



WEST UNIVERSITY OF TIMISOARA

FACULTY OF MATHEMATICS AND COMPUTER SCIENCE

STUDY PROGRAM:  
COMPUTER SCIENCE IN ENGLISH

## MASTER DISSERTATION

**COORDINATOR:**

Associate Prof. Marc Eduard  
FRÎNCU

**GRADUATE:**

Maria Minerva VONICA

Timișoara  
2021

WEST UNIVERSITY OF TIMIȘOARA

FACULTY OF MATHEMATICS AND COMPUTER SCIENCE

STUDY PROGRAM:  
COMPUTER SCIENCE IN ENGLISH

## MASTER DISSERTATION

Name

**COORDINATOR:**

Associate Prof. Marc Eduard  
FRÎNCU

**GRADUATE:**

Maria-Minerva VONICA

Timișoara  
2021

# Chapter 1

## Application

### 1.1 Dataset

In order to build the dataset, we made use of the freely available data from the Landsat Archive, specifically from collection 1, level 1. This consists of data products generated from Landsat 8 Operational Land Imager/Thermal Infrared Sensor, Landsat 7 Enhanced Thematic Mapper Plus, Landsat 4-5 Thematic Mapper, and Landsat 1-5 Multispectral Scanner instruments [C1L]. For the purpose of this paper, we will focus only on images collected from the Landsat 8 satellite.

#### 1.1.1 Landsat 8

For the purpose of this paper, we will use images collected by the Landsat 8 satellite (Figure 1.1), which was launched on an Atlas-V rocket from Vandenberg Air Force Base, California on February 11, 2013. Landsat 8 is the most recently launched Landsat satellite and carries the Operational Land Imager (OLI) and the Thermal Infrared Sensor (TIRS) instruments. Landsat 8 orbits the Earth in a sun-synchronous, near-polar orbit, at an altitude of 705 km, inclined at 98.2 degrees, and completes one Earth orbit every 99 minutes. The satellite has a 16-day repeat cycle with an equatorial crossing time: 10:00 a.m. +/- 15 minutes. It acquires about 740 scenes a day on the Worldwide Reference System-2 (WRS-2) path/row system, with a swath overlap (or sidelap) varying from 7% at the equator to a maximum of approximately 85% at extreme latitudes. A Landsat 8 scene size is 185 km x 180 km [LANa].

**Worldwide Reference System-2** We will be referring to each scene's location based on its path and row coordinates from the worldwide reference system-2, which is a global notation system for Landsat data. It enables a user to inquire about satellite imagery over any portion of the world by specifying a nominal scene center designated by path and row numbers. The WRS has proven valuable for the cataloguing, referencing, and day-to-day use of imagery transmitted

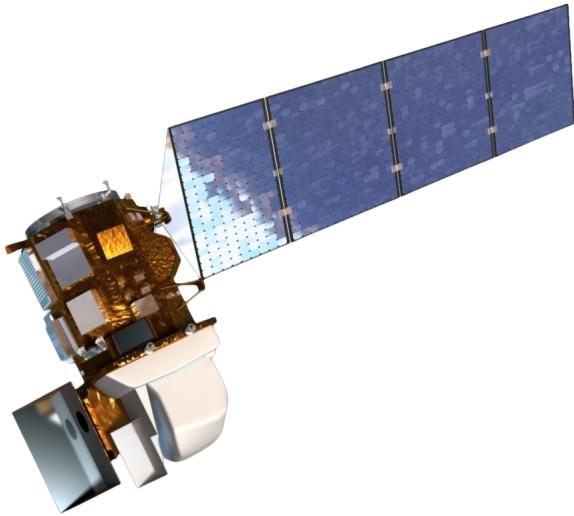


Figure 1.1: LandsatSatellite [LANb]

from the Landsat sensors [wrs]. Landsat’s trajectory projected on the world map can be seen in figure 1.2.

### Landsat scene naming convention

Each Landsat scene is named after a well-defined convention in order to easily check information such as the WRS path, row and the date of acquisition. Having access to this information without the need to download the scene itself or the metadata file which holds this information represents a valuable asset, since we can easily filter the data based on the naming convention itself. In the table 1.1.1 below we represent what each part of a Landsat scene of the form **LXS PPPRRR YYYYDDD GSIVV** means.

### Operational Land Imager

The Operational Land Imager (OLI) is a remote sensing instrument aboard Landsat 8, built by Ball Aerospace & Technologies. The sensor collects moderate resolution data that is used to monitor changing trends on the surface and evaluate how land usage changes over time. The images and data that OLI has helped collect have practical applications today in agriculture, mapping, and monitoring changes in snow, ice, and water [KK14]. The OLI operates in the visible (VIS) and short wave infrared (SWIR) spectral region, having a width of 185 km. It uses nine channels, which range from wavelengths of 443 nm to 2,200 nm. Of these nine channels, eight are multispectral and one is panchromatic. The eight multispectral channels have a 30-meter spatial resolution, and the panchromatic channel has a 15 meters one. The OLI generates

