Indian Institute Of Technology Bombay

GNR618 : Remote Sensing and GIS Applications to Cryosphere

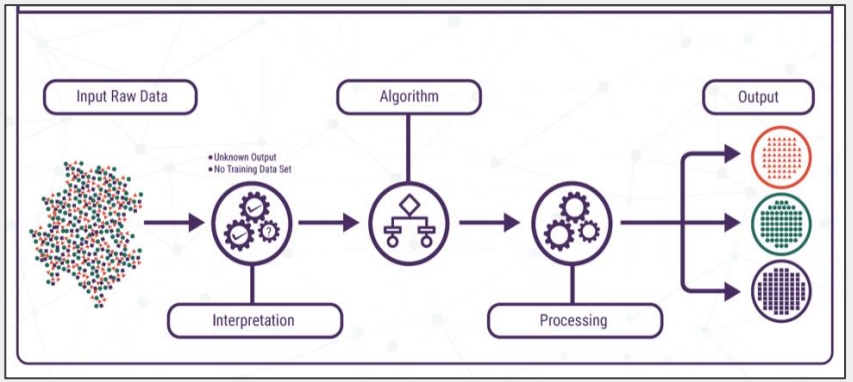
Glacier features classification using fully polarimetric SAR (POLSAR) Data Science and Machine Learning

Group 35

Gaurav Anand, 170070034, Department Of Electrical Engineering
Mayur Ware, 19D070070, Department Of Electrical Engineering
Aayush Shrivastava, 19D070002, Department Of Electrical Engineering
Tanmay Hiremath, 19D070061, Department Of Electrical Engineering
Siddhant Shingade, 19D070057, Department Of Electrical Engineering

Index

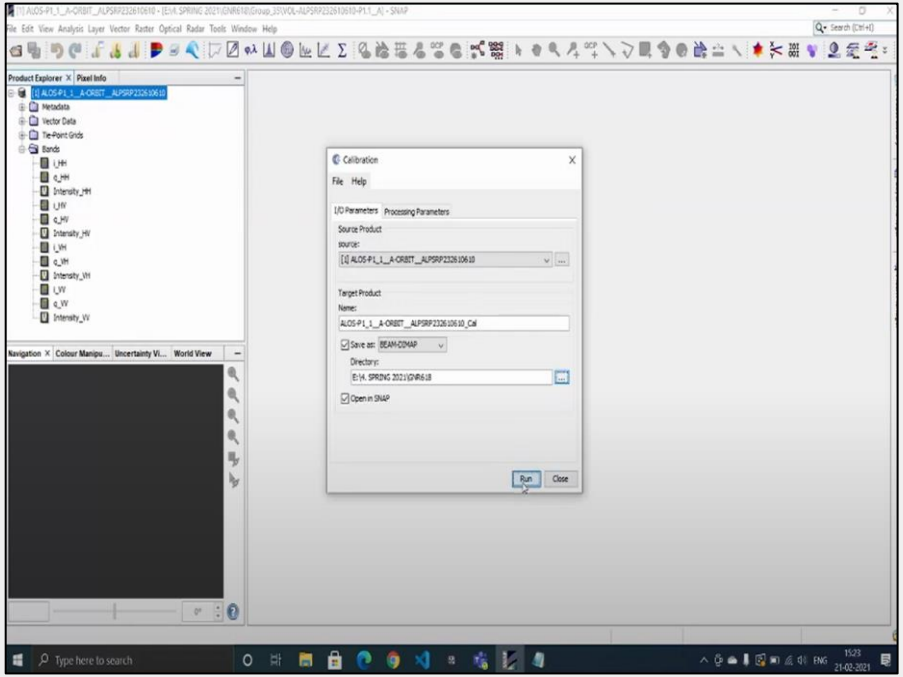
Introduction



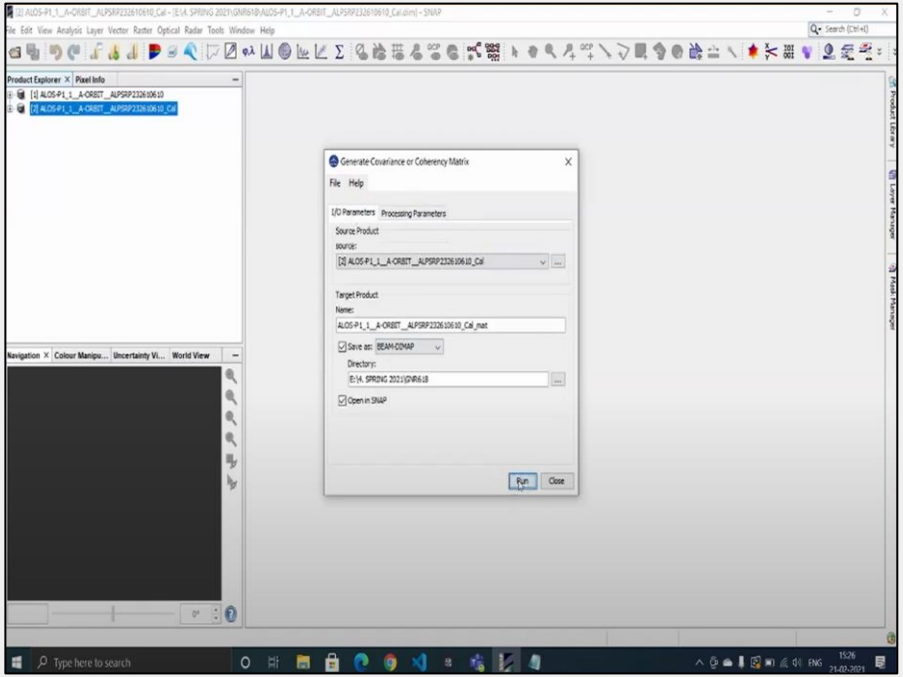
In the field of glaciology, satellite remote sensing allows us to conduct research on remote areas that are hard to access, and to gather data on a spatial extent not possible by eldwork alone. Thus we use POL-SAR data of the Gangotri Glacier located in Himalayas to classify the various features of the glacier.
Data acquired by polarimetric SAR (POLSAR) are directly related to physical properties of natural media.

