Answers - GLaSS Training material, Lesson #9

Part 1 Choice of data and processing level

Many mine tailing ponds will probably be smaller than the pixel size of medium resolution imagery (e.g. MERIS/Sentinel-3 pixels of 300*300 m are too large). Therefore, high resolution data such as Landsat or Sentinel-2 are required.

A disadvantage of high resolution imagery is the lower sensitivity, which can be too low for monitoring water quality (see lesson 8). However, because the ponds are often very bright in comparison to natural water, the sensitivity is not expected to be a problem. So high resolution data should be used.

Other things that should be considered to make the choice for a specific satellite data type are for example the revisit time (lesson 1), availability, and band settings of the imagery. A low revisit time makes it difficult to monitor quick changes. However, to locate mine tailing ponds a frequent revisit time is not needed, the ponds will usually be in place for years or even decades. We will need to process large amounts of images to cover complete continents. Therefore, imagery should preferably be freely available and easy to download. Last, more bands allow to trace more details in content. The differences in intensity and colour between mine tailing ponds and water generally also expected to be so large that probably no detailed bands are needed to locate them. However, to trace more details such as chemical content, more spectral bands are needed.

In principle, it would be best to use atmospherically corrected (L2) imagery. First, because the atmospheric correction removes a large portion of the recorded light that is not of interest. Small changes in reflectance at ground level might not be significant when working on TOA data, while it might reflect important differences at ground level. Second, by removing the atmosphere and working with reflectance, the data are comparable between different images. A higher reflectance at ground level indeed means a brighter surface, while a higher signal at TOA could be caused by another type of atmosphere (e.g. reflectance by fog).

However, atmospheric correction schemes relay either on assumptions on water properties (e.g. black pixel approach, or methods that include properties of the water and its contents to solve the complete radiative transfer including atmosphere and water), or they are trained on specific water types (e.g. neural networks for atmosphere, which are trained with combinations of TOA and water leaving reflectance), or they need detailed information on the atmosphere at the time of image acquisition (e.g. models that solve the radiative transfer using actual aerosol types etc). Neither of these are applicable for this procedure: water properties of mine tailing ponds are largely out of the range of standard assumptions or training ranges and there are no atmospheric details available for each image when you are processing a complete continent. Therefore, we need to work with TOA (L1) data.

Based on the above, we will work with Landsat 8 data. This is high resolution data (30 m), the data is freely available and relatively easy to download. By the time of writing this exercise, Sentinel 2 data just becomes available. Its specifications are even better than for Landsat 8 (smaller spatial resolution of 10 m, also freely and easily available and more spectral bands). Therefore, Sentinel 2 data will in future probably be an even better option.

Because we will need to work without atmospheric correction, we will first calculate TOA reflectance from the TOA radiance. This transformation corrects the measured radiance for the 'standard solar irradiance', so it corrects for the amount of sunlight that is send to a specific spot on Earth based on the time of the day and the season, and therefore makes the images somewhat more inter-comparable. Information to do this is included in the meta data.

Unless you know the area, it would be very hard to locate the mine tailing pond in the image just by scrolling around.

The radiance units are W/(m2*sr*µm), which is Watt (intensity) per square meter (surface area) per sterradian ('how broad the ray is') per micrometer (within the spectral band of the satellite)

To calculate reflectance, the equation for a band X is:

