University of Tartu
Faculty of Science and Technology
Institute of Ecology and Earth Sciences
Department of Geography

Project Report on Data Science in Remote Sensing
Topic: Monitoring Beaver-induced flooding in Estonia

Students: Lilian Akudo Akanazu
Sourav Karmakar
Mizanur Rahman
Supervisor: Mihkel Kaha

Tartu 2021

Table of contents

1 Introduction

Beavers are said to be ecosystem engineers and are essential for sustaining wetland and floodplain habitats in many landscapes (Nick, 2017). In large natural landscapes, beavers help other species like elks and salmons to thrive when the floodplains caused by beavers sustain marshes and willow thickets, and wolves that feed on elks are also found in similar areas. Thus, beavers also increase the biodiversity along rivers and wetlands (Nick, 2017) but, there is more to the beaver issues. Creating dams strongly influences the flow of water and the increased deposition of water in forest ecosystems causes floods that damage natural vegetation. They also target the largest trees in willows and most expensive birch trees to build their lodge and provide food for the winter. Because beavers are well-known agents of disturbance and wetlands formation in their distribution, it is important to understand how and where they are in the forest and what impacts their modification has on natural vegetation.

1.1 Research Problem

Vegetation change in wetlands can be a function of water level fluctuations resulting from natural or anthropogenic disturbances over various time scales. Research has shown that the natural causes of vegetation change may be related to climatic shifts accompanied by increases or decreases in precipitation or faunal activity, specifically, that of beaver and muskrat (Nick, 2017). It is generally recognized that anchored bog vegetation is usually killed because of prolonged flooding (Nick, 2017). Based on an evolutional study of beaver-induced flooding in the arctic, Tape et. al (2018) stated that beaver pond formation increases winter water temperatures in the pond and downstream, likely creating new and more varied aquatic habitat, but the specific biological implications of this ponding in forests are unknown. It is based on this premise that this project would focus on accessing the damages on forests caused by beaver-induced flooding.

1.2 Study Area

The study area is the forested landscape of Estonia, and the southern part of the country has been selected as the focus for this research. This area was selected because of the abundance of lakes and ponds in the area as it has been established earlier that waterbodies are beavers' natural habitat. Figure 1 below shows a map of the study area highlighting the region of interest in Estonia this project will focus on. This area was chosen because of the known clustering of beaver points in the area based on the derived beaver location data obtained for this research.