Method and Data Processing

Data Calibration



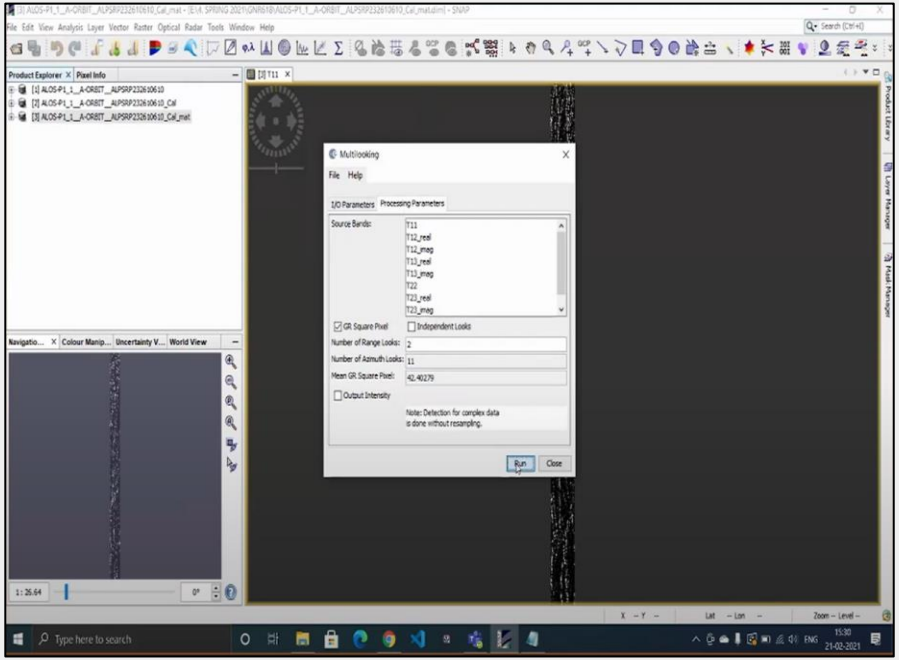
T Matrix Generation



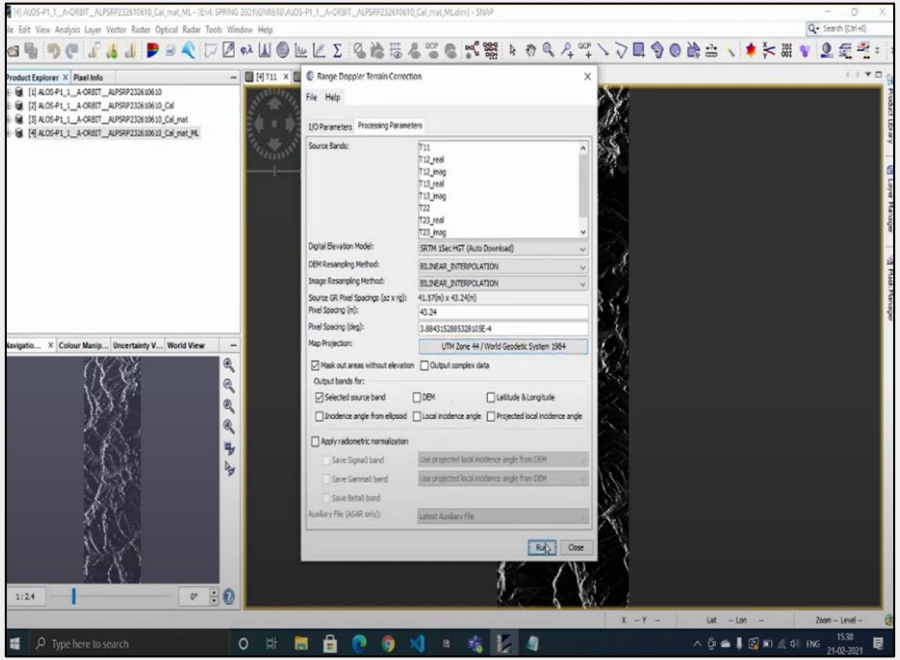
We select the data and select radar for radio calibration. The calibrated data is just the original data multiplied with calibration factor. T matrix has direct interpretation of target.

Method and Data Processing

Multilooking



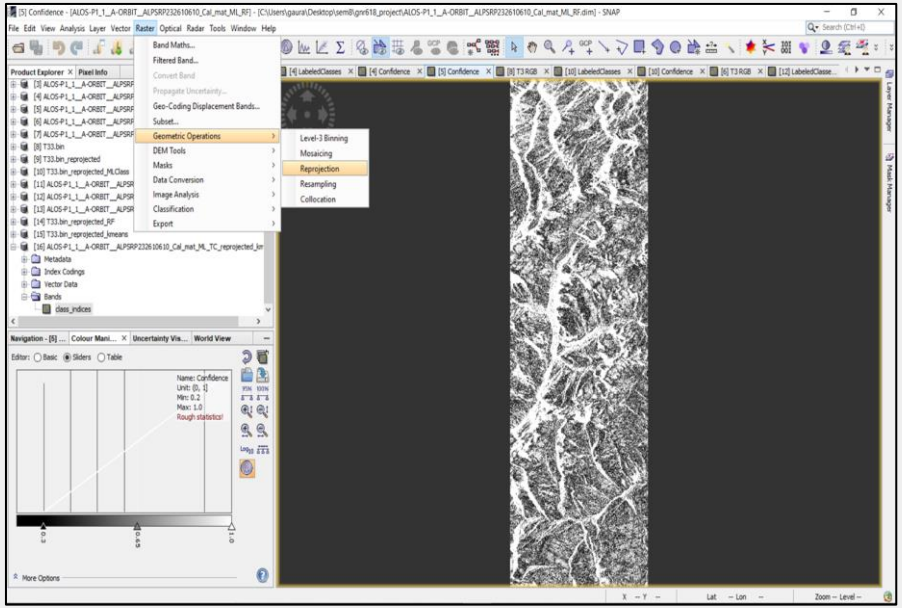
Range Doppler Terrain Correction



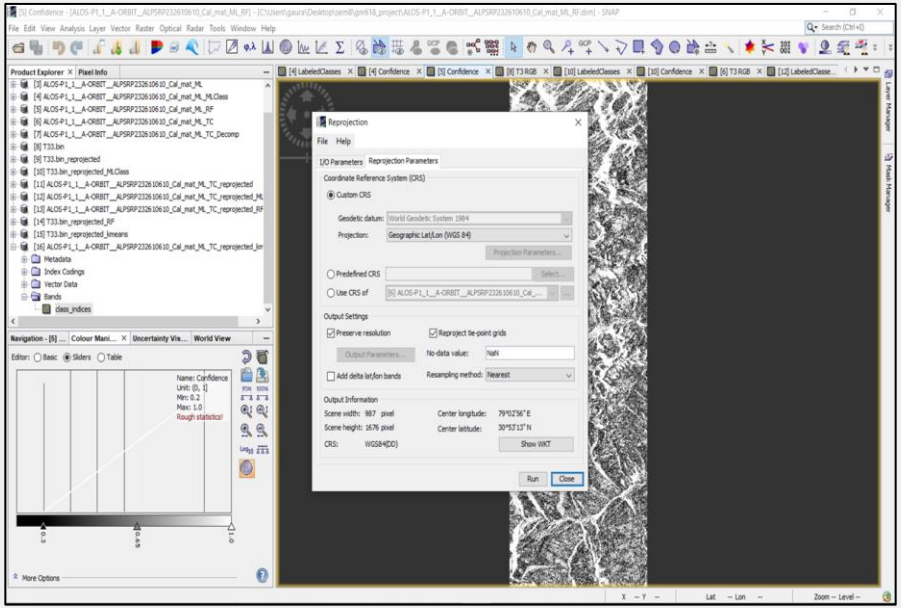
We select multilooking from the radar tab. Multilook data is more clear than single look data. Then we select Radar->Geometric->Terrain Correction->Range Doppler Terrain Correction. Terrain correction is needed since we need to correct the geometric effect as the radar is looking sideways.