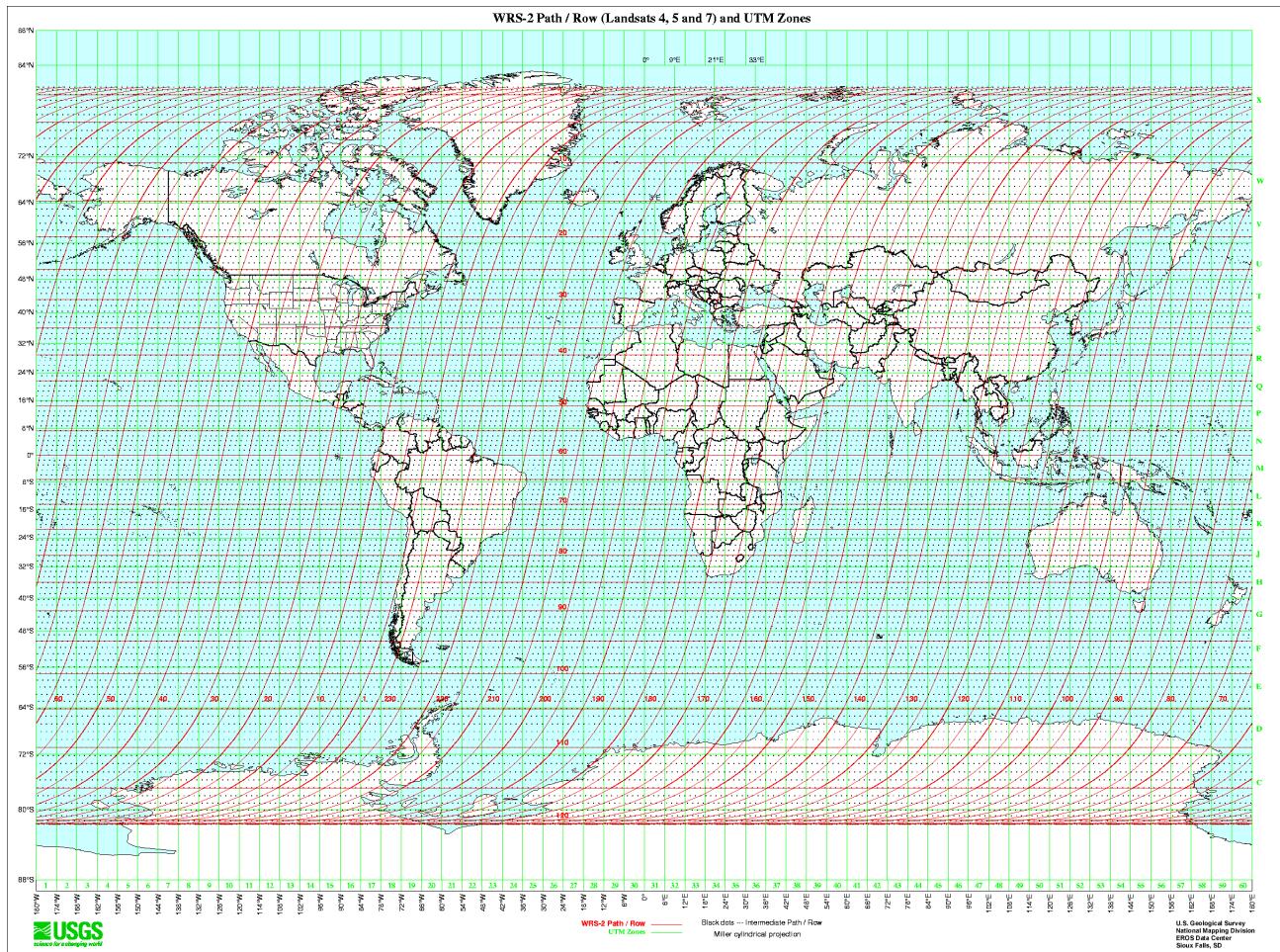


Figure 1.2: WRS-2 Path/Row for Landsat [wrs].

9 bands for Landsat as shown in 1.3.

### Thermal Infrared Sensor

The Thermal Infrared Sensor (TIRS) measures land surface temperature in two thermal bands with a new technology that uses Quantum Well Infrared Photodetectors to detect long wavelengths of light emitted by the Earth whose intensity depends on surface temperature. These wavelengths, called thermal infrared, are well beyond the range of human vision [lgn].

The thermal infrared sensor generates 2 bands for Landsat as shown in 1.4.

### OLI and TIRS bands

Landsat 8 acquires data from these sensors in 11 bands, as following in table 1.2.

<b>L</b>	Landsat
<b>X</b>	Sensor
<b>SS</b>	Satellite
<b>PPP</b>	WRS path
<b>RRR</b>	WRS row
<b>YYYY</b>	Acquisition year
<b>DDD</b>	Julian day of the acquisition year
<b>GSI</b>	Ground station identifier
<b>VV</b>	Archive version number

Table 1.1: Landsat 8 scene naming convention [sn].