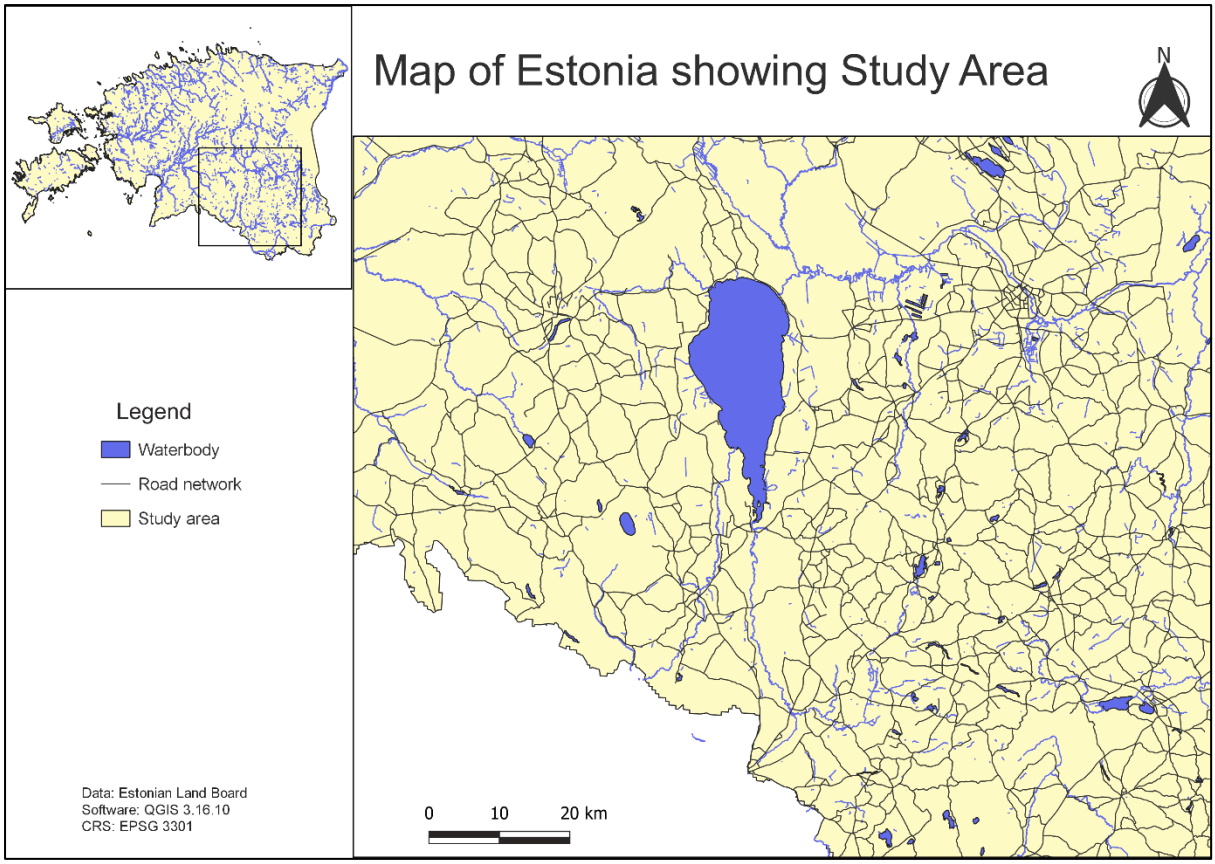


Figure 1 – Map of Estonia showing location of study area

1.3 Research Methodology

This study is based on empirical methods which include satellite imagery of the study area, orthophotos of the area of interest, collection of known beaver habitat points from hunters, established theoretical frameworks by other researchers, the spatial contour of the study area, and the water and land data downloaded from the Estonian Land Board. The satellite imagery was used to detect the location of floods damage and to access the pattern of forest damage around water bodies, the contour of the study area was used to extract the total area of interest and the known beaver points served as a reference in the identification of active beaver locations for the collection of training samples for the algorithm. This project employed the unsupervised and supervised classification methods to classify the forests as damaged or healthy using the training polygons as signatures for classification.

1.3.1 Data Acquisition

10 sentinel 2 satellite imageries of the study area were downloaded from Copernicus sci-hub with dates ranging from 2017-2020. The beaver location points were gathered by the hunters

in previous years and this data was shared by the project supervisor. The orthophotos of the study area were also downloaded to identify damaged forests from space and for this project, the orthophoto used was the 'ajooline' for the year 2019 and 2021, which is a historical RGB orthophoto that can show the change in time, the condition of the forests in the region of interest. The land and water shapefiles were downloaded from the Estonian Land Board to identify the locations of water bodies in the area as well as to create buffers around the area of interest.

1.3.2 Data pre-processing

The downloaded satellite imageries were inspected for visibility and cloud-free cover and the best 5 of the satellite imageries were selected for the project. These satellite imageries were resampled using the sentinel application platform (SNAP) to adjust them to the same resolution of 10 meters, against the 10m, 20m, and 60m they had during download. Bands 2, 3, 4, 5, 6, 7, 8, 8A, 11, and 12 were selected during the resampling which was used to calculate the spectral indices of the imageries. The downloaded orthophoto was used to map out damaged forests by beavers and healthy forest polygons using the beaver points gathered as a guide, to generate training polygons for our model before processing.

2 Data Processing

Step 1 - Calculation of spectral indices: Three spectral indices were selected for the calculation to understand the condition of the vegetation cover in the region of interest and its characteristics. The three spectral indices were the NDVI, NDWI, and MSI, and these indices were calculated for all 5 satellite imageries that were resampled in SNAP.