Method and Data Processing

Reprojection Step 1



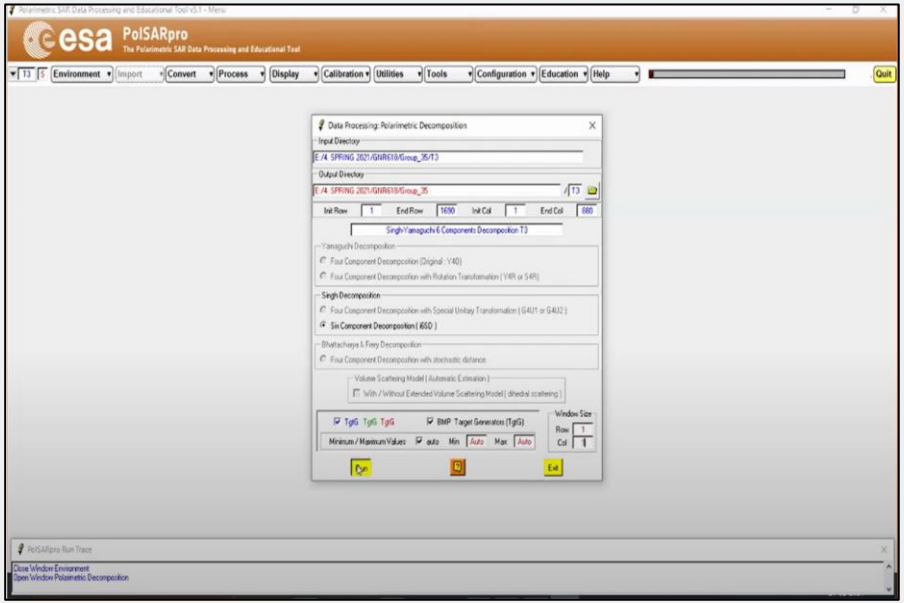
Reprojection Step 2



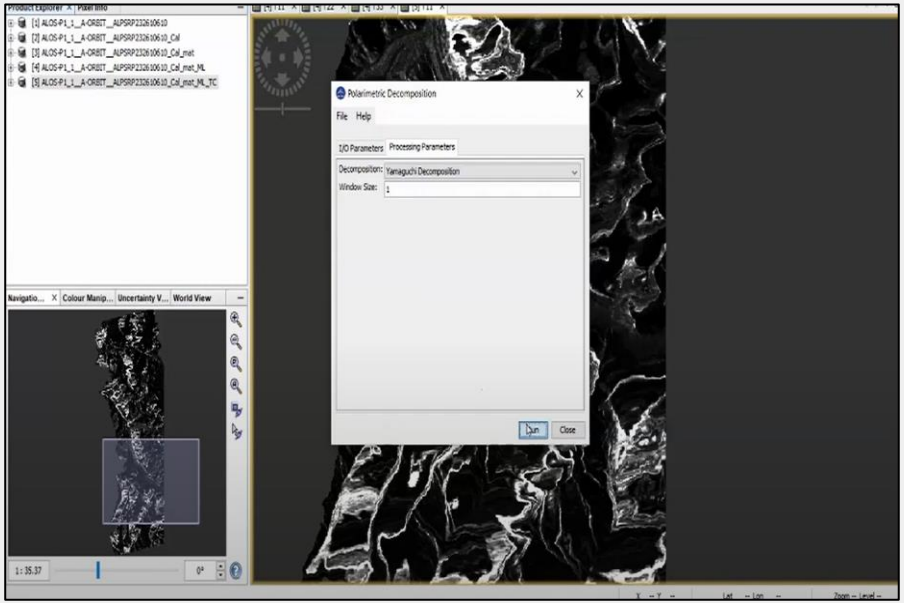
There are multiple CRS, and the product can be projected to another system, even to define a new custom coordinate reference system.
For this we go to Radar -> Geometric Operations -> Reprojection

Method and Data Processing

Decomposition in PolSARpro



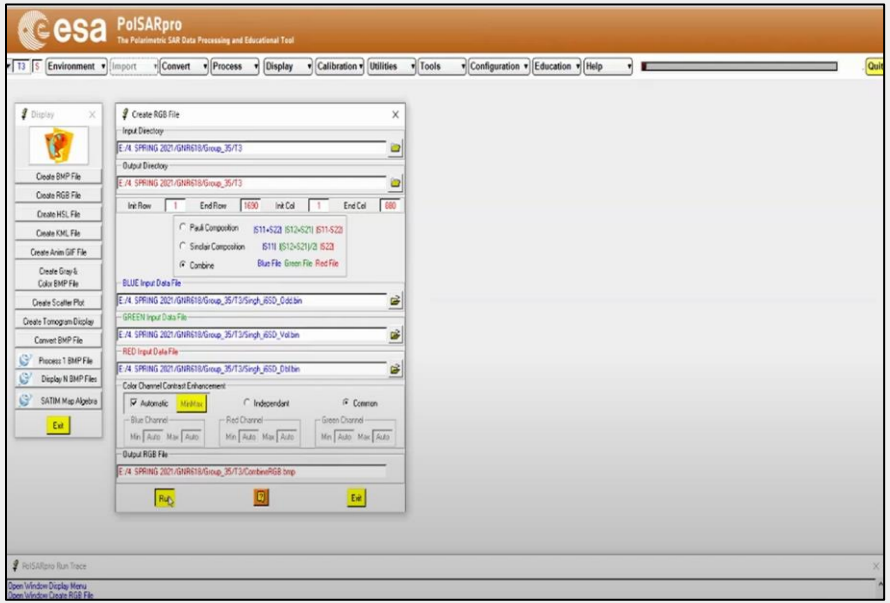
Decomposition in Snap



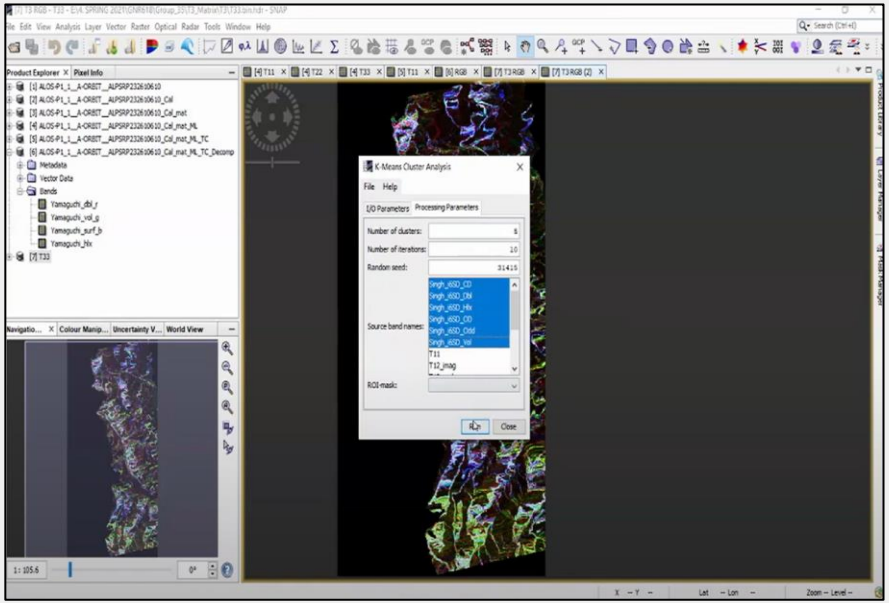
In PolSARpro, we select polarimetric decomposition from the Process tab and select iSSD. We then select TgtG and BMP and keeping row and column size as 1, we run it.

Method and Data Processing

RGB PolSARpro



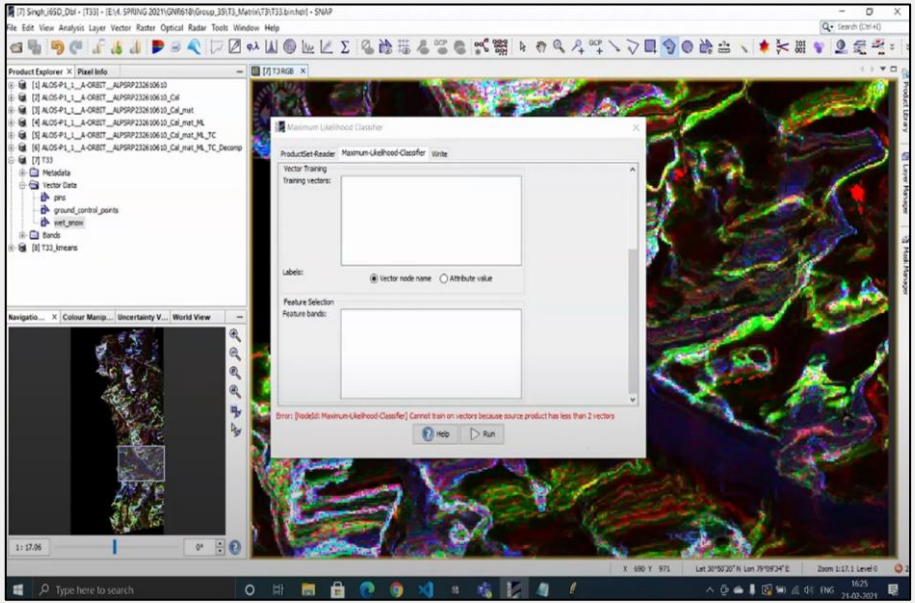
K Means Cluster Analysis



After opening RGB window we give the same extensions to Red, Blue and Green as earlier.
Then in Raster->Classification->Unsupervised Classification->K-Means Cluster Analysis we select 5 clusters,10 iterations we select all the 6 files.

