Elena Chambers

FNRM 3462

04.27.21

**Changes in Icelandic glacial extent measured from satellite images in Jupyter notebooks**

**Abstract**

Glaciers are a valuable source of information in climate change status and Icelandic glaciers have been retreating steadily over the recent past. Using a combination of segmentation and k-means clustering classification in a Python Jupyter notebook, a rough estimate of glacial extent was calculated and mapped in ArcGIS Pro. Landsat 8 data from summer of 2016 and 2020 were used to create composite red-green-blue (RGB) images that were segmented into 50x50 pixel grids and a 5-cluster k-means algorithm was executed to produce rudimentary classifications. It was determined that the Hifsjokull glacier has decreased in area by 34.5 km2, which a reasonable estimate based on previous studies of the region.

**Introduction**

As human contributions to global climate change becomes more prominent even in our daily lives, increased greenhouse gas emissions have resulted in major changes in the glaciated regions on the planet. Even minute chemical changes in the atmosphere can result in regional changes in both temperature and precipitation and are key factors in determining the mass balance in glaciers (Jóhannesson, 1997). While glacial advance and retreat cycles occur naturally, the present retreat rate of glaciers in Iceland is faster than can be explained by natural processes alone; since 1995, Icelandic glaciers have lost roughly 7% of their total volume (Hannesdóttir et. al., 2019). All of Iceland’s glaciers contain ~3,600 km3 of water, and once completely melted, would result in a global sea level rise of 1 cm (Björnsson et. al, 2008). The premise of this project is to classify the glacial extent of the Hifsjokull glacier in central Iceland and to calculate glacial extent area changes from 2016 to 2020.

A picture containing text, blue

Description automatically generated*Study Area*

The Hifsjokull glacier (labeled with a red “A” pop-up in the image to the right) is the third largest glacier in Iceland and is located atop a dormant volcano. While geothermal heat flux derived from subglacial volcanos can result in non-negligible contributions to glacial retreat, geothermal heat flux beneath Hifsjokull glacier has little implications with faster glacial retreat rates within the past 20 years (Jóhannesson, et. al, 2020). As such, the Hifsjokull icecap is an ideal glacier for the purposes of this project.

maps.google.com

**Methods**

*Data Collection*

Landsat 8 scene tiles were downloaded from the USGS Earth Explorer portal. Search criteria were restricted between June and August months to minimize fresh snow cover over the glacier and covered the full extent of Hifsjokull glacier in central Iceland.

The first image is a tile from the summer of 2016 (collected 8/24/2016): WRS path 219, row 15, 23.44% total cloud cover, 22.10% cloud cover over land features, and in the UTM WSG84 Zone 27 projection. The second image is a Landsat 8 image from the summer of 2020 (collected 7/2/2020), WRS path 219, row 15, 8.50% total cloud cover, 5.95% cloud cover over land features. These tiles include the Myrdalsjokull, Langjokull, and Eyjafjallajokull glaciers in the southwest portion of Iceland, tile image extent shown in *Fig 1.* The full Level-2 bundle was downloaded for both scenes.

A map of the world

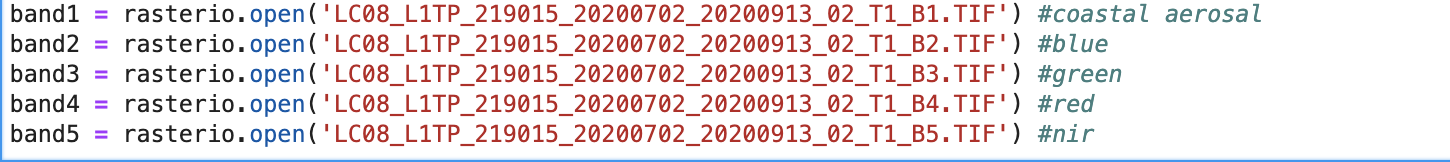
Description automatically generated with medium confidence

**Figure 1** shows the geographic location of the 2016 and 2020 Landsat image tiles. Myrdalsjokull, Langjokull, Eyjafjallajokull, and Hifsjokull are fully enclosed within the tile, while Vatnajokull glacier in the northeast corner is only partially included. The tiles are 185 km x 180 km. Image screen captured via EarthExplorer.

