasets, and thus less representative. Moreover, there are multiple difficulties using the Landsat ETM+ images for land cover classification. The first difficulty is that only the reflection of the upper most layer can be measured. It is for example possible that a sandstone is overlain by a small layer of clay. In this case, this sandstone would be classified as clay. Another example is rocks or soils overlain by thin layers of water. This small layer of water will show different responses than the bare rocks or soils would.

Another difficulty is that the Landsat images show average reflections for each 30 meter by 30 meter tile. Within these 900 m^2 a lot of variation can be found. All this variation is averaged out. In addition, because all rocks consist of multiple minerals, the reflection of the separate minerals is often averaged out. In fact, a certain rock type is only clearly distinguishable if the 30x30 meter tiles are only filled with this one rock type. An effective spectral signature for a specific rock type can only be created if this is the case.

Classification in QGIS is not possible using only a global geological map. For example, the scale of the GeoBolivia geological map is much larger than the scale of the local differences. This is illustrated in Figure 18, the geological features identified by GeoBolivia include multiple distinguishable regions. This is probably also the reason the semi-automatic classification identified almost the whole region as the inputted slate or sandstone. The spectral signatures might have high standard deviations, and might thus overlap with signatures of other rock types.

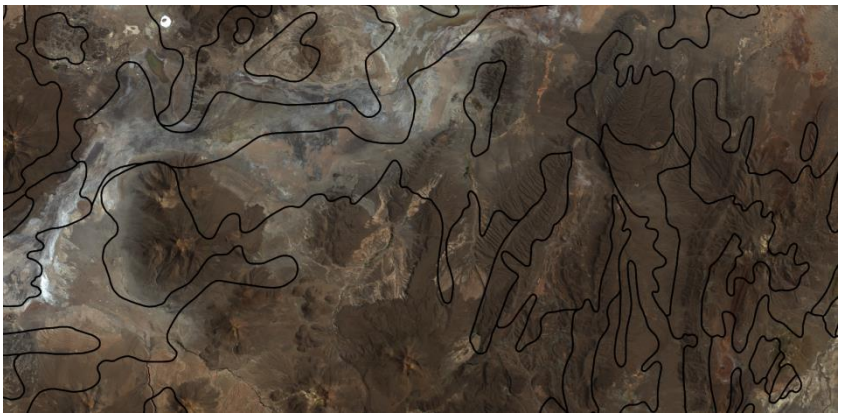

Figure 19: The Landsat ETM+ natural color image, overlapped by the GeoBolivia geological map. (Note that the GeoBolivia map is somewhat shifted to the left because of a small difference in reference system)

The difference map of band 5 and 7 and the semi-automatic classification might contain errors because of the atmospheric correction done in QGIS. The atmospheric correction in QGIS gives the same results for different settings. This suggests the correction does not take into account local differences, for example local climate, accurately. For better classification of rock types, not only more detailed data is needed, but probably also a better atmospheric correction module.

Chapter 6. Conclusion and Recommendations

It is possible to calculate rather accurate watersheds in QGIS for the area around the Salar de Uyuni. However, a map including local depressions that do not overflow into the Salar the Uyuni has to be fed into the *r.watershed* function. Without this map, the function assumes that these lakes do overflow into the basin and the watershed that will be created becomes way too large. For high relief areas, *r.watershed* outputs an accurate watershed border. In low relief, flat areas, the results need to be looked at. Here the errors in the Aster GDEM files seem to be the main reason to incorrectly divide the area into watersheds.

For the results of *r.terraflow* the same principle as for *r.watershed* applies. In high relief areas, the SFD and MFD river paths frequently correspond to actual river paths seen in Google Earth Pro. For most areas, MFD gives more realistic results than SFD. In low relief areas, the results for both SFD and MFD are inaccurate and do not correspond with the reality at all. Digital elevation data with higher resolution and less errors, for example Airborne LIDAR images, might improve the results from both *r.watershed* and *r.terraflow*.

Landsat ETM+ satellite images can be used to determine river paths. However, the resolution of the Landsat ETM+ data is too low to distinguish small river paths. Only river paths that are as wide or wider than the 30 meter spatial resolution can be distinguished. Band 1 (blue) shows higher absorbance for active river paths that have water flowing through it. On the thermal band, band 6, these rivers can also be distinguished. These rivers are colored black due to higher absorption of water in this band. Band 3 and 4 show different reflections for different types of rivers. These differences seem to be caused by differences in vegetation and types of deposits in and around the river. No direct relationship between river type and reflectance in band 3 and band 4 is discovered.

Detailed land cover classification using the semi-automatic classification plugin in QGIS using global geological maps is impossible, because the spectral signatures of most land covers are too much alike. However, salt and clay can be easily distinguished. Salt has a high reflectance in the visual spectrum (band 1, 2 and 3), and can hereby be recognized. A difference map between band 5, which has a constant reflectance for most minerals, and band 7, which has a significantly lower reflectance for clay only, can be used to indicate amounts of clay. Because this map shows the ratios of the reflections instead of the level of reflectance, this method is relatively safe and applicable.

For more accurate land cover classification, more detailed information is needed. For example, this can be acquired by hyper spectral sensors. Multispectral images have a much higher resolution than Landsat images and have a much wider spectral range. These sensors have a smaller spectral bandwidth and thus can distinguish separate peaks in land cover reflectance. In fact, the abilities of these sensors come close to those of spectroscopes used in laboratories. Another possibility is local data acquisition using GPS-systems. Recognized land covers are entered in the GPS-system and later loaded into GIS-systems to calculate spectral signatures.

During the research, a lot of errors were encountered in QGIS. A significant part of the functions in QGIS do not work well on the Windows interface. For some of these functions, the GRASS GIS provided with QGIS can be used to solve the problem. Unfortunately, another interface had to be used in order to get others working. For further research usage of Linux or any other operating system than Windows working with QGIS is recommended. Another option is to purchase more sophisticated software, for example ArcGIS or ENVI, as these programs claim to have a better functioning interface.

References

Clark, R.N., 1999. Spectroscopy of Rocks and Minerals and Principles of Spectroscopy. In: A.N. Rencz, ed. 1999. *Manual of Remote Sensing Volume 3: Remote Sensing for Earth Sciences*. New York: John Wiley & Sons Inc, pp. 3-58.

Congedo, L., 2014. Land Cover Classification of Cropland: a Tutorial Using the Semi-Automatic Classification Plugin for QGIS. [Available at: www.directionsmag.com].

Ehlschlaeger, C., 1989. Using the A^T Search Algorithm to Develop Hydrologic Models from Digital Elevation Data. In: [Author unknown], 1989, *Proceedings of International Geographic Information Systems (IGIS) Symposium '89*, Baltimore, MD, pp 275-281.

Lang, H.R., 1999. Stratigraphy. In: A.N. Rencz, ed. 1999. *Manual of Remote Sensing Volume 3: Remote Sensing for Earth Sciences*. New York: John Wiley & Sons Inc, pp. 357-374.

Lehner, B., 2013. *HydroSHEDS Technical Documentation*, 2nd edition. Washington: World Wildlife Fund US.

Meyer, D., et al. 2011. *ASTER Global Digital Elevation Model Version 2 – Summary of Validation Results*. NASA Land Processes Distributed Active Archive Center and the Joint Japan-US ASTER Science Team.

Orris, G.K., 1995. Bibliography and summary of date available for the Salar de Uyuni, Bolivia. USGS

Sabine, C., 1999. Strategies for Mineral Exploration. In: A.N. Rencz, ed. 1999. *Manual of Remote Sensing Volume 3: Remote Sensing for Earth Sciences*. New York: John Wiley & Sons Inc, pp. 375-448.

Toma, L., et al, 2001. *Flow Computation on Massive Grids.* Atlanta, USA: ACM-GIS.

Tibaldi, A., Corazzato C. & Rovida, A., 2008. Miocene–Quaternary structural evolution of the Uyuni–Atacama region, Andes of Chile and Bolivia. In: Gudmundsson, A., et al., 2008, *Tectonophysics 471, Understanding stress and deformation in active volcanoes*. Amsterdam, Elsevier B.V., pp. 114-135.

USGS. *Landsat 7 Science Data Users Handbook*. Landsat Project Science Office at NASA's Goddard Space Flight Center in Greenbelt, Maryland.[Availavle at: landsathandbook.gsfc.nasa.gov]

List of Abbreviations

ASTER - Advanced Spaceborne Thermal Emission and Reflection Radiometer

DEM – digital elevation model

ETM+ - enhanced thematic mapper plus

GDEM – global digital elevation model

GIS – geographic information system

GPS – global positioning system

HydroSHEDS - hydrological data and maps based on shuttle elevation derivatives at multiple scales

MFD – multiple flow direction

NASA - National Aeronautics and Space Administration

ROI – regions of interest

SFD – single flow direction

STRM - Shuttle Radar Topography Mission

USGS – United States Geological Survey

WWF – World Wildlife Fund

Acknowledgements

I want to thank everyone who helped me with this bachelor thesis. In special, I want to thank my supervisor Ir. Koen van Toorenenburg, for all his time and help during this bachelor thesis, and Adriaan van Natijne, for continuously helping me out with all my QGIS and computer problem.

APPENDIX A
WATERSHED DEFINITION RESULTS

Figure A1: *The 'true' watershed (black) and sub basins (white), defined using Google Earth Pro.*

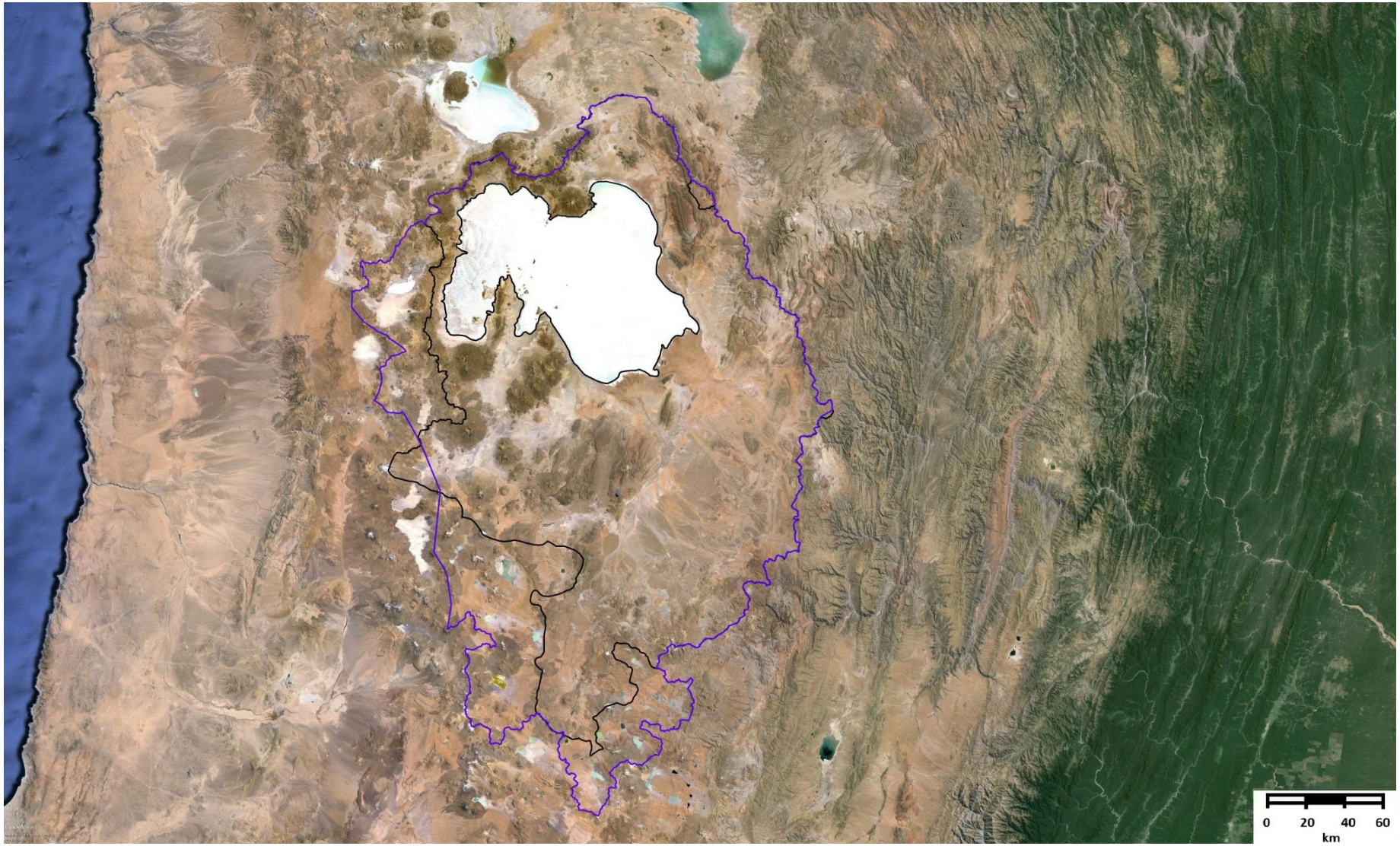

Figure A2: *The main watershed of the area as provided by GeoBolivia (blue). The black line indicate the 'true' watershed as defined using Google Earth Pro.*