Method and Data Processing

Vectors



Coherent and Incoherent matrices

Coherent

$$[S]_{relative} = \begin{vmatrix} |S_{HH}| & |S_{HV}|\, e^{j(\varphi^{HV} - \varphi^{HH})} \\ |S_{VH}|\, e^{j(\varphi^{VH} - \varphi^{HH})} & |S_{VV}|\, e^{j(\varphi^{VV} - \varphi^{HH})} \end{vmatrix}$$ 5

Scattering matrix

$$k_p = (1/2)\,\mathrm{Trace}\,\{S[\psi]\} = \frac{1}{\sqrt{2}} \begin{bmatrix} S_{HH} + S_{VV} \\ S_{HH} - S_{VV} \\ 2S_{HV} \end{bmatrix}$$

Scattering vector

$$\Psi = \left\{ \sqrt{2}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \sqrt{2}\begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}, \sqrt{2}\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \right\}$$

Incoherent

Coherency or Covariance matrix

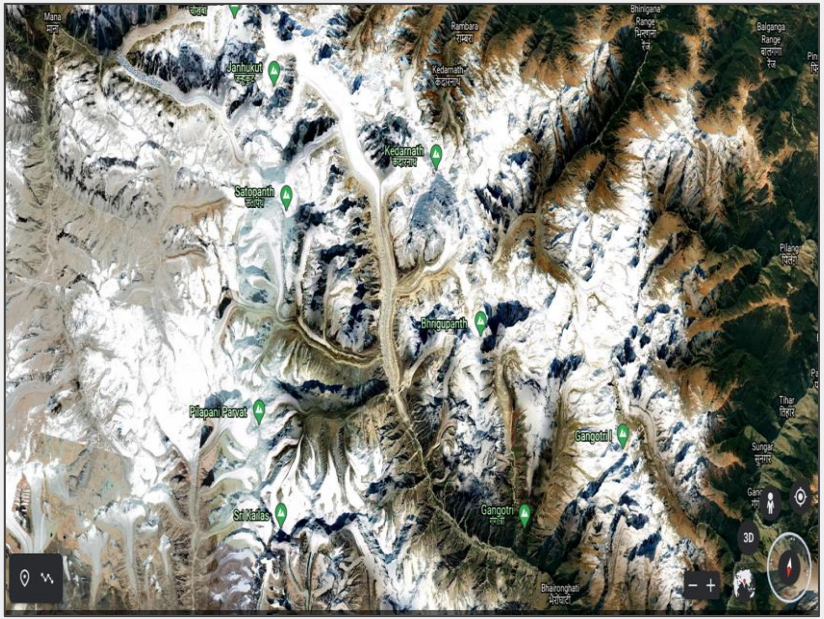
$$\langle [T] \rangle = \frac{1}{n} \sum^{n} k_p\, k_p^{\dagger} = \frac{1}{2} \begin{vmatrix} \langle |S_{HH} + S_{VV}|^2 \rangle & \langle (S_{HH} + S_{VV})(S_{HH} - S_{VV})^* \rangle & \langle 2S_{HV}^* (S_{HH} + S_{VV}) \rangle \\ \langle (S_{HH} - S_{VV})(S_{HH} + S_{VV})^* \rangle & \langle |S_{HH} - S_{VV}|^2 \rangle & \langle 2S_{HV}^* (S_{HH} - S_{VV}) \rangle \\ \langle 2S_{HV}(S_{HH} + S_{VV})^* \rangle & \langle 2S_{HV}(S_{HH} - S_{VV})^* \rangle & \langle 4|S_{HV}|^2 \rangle \end{vmatrix} = \begin{bmatrix} T_{11} & T_{12} & T_{13} \\ T_{12}^* & T_{22} & T_{23} \\ T_{13}^* & T_{23}^* & T_{33} \end{bmatrix}$$ 9

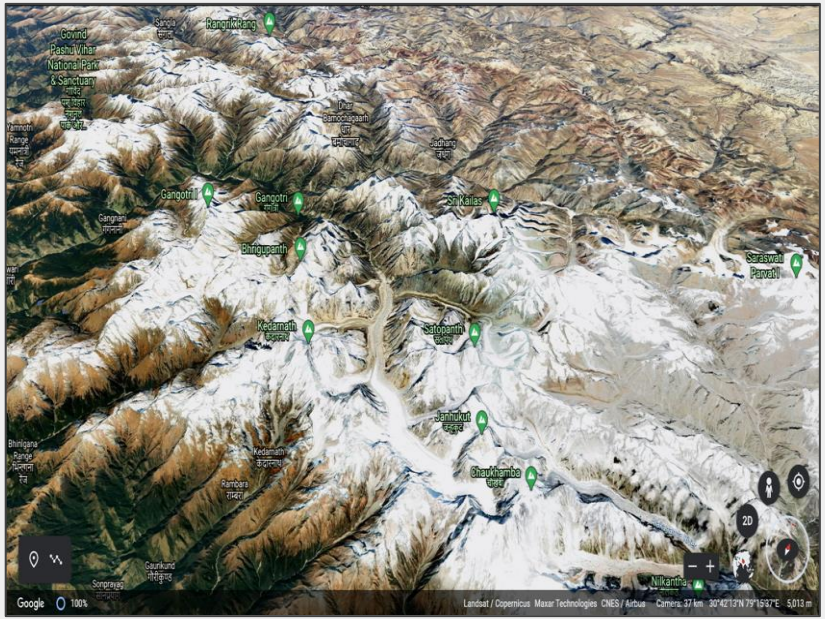
Re Im Re Im

For supervised classification we use vector and in new vector data we give name to it. After doing this for all areas we do supervised classification in raster. Coherent and Incoherent matrices are defined as shown in the picture. As a result, the elements of the T3 matrix are generated.

Study Area

Study area of our project is Gangotri Glacier, the longest glacier in the Indian Himalaya. It is located in Uttarkashi District, Uttarakhand, India in a region bordering Tibet. This glacier is bound between 30°43'22"{30°55'49" (lat.) and 79°4'41"{79°16'34" (long.), extending in height from 4120m to 7000m above sea level.
The overall geosystem is made up of various small glaciers as well as mountains, rivers and many valleys. Other important glaciers in the area are Chaukhamba, Janhukut, Satopanth, Bhrigupanth and Rangrik Rang. Major rivers in the area which contribute to the Bhagirathi river are Maindi, Swachand, Ganohim, Kirti, Raktavarn, Chaturangi and Medu.





Results and Discussion

We have obtained three types of datasets from the given data :

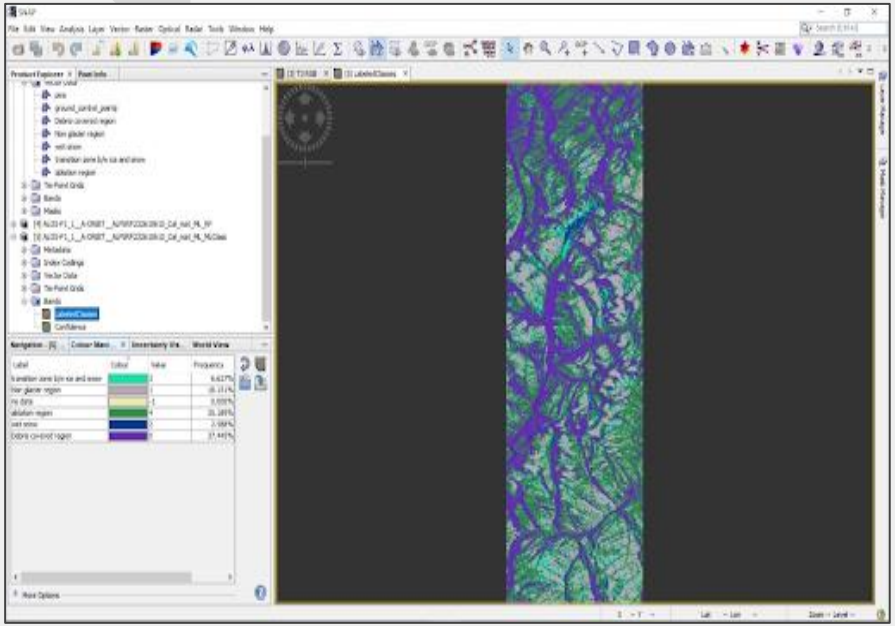
1) Multilook Data
2) Terrain Corrected Data without Singh-Yamaguchi Decomposition
3) Terrain Corrected Data without Singh-Yamaguchi Decomposition