*Converting the downloaded dataset folder to workable files*

Landsat 8 imagery downloaded from EarthExplorer is automatically stored in .tar folders. Within the .tar folder, each spectral band is stored as an individual .tif file and must be extracted and imported into the Jupyter Notebook library.

In Python, the Rasterio package was utilized to designate each file as variables, and each band can be plotted individually, for example:

**

A screenshot of a map

Description automatically generated with low confidence

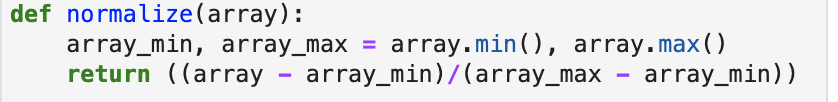
**Figure 2:** shows the red, green, and blue channels of the 2020 Image. Areas with the highest intensity/values are highly saturated. Glacial/snowy areas have high very high surface reflectance in these bands.

In order to perform multispectral analysis, individual bands must be concatenated into larger (and multidimensional) arrays. The caveat, the size of the array produced by combining multiple bands exceeds the computational limits of the Jupyter kernel when trying to produce an image. To overcome this issue, a polygon GeoJSON feature was constructed around the location of interest (Hifsjokull) using latitude/longitude coordinates. The polygon was then used to mask and clip all array values outside of the area of interest and drastically reduced the computational load.

*Preprocessing*

Georeferencing, radiometric, and atmospheric corrections were computed prior to data download via USGS software, and further calibrations would have most likely jeopardized the quality of the data. The clipped images from both 2016 and 2020 had no cloud cover, so cloud cover masking was not necessary.

In order to use plotting packages, each spectral band was normalized such that all values were within 0.0-1.0 scale for image display.



Each clipped Landsat scene was segmented into a 50x50 pixel grid in order to prevent the classifier from overfitting the data using the Scikit Image package.

*Classifier*

Considering the focus of the project was to classify glacial extent, the simple and effective K-means unsupervised classification technique was utilized. It was not necessary to perfectly classify waterbodies, vegetation, or anything other than snow and ice which have very strong spectral characteristics.

The K-means algorithm is included within the Scikit-Learn package and works by separating data points into a defined number of groups “k”. In order to execute this algorithm, the number of clusters must be first be defined. To identify the ideal number of clusters, the k-means algorithm was performed using 3 clusters, 5 cluster, and 10 clusters on the RBG composite images. The resulting classifications were compared, and the algorithm performed the best when using 5 clusters (fig 4).

*Converting clusters to raster layers with geospatial extent*

To compute area glacier area and outline the glacial extent the k-means cluster plot was converted into a raster layer that contained a cluster number label for each segment of the image. Cluster groups were classified as glacier, snow, and everything else. Cluster group classifications that represented glacial extent were merged to create a polygon feature.

The polygon features were exported from the Jupyter notebook and uploaded into ArcGIS Pro to quickly compute polygon area and plot glacial extent (fig 5).

*Accuracy Assessment*

To estimate the accuracy of the classifier, the results of the K-means clusters were visually inspected and compared to the RGB scene images (fig. 3). Quantitively describing the accuracy of the classifier via confusion matrix or extraordinarily difficult, as building a training dataset in Python was beyond my ability. However, in addition to visual inspection, the glacial extent mapped in the project was compared to previous studies.

**Results & Discussion**

Graphical user interface, application, website

Description automatically generated

**Figure 3:** shows the red, green, and blue (RGB) composite of the 2016 (left) and 2020 (right) Landsat images of Hifsjokull. Iceland inset on the upper left of each composite image includes a mini-RGB scene for reference. \*RGB images were produced in Jupyter and imported into ArcGIS Pro to create the figure.

Due to the nature of this glacier being situated over a dormant volcano, there are elevation changes from the center of Hifsjokull to the glacial margins. Because of this, there is a degree of distortion when viewing this flattened image. Slope area presented in this project might be less than true slope area. Additionally, the 2016 RGB composite image appears to have a greater snow cover than that of 2020, despite having been captured in Late August. The additional snow in 2016 may have resulted in more opportunities for the classifier to misclassify snow as glacier than for the 2020 image.