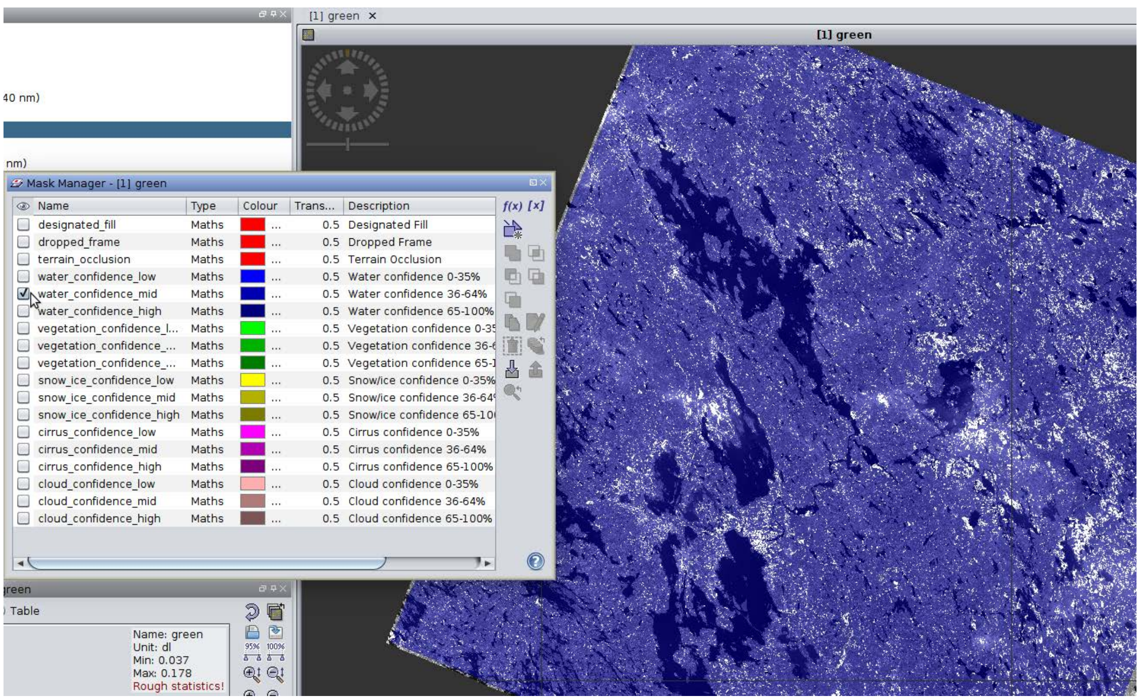
reflectance band X = band_X (DN) * RADIANCE_MULT_BAND_X + RADIANCE_ADD_BAND_X

By eye, you will not see a difference between the radiance and the reflectance data. However, the units are now 'dl', which stands for 'dimensionless'. Reflectance is a ratio between upwelling and downwelling light so this makes sense.

Part 2 Flagging of unnecessary data

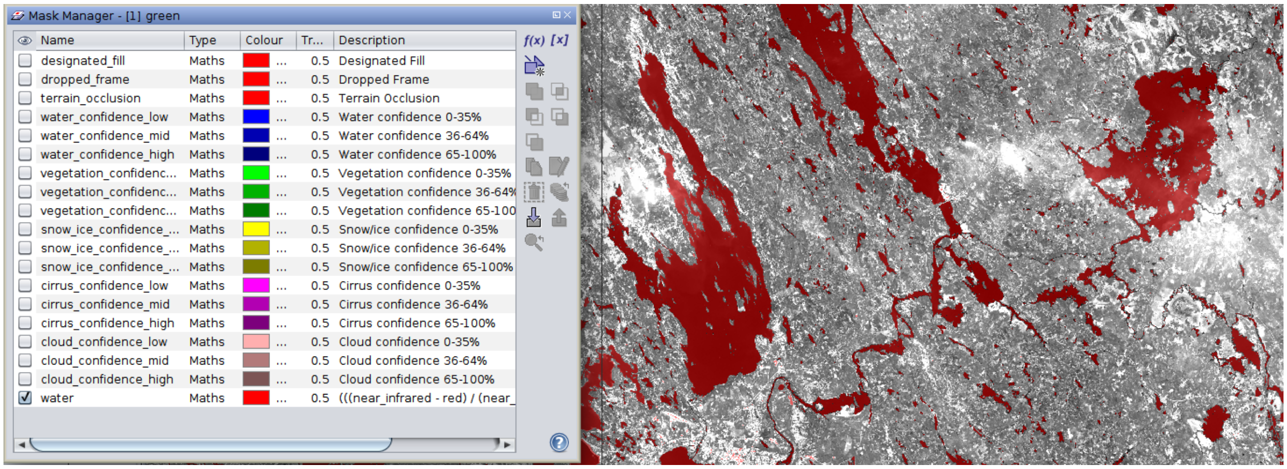
To remove large amounts of data that does not need further processing, it should – theoretically - be the best to first remove land and clouds. The remaining pixels should be water. Within the water pixels we could start looking for mine tailing ponds.