The normalized difference vegetation index (NDVI) is used to quantify vegetation greenness and is useful in understanding vegetation density and assessing changes in plant health and plant cover. This index was calculated to access the changes in forest cover and to see the pattern of forest loss in areas of beaver colonization. NDVI is calculated as a ratio between the red (R) and near-infrared (NIR) values in traditional fashion: NDVI= (NIR - R) / (NIR + R). In Annex 1 figure 1 below, the NDVI of one spectral image shows the pattern of vegetation cover in the study area. The darker areas represent more vegetation cover while the lighter areas show less vegetation cover. Along water bodies, it was observed that forest cover

becomes reduced however, it is not logical to conclude at this point that they have been destroyed by beaver until further analysis.

The NDWI is a remote sensing-based indicator sensitive to the change in the water content of leaves (Gao, 1996). The Normalized Difference Water Index (NDWI) is known to be strongly related to the plant water content and a very good proxy for plant water stress. This index was calculated to access the change in water content of the vegetation to see the spatial distribution of damaged trees. This index is derived from Near-Infrared (NIR) and Short Wave Infrared (SWIR). NDWI= (NIR-SWIR)/ (NIR+SWIR). The figure in annex 1 figure 2 below shows the water index calculated image which is a reflectance of the water content in forested vegetations in the area.

The moisture stress index (MSI) uses the NIR and SWIR channels to measure healthy moisture, pixel by pixel with a simple algorithm (Ceccato et al, 2001). The MSI is a reflectance measurement, sensitive to increases in leaf water content, and was calculated to see the level of stress in the plant canopies because of changes in the water content of the plants. This index is derived by dividing the short-wave infrared by the near-infrared spectral bands. MSI= (SWIR/NIR). In annex 1 figure 3 below, the image shows the lesser plant water stress and more soil moisture content as the values were within the expected range.

Step 2 – Creating Virtual Raster: A virtual raster is a special type of raster whose pixel values are created upon request during data processing or visualization. Its memory footprint includes a reference to a source raster along with details of how the source pixel values will be modified to create new pixels. The virtual raster was created using the calculated spectral indices, the mask layer containing the 100m buffer around flowing water bodies and healthy forests, and the resampled images from SNAP. This process was done to conserve time and disc space and to calculate only the pixels of the bands needed for this analysis instead of the entire image. Figure 2 below shows a sample of the virtual raster layer generated using a combination of the spectral indices and the resamples imageries.