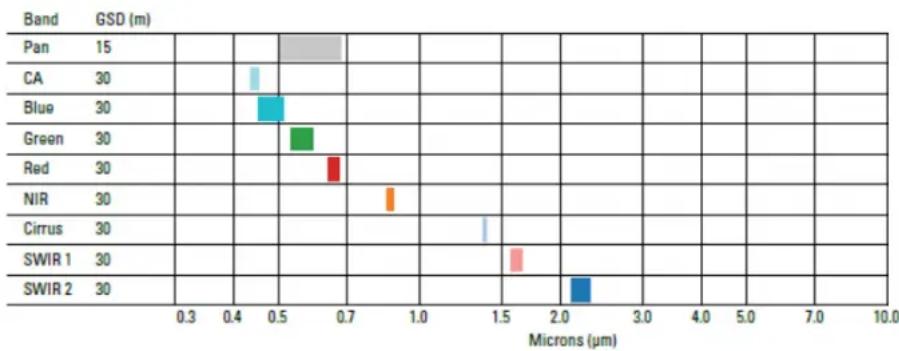


Figure 1.3: Landsat 8 OLI generated bands [l8o]

### 1.1.2 World Glacier Inventory

The World Glacier Inventory (WGI) proves to be a useful resource for building our dataset, since it contains information for over 130,000 glaciers. Inventory parameters include geographic location, area, length, orientation, elevation, and classification. The WGI is based primarily on aerial photographs and maps with most glaciers having one data entry only. Hence, the data set can be viewed as a snapshot of the glacier distribution in the second half of the 20th century. It is based on the original WGI (WGMS 1989) from the World Glacier Monitoring Service [WGI].

There are a number of ways to retrieve data from the inventory:

- download the entire database in a single ASCII text file (wgi\_feb2012.csv);
- search by parameter using the Search Inventory interface;

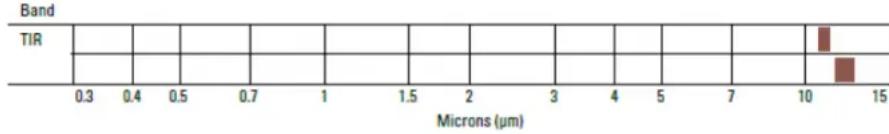


Figure 1.4: Landsat 8 TIRS generated bands [l8o]

- extract regions through the Extract Selected Regions interface.

The ASCII text file will be used with the purpose to define which are the glaciers to be included in the dataset to be built. An example of how this file looks like can be found in Figure 1.1.

1	wgi_glacier_id	political_unit	continent_code	drainage_code	free_position_code	local_glacier_code	glacier_name	lat	lon
2	SU5X1430909 <del>P</del> <del>SU</del>		5	X143		9	90 Zyryuzamjo	38.92	71.272
3	AT4J143OE00 <del>P</del> <del>AT</del>		4	J143	OE		6 ZWISELBACH W	47.112	11.038
4	AT4J143OE00 <del>P</del> <del>AT</del>		4	J143	OE		5 ZWISELBACH	47.11	11.052
5	CH4L01200008 CH		4	L012		0	8 ZWISCHBERGEN GL	46.108	8.041
6	CN5N23610001 CN		5	N236	I0		1 Zuxuehui	31.828	94.675
7	CH4J143040001 CH		4	J143		4	1 ZUORT VADRET DA	46.738	10.271
8	CN5O282B002 <del>P</del> <del>CN</del>		5	O282	B0		23 Zuogipu	29.212	96.893
9	CN5O282A047 <del>P</del> <del>CN</del>		5	O282	A0		476 Zuoguzasan	29.958	95.92
10	SU5X1430831 <del>P</del> <del>SU</del>		5	X143		8	310 ZULUMART	39.13	72.78
11	CN5N224E001 <del>P</del> <del>CN</del>		5	N224	E0		12 Zuma	29.839	96.456
12	SU5X1430948 <del>P</del> <del>SU</del>		5	X143		9	489 Zotkin	38.649	71.244
13	SU5X1430949 <del>P</del> <del>SU</del>		5	X143		9	490 Zotkin	38.649	71.244
14	SU5X1430932 <del>P</del> <del>SU</del>		5	X143		9	326 Zordi-Birauso	38.673	71.664
15	NZ6B868B000 <del>P</del> <del>NZ</del>		6	B868	B0		7 ZORA	-43.739	169.823
16	SU5T09106366 <del>P</del> <del>SU</del>		5	T091		6	366 ZOPKHITO	42.88	43.43
17	IT4L01104020 IT		4	L011		4	20 ZOCCA S	46.285	9.647
18	IT4L01104021 IT		4	L011		4	21 ZOCCA E	46.292	9.653
19	AQ7SSI00012 <del>P</del> <del>AO</del>		7	SSI0		0	125 Znosko Glacier	-62.1005	-58.4865
20	SU4X0300190 <del>P</del> <del>SU</del>		4	X030		1	903 ZNAMENITYY	80.53	61.02

Figure 1.5: WGI ASCII

The *parameters* which will be extracted for the dataset construction are the following:

- **wgi\_glacier\_id**: unique id representing one glacier (or part of it, if the coverage area is larger);
- **glacier\_name**: name of the glacier (if it has one);
- **lat**: latitude of the glacier;
- **lon**: longitude of the glacier.

### 1.1.3 Asset Acquisition

Since Landsat 8 acquires over 700 scenes per day, this means that there are over two million scenes available for download, either making use of already built user friendly tools or by simply querying for them directly.

## USGS Earth Explorer

One of the most popular services for satellite imagery downloading is USGS Earth Explorer. This is used for querying and ordering of satellite images, aerial photographs, and cartographic products through the U.S. Geological Survey. The tool is particularly useful when the main focus is to analyse a specific area rather than trying to acquire a large dataset of scenes. One can easily search for assets based on criteria such as world reference system path and row variables, latitude, longitude, cloud coverage, capture date and many others [USG].