The standard masks are not sufficient. Especially the 'water' masks do not seem to be able to make a proper distinction between land and water. The masks with the lowest and highest water confidence levels do not locate any of the lakes, while the 34-64 % water confidence mask covers almost the complete image.



When we tried more images, we found that for flagging clouds (and cirrus) similar issues occur. Especially the cirrus and cloud flags with medium and low confidence sometimes mask complete images.

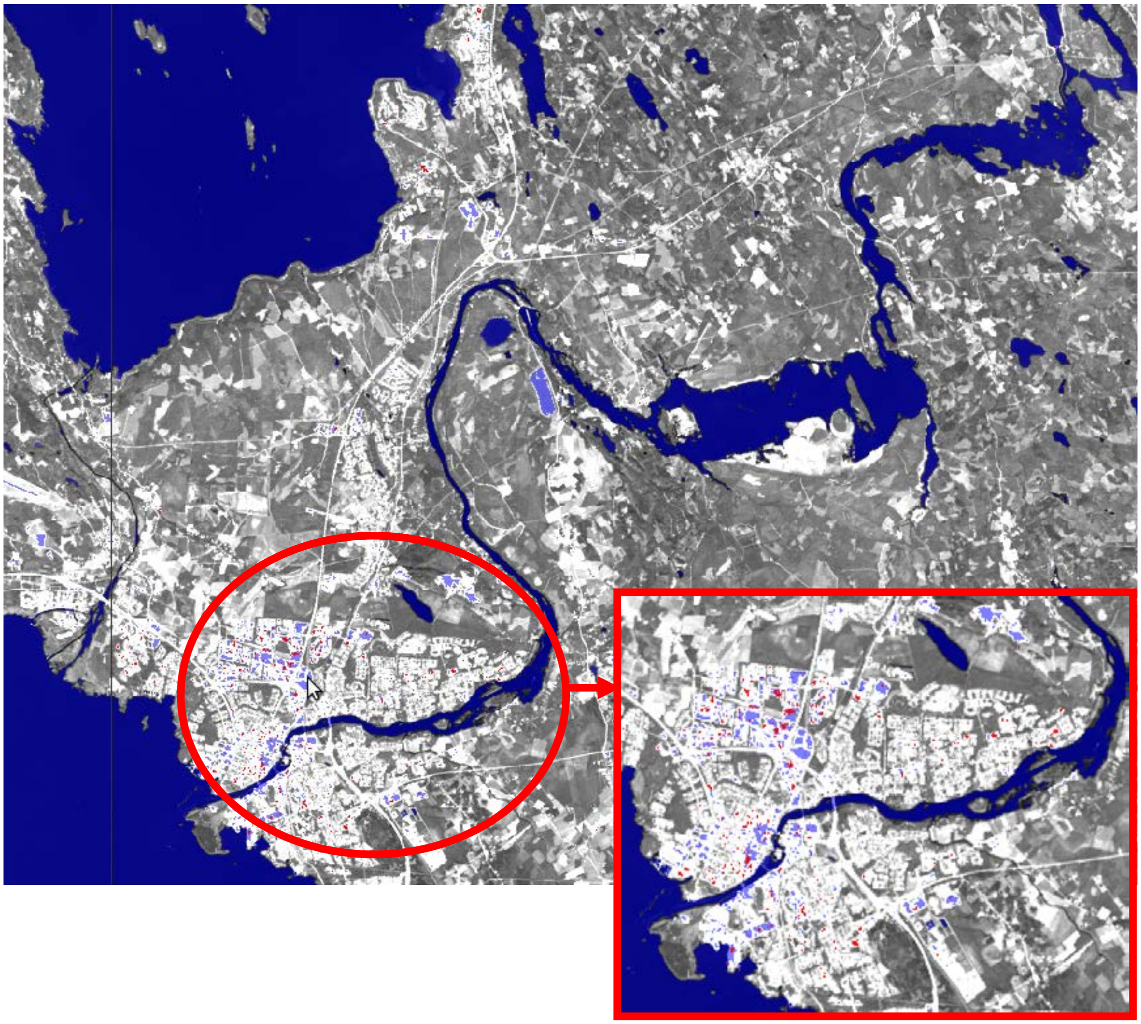
As you can see in the image, our new mask flags out the land, what we keep within our mask is the water.