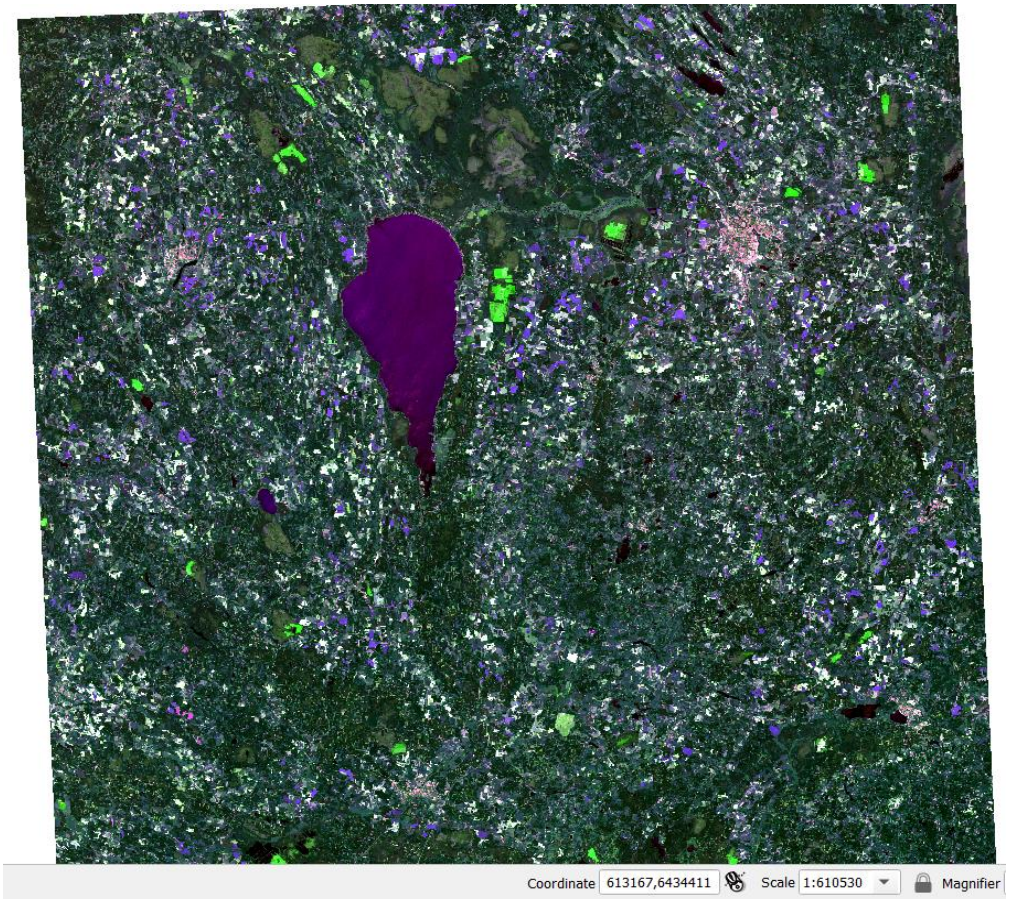


Figure 2 – Sample of Virtual Raster imagery

Step 3 - Clip and Mask from virtual raster: A mask layer was generated after creating the virtual raster by clipping the polygons of the beaver areas and the healthy forest to the virtual raster image. This was done to get rid of a large amount of data not needed for the training model and to narrow the focus on the regions of interest. These clipped raster imageries were merged, and this procedure was repeated for both beaver polygons and the healthy forest polygons independently.

Step 4 - K-means Clustering: A cluster of polygons was generated after the clipped raster polygons were merged to group the dataset into homogenous groups. The K-means algorithm is an iterative algorithm that tries to partition the dataset into *K*-pre-defined distinct non-overlapping subgroups (clusters) where each data point belongs to **only one group**. It tries to make the intra-cluster data points as similar as possible while also keeping the clusters as different (far) as possible (Imad, 2017). It assigns data points to a cluster such that the sum of the squared distance between the data points and the cluster's centroid (arithmetic mean of all the data points that belong to that cluster) is at the minimum. The less variation we have within clusters, the more homogeneous (similar) the data points are within the same cluster (Imad,

2017). Figure 4 in annex 1 below shows the classes for the clustering process. In healthy forest and beaver classes, there were some variations based on the reflectance signature as different type of forests from clear cut, bare soils, coniferous forests, deciduous vegetation, etc. were considered as healthy forest. In order to reduce the reflectance variation, subclasses were done by computer based on their reflectance signature and in the end of the analysis these subclasses were again converted into masterclass. For this project processing, 20 clusters of healthy forests and 5 clusters of beaver-damaged forests were generated. The subclasses have similarity in spectral signature thereby more distinctly separable during the classification process. The Semi-automatic plugin (SCP) in QGIS was used to perform this operation.

Step 5 – Vectorization, Cleaning, and Dissolving: The SCP plugin uses vector layers to process unsupervised random forest classification and as a result, the clustered data which is in raster format was transformed to vector layer using the raster to vector tool in QGIS. During this transformation process, some errors were generated as with all vectorizations. To remove the errors, the data was cleaned using the v.clean tool and the fix geometries tool in QGIS. The vectorization process generated small vector polygons for the clustered raster pixels and to merge them into homogenous polygons, the dissolve tool in QGIS was used to merge polygons of similar values into one class. Figures 5 and 6 in annex 1 below show the dissolved images of the forest and beaver polygon layers which were used as training input for the algorithm.