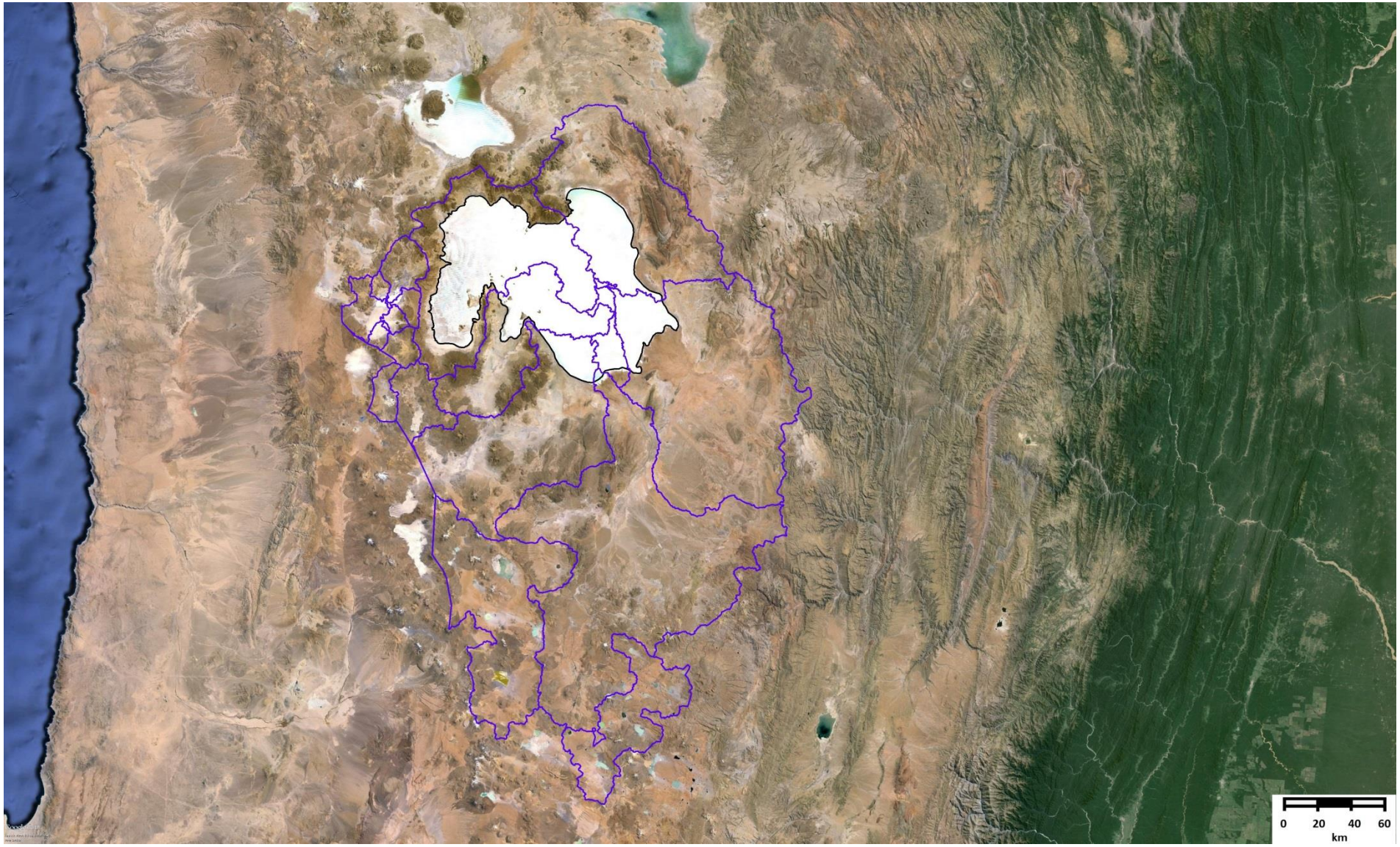

Figure A3: *The subdivision of the watershed (niveau 4) as provided by GeoBolivia.*

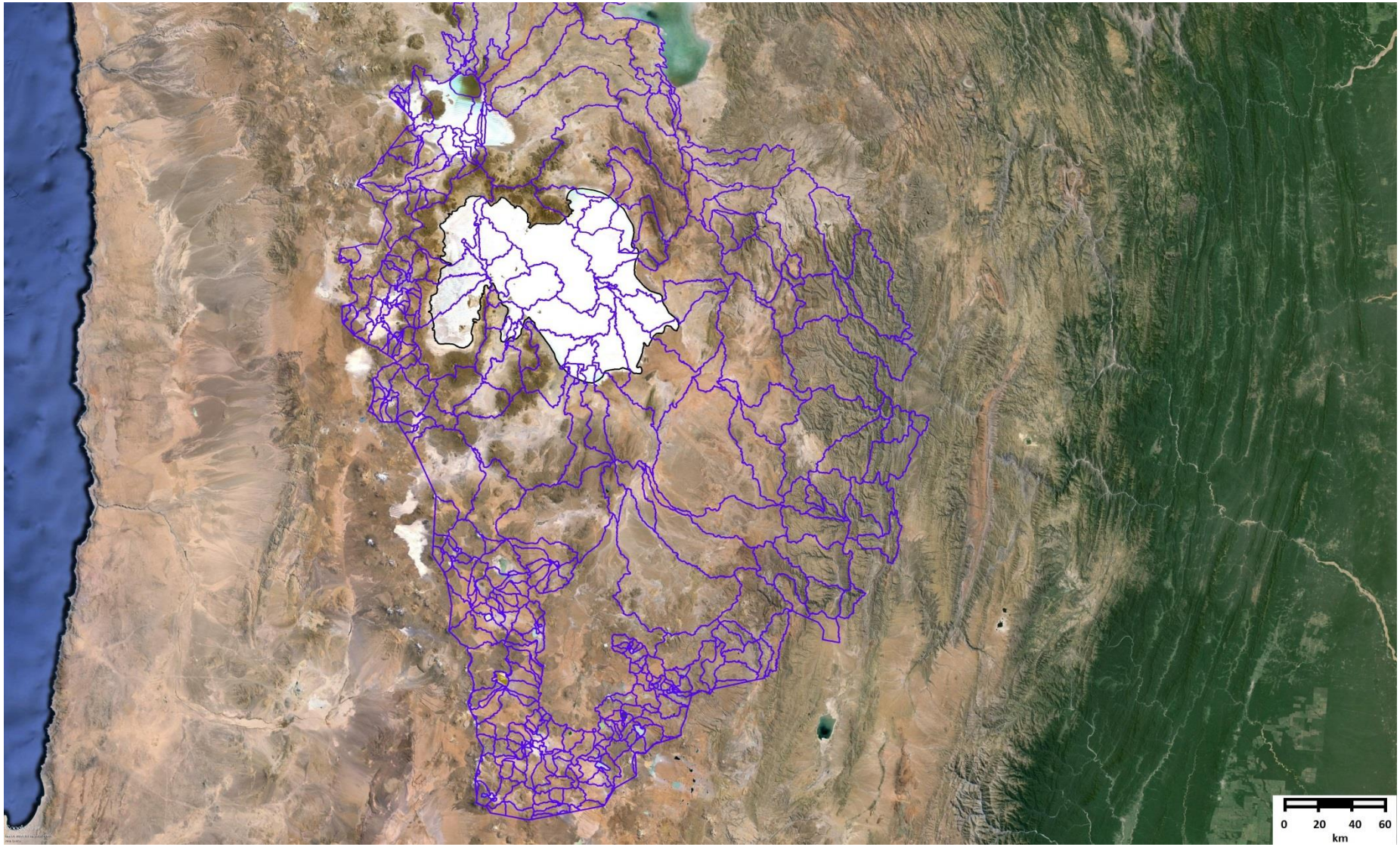

Figure A4: *The subdivision of the watershed (niveau 5, smallest subdivision available) as provided by GeoBolivia.*

Figure A5: *The main watershed of the area as provided by WWF.*

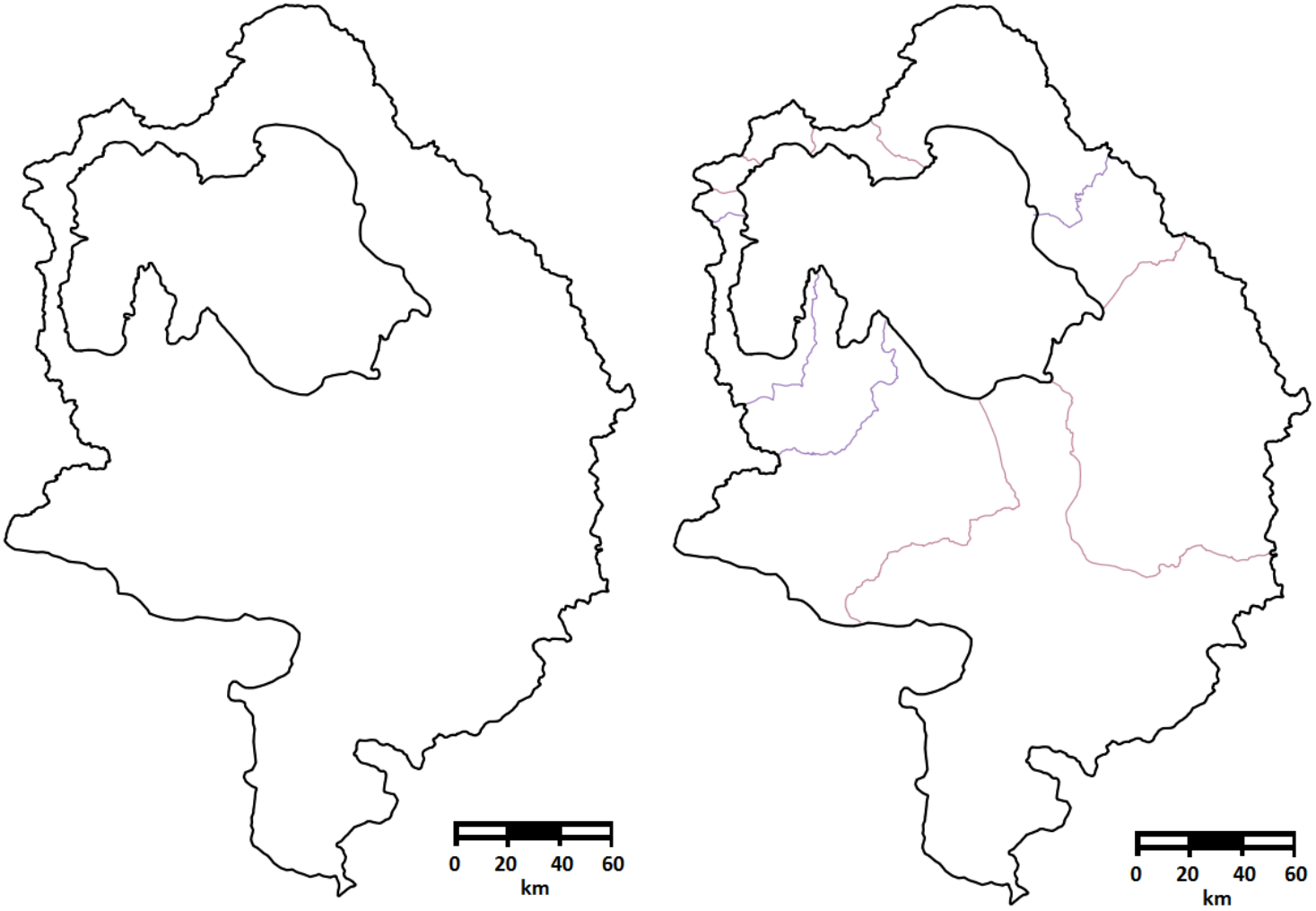

Figure A6: *The total watershed [left] and sub basins [right] created in Google Earth Pro as loaded in QGIS.*

