3. Materials and Methods

3.1 Study area

Lake Leman is located on the north side of the Alps (46º26'N, 6º23'E), at an altitude of 372 m, on the border between Switzerland and France. It constitutes the biggest and most important inland water basin in Western Europe with a surface of 580 km^2, maximum length of 72 km and maximum width of 14 km (CIPEL, 2016). The maximum and mean depths are 309 m and 152 m respectively. The lake contains a quantity of 89 km^3 of water and supplies with drinking water more than 800'000 people. It counts many tributaries and receives water from both Switzerland and France, with its main tributary being the Rhone; it alone contributes 75 % of the total inflow in the lake.

Lake Leman underwent a strong cultural eutrophication in the early 1960s and it has been monitored since 1971. Phosphorous concentrations were at 90 µg L^{-1}, however with the implementation of phosphorous reduction measures in 1972, levels dropped to almost 30 µg L^{-1} in 2005 (Tandoléké, et al., 2009). Long-term monitoring of the lake for almost 45 years has generated a unique in situ measurements database, available to examine changes in water quality status of the lake. Many studies have analysed long-term changes in lakes productivity and trophic level status (Anneville & Pelletier, 2000; Tandoléké, et al., 2009). However, this study as it will be seen in detail hereafter was based on short-term observations due to the recent availability of our examined satellite dataset.

3.2 Aimed optical active components

As analyzed in 2.1 there are several optical active water quality variables, able to be retrieved by satellites datasets. In the current study, which is based on a short-term satellite dataset, we aimed mainly to observe general trends of phytoplankton density (represented by Chl-a concentrations), which is generally influenced by higher content of lakes' main nutrients (phosphates and nitrates). The choice of aimed optical active components was basically limited by the low availability of in-situ data due to the recent date of analyzed satellite images.

3.3 Validation with in situ measurements

Validation of satellite images treatment results was made using an independent unpublished in situ dataset provided by Moisset (Moisset , 2016). It is important here to notice that since this is a rather qualitative approach study than quantitative, available data will be used for simple comparisons and not extensive and precise correlation.

Provided in situ dataset, consisted of two different sampling points in the lake, SHL2 and GE3 (Figure 12). It included measurements of Chl-a, phosphates (PO_4) and nitrates (NO_3) at several lake depths and in two different dates in August 2015, on the 3rd and the 14th (full table of data is given in Appendix A1). As analyzed in 2.1, light penetration gets attenuated after the first meters of water depth, hence any validation was made taking into account mean values of measurements included in the first 10 m of depth.

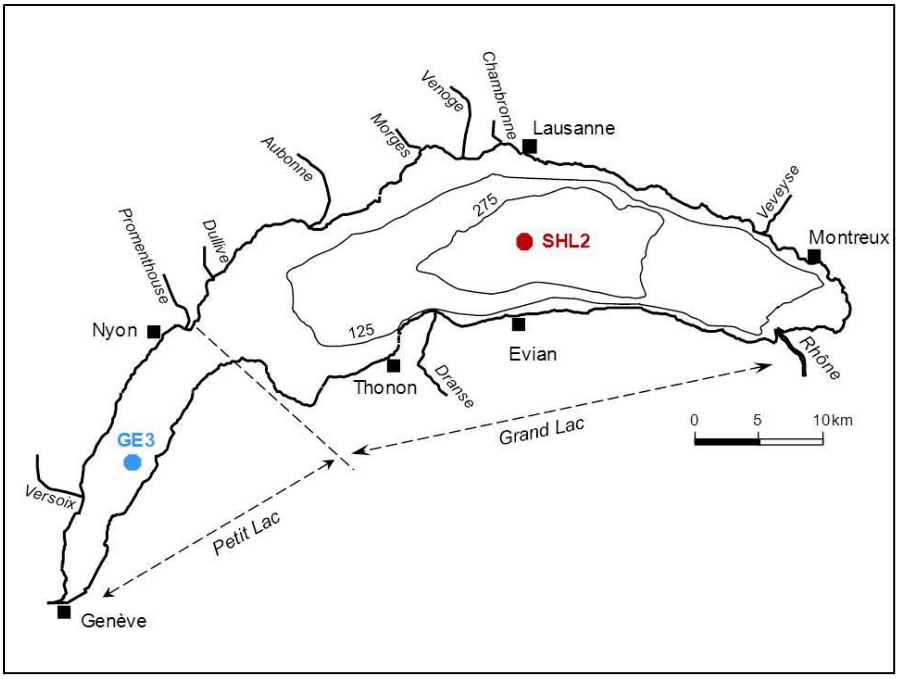

Figure 12: In situ sampling sites in Lake Geneva (CIPEL, 2016).