Step 6 - Preparation of Training Data: To generate the training data for the model, the dissolved layer was added to the SCP plugin to train the machine and create signatures and a unique id for each class. The first classification was done using the polygons of the beaver damaged areas, then the healthy forest polygons were also added in the same training data to train the machine on identifying the healthy forests in the study area. The result was a classification of beaver damaged areas and healthy forests with unique IDs assigned by the classification, such that the machine can classify in future, sample areas into either healthy forests or beaver-damaged areas by recognizing the attributes of the trained data. Figure 7 in annex one below shows the classification and the unique ID assigned to each type of forest and beaver sampled area.

Step 7 - Random Forest Classification: This process focused on classifying the pixels in the study area and identifying them as either damaged forests or healthy forests based on the signatures of the established training model. At first, the SCP plugin was used but produced

errors in the classification after lengthy hours so, the dzetsaka plugin in QGIS was used as an alternative in the random forest classification. The input data for this process was the dissolved layer containing the beaver and forest region of interest, the virtual raster generated by a combination of all spectral indices and imagery bands, and the mask layer containing the 100m buffer around flowing water body and healthy forest. The result was the output of the random forest classification, confidence map, and confusion matrix. The confidence map (as seen in annex 1 figure 8 below) shows how confident the machine was when it detected a beaver damaged area and a healthy forest area through graduation of the color intensity, that is, the darker the pixel, the more confident the machine was in identifying that area as either a beaver-damaged area or a healthy forest. From the confusion matrix, several indices were calculated to access the accuracy of the model and how precise it was in detecting these beaver-damaged forests. The random forest classification was iterated three times for each imagery with the addition of new polygons for the healthy forest to give the machine more sample areas for correct classification. After the three iterations, the whole polygons were combined into one layer and the forest classification was iterated one more time to have more sample areas to feed the machine as a healthy forest so that the classified result would be as precise as possible.