We have used three types of models to train out datasets:

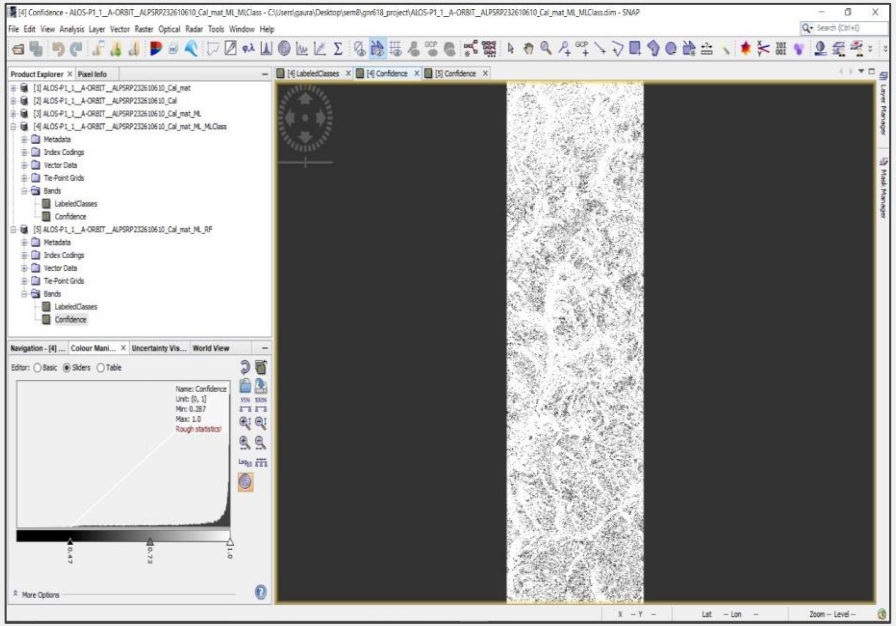
1) Maximum Likelihood Classifier
2) Random Forest Classifier
3) Unsupervised Classifier

Results and Discussion

1) Multilook Data



Maximum Likelihood Classifier on Multilook Data

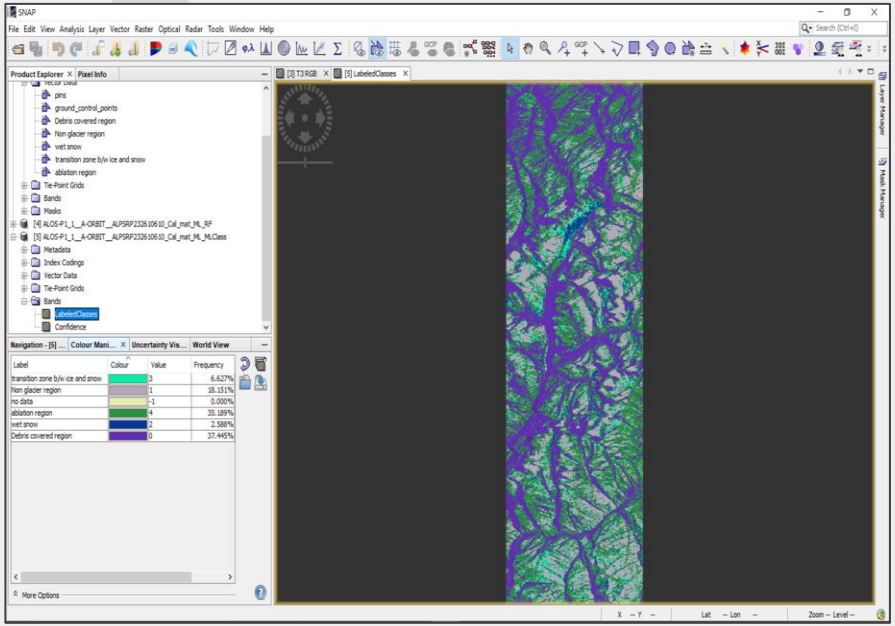


Confidence : Maximum Likelihood Classifier on Multilook Data

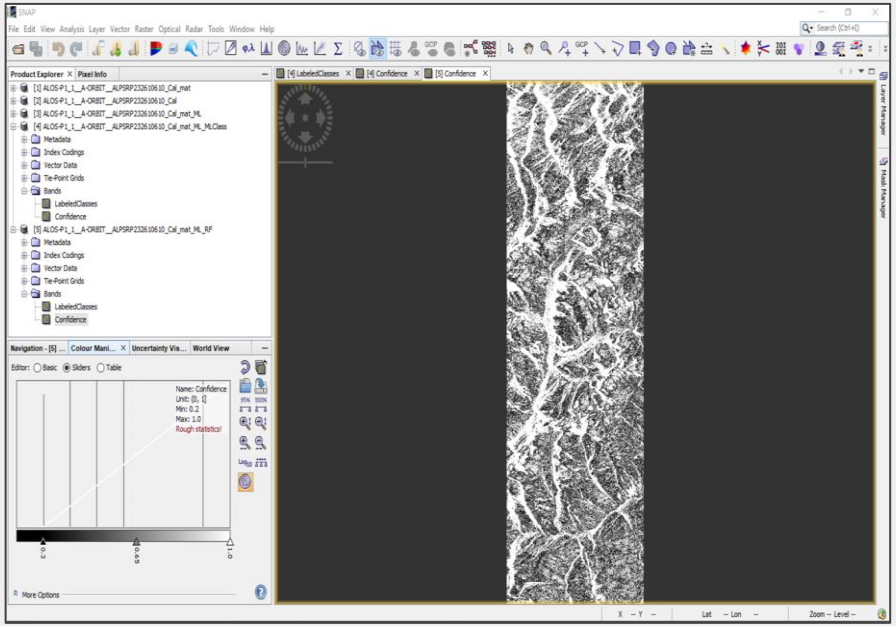
Maximum Likelihood Classifier model predicts almost the same frequencies for Ablation Region and Debris Covered Area, higher than others. The min confidence we get is 0.287 and the max confidence we get is 1.00. The most of the region is above 0.9 as we can see in the graph that the peak is very close to 1

Results and Discussion

1) Multilook Data



Random Forest Classifier on Multilook Data

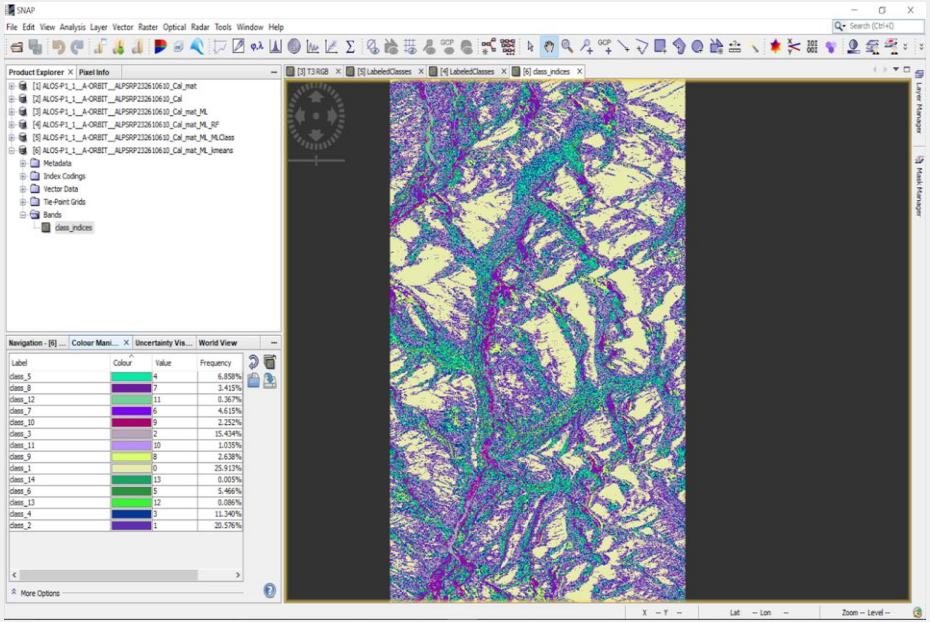


Confidence : Random Forest Classifier on Multilook Data

Random Forest Classifier model predicts higher frequency for Debris Covered Area than others. Confidence levels are low for most of the regions. Moreover the confidence levels are discrete. So we can say that the model is not that confident in the class which it predicts. The max confidence we get is 1.00 and the min confidence we get is 0.2.