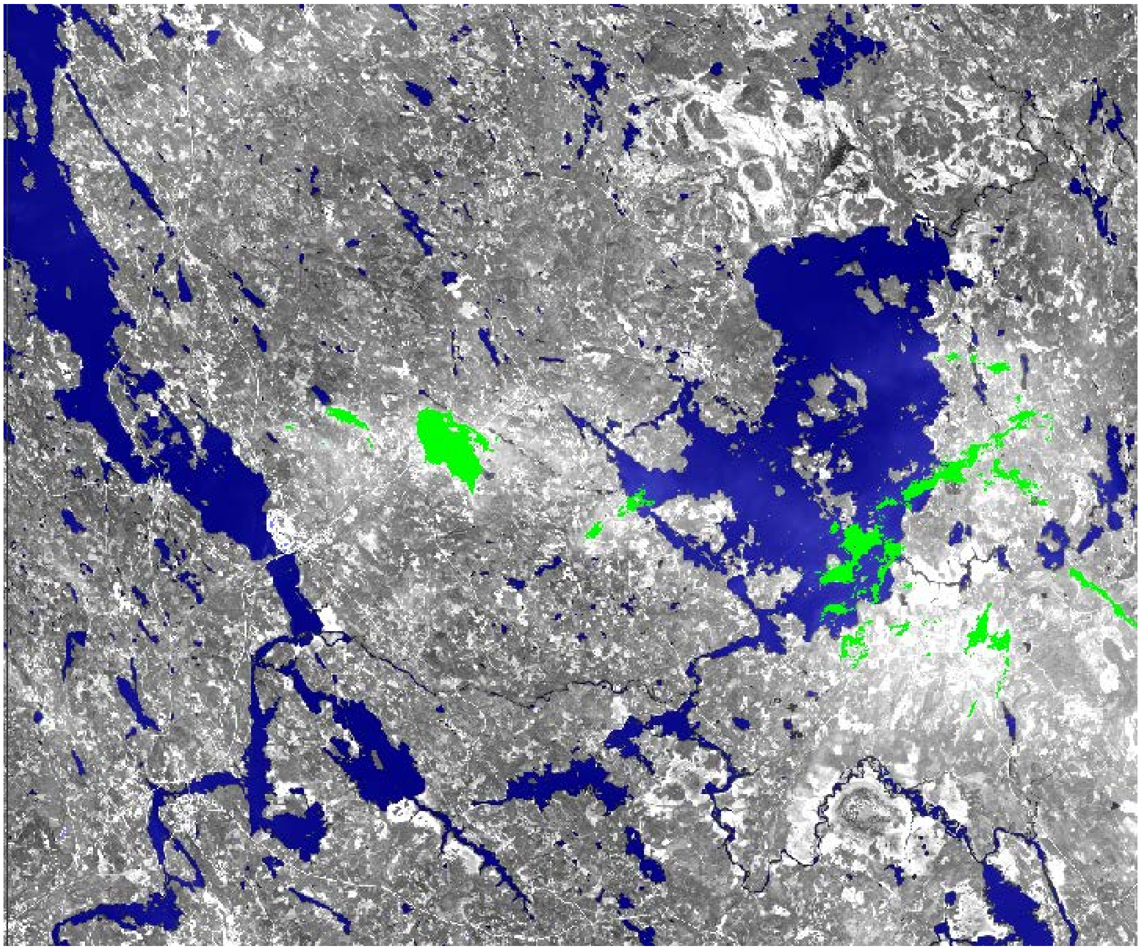
Our new water mask finds water based on the assumption that the reflectance of water in the NIR is lower than in the red, because of the high absorption of water in the NIR. The blue/green part of the spectrum is much more influenced by atmospheric effects and can therefore not be used here. Additionally, the new water mask assumes that the SWIR (short-wave infrared) has to be positive and smaller than 5 (for quality control). The new mask generally selects the darkest pixels, which makes sense, considering that the reflectance of water is generally lower than that of land or clouds. However, there are exceptions.