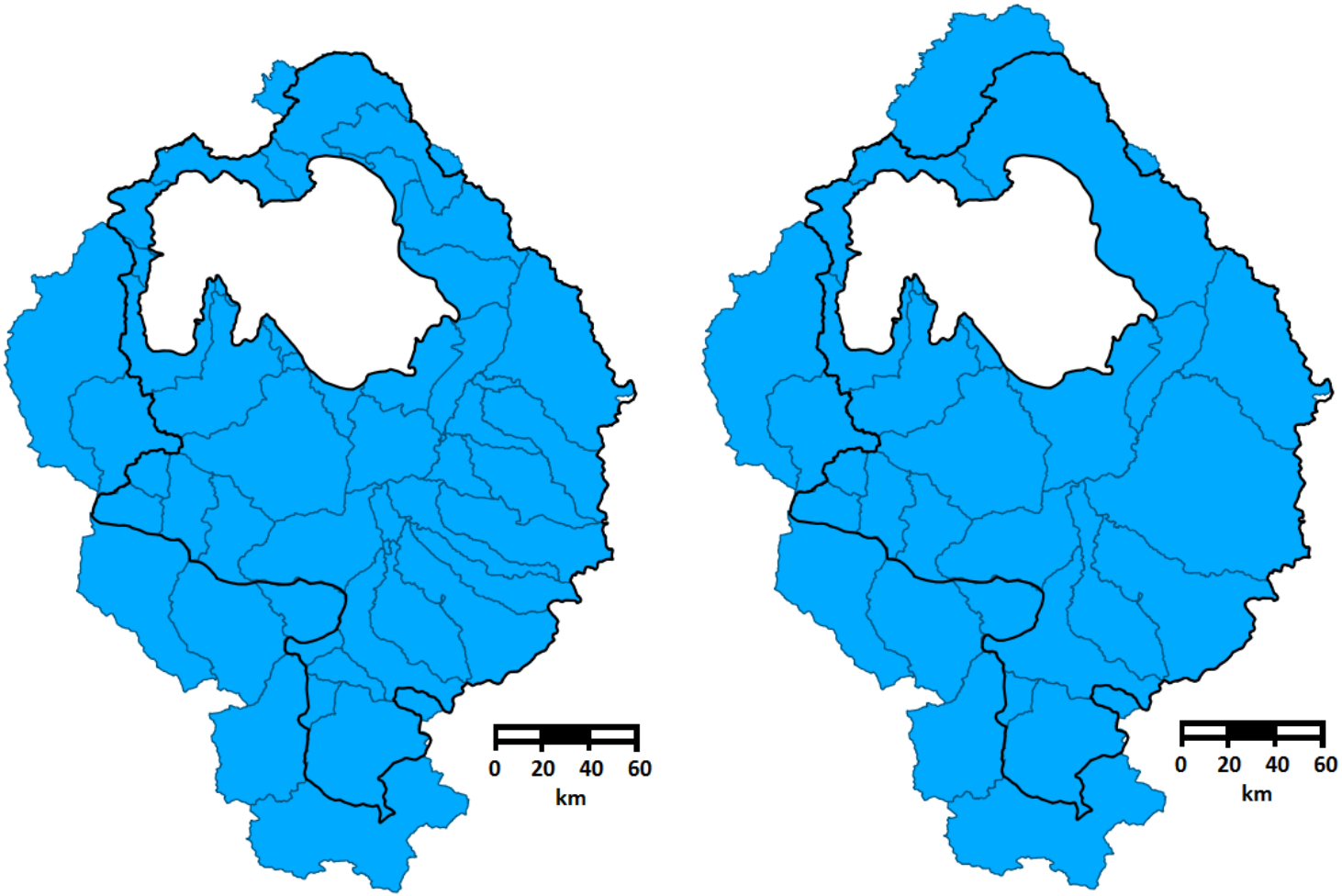

Figure A7: *The resulting watershed as calculated by the r.watershed function in QGIS for threshold values of 1 million [left] and 2 million [right]. The black lines indicate the 'true' watershed created in Google Earth Pro.*

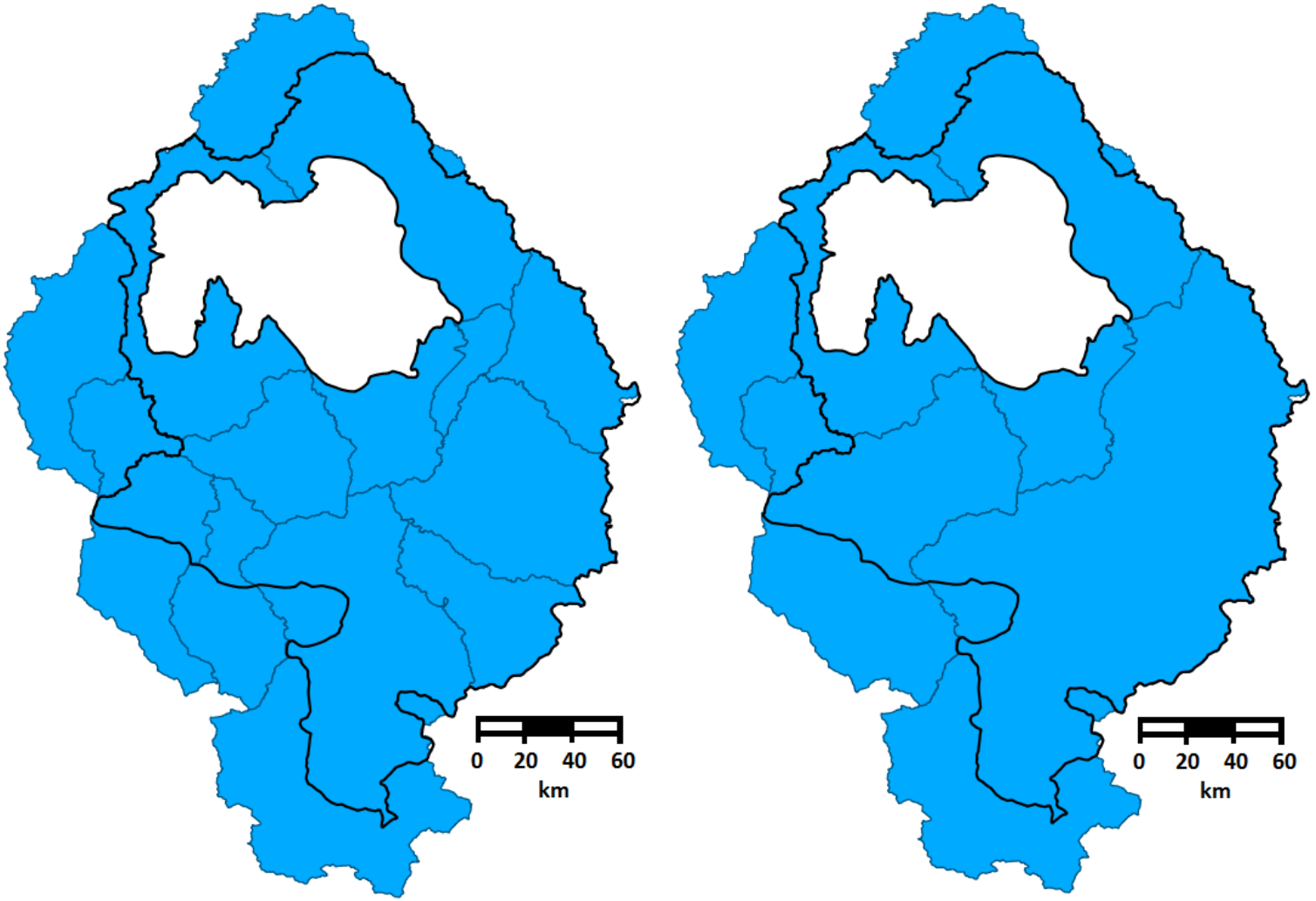

Figure A8: *The resulting watershed as calculated by the r.watershed function in QGIS for threshold values of 3 million [left] and 4 million [right]. The black lines indicate the 'true' watershed created in Google Earth Pro.*