However, downloading a large set of assets proves to be rather difficult by using this tool alone, since the parameters for each scene need to be manually set. On top of this, the query results have to be picked by hand and then passed for downloading through another application which handles their bulk download. This makes the process of building the dataset rather slow, frustrating and error prone. Such an example can be viewed in Figure 1.2, for the Belvedere glacier (45.9830, 6.888).

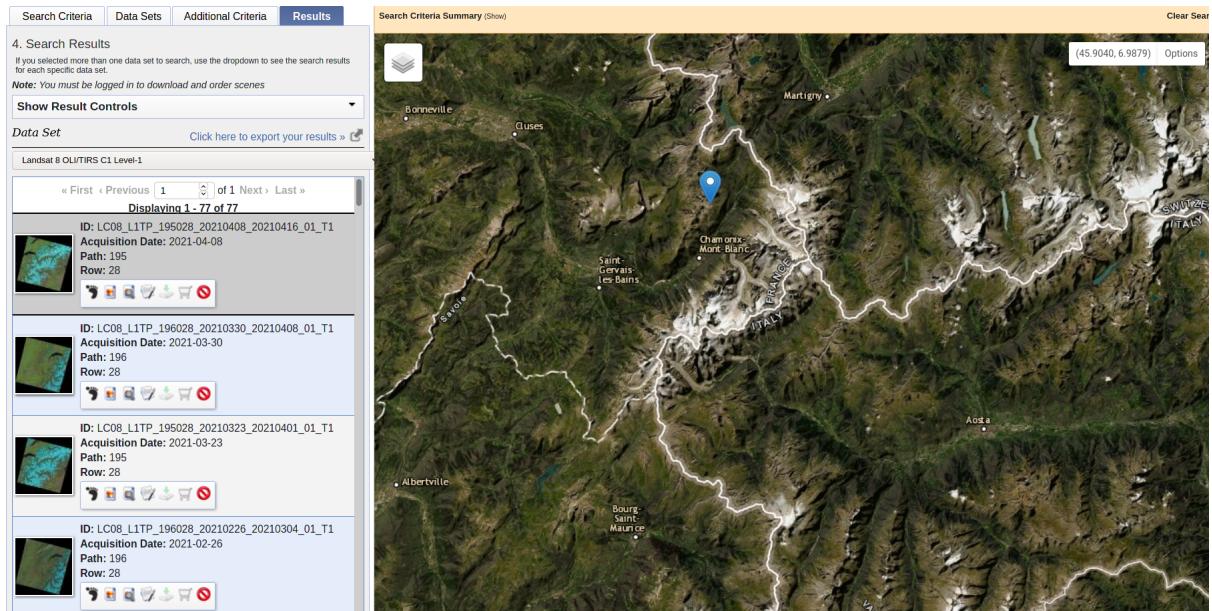


Figure 1.6: EarthExplorer

## SpatioTemporal Asset Catalog API

In order to fix the problem of excessive manual labour which appeared by using the USGS Earth Explorer, we rather implemented an endpoint of the SpatioTemporal Asset Catalog API, specifically, the following: [http://nsidc.org/data/glacier\\_inventory/index.html](http://nsidc.org/data/glacier_inventory/index.html) [STA]. The main idea of searching by using parameters still remains, but instead of manual inputting data for the search data, we rely on using the above-mentioned World Glacier Inventory ASCII

text file, since it already has all the required information for each glacier.

By using this method we can pick which glaciers we want to download based on their coordinates and calculate a bounding box representing the area we want to search, required for the STAC API query. Since there might be clouds which could obfuscate the area of interest in the image, we also add a maximum allowed cloud coverage along the bounding box.

The STAC API query also requires a name for the collection of assets we want our queries to be made on, which for us is landsat-8-l1 (Landsat 8 Collection 1, Level 1). Using these three parameters we can now easily acquire a large number of assets with minimal manual labour, as compared to the more user friendly tool provided by USGS.

The downloaded assets will be stored at a user specified disk location and they will be structured as shown in the Figure 1.3.

### Landsat 8 Collection 1, Tier 1

To support analysis of the Landsat long-term data record, the Landsat data archive was reorganized into a formal tiered data collection structure, which ensures that all Landsat Level 1 products provide a consistent archive of data quality to support time-series analysis and data “stacking”, while controlling continuous improvement of the archive, and access to all data as they are acquired. The implementation of collections ensures consistent and known radiometric and geometric quality through time and across instruments while improving control in the calibration and processing parameters [lc1]. By using data from this collection, we ensure that our images are fit for accurate pixel-to-pixel processing.