Assuming that mine tailing ponds have a much higher reflectance than water, we are interested in the locations where there are bright pixels included in the water mask.

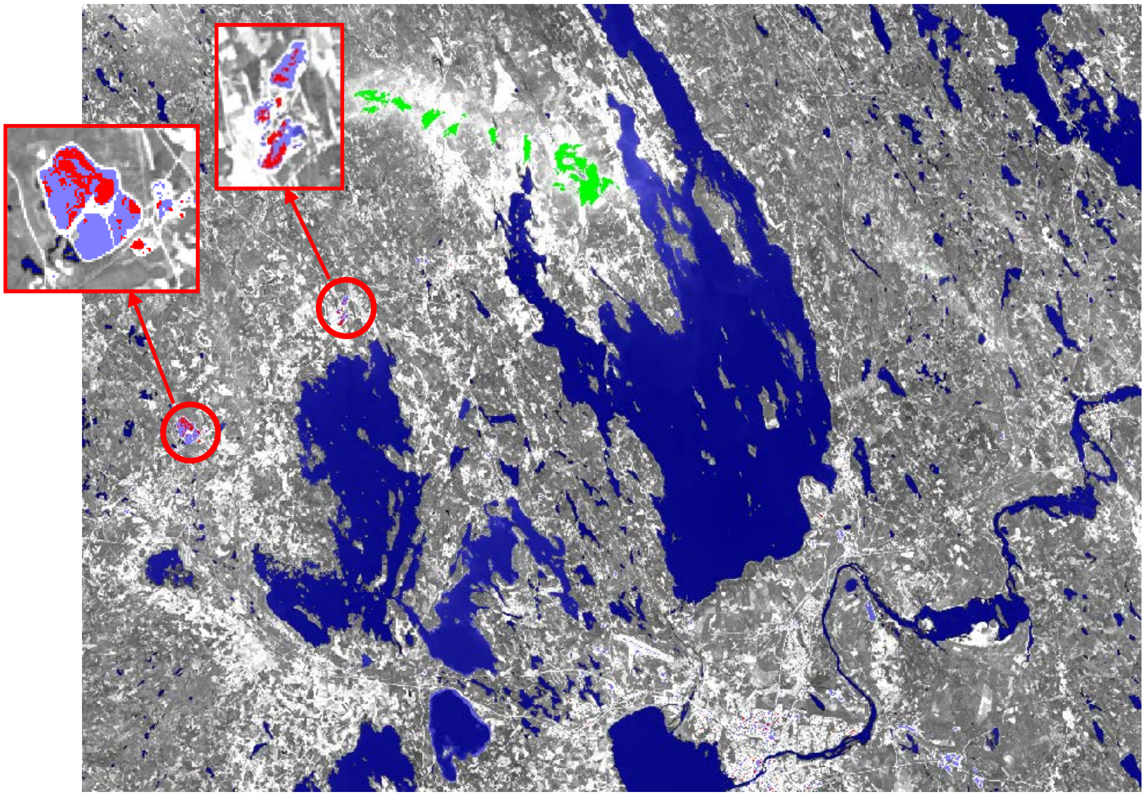
For the area at the south-western end of the river, find the screen shot below. Taken the masks seriously, the area in the circle would be a mosaic of small water and small cloud patches. However, this is probably a city. The roofs of the buildings have a non-natural spectral signature that could fall in water as well as in cloud masks. These are false positives for both the cloud as the water mask.