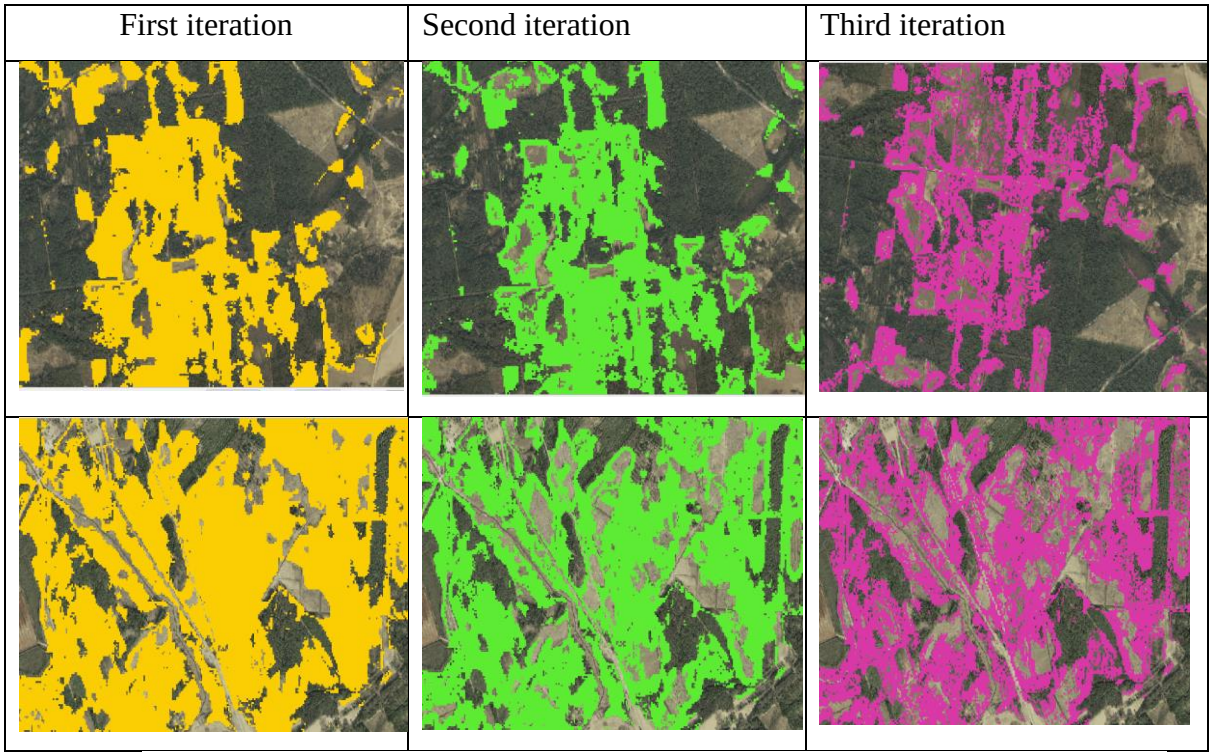


Figure 2 – Sequence of iterations of random forest classification carried out for training algorithm

Figure 2 above shows the three times iteration results of two places where the model falsely classified deforested lands as beaver damaged areas. On the first iteration, large areas were misclassified as beaver-damaged areas. Then some new healthy forest polygons were drawn in the misclassified areas and were added to the machine as a new sample of forests to train the model again. On the second iteration, as the model understood some deforested lands as forest areas, it classified those areas as forest as a result, beaver-damaged areas were reduced. The third iteration gave the best result among these three iterations showing that falsely classified areas would be reduced with continuous iteration.

Step 8 – Data Cleaning: After getting the confusion matrix, validation of the classification was done. Before the validation, data cleaning was conducted to get rid of small classification error that was produced at the time of processing. For that, at first the areas having confidence of at least 75% was selected through raster calculation. Then all healthy forest polygons were reclassified to one masterclass and then removed as beaver damaged areas was the only focus. In the meantime, all the beaver damaged classes were reclassified into one class. Following this, the raster image containing the beaver damaged classes was vectorized. After that, a 20m buffer was drawn around the beaver damaged vectorized layer. To avoid the overlapping, buffer was later dissolved. In order to get rid of minor errors in the dissolved layer, it was undergone cleaning process with 20m threshold with "v.clean" tool in QGIS.

Step 9 – Validation of the Results: After the cleaning, the result of the model was validated by selecting sample polygons and confirming their current state using the orthophotos of the areas. For that, 1% of the classified polygons of cleaned areas was randomly generated and checked one by one manually by comparing with the orthophotos whether these were beaver damaged areas or not. The result of the visual validation accesses the true accuracy of the model in predicting beaver-damaged areas.

3 – Analysis of Result

3.1 Classification Results: The classification result gave credence to our initial position that, beaver-damaged forests were found in the study area and they are more active in forest areas close to the water bodies. However, in the random forest classification, some deforested and anthropogenically damaged areas were also counted as beaver damaged areas. While iterating the result multiple times by adding new healthy forest polygons to the training data, the result became better after each iteration with the removal of forested areas from being classified as beaver-damaged areas. On every iteration, falsely classified forest areas were diminishing and

thus beaver damaged areas were being more realistic compared to the previous iteration. Figure 3 below shows a representation of the result of the third classification and orthophoto imagery of the classified area.