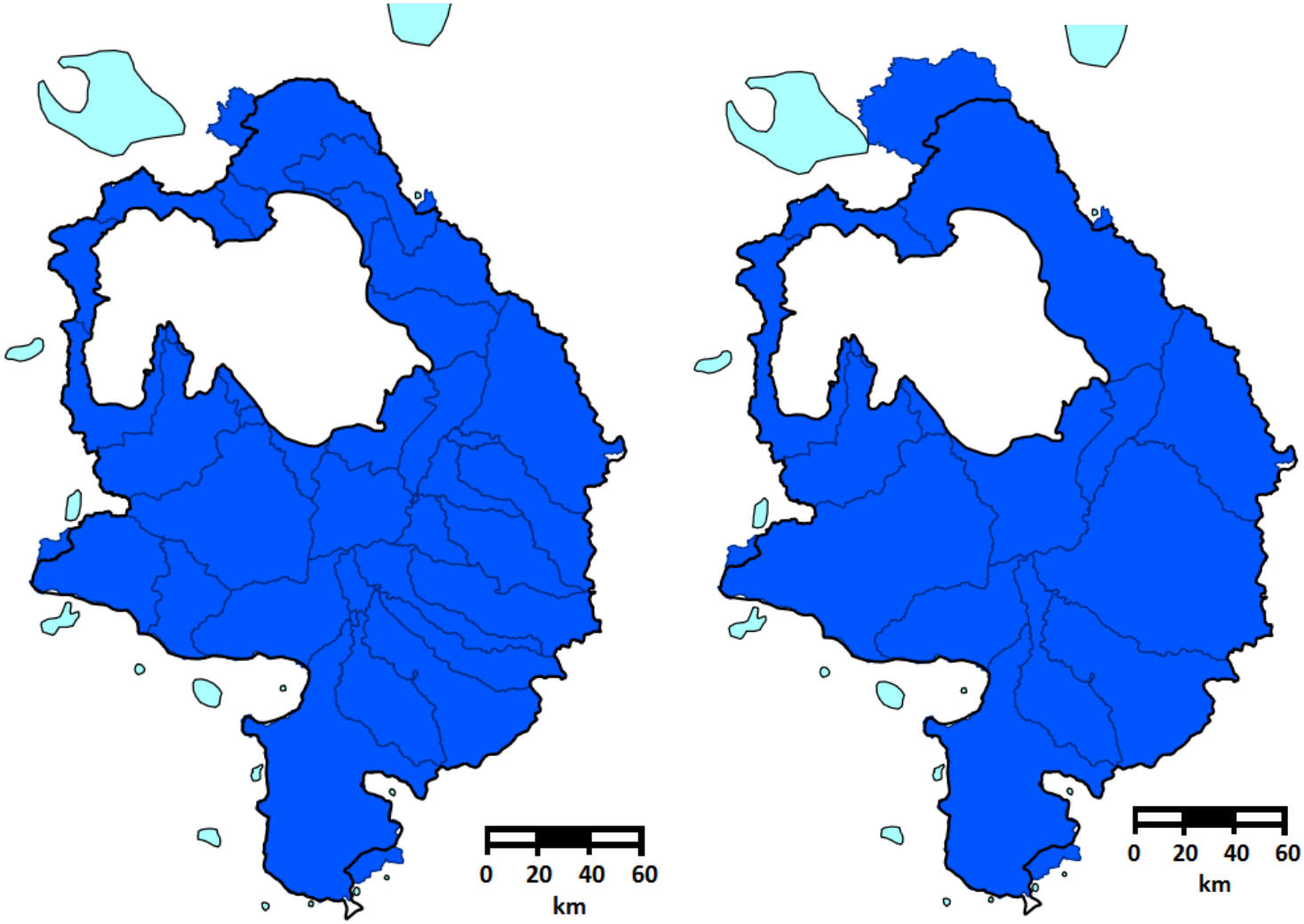

Figure A9: *The resulting watershed as calculated by the r.watershed function in QGIS for threshold values of 1 million [left] and 2 million [right] using the depression location map (light blue) as input. The black lines indicate the 'true' watershed created in Google Earth Pro.*

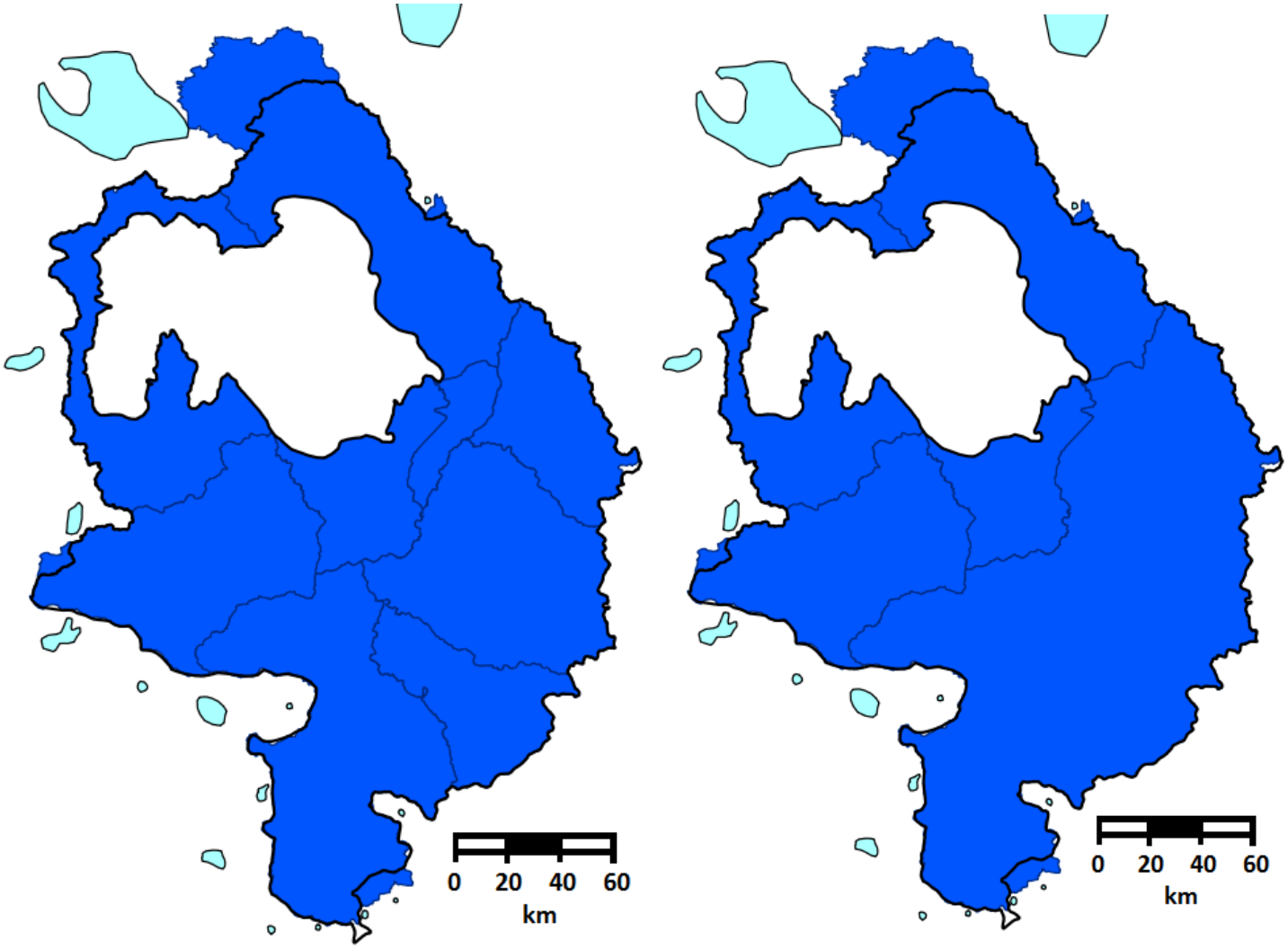

Figure A10: *The resulting watershed as calculated by the r.watershed function in QGIS for threshold values 3 million [left] and 4 million [right] using the depression location map (light blue) as input. The black lines indicate the 'true' watershed created in Google Earth Pro.*

APPENDIX B
RIVER PATH DEFINITION RESULTS

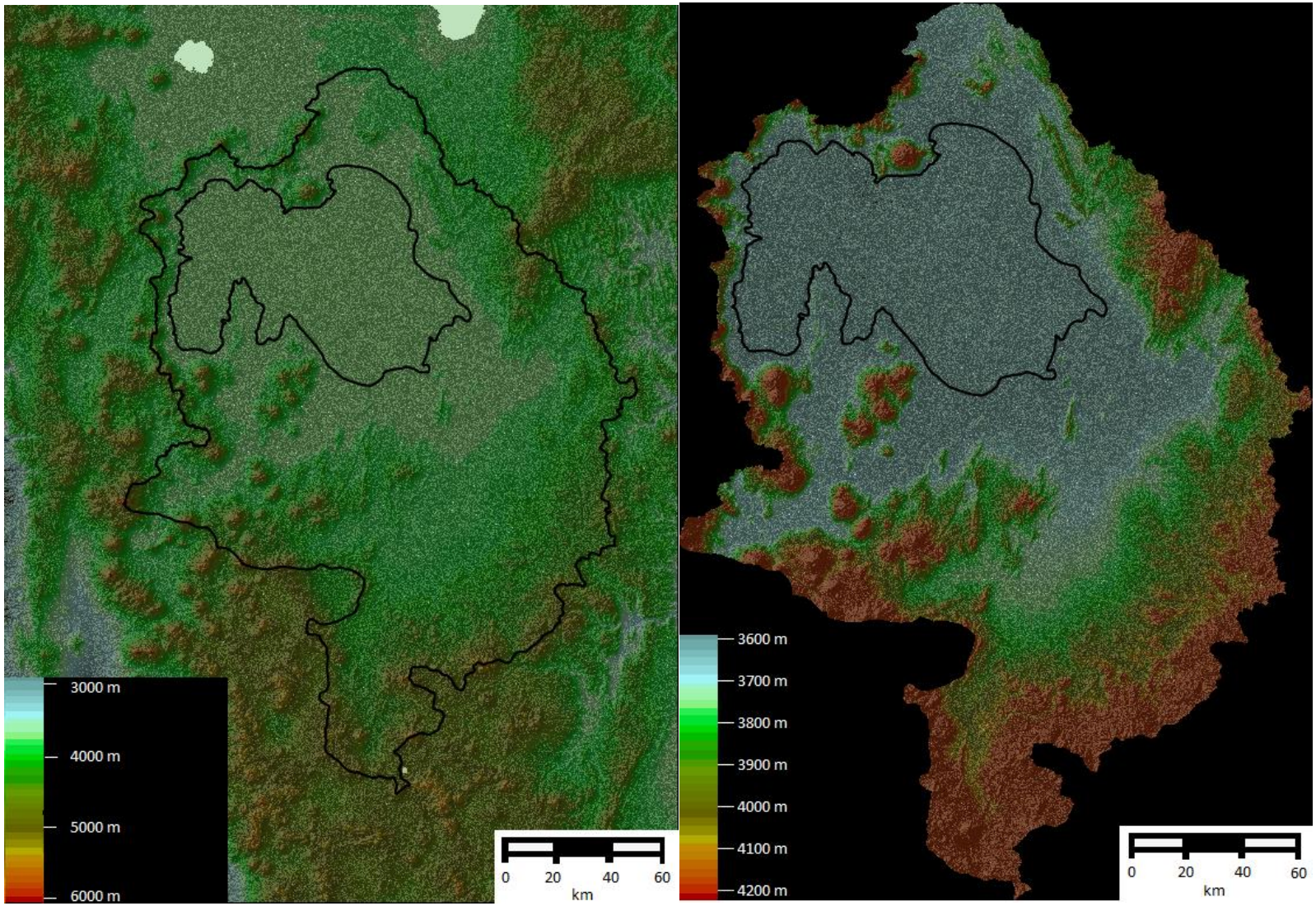

Figure B1: *Relief maps of the area. The left map shows the relief in and around the basin. This map is generated using intervals of 100 m. The right map shows a more detailed relief map within the watershed of the basin. For this map, intervals of 20 meters are used.*