Results and Discussion

1) Multilook Data

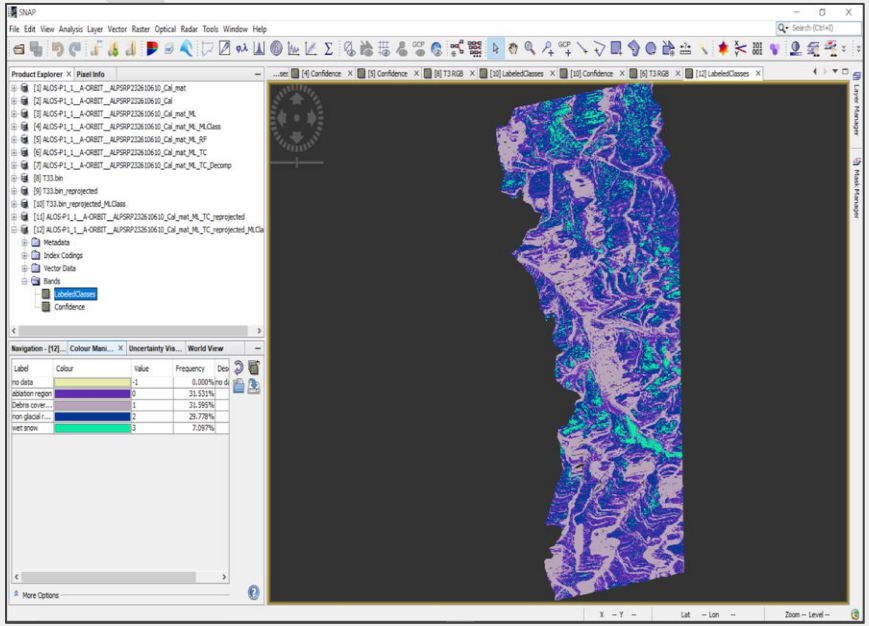


Unsupervised Classifier on Multilook Data

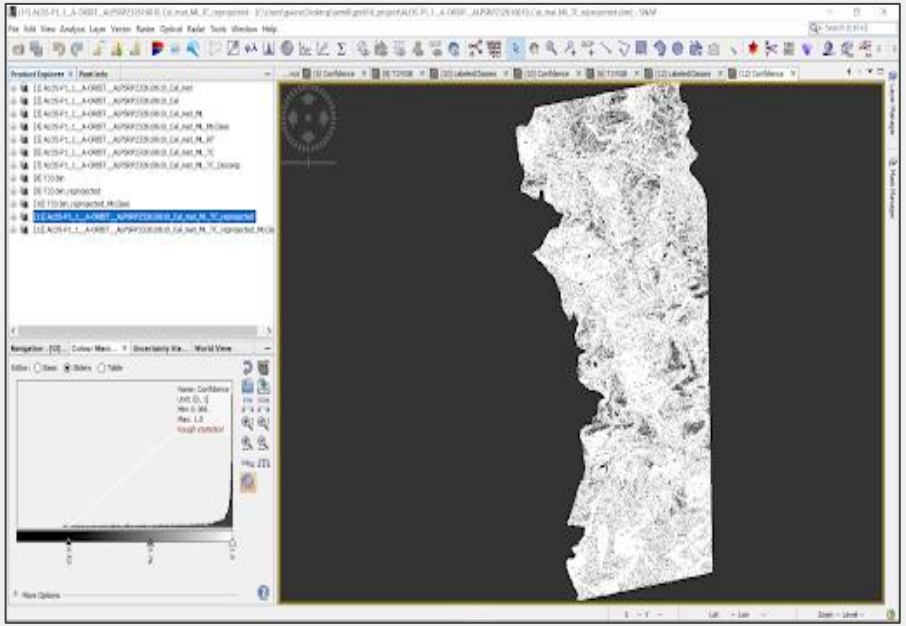
We can observe that this classier has classified the data in multiple classes generated by itself. It has further breaked the supervised classes into two or three classes. Unsupervised Classifier predicts more classes according to diversities in supervised classes.