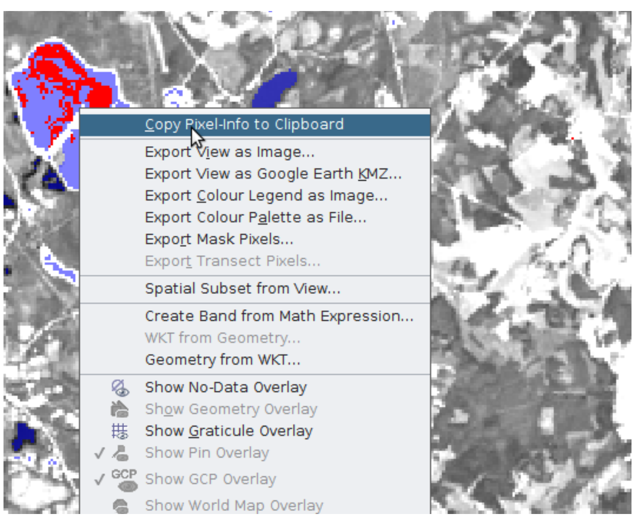
There are some areas for which there is an overlap between water (dark pixels) and high confidence cirrus clouds. Those look like (cirrus)clouds indeed (see below). Applying a cloud flag based on this mask might help to remove unnecessary pixels. However, by far not all of the clouds get removed with this flag. The high confidence cirrus flag seems conservative: it does not seem to show false positives, but has a lot of false negatives.



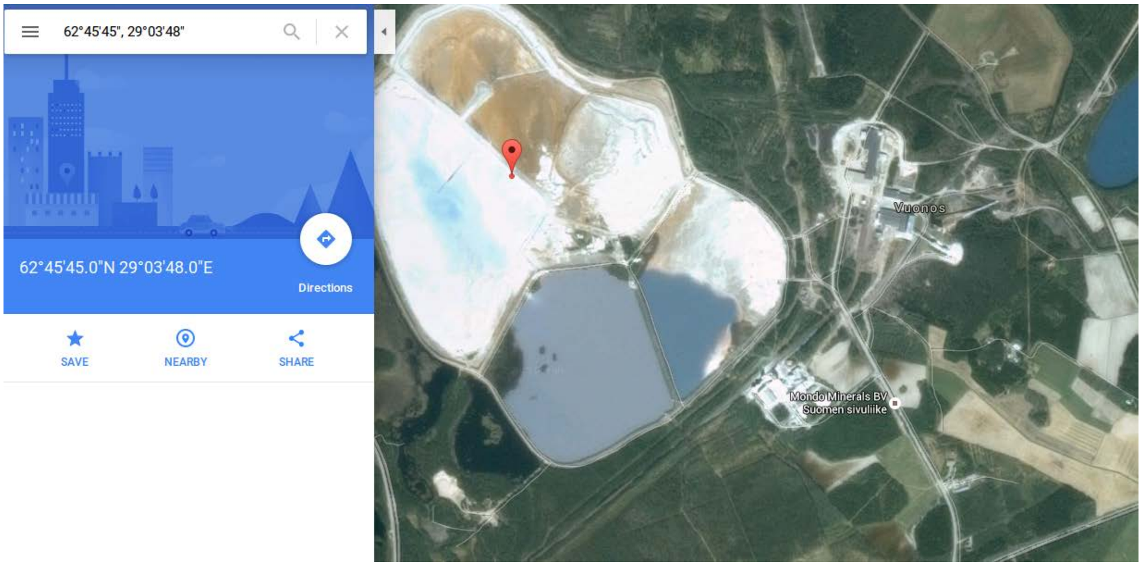
There also some other areas for which there is an overlap between bright pixels in the water mask and high confidence of clouds. Those features do not look like clouds (see below for two examples). It looks like we are also removing other features would we remove these 'cloud' pixels.



Knowing that for Finland there is very detailed spatial data available, we can use external sources. A practical solution is to use a map to look up what features show up as cloud here. Go the 'Pixel Info' tab or click with the right mouse button at the location to copy the pixel info the the clipboard and paste this in a text file. Now you know the GNSS location (lat/lon) of the two locations in the circles of the image above.



Use for example Google maps (select the “Earth” function) to find out what these features are. The feature in the left bottom circle for example is located around Longitude: 29°03'48" E, Latitude: 62°45'45" N.



Checking the website of the company tells us that this is a talc mining company.

The other location in the circle (62°49'33", 29°14'32") is close to the village Horsmanaho and with some additional searches on Internet based on this name this also appears to be a talc mine.

We better not use the cloud flag, as it seems that the cloud flag even shows specifically the mine tailing ponds.

The snow/ice mask masks some pixels in one of the two mines we have found so far. The other mine has a pond connected to it, but that has the colour of natural water, so we could not expect to find it via any algorithm based on the optical signature.