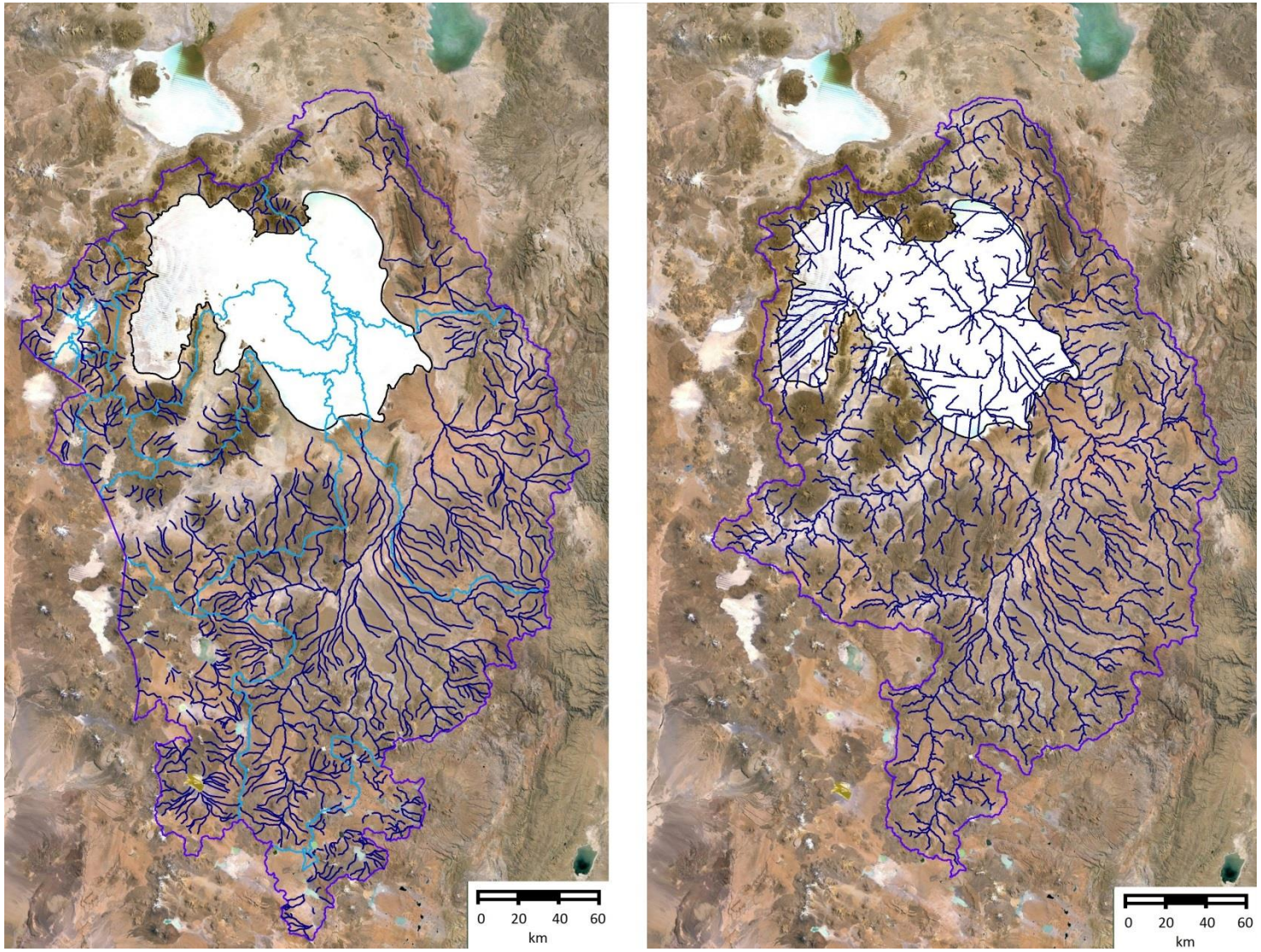

Figure B2: *The watershed and the river paths, as obtained from GeoBolivia [left] and WWF [right].The dark blue lines indicate the river paths, the lighter blue lines the (subdivision of the) watershed.*

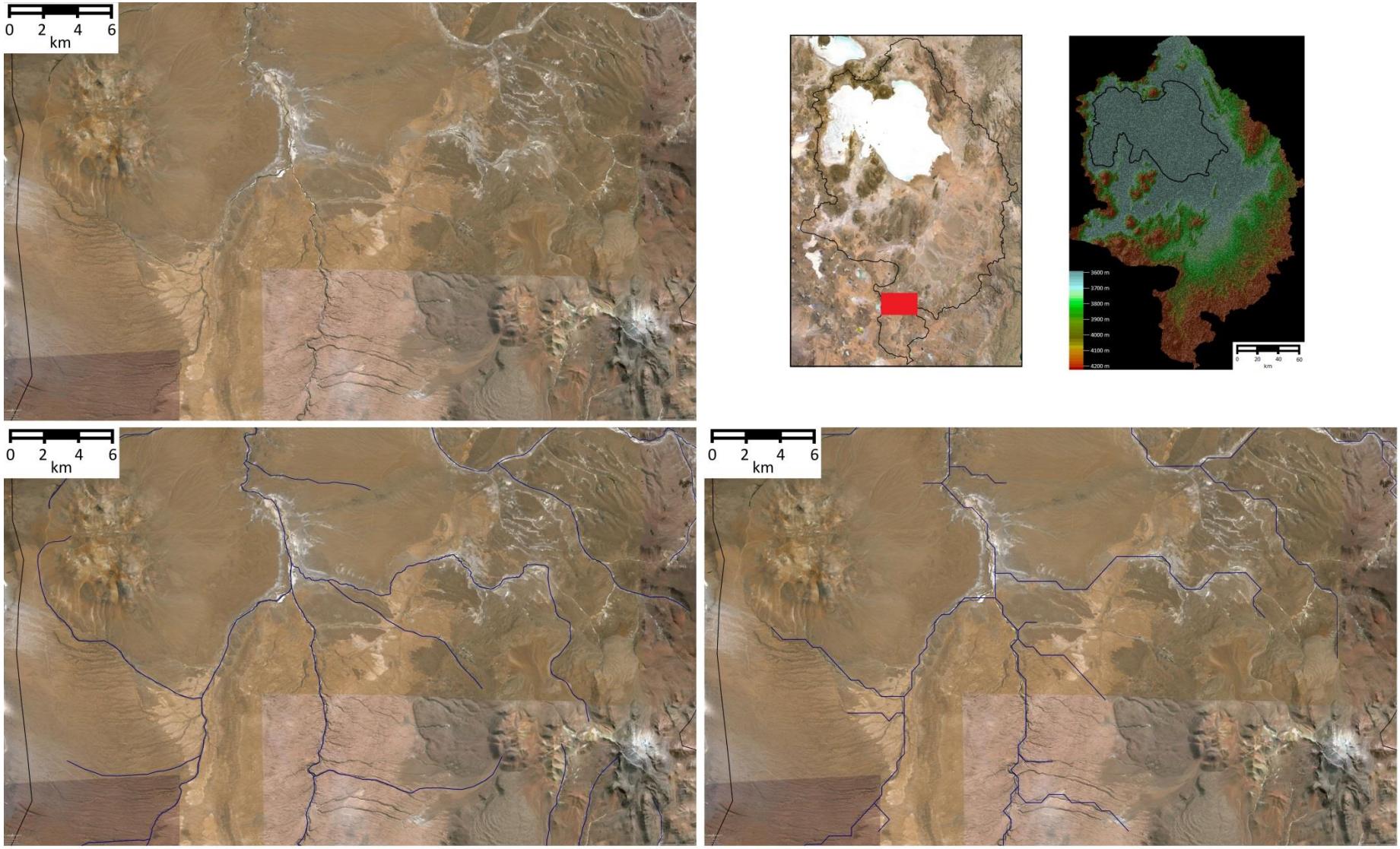

Figure B3: *Comparison of GeoBolivia [bottom-left] and WWF [bottom-right] rivers using Google Earth Pro [top-left], in high relief area. The studied area is marked on the basin map [top-center] in red.*

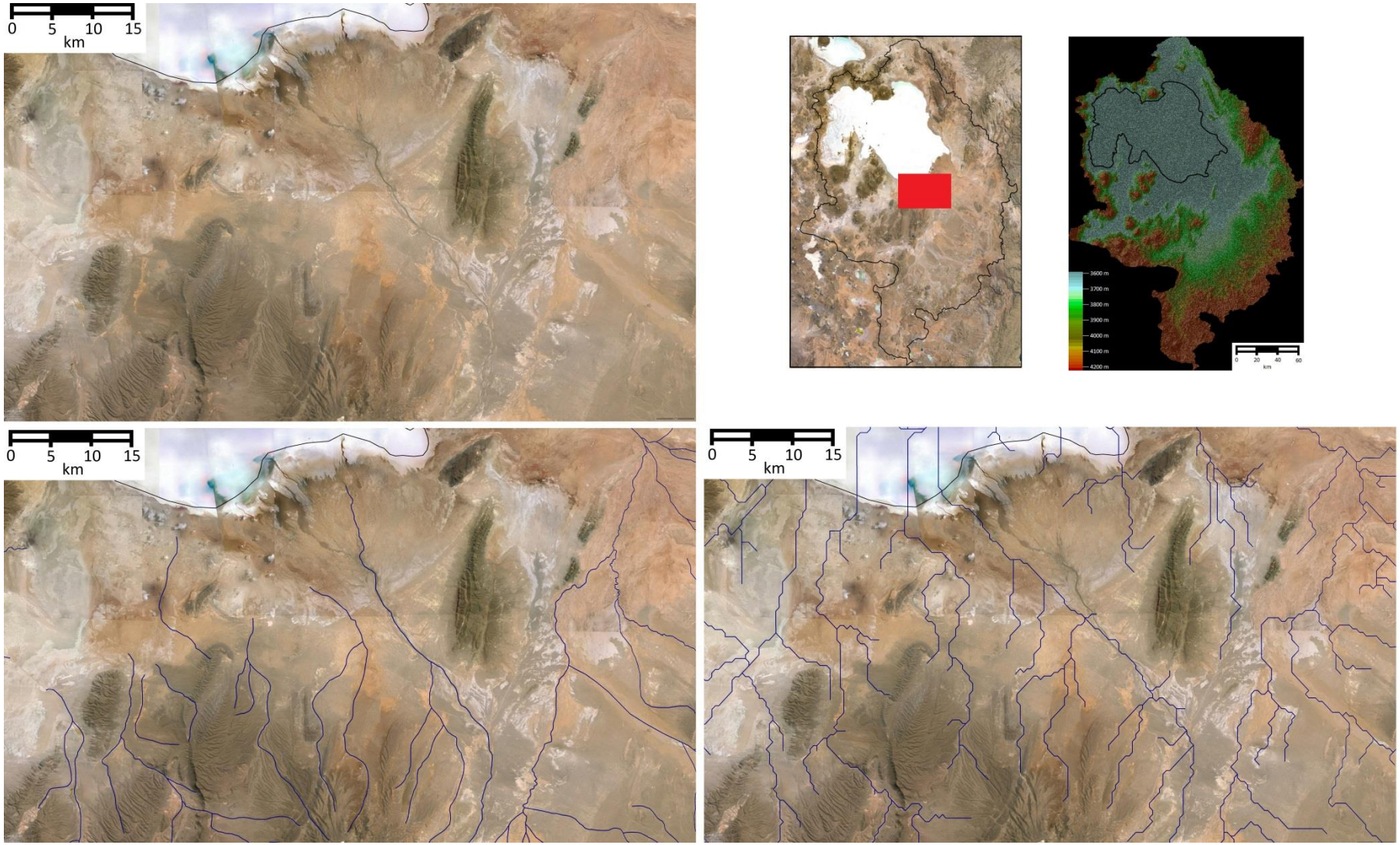

Figure B4: *Comparison of GeoBolivia [bottom-left] and WWF [bottom-right] rivers using Google Earth Pro [top-left], in low relief area (delta of the Rio Grande de Lipez). The studied area is marked on the basin map [top-center] in red.*