Results and Discussion

2) Terrain Corrected Data without Decomposition



Maximum Likelihood Classifier on TC Data without Decomposition Decomposition

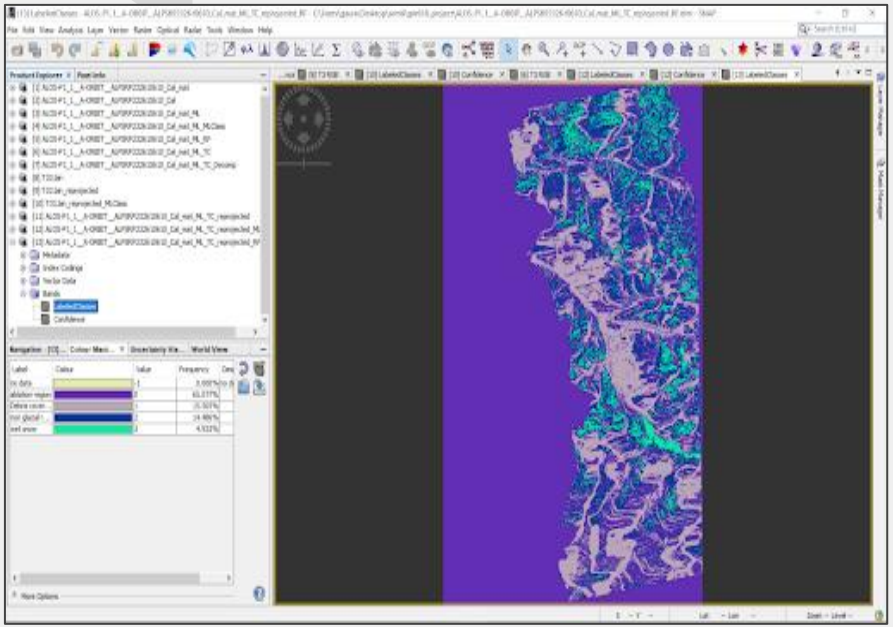


Confidence : Maximum Likelihood Classifier on TC Data without

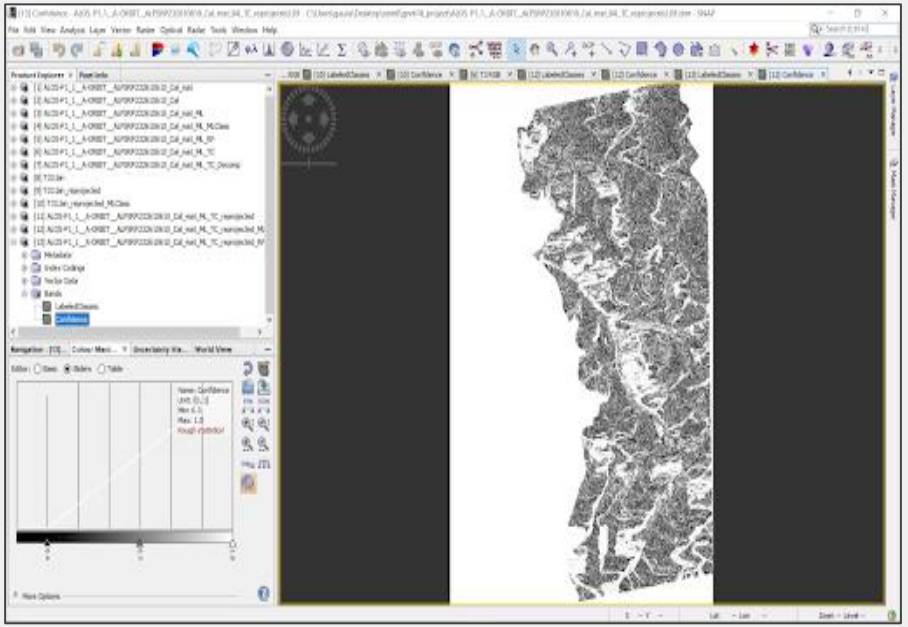
Maximum Likelihood Classifier model predicts almost the same frequencies for Ablation Region, Debris Covered Area and Non Glacial Region.
The min confidence we get is 0.366 and the max confidence we get is 1.00. The most of the region is above 0.9 as we can see in the graph that peak is very near to 1.

Results and Discussion

2) Terrain Corrected Data without Decomposition



Random Forest Classifier on TC Data without Decomposition

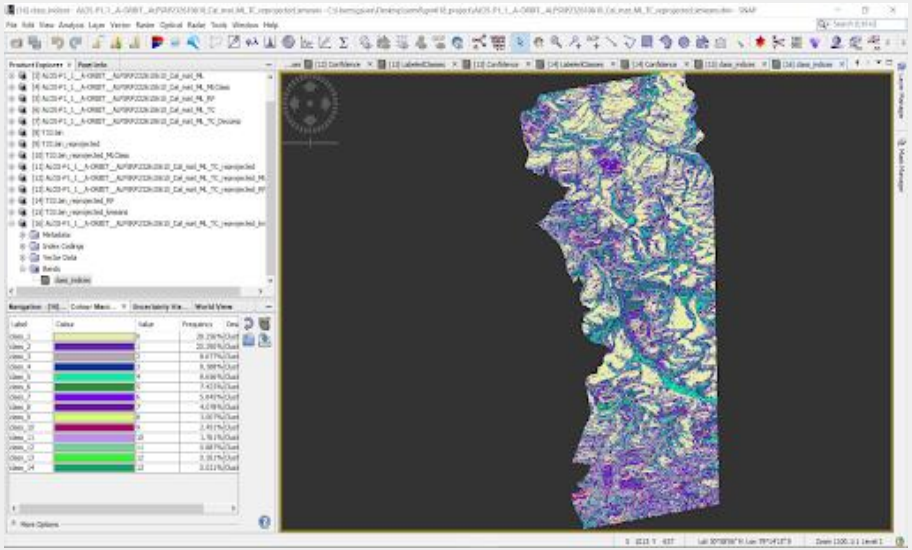


Confidence : Random Forest Classifier on TC Data without Decomposition

Random Forest Classifier model predicts higher frequency for the Ablation Region than other regions.
In this model we can see the confidence levels are low for most of the regions. Moreover the confidence levels are discrete. So we can say that the model is not that confident in the class which it predicts. The max confidence we get is 1.00 and the min confidence we get is 0.3.

Results and Discussion

2) Terrain Corrected Data without Decomposition

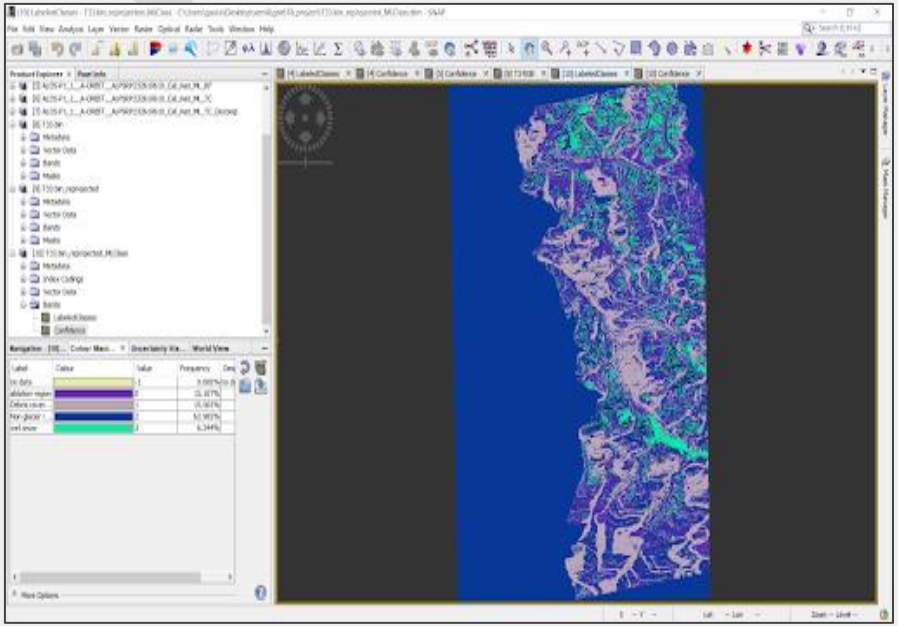


Unsupervised Classifier on Terrain Corrected Data without Decomposition

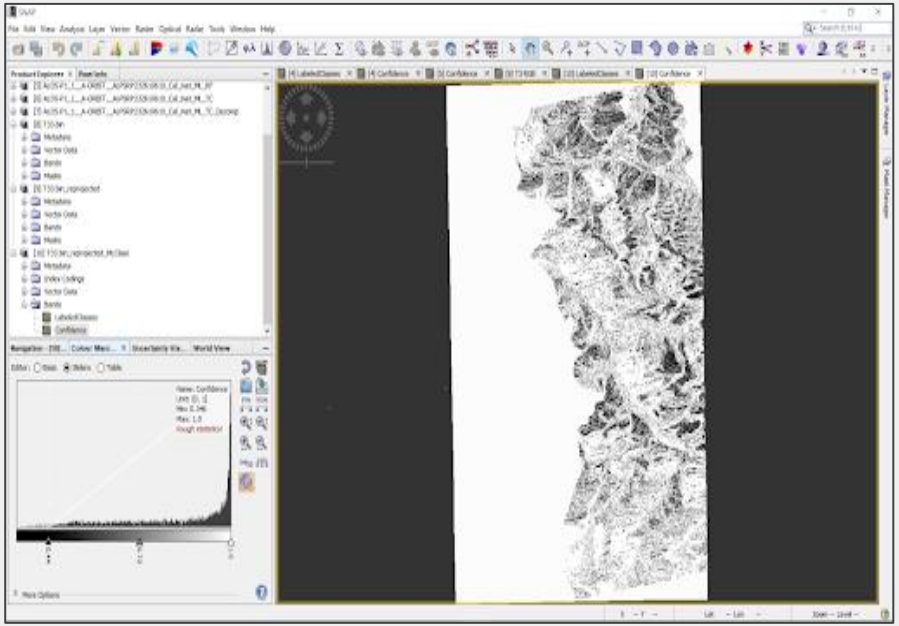
We can observe that this classier has classified the data in multiple classes generated by itself. It has further breaked the supervised classes into two or three classes. Unsupervised Classifier predicts more classes according to diversities in supervised classes.