3.4 Satellite datasets selection

Our satellite dataset consisted of two different satellite missions' images, Sentinel-2A and Landsat 7. According to the availability of in situ data (in August 2015) and the simultaneous availability of both Sentinel and Landsat captures, we concluded to a common image selection on the 29th of August 2015. To be more precise, available captures of Sentinel-2 in August 2015 were for the 12th, 19th, 22nd and the 29th of August. However, images of the 12th and 22nd were unsuitable due to the fact that they did not include the lake in total and capture on the 19th of August was covered by clouds (relevant information illustrations in Appendix A2). Available Landsat captures in August 2015 were, for Landsat 7 on the 13th and the 29th of August and for Landsat 8 on the 5th and 21st of August (Appendix A3). Thus, available common captures for both satellite missions was that of the 29th of August for Sentinel-2A and Landsat 7 satellites. In situ data, as explained in 3.3, dated on the 3rd and 14th of August, which has a time discrepancy with our satellite dataset but is still valid to draw conclusions.

Data acquisition for Sentinel-2A scenes is possible in the Scientific Data Hub platform, which provides complete, free and open access to all Sentinel products (ESA, Sentinels Scientific Data Hub, 2016). Sentinel data policy aims for maximum availability of data and indeed this is a free access platform with friendly interface. Data format after downloading is in SENTINEL-SAFE format, which has been designed as a common archiving and conveying data within ESA. It basically wraps a folder containing image data in binary data format and a metadata file in XML. Figure 13 shows Sentinel-2 product format. The folder contains a manifest.safe file which holds the information in XML, a preview image in JPEG2000 format, subfolders for measuring datasets (granules/tiles) in GML-JPEG2000 and datastrip level information, auxiliary data subfolders and HTML previews. Selected Sentinel-2A scene's details can be found in Appendix A4.

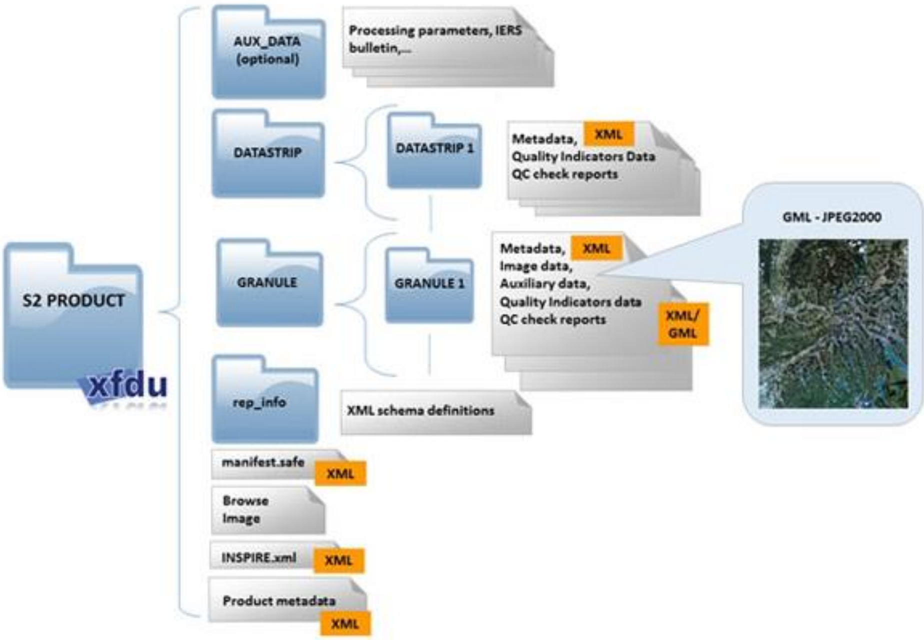

Figure 13: Sentinel-2 products folder format (ESA, European Space Agency , 2016).

USGS decided an open and free availability of its archive since 2008. Landsat data are available for immediate download, mainly in EarthExplorer platform (USGS, EarthExplorer, 2016). Landsat scenes products folders contain several files, including band and metadata files. Landsat 7 folders include Gap Mask files for each band. An overview of Landsat scenes folders format is shown in Figure 14. Basic details of Landsat 7 ETM+ scene selected can be browsed in Appendix A5.

	Individual Band Files (_B1.TIF)	Quality Band File (BQA.TIF)	Metadata File (MTL.txt)	Work Order file (WO.txt)	Ground Control Points File (GCP.txt)	Gap Mask Files (gap_mask)	Verify Image File (VER.jpg)	Verify Image Report (VER.txt)	README File (README.GTF)
Landsat 8 Operational Land Imager/ Thermal Infrared Sensor	11 files	1 file	1 file						
Landsat 7 Enhanced Thematic Mapper	9 files		1 file		1 file	1 directory (9 files)			1 file

Figure 14: Available file formats of Landsat 7 and 8 scenes (USGS, 2016).