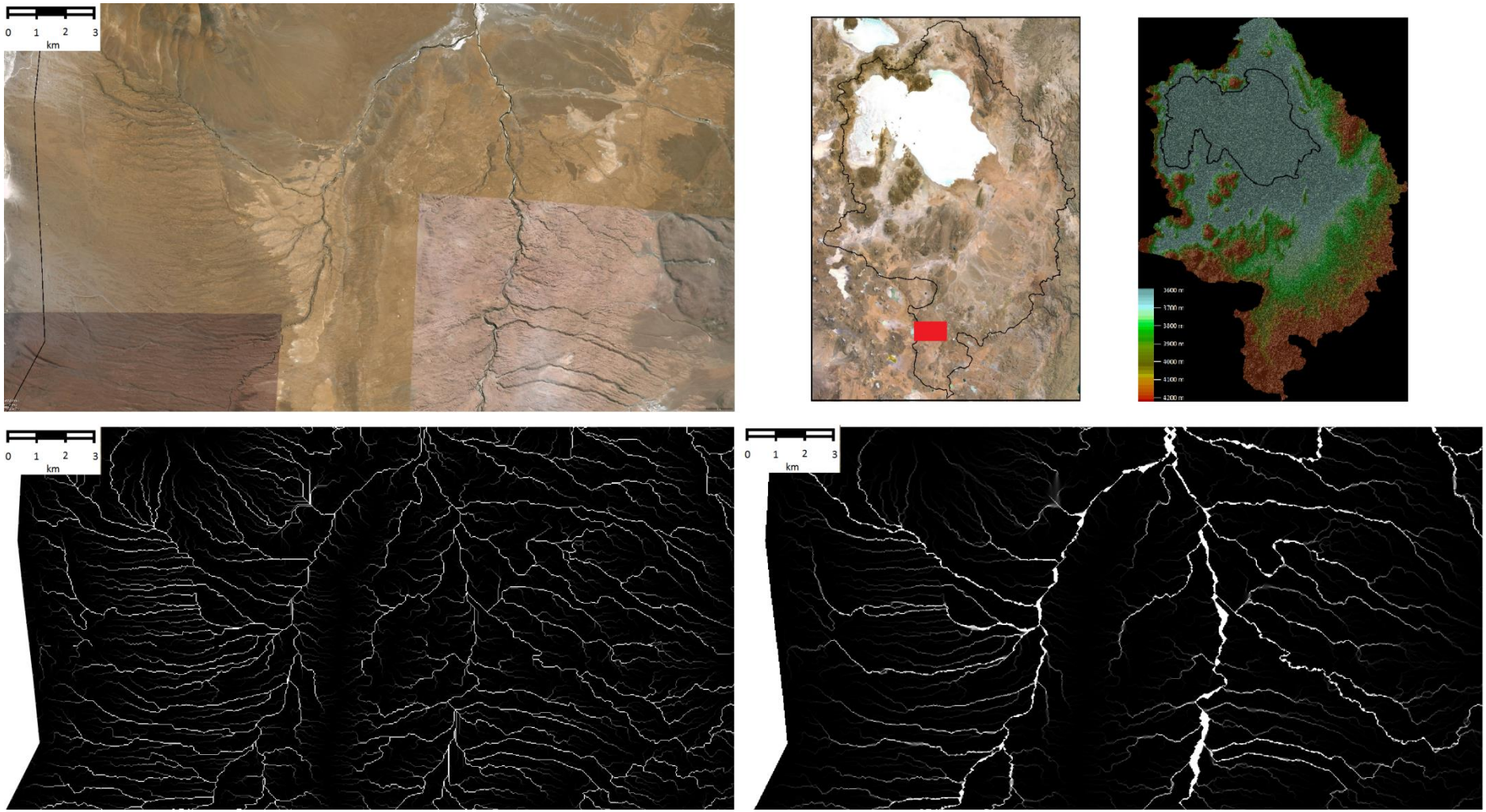

Figure B5: *Comparison of SFD [bottom-left] and MFD [bottom-right]results using Google Earth Pro [top-left] in a high relief area. The studied area is marked on the basin map [top-center] in red.*

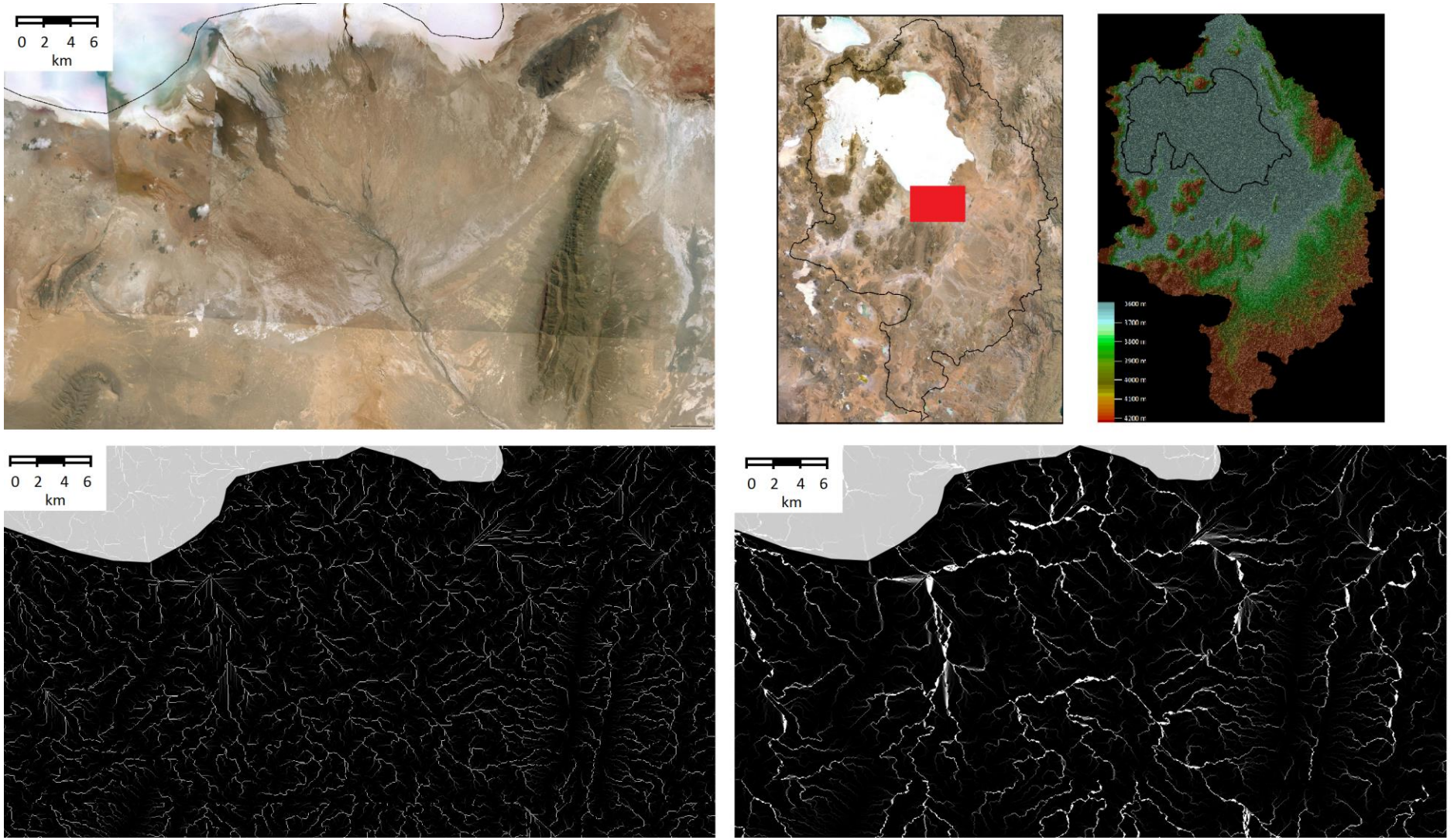

Figure B6: *Comparison of SFD [bottom-left] and MFD [bottom-right] results using Google Earth Pro [top-left] at the delta of the Rio Grande de Lipez river. The studied area is marked on the basin map [top-center] in red.*

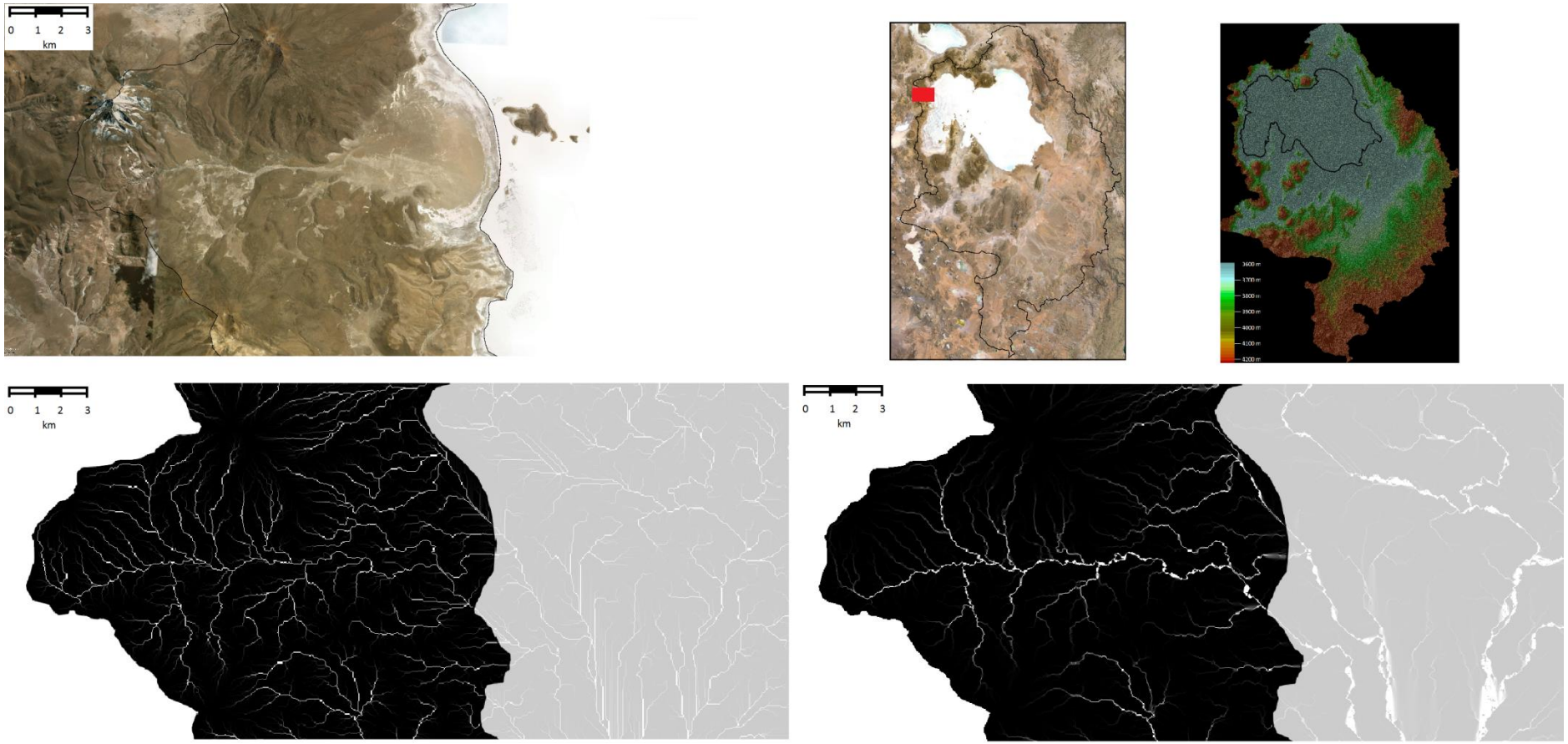

***Figure B7**: Comparison of SFD [bottom-left] and MFD [bottom-right] results using Google Earth Pro [top-left] at the western side of Salar de Uyuni. The studied area is marked on the basin map [top-center] in red.*