2016: K-means classifier

Graphical user interface, application

Description automatically generated

2020: K-means classifier

Graphical user interface, application

Description automatically generated

**Figure 4:** the 2016 and 2020 scenes were classified using the k-means algorithm in Scikit-Learn. In both the 2016 and 2020 images, glacial extent was underrepresented using only 3 clusters, are well represented using 5 clusters, and overclassified using 10 clusters.

While the k-means cluster is simple to program and produces reasonable classification, the algorithm is not always best suited for classifying areas with similar spectral behaviors (ie, snow and glacier). The main benefit of a completely unsupervised classifier was that a training set did not need to be prepared. When working completely from a Jupyter notebook (up until detailed figure production), creating a training set is extremely challenging and requires more coding experience that I presently have. Conversely, the primary disadvantage to the k-means classifier is that the algorithm with cluster pixels in unknown classes that are not easily constrained and are highly context specific. And, defining threshold parameters to separate snow versus glacial pixels is not possible.

Map

Description automatically generated

**Figure 5:** shows the outline of the glacier polygons created from the k-means clusters and subsequent merging. The blue extent refers to the 2016 glacial margin, while the pink represents the 2020 glacial margin. Note, in the 2016 extent, there is a speckling of extraneous polygons that were misclassified as glacier, when they should have been associated with snow cover classes. The areas presented in Table 1 are of the large central polygons.

In fig. 5, the 2020 glacial extent has been plotted over the 2016 glacial extent. The most prominent glacial retreats from the results appear to be in the north eastern region of Hifsjokull. However, that particular decrease is not well represented in the RGB composite images in figure 3.

|  |  |
| --- | --- |
| **Year** | **Glacier Polygon Area** |
| 2016 | 815.9 km2 |
| 2020 | 781.4 km2 |
| Change in area | Decrease of 34.5 km2 |

**Table 1:** presents the area of the classified glaciers. There isa decrease in glacial area by 34.5 km2 from the year 2016 to 2020. Polygon areas were found in the Attribute Table in ArcGIS Pro.

In 2002, Hofsjökull’s area was measured at 890 km2, suggesting that the measurement produced in the project using the k-means classifier was not entirely unreasonable at 815.9 km2 in 2016 (Sigurðsson, et. al, 2002) However, it is likely that the area values calculated here are underestimates of the true area of the glacier.

**Conclusion**

The project has been optimistically successful. While the k-means clustering classification algorithm is not the most sophisticated and customizable classifier, it certainly produces a reasonable first pass. Future work could include building a training dataset with ground truth points to calculate error and confusion matrices that will better support the final classification. Additionally, incorporating elevation data could address area underestimates along the margins of the glacier where there are steeper elevation changes,

**References**

Bjornson, Helgi & Pálsson, Finnur. (2008). *Icelandic glaciers.* Jökull. 58.

Hrafnhildur Hannesdóttir, Finnur Pálsson, Þorsteinn Þorsteinsson, Benedikt G. Ófeigsson, Snævarr

Guðmundsson, Kieran BaxterIcelandic. Meteorological Office, The Institute of Earth Sciences,

University of Iceland and the Southeast Iceland Nature Centre (2020). *Overview of Icelandic*

Jóhannesson T, Pálmason B, Hjartarson Á, Jarosch AH, Magnússon E, Belart JMC, Gudmundsson MT

(2020). Non-surface mass balance of glaciers in Iceland. Journal of Glaciology 66(258), 685–697.

https://doi.org/ 10.1017/jog.2020.37

Jóhannesson, T. (1997). The response of two Icelandic glaciers to climatic warming computed with a

degree-day glacier mass-balance model coupled to a dynamic glacier model. Journal of

Glaciology, 43(144), 321-327. doi:10.3189/S0022143000003270

Sigurðsson, O., T. Jóhannesson, Th. Thorsteinsson & G. Larsen (2002). Ice core drilling on the Hofsjökull

ice cap for measuring glacier mass balance. Nordic Hydrological Programme Report, 47.

Proceedings of the XXII Nordic Hydrological Conference, Røros, Norway, August 4-7 2002.