3.5 Image preparations

Image preparation and treatments were done using exclusively ArcGIS software (version 10.3) (ESRI ArcGIS, 2014). We first created a geodatabase named Leman SMakri.gdb. We then proceeded to images preparations and additional layers creation that we next used in images treatments. The main steps of images preparation and additional layers creation were as follows:

1) Using the "Editor" tool, we created a polygon defining our study area, which we used next to extract the area of interest from each bands' raster files. For that we used the tool "Extract by Mask", in order to cut all bands' raster files of both Sentinel-2A and Landsat 7.

2) For each cut scenes' raster we preceded to radiometric enhancements of levels of grey, using and modifying the histograms in each layer's "Properties" window. This is an image modification by changing pixel brightness values in order to improve its visual impact.

3) We inserted two coordinates' point layers for the two different sampling sites GE3 and SHL2. For that we first created two excel files containing the coordinates of the two different points. In ArcGIS we used the "Excel to Table" conversion tool to convert the excel file to a table in ArcGIS. Using the path: File→ Add Data→ Add XY Data, defined the x, y attributes according to the coordinates in the tables we created. It is also important here to define the projection coordinate system using the Edit option, in our case WGS_1984. This created two point layers containing our coordinates which we exported by right clicking on the layer and by using the path: Data→ Export Data. We saved the file as shapefile.

4) We created two buffer layers around the two different sampling sites. These layers were used for pixels statistics, as we will in detail next. To create the two buffer layers we used the "Buffer" tool. For Sentinel-2A images we created a buffer of 10 m, taking into account used bands spatial resolution of 10 m, in order to measure at least 1 pixel's statistics. For Landsat 7 scene we used a 30 m buffer around the sampling sites points.

Further details including screen captures of main image preparation made, can be found in Appendix A6. In the next chapter we can see incorporated treatments, in order to transform and reveal measured spectral reflectance of our images.

3.6 Image transformations

Satellite image transformations involve manipulations of multiple bands data in order to highlight particular properties or features of interest within the study area, in a better and more effective way than the original input images. These transformations produce "new" images that have the potential to reveal better the electromagnetic radiation captured by the satellite sensors. The multi-spectral nature of our images allows their transformation to new images with a different set of components or bands. The idea here is to reduce the number of dimensions and keep only those useful and more pertinent to get some more evident information about the zones of interest. In our study we have incorporated image transformation involving color composites, band ratios and indexes such as NDVI, NDWI and Chl-a, which are all analyzed hereafter.

3.6.1 Color composites

This transformation refers to the creation of new multispectral images that consist of the three primary color bands (red, green and blue). Composites can be true, meaning that corresponding to red, green and blue bands of the satellite image were assigned respectively to the R, G, B colors for display, or they can also be false. False composites use an arbitrary display assignment. In this later case the display color of the target feature does not correspond to its true color; however some display schemes may be more suitable for detecting certain objects. Both composite have as aim to attribute to different features different colors and reveal more information.

As discussed in 2.2, with respects to water remote sensing and vegetation observations we are particularly interested in the visible and NIR bands, where we can observe the biggest reflectance differences between vegetation, water and other constituents. We discussed that clear water has a maximum reflectance in the blue-end part appearing dark-bluish. Turbid water has increased reflectance in the red-end region appearing more brownish. Chl-a in water would give low reflectance in the blue and red and higher in the green part.

According to our aimed features we used, for the Sentinel-2A scene, we took the corresponding bands 2 (blue), 3 (green), 4 (red) and 8 (NIR) (Figure 11). We performed the true color composition R: 4 G: 3 B: 2 and the false NIR color composition R: 4 G: 3 B: 8, using the tool “Composite Bands”, illustrations in Appendix A7.

For the Landsat image, we took the bands 1, 2, 3 and 4 which correspond, as seen in Table 2 in 2.4, to the blue, green, red and NIR region respectively. In ArcGIS, using again the tool “Composite Bands”, we realized the true color composition R: 3 G: 2 B: 1 and the false NIR color composition R: 3 G: 2 B: 4.

3.6.2 Band ratios

Another image manipulation technique is band ratioing. This refers to different bands’ division, according to the desired features to observe. For each pixel we divide the digital number value of one band to another band. This yields to a new set of pixel numbers for the resulting rasters from 0 to 255 for Landsat 7 images and 0 to 4095 for Sentinel-2 images, however the majority of resulting numbers are fractional. The aim of this manipulation is to reveal differences between spectral reflectance of the features we intent to observe.

In the current study and according to aimed features, for the Landsat 7 scene we realized the band ratios 4/1, 1/4, 2/3, 3/1, 3/2. In ArcGIS, this is done with the use of the tool “Raster Calculator”. As the most possible is to have fractional numbers results, we incorporated the option “Float” in the mathematical equation used (Appendix A8). For the Sentinal-2A scene, we realized the corresponding ratios 8/2, 2/8, 3/4, 4/2, 4/3.