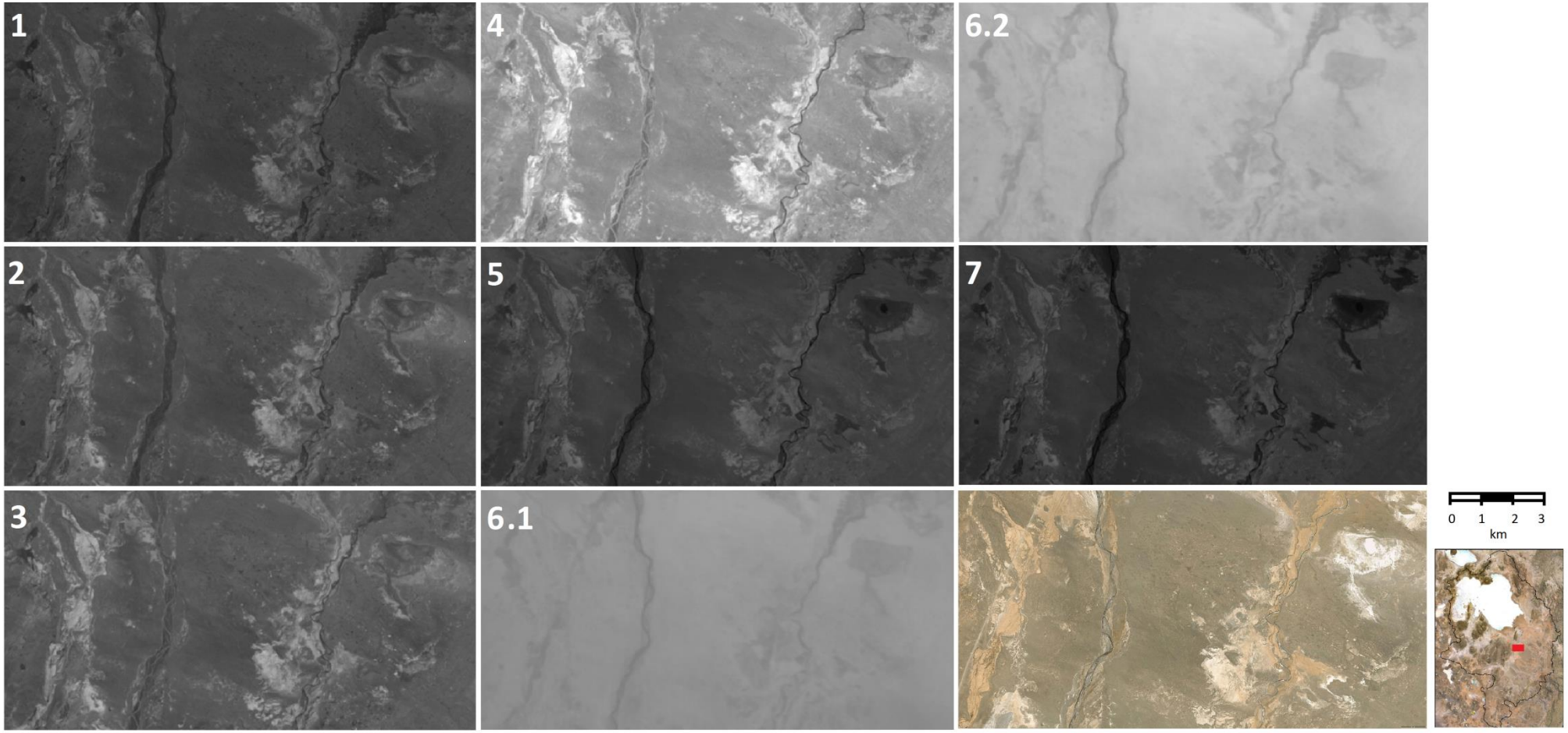

Figure B8: *River response of Landsat ETM+ bands 1 to 7 and in Google Earth Pro [bottom-right].*

APPENDIX C
LANDSAT INTERPRETATION RESULTS

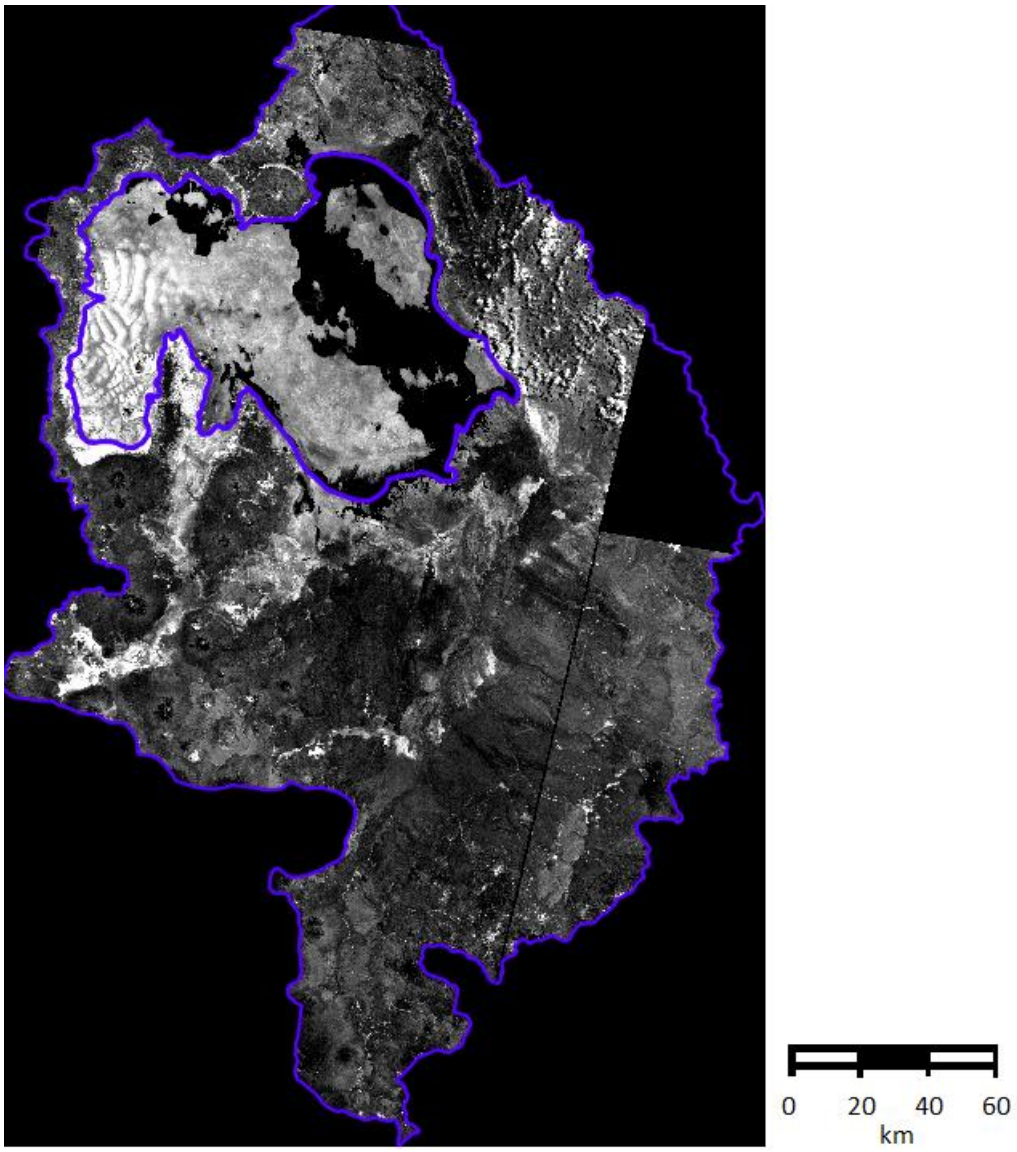

Figure C1: *The difference map between band 5 and 7. White areas indicate areas with high amounts of clay. The watershed and the salt lake outline are given in blue.*

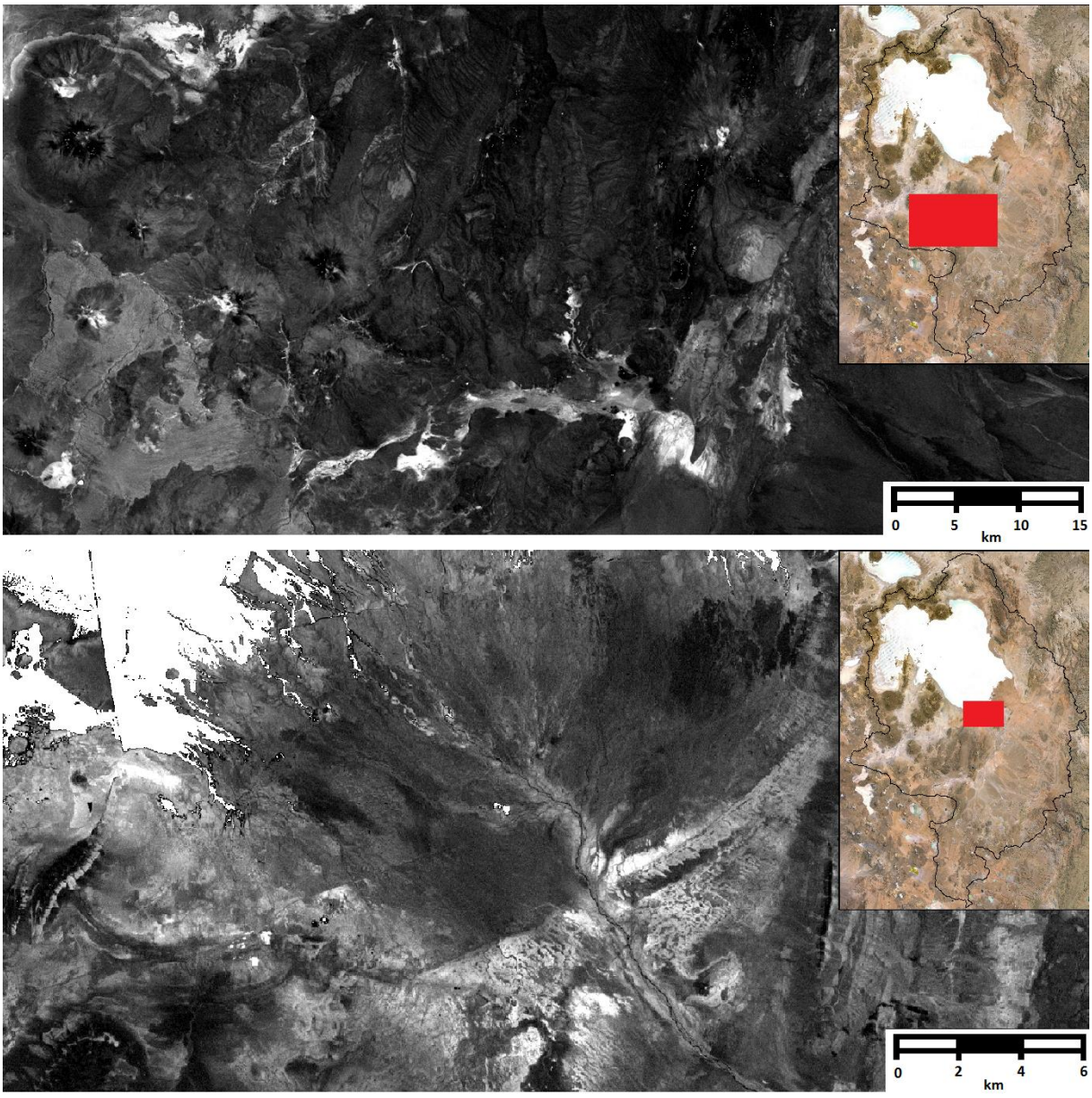

Figure C2: *Details of the difference map between band 5 and 7. White areas indicate areas with high amounts of clay.*