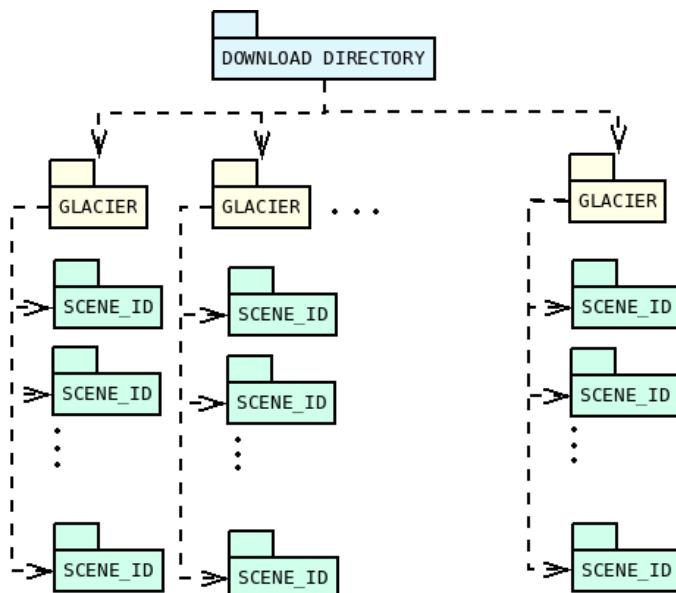


Figure 1.7: Download Directory

## 1.2 Dataset entities

TODO better make this intro... We will further describe which are the bands necessary for the image generation and what are their uses. On top of this, we will further explain what is the normalized snow difference index and what is the optical flow used for.

### 1.2.1 Bands

Each Landsat 8 band is represented by a 16 bit grayscale image with a resolution between 7000 and 10000 pixels, each pixel representing 30 meters. We can conclude, therefore, that one scene covers around 200 and 300 km of Earth. Only the green and SWIR1 bands will be used for the purpose of this thesis and below we will discuss the specifications of each.

#### Band 3 - Green Band

The green band, alongside with the red and blue ones, fall in the visible spectrum and it is usually used for mapping peak vegetation. Figure 1.8 is an example of the green band for the Belvedere glacier, specifically taken from scene LC81950282015098LGN01.

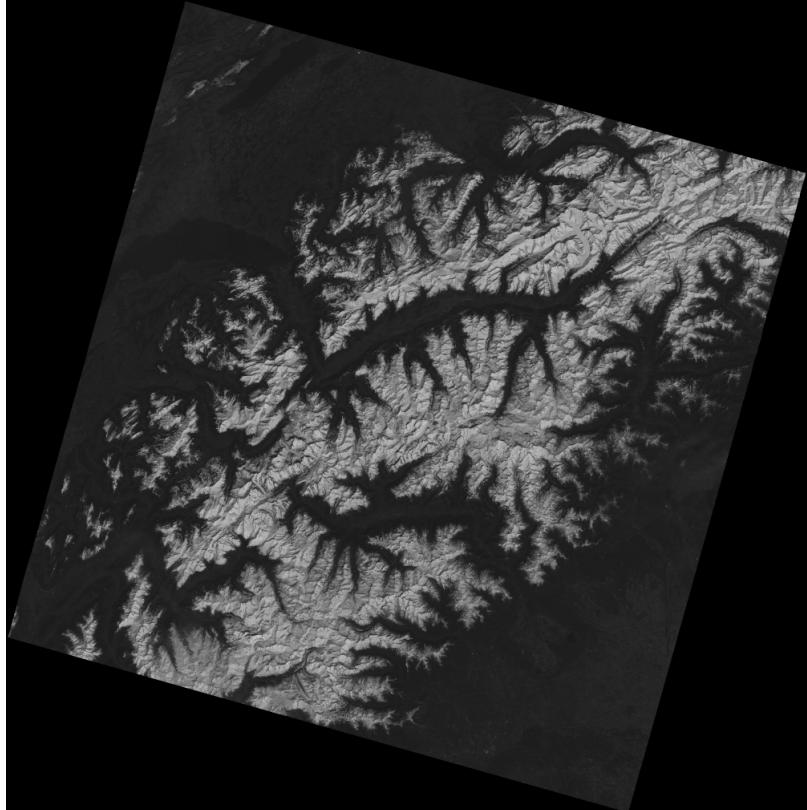


Figure 1.8: Green band of scene LC81950282015098LGN01 from the Belvedere glacier, Italy.

## Band 6 - SWIR1 Band

The shortwave infrared 1 band is particularly useful for enhancing geology objects such as rocks and soils which look similar in other bands [bd]. Alongside this, it also discriminates moisture content of soil and vegetation and penetrates thin clouds [ban]. Figure 1.9 is illustrated below as an example of a SWIR1 band taken from the same scene as the green one above.