In order to evaluate ratios’ resulting values we realized zonal statistics calculations to the new resulting band ratio rasters. In order to define the zone of calculations, we used the buffer layers (as explained in 3.5), which defined a zone of 10 m and 30 m around the sampling sites, for Sentinel-2A and Landsat 7 scenes respectively, according to their spatial resolution. In ArcGIS for implementing zonal statistics we used the tool “Zonal statistics as table”, which results to an excel table output (an example can be seen in Appendix A8).

3.6.3 Index calculations (NDVI, NDWI, Chl-a)

Image ratioing may involve more complex ratios, than just the division of two bands. These more complex ratios, called indexes, which may involve the sums or subtraction between bands, have been extensively developed for monitoring various observed features. In the current study we incorporated the NDVI index widely used for monitoring vegetation conditions and the NDWI for revealing differences between water and other features, both widely used in satellite remote sensing, as well as some Chl-a indexes, which are presented hereafter.

The NDVI index is mostly used to reveal vegetation coverage of soils. It is sensitive to chlorophyll contents in plants and given that plant leafs absorb the red and reflect the NIR; the index tries to reveal the high NIR reflectance areas. The index tales values from -1 to 1, with areas of low vegetation having negative values and the areas of high vegetation having high values. The NDVI index is given by the following equation:

$$\boldsymbol{NDVI_{SRS} = (NIR - Red / NIR + Red)} \quad \textbf{Eq.1}$$

For the Landsat 7 scene we used the bands 3 (red) and 4 (NIR). For the Sentinel-2A scene we used the corresponding bands 4 (red) and 8 (NIR). In ArcGIS, we used the specialized tool for the calculation of the NDVI index, which can be found from the path Windows→ Image Analysis and after having selected the relevant bands, clicking on the NDVI calculation tool. We used the "Scientific Output" option to obtain values from -1 to 1 (Appendix A9).

The NDWI index is used for estimation of the water content and for soil moisture estimation. As discussed in 2.2, water has a peak in reflectance in the green part of the visible spectrum and its opposition to the NIR has the potential to reveal water covered areas. The NDWI index is given by the equation:

$$\boldsymbol{NDWI_{SRS} = (Green - NIR / Green + NIR)} \quad \textbf{Eq. 2}$$

Values are similar to NDVI index, with low values 0 or -1, resulting to bright surfaces of no vegetation and high values +1, to water content. This index was mainly used to validate the accuracy and effectiveness of bands used for Landsat 7 and Sentinel-2A, for detecting water content.

Besides these widely used satellite remote sensing indexes, several studies have worked on developing specific Chl-a indexes by in situ radiometric measurements. Table 4 summarizes some key, in situ radiometric defined indexes, chosen to be presented in this study, along with the standard satellite indexes presented before. In table 4 we can find the details and the equations that define each index, as well as the corresponding bands for each satellite. Numbers in indexes equation represent the reflectance wavelengths defined by the in situ reflectance. References are given in the table.

Table 4: In situ radiometric defined indexes used to evaluate satellite reflectance accuracy of Sentinel-2 and Landsat.

Index details and equation	Landsat 7 bands	Sentinel-2A bands	References
$NDCI_{RIS}$= R(709)-R(665) / R(709)+R(665) ***Eq. 3*** *Normalized Difference Chlorophyll Index*	4-3 / 4+3	5-4 / 5+4	(Mishraa & Mishraa, 2012; Augusto-Silva , et al., 2014)
$NDVI_{RIS}$= R(748)-R(675) / R(748)+R(675) ***Eq. 4*** *Normalized Difference Vegetation Index*	4-3 / 4+3	6-4 / 6+4	(Rouse, et al., 1974)
$NDBI_{RIS}$= R(550)-R(675) / R(550)+R(675) ***Eq. 5*** *Normalized Difference Algal Bloom Index*	2-3 / 2+3	3-4 / 3+4	(Xue, et al., 2015)

Indexes calculations were made using the “Raster Calculator” tool in ArcGIS, creating new raster layers. An example of index raster calculation can be found in Appendix A10.

It is important to notice here the variability of bands in Sentinel-2A indexes bands. As discussed in 2.3, for Sentinel-2 there are four narrow bands for vegetation characterization (705nm, 740nm, 783 nm and 865 nm) (Figure 9). On the other hand, less variable, wider bands of Landsat 7, resulted to the use of same bands for most indexes. The $NDVI_{SRS}$, presented before, the $NDVI_{RIS}$ and the $NDCI_{RIS}$ used bands coincide. In addition, it is importance remark that this a qualitative approach study and there was not any atmospheric correction made. Hence, we cannot expect any precise results on the implementation of these indexes. Our results are presented in the next chapter.