The image was taken in July (middle of the summer), at not very high latitude. There are no tall mountains in this area, so there will not be any real snow or ice.

In the 'Answer data' folder is a spreadsheet with a plot of example reflectance spectra for water, land, cloud and ice/snow taken from this image.

If we check the two equations, the water is traced with ((NIR - red) / (NIR + red) < 0.1, or, in other words: for water the reflectance in the red band is lower than the reflectance in the NIR. This is correct and makes sense. However, this also applies to snow/ice.

The reflectance in the NIR and SWIR_1 bands are similar for snow/ice and water. However, for snow/ice the reflectances in the visible part of the spectrum are much higher than for water.

Checking the NDSI ((green - swir_1) / (green + swir_1)) in some more detail, this equation can indeed be used to distinguish snow/ice from clouds. For clouds, the reflectance in the green is lower than the reflectance in swir_1 (so the NDSI will be small), while for ice/snow the reflectance in the green is higher than the reflectance in swir_1 (so the NDSI will be larger). In our example in the spreadsheet, the NDSI is even negative for the cloudy pixel, so it will not be masked as clouds if this was done just based on this equation (it is also based on the weighting with the other mask functions). It does illustrate why the cloud mask provided by Landsat 8 shows so many false negatives. However, for larger, thicker clouds the equation 0.25 < NDSI < 0.7 works.

If we expect mine tailing ponds to show up as water but more bright, it makes sense that the pixels we are interested in are often masked as snow. Three of our four mines with ponds show up with some pixels classified as snow/ice in this image. The talc mine close to Horsmanaho has some water/ponds nearby but the colour of that is very similar to natural water and therefore not recognised.

Part 3 Locating mines

There are some other pixels that appear in the both the our water mask and the snow/ice confidence high mask. Checking the coordinates and Internet learns us that those pixels are located in the areas of:

62°56'08", 28°42'53", which is a Copper mine. The copper mine shows as water and snow/ice, while there is no water: the bare rocks resemble water with a high reflection.

and

63°10'18", 29°26'43", a Soapstone mine.

So far, the combined masks (water and snow/ice confidence high) seems to work to locate mine (tailing pond)s. Two mines with ponds are properly located. One mine without pond also shows, but that is not a problem. One mine with a pond does not show up (false negative) but although this is a problem for a user of this algorithm, it cannot be expected that a tailing pond with a natural water colour can be traced as mine tailing pond.

The first answer on the Mongolian case depends on the locations you are checking.

At 48.54623 N, 105.27106 E there is no mine. It is a natural lake, and therefore for our purpose, a false positive. However, it has optical properties that resemble mine tailing ponds (water, highly reflecting). So an algorithm cannot be prevented to locate natural lakes or ponds like this.

In Mongolia, the combination of the two masks seems to work, generally, to locate mines with tailing ponds. However, there are some false positives (highly reflecting lakes). We have also found at least one false negative (a part of a mine with tailing ponds that does not show up), however, those are hard to count because that requires to know where are mine tailing ponds are in this area.

Part 4 Options for further improvement

In Azerbaijan, for the original water mask, a lot of land pixels show up in the mask. Somehow, areas with bare rocks tend to show up as water in our mask, and when this is combined with something bright, the algorithm leads to many false positives (potential mine tailing ponds in areas where there are no mines).

The new water2 mask works better for this image.

However, the pixels that are now in the combined mask as 'potential mine tailing ponds' are not mine tailing ponds. The plumes of the river Kür where it enters the oceanic waters show as potential mine tailings. They are water and highly reflecting, but these are certainly false positives. Another category of pixels that show up are a combination of real snow/ice on mountain tops with clouds.

Options to improve the screening for potential mine tailing ponds are:

- Improve the water mask (see example for the image from Azerbaijan)
- Improve the cloud mask. Or: use several images of the same location. Mine tailing ponds will show up in all images, but clouds will only show up occasionally.
- Improve the mask that looks for bright waters, instead of using the snow/ice mask
- Use higher resolution imagery to find also smaller ponds
- Use a sensor with more spectral bands, to improve the mask that selects bright waters. This might even allow locating ponds with specific elements dissolved in it.