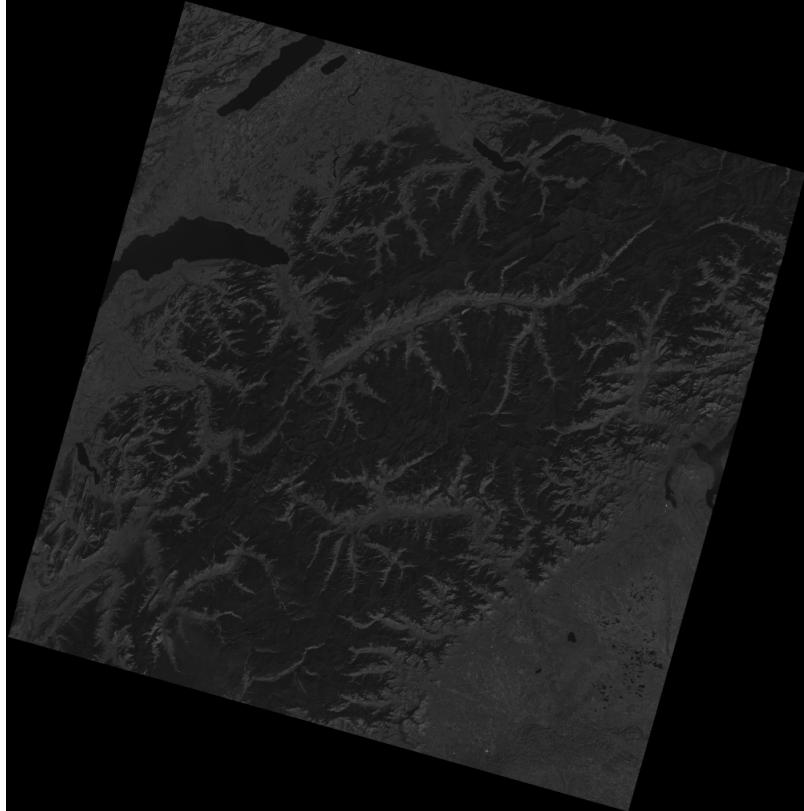


Figure 1.9: SWIR1 band of scene LC81950282015098LGN01 from the Belvedere glacier, Italy.

## Normalized Snow Difference Index

The normalized snow difference index (NDSI) is an index which relates to the presence of snow/ice in a pixel. Snow and ice usually have a very high reflectance in the visible spectrum and very low one in the shortwave infrared one, which is useful for mapping out most types of clouds from the scene [nds]. We can therefore use the formula from equation 1.1 in order to highlight the snow and ice pixels from a Landsat 8 image.

$$NDSI = \frac{green - SWIR1}{green + SWIR1} \quad (1.1)$$

By combining the snow reflectance behaviour of snow and ice for each of these bands, we can create a normalized snow difference index image which has values in the range [-1, 1]. We can

then apply a threshold on the pixels which states that if the NDSI value for a pixel is larger than 0.0, then that pixel represents snow/ice covered land; similarly, if its value is smaller than 0.0, that pixel represents snow/ice free land [nds], [Hal15], as represented in equation 1.2. Of course, this represents the general value for the threshold, but with the increase of its value, we can more accurately differentiate between snow and ice. With larger threshold values we can state that a pixel represents ice covered land rather than just snow fall land [Hal15]. We have tried different values for the threshold and came to the conclusion that a value of 0.3 is best suited for our needs.

$$\begin{cases} \text{snow/ice land} & \text{if } NDSI \geq 0.0 \\ \text{snow-free land,} & NDSI < 0.0 \end{cases} \quad (1.2)$$

An example of a generated NDSI for the Belvedere glacier, scene LC81950282015098LGN01, can be viewed in Figure 1.10.



Figure 1.10: NDSI of scene LC81950282015098LGN01 from the Belvedere glacier, Italy.

### 1.3 Alignment

Landsat's trajectory orbiting Earth is not fully precise, therefore not all images will be pixel-to-pixel aligned, which is a problem for image processing. Since we want to track each pixel's movement, we must be sure that its coordinates in the image do not change between any two

given scenes. Figures 1.3 highlights the discrepancy from scene TODO to scene TODO, as an example. Since the bands we work with have a spacial resolution of 30 meters, which means

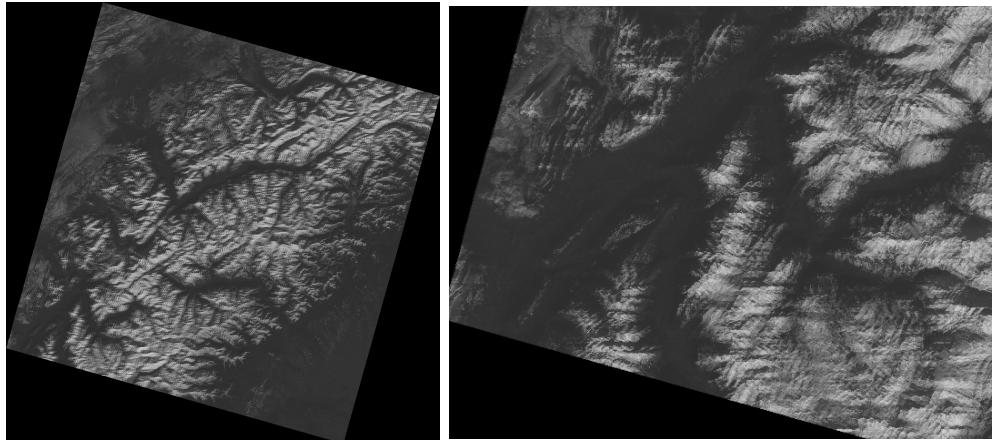


Figure 1.11: Overlapped unaligned green band of scene LC81950282013316 and reference LC81950282013364 of the Belvedere glacier (195/028 WRS-2).

that each pixel in the image represents 30 meters. Even with a misalignment of just 50 pixels we would end up with a 1.5 km difference between two scenes. Tracking pixels through the image without aligning them first would mean that we could never be sure that what we are looking at is indeed the same location. We have used different approaches on alignment which will be described in 2.3.1. Figure 1.12 highlights the alignment corrected scene from 1.3.