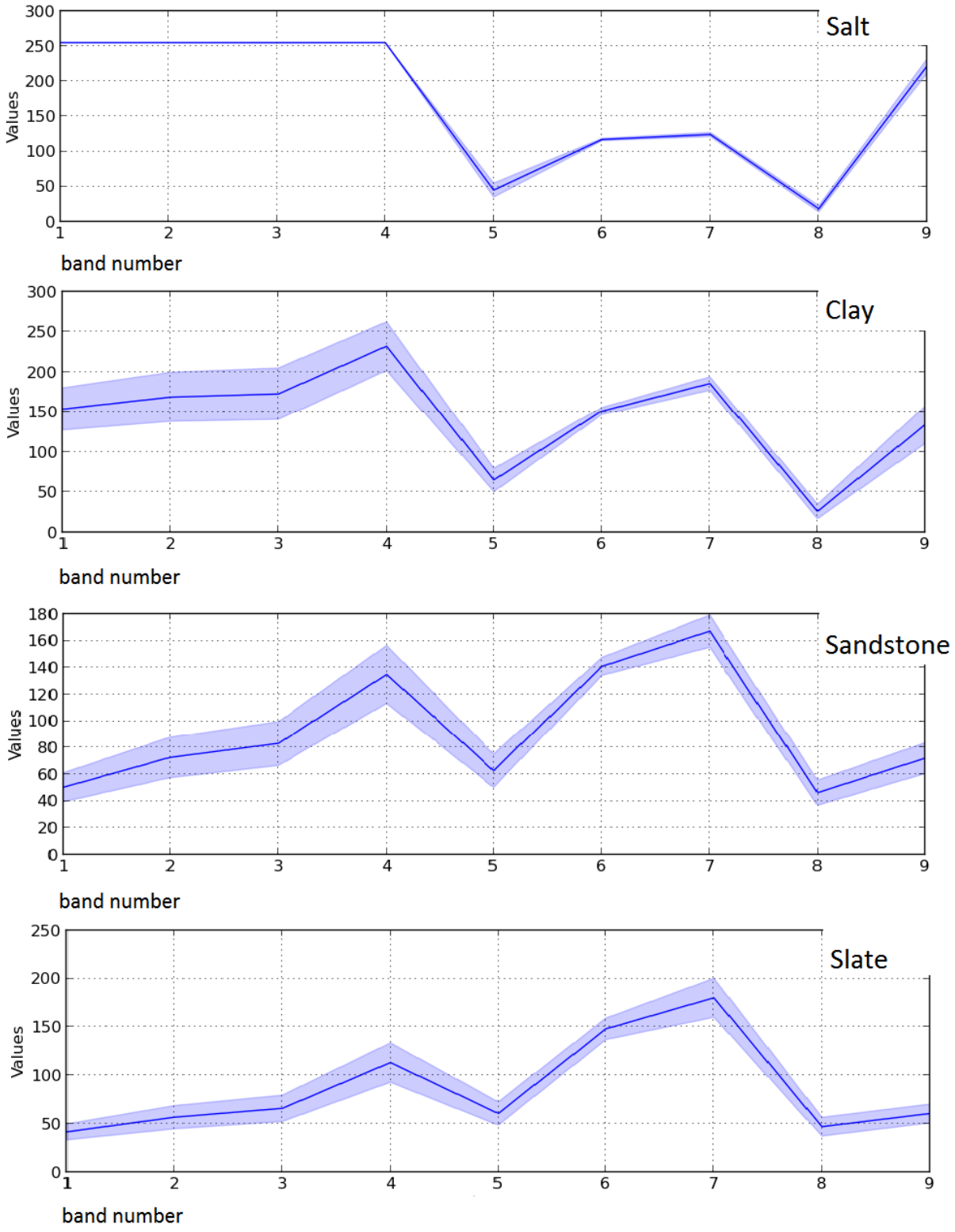

Figure C3: *The spectral signatures of salt, clay, sandstone and slate, as calculated by QGIS. The blue line is the average response, the light blue areas gives the standard deviation. The clay is recognized using the difference map between band 5 and 7, the slate and sandstone using the geological map provided by GeoBolivia.*

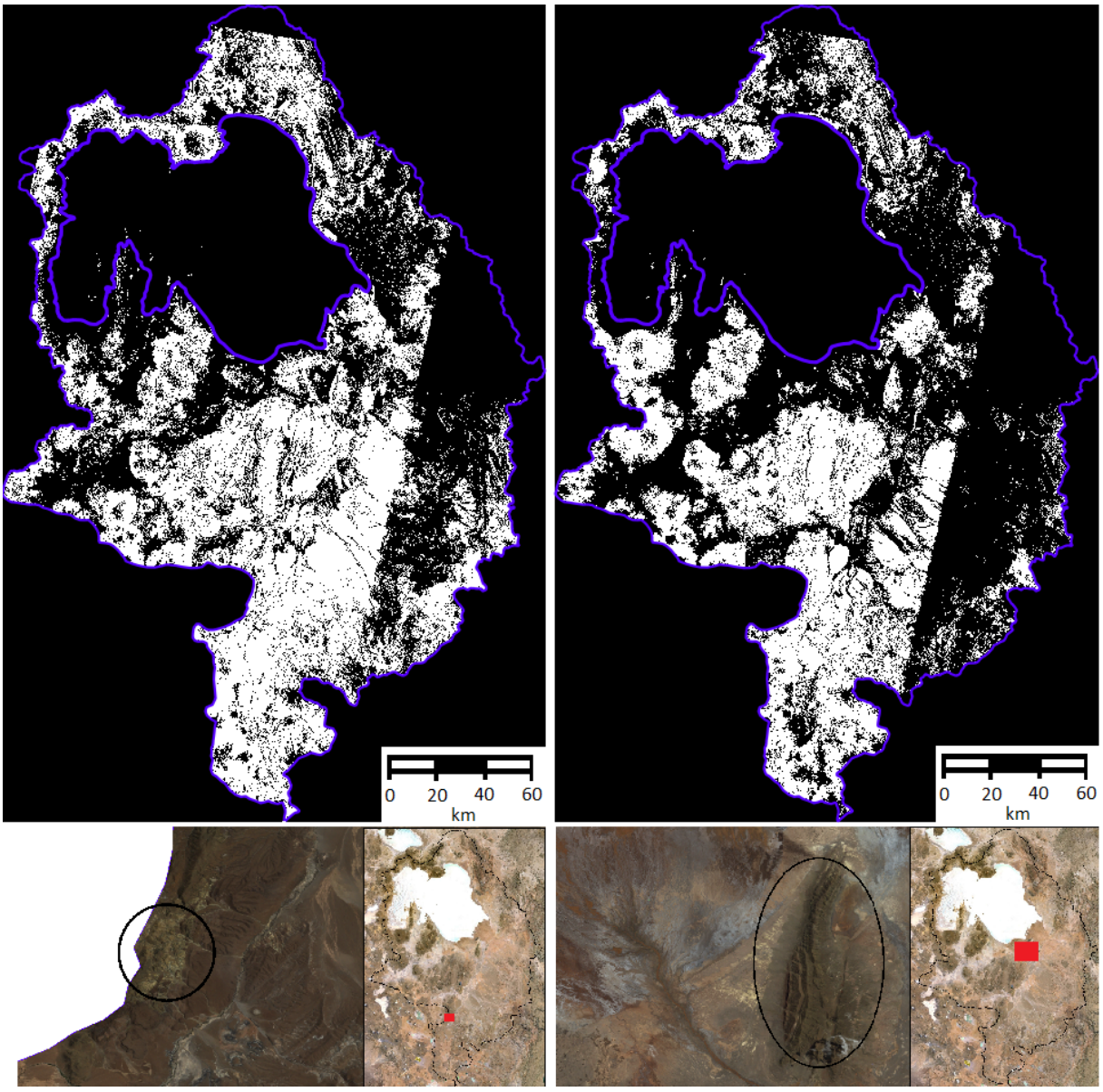

Figure C4: *The result of the semi-automatic classification plugin for sandstone [up left] and slate [up right] using the maximum likelihood algorithm at a threshold value of 99. The sandstone [bottom left] and slate [bottom right] selected as ROI are shown on the bottom most images.*

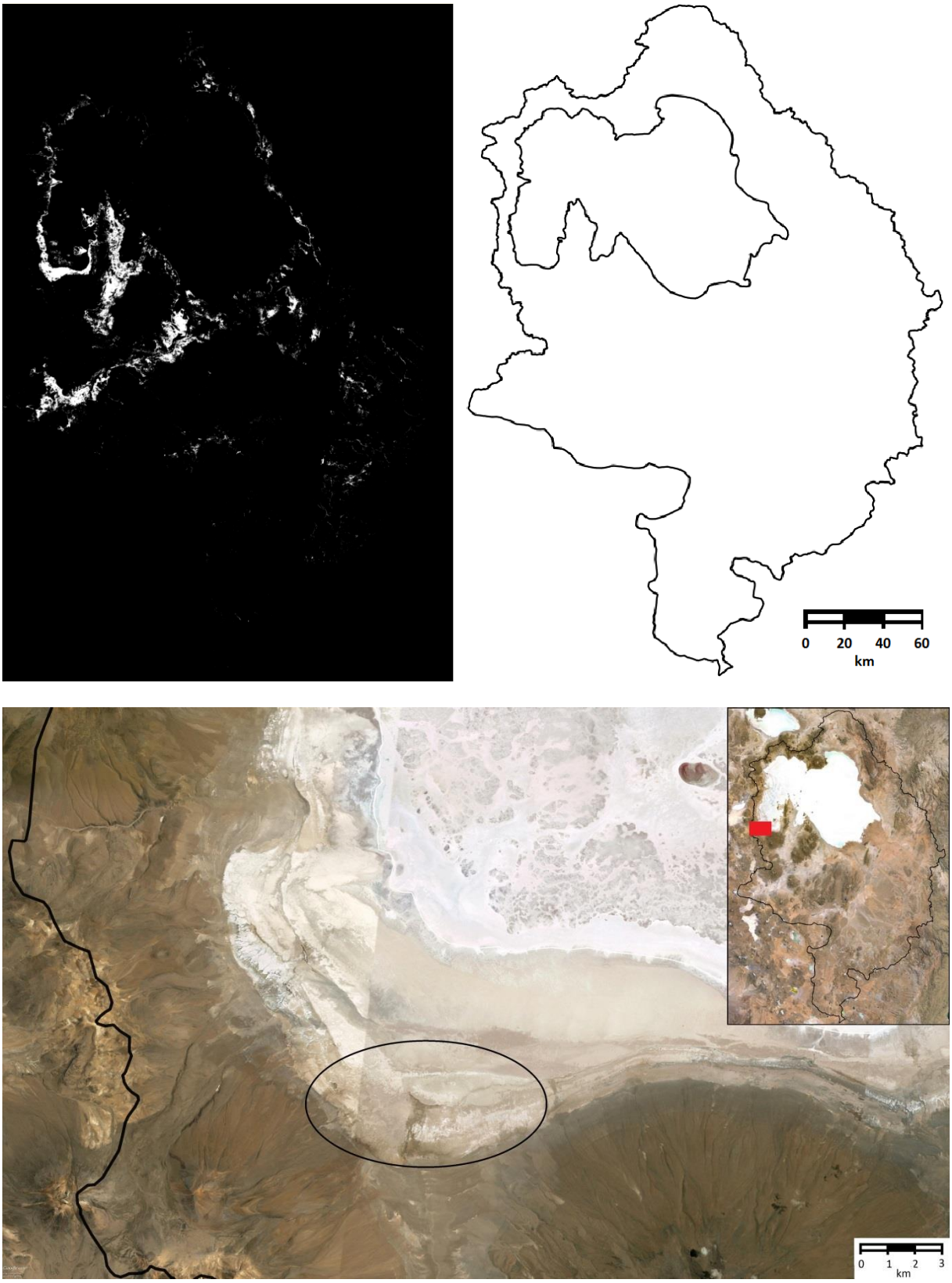

Figure C5: *The result of the semi-automatic classification plugin for clay using the maximum likelihood algorithm at a threshold value of 50 [up left]. The outline of the salt lake and the watershed are also shown [up right]. The clay selected as ROI is shown given in the bottom picture.*