Results and Discussion

3) Terrain Corrected Data with Decomposition



Maximum Likelihood Classifier on TC Data with Decomposition Decomposition

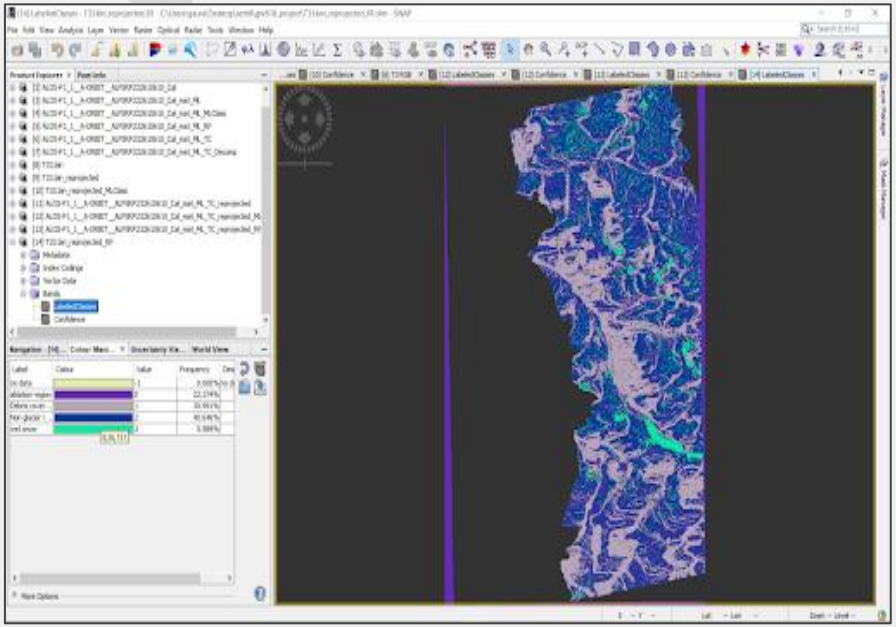


Confidence : Maximum Likelihood Classifier on TC Data with

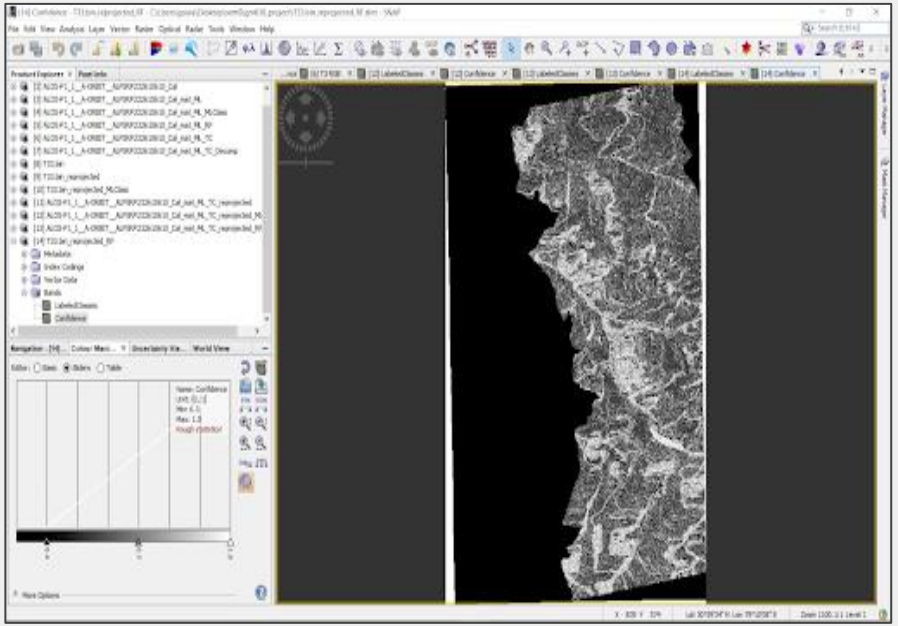
Maximum Likelihood Classifier model predicts higher frequency for the Non Glacial Region than other regions.
Now this time the maximum likelihood classifier predicts the class with a not as good confidence as before the decomposition as we can see in the image, a lot of black areas. The minimum confidence we get is 0.346 and maximum confidence is 1.00.

Results and Discussion

3) Terrain Corrected Data with Decomposition



Random Forest Classifier on TC Data with Decomposition

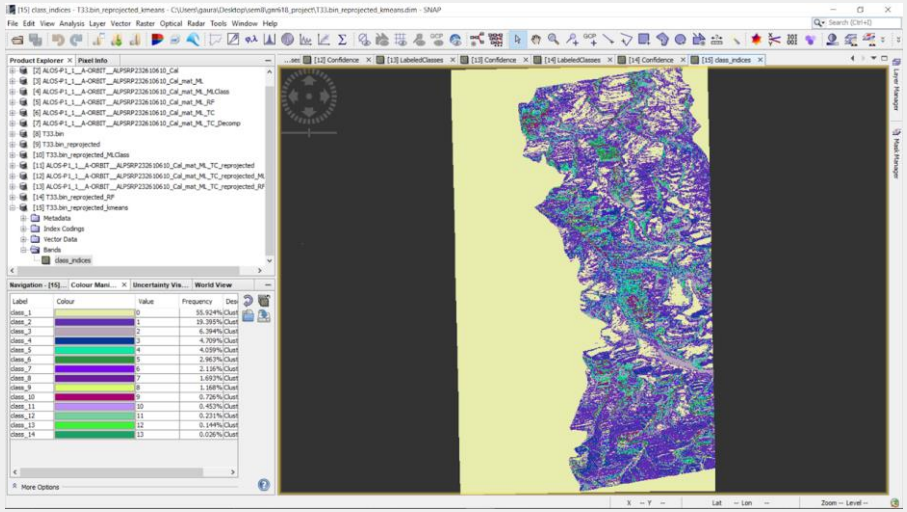


Confidence : Random Forest Classifier on TC Data with Decomposition

Random Forest Classifier model predicts lower frequency for the Wet Snow Region than other regions.
In this model we can see the confidence levels are low for most of the regions. Moreover the confidence levels are discrete. So we can say that the model is not that confident in the class which it predicts. The max confidence we get is 1.00 and the min confidence we get is 0.2.

Results and Discussion

3) Terrain Corrected Data with Decomposition :



Unsupervised Classifier on Terrain Corrected Data with Decomposition

We can observe that this classier has classified the data in multiple classes generated by itself. It has further breaked the supervised classes into two or three classes. Unsupervised Classifier predicts more classes according to diversities in supervised classes.

Conclusion

1) Remote sensing helps us locate glaciers and gather data related to them which we can use to implement other techniques to extract information about them.
2) We in this project, applied Machine Learning techniques after processing the data. We saw how machine learning can help in classifying various features of glaciers. We tried two supervised learning models and one unsupervised learning model to segregate the features of the glacier.
3) First we trained on the multilook data, in the multilook data we get the models get the best confidence, after that we trained on the terrain corrected data as multilook had no information regarding the 3rd dimension in this result we still get good confidence from the models but as we go to the Singh-Yamaguchi decomposed data the condence of the models drop more down on the predictions it make.
4) After this we also tried to train an unsupervised learning model K-means clustering which gave us a total of 12 classes which tell us that there can be a lot more features of the glacier that can be predicted using Machine Learning if we give proper labelled data.
5) Thus machine learning gives us a lot of useful techniques through which we can extract the information of not only Glaciers but also Ice Sheets and Icebergs from the PolSAR data or optical data.

References

1) Recorded Video Lectures of GNR 618 : Remote Sensing and GIS applications to Cryosphere by Prof. Gulab Singh, CSRE, IIT Bombay

2) Rees,W.G., Remote Sensing of Snow and Ice (2006)

3) (Optical Science and Engineering) Jong-Sen Lee, Eric Pottier - Polarimetric Radar Imaging From Basics to Applications -CRC Press (2009)

4) L_12 Glacier Facies Mapping HKH Region Naveen Tripathi, Vedas, ISRO

5) Wikipedia [en.wikipedia.org]

6) Google Earth Engine

7) POLSARPro Software

8) SNAP Software

THANK YOU!