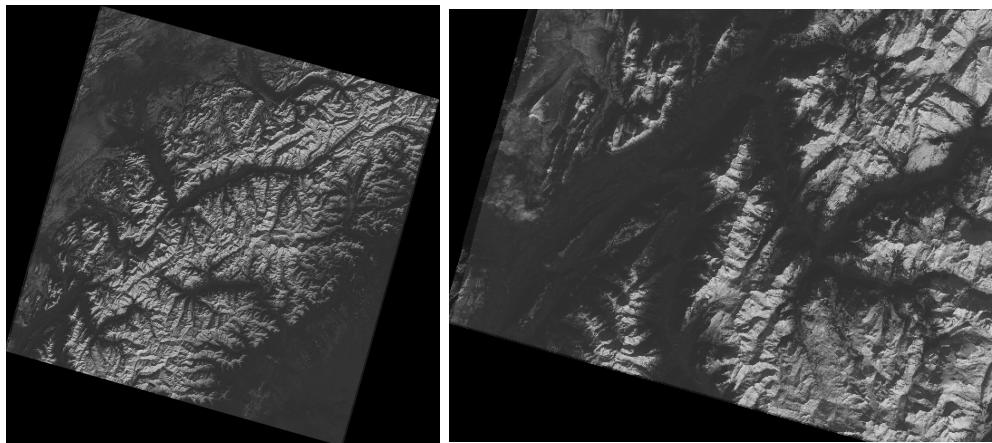


Figure 1.12: Overlapped aligned green band of scene LC81950282013316 and reference LC81950282013364 of the Belvedere glacier (195/028 WRS-2).

**Optical Flow Vector**

**Optical Flow Generated Image**

## 1.4 Graphical User Interface

<b>Bands</b>	<b>Wavelength (microme- ters)</b>	<b>Resolution (meters)</b>	<b>Usual mapping uses</b>
Band 1 (Coastal aerosol)	0.43-0.45	30	Coastal and aerosol studies
Band 2 (Blue)	0.45-0.51	30	Bathymetric mapping, distinguishing soil from vegetation and deciduous from coniferous vegetation
Band 3 (Green)	0.53-0.59	30	Emphasizes peak vegetation, which is useful for assessing plant vigor
Band 4 (Red)	0.64-0.67	30	Discriminates vegetation slopes
Band 5 (Near Infrared)	0.85-0.88	30	Emphasizes biomass content and shorelines
Band 6 (SWIR1)	1.57-1.65	30	Discriminates moisture content of soil and vegetation; penetrates thin clouds
Band 7 (SWIR2)	2.11-2.29	30	Improved moisture content of soil and vegetation; penetrates thin clouds
Band 8 (Panchromatic)	0.50-0.68	15	Sharper image definition
Band 9 (Cirrus)	1.36-1.38	30	Improved detection of cirrus cloud contamination
Band 10 (Thermal Infrared) 1	10.6-11.19	100	Thermal mapping and estimated soil moisture
Band 11 (Thermal Infrared) 2	11.50-12.51	100	Improved thermal mapping and estimated soil moisture

Table 1.2: Landsat 8-9 Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) bands [ban].

<b>Wavelength</b>	0.53- 0.59 micrometers
<b>Spacial resolution</b>	30 meters
<b>Resolution</b>	between 7000x7000 pixels and 10000x10000 pixels
<b>Depth</b>	16-bit
<b>Format</b>	grayscale

Table 1.3: Landsat 8 green band specifications.

<b>Wavelength</b>	1.57 - 1.65 micrometers
<b>Spacial resolution</b>	30 meters
<b>Resolution</b>	between 7000x7000 pixels and 10000x10000 pixels
<b>Depth</b>	16-bit
<b>Format</b>	grayscale

Table 1.4: Landsat 8 SWIR1 band specifications.

# Chapter 2

# Design and Implementation

The design of the application will be discussed while focusing on two different aspects:

- **search and download**
- **processing**

The first section is optional and can be run independently from the processing, since the user can have an already created database of images. However, we have focused on simplifying the process of satellite imagery downloading as described in 1.1.3 through using the sat-search library as a helper for searching and downloading assets. More information on querying can be found in 2.1. Given an existing set of data, the next step is to pass the root folder which contains the glacier directories to the processing unit, also designed as a plug-in mechanism which will trigger the graphical user interface to pop up and allow for the processing options to appear. More information on this section can be found at 2.3. From there, one can start processing by simply making use of the predefined buttons described in 1.4.

## 2.1 Search and Download

## 2.2 Normalized Snow Difference Index

The NDSI is generated by using the formula 1.1 described in section 1.2.1 on the green and shortwave infrared bands from a specific scene. Each band pixel is converted to **float32** in order to increase the **precision** of calculation; as each band is read as a n-dimensional array we can use numpy's optimised processing for generating the NDSI index image, since it converts the data and runs on native code rather than going through Python's interpreter. We then **filter** the generated image such that everything which is snow-free land will be 0 (**black**), but this is

only applied for better visual analysis. In the background, we work with the full range of the image for better precision. Figure 2.1 holds the technical diagram for the NDSI.



Figure 2.1: Technical diagram for the NDSI class.

# Contents

<b>1 Application</b>	<b>1</b>
1.1 Dataset . . . . .	1
1.1.1 Landsat 8 . . . . .	1
1.1.2 World Glacier Inventory . . . . .	4
1.1.3 Asset Acquisition . . . . .	5
1.2 Dataset entities . . . . .	8
1.2.1 Bands . . . . .	8
1.3 Alignment . . . . .	10
1.4 Graphical User Interface . . . . .	12
<b>2 Design and Implementation</b>	<b>15</b>
2.1 Search and Download . . . . .	15
2.2 Normalized Snow Difference Index . . . . .	15
<b>List Of Figures</b>	<b>17</b>
<b>List Of Tables</b>	<b>18</b>
2.3 Processing . . . . .	21
2.3.1 Alignment . . . . .	21
<b>3 Glossary</b>	<b>22</b>
3.1 Acronyms . . . . .	22

# List of Figures

1.1	LandsatSatellite [LANb]	2
1.2	WRS-2 Path/Row for Landsat [wrs]	3
1.3	Landsat 8 OLI generated bands [l8o]	4
1.4	Landsat 8 TIRS generated bands [l8o]	5
1.5	WGI ASCII	5
1.6	EarthExplorer	6
1.7	Download Directory	7
1.8	Green band of scene LC81950282015098LGN01 from the Belvedere glacier, Italy.	8
1.9	SWIR1 band of scene LC81950282015098LGN01 from the Belvedere glacier, Italy.	9
1.10	NDSI of scene LC81950282015098LGN01 from the Belvedere glacier, Italy.	10
1.11	Overlapped unaligned green band of scene LC81950282013316 and reference LC81950282013364 of the Belvedere glacier (195/028 WRS-2)	11
1.12	Overlapped aligned green band of scene LC81950282013316 and reference LC81950282013364 of the Belvedere glacier (195/028 WRS-2)	11
2.1	Technical diagram for the NDSI class.	16

# List of Tables

1.1	Landsat 8 scene naming convention [sn]. . . . .	4
1.2	Landsat 8-9 Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) bands [ban]. . . . .	13
1.3	Landsat 8 green band specifications. . . . .	14
1.4	Landsat 8 SWIR1 band specifications. . . . .	14
3.1	Acronyms table . . . . .	22

# Bibliography

- [ban] bands. <https://www.usgs.gov/faqs/what-are-best-landsat-spectral-bands-use-my-research>. Accessed: 2021-06-21.
- [bd] bd. <https://landsat.gsfc.nasa.gov/landsat-8/landsat-8-bands/>. Accessed: 2021-06-22.
- [C1L] Landsat collection 1. [https://www.usgs.gov/core-science-systems/nli/landsat/landsat-collection-1?qt-science\\_support\\_page\\_related\\_con=1#qt-science\\_support\\_page\\_related\\_con](https://www.usgs.gov/core-science-systems/nli/landsat/landsat-collection-1?qt-science_support_page_related_con=1#qt-science_support_page_related_con). Accessed: 2021-06-20.
- [Hal15] Dr. Dorothy K Hall. Viirs snow cover algorithm theoretical basis document (atbd). 2015.
- [KK14] Edward J Knight and Geir Kvaran. Landsat-8 operational land imager design, characterization and performance. In *Proceedings of the ECAI 94 Workshop on Applications of Evolutionary Algorithms*. Landsat-8 Sensor Characterization and Calibration, 2014.
- [l8o] l8otb. <https://gisgeography.com/landsat-8-bands-combinations/>. Accessed: 2021-06-22.
- [LANa] Landsat. [https://www.usgs.gov/core-science-systems/nli/landsat/landsat-8?qt-science\\_support\\_page\\_related\\_con=0#](https://www.usgs.gov/core-science-systems/nli/landsat/landsat-8?qt-science_support_page_related_con=0#). Accessed: 2021-06-21.
- [LANb] Landsatpic. <https://landsat.gsfc.nasa.gov/sites/landsat/files/2013/01>. Accessed: 2021-06-21.
- [lc1] lc111. [https://prd-wret.s3.us-west-2.amazonaws.com/assets/palladium/production/atoms/files/LSDS-1656\\_%20Landsat\\_Collection1\\_L1\\_Product\\_Definition-v2.pdf](https://prd-wret.s3.us-west-2.amazonaws.com/assets/palladium/production/atoms/files/LSDS-1656_%20Landsat_Collection1_L1_Product_Definition-v2.pdf). Accessed: 2021-06-22.
- [lgm] lgng. <https://landsat.gsfc.nasa.gov/>. Accessed: 2021-06-21.

- [nds] ndsi. <https://nsidc.org/support/faq/what-ndsi-snow-cover-and-how-does-it-compare-fsc>. Accessed: 2021-06-22.
- [sn] sn. <https://gisgeography.com/landsat-file-naming-convention/>. Accessed: 2021-06-22.
- [STA] Stac. <https://stacspec.org/STAC-api.html>. Accessed: 2021-06-20.
- [USG] Usgs. <https://earthexplorer.usgs.gov/>. Accessed: 2021-06-21.
- [WGI] World glacier inventory. [http://nsidc.org/data/glacier\\_inventory/index.html](http://nsidc.org/data/glacier_inventory/index.html). Accessed: 2021-06-20.
- [wrs] wrs. <https://landsat.gsfc.nasa.gov/about/worldwide-reference-system/>. Accessed: 2021-06-22.

## 2.3 Processing

### 2.3.1 Alignment

# Chapter 3

## Glossary

### 3.1 Acronyms

WGI	World Glacier Inventory
NDSI	Normalized Snow Difference Index

Table 3.1: Acronyms table