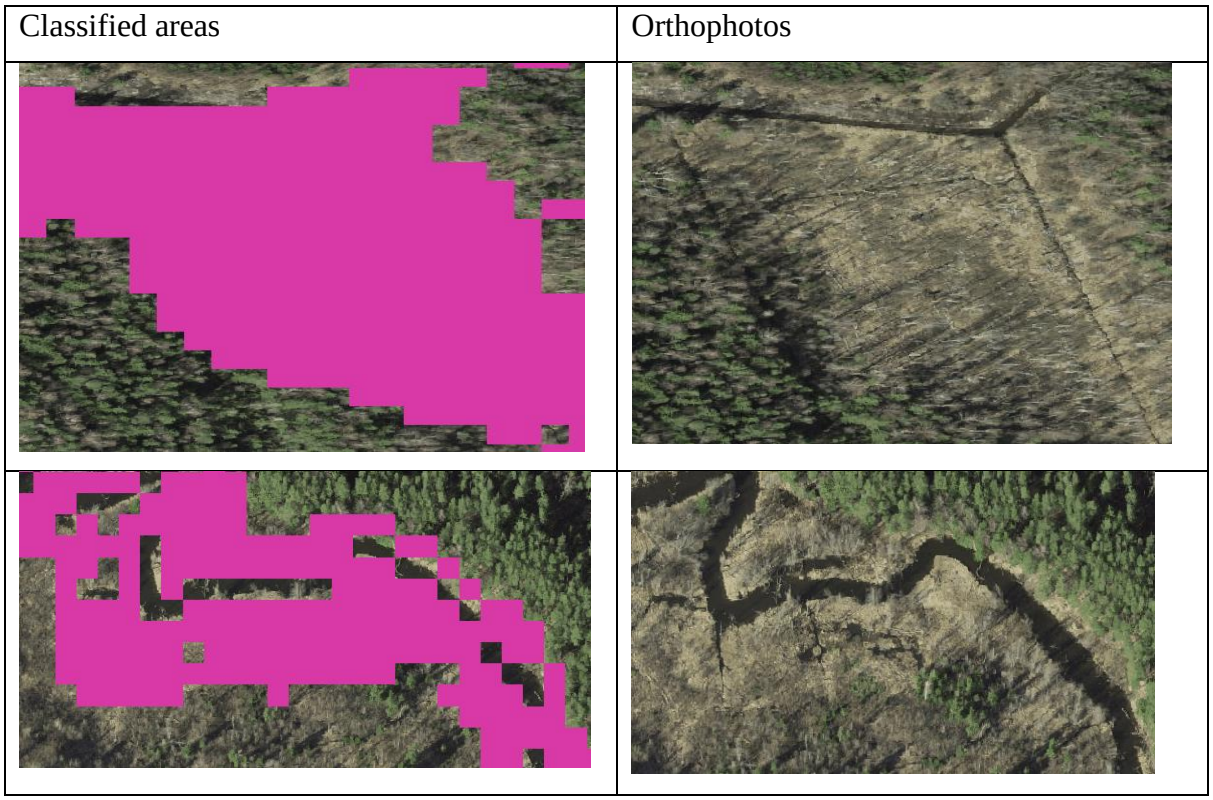


Figure 3 – Third iteration classification and orthophoto comparison of the classified area

The figure above is a glimpse of truly classified areas for beaver-damaged forest and a comparison with orthophotos. It is observed that a substantial number of beaver-damaged areas were identified through this study but not the all and in the third iteration, waterbodies were eliminated from the classification as beaver-damaged areas. Although some areas were mixed with deforested forests and human-damaged forest areas as both beaver damaged and deforested areas appeared almost similar in satellite imageries and orthophotos, some beaver damaged areas were correctly identified as shown in the figure above.

3.2 Analysis of classification Indices

The confusion matrix of the third iteration was reduced to a binary classifier, basically, a 4*4 matrix and four key values were identified: **true positive, true negative, false positive**, and **false negative**. As the interest of this study was on the beaver damaged areas, the true positive represented beavers, and the true negative represented the healthy forest areas. That is, the

matrix would be a Yes when it is a beaver classified area and a No when it classified an area as forest.

Basic terms of the confusion matrix

- **true positives (TP):** These are cases in which we predicted yes (they are beaver damaged areas), and they are beaver damaged areas.
- **true negatives (TN):** We predicted no, that is they are healthy forests, and they are healthy forests.
- **false positives (FP):** We predicted yes, as beaver-damaged areas but they are forests (Also known as a "Type I error.")
- **false negatives (FN):** We predicted no that they are forests, but they are beaver-damaged areas. (Also known as a "Type II error.")

Accuracy: Measures the overall correctness of the model

- (total positive +total negative) / total

Misclassification Rate: Measures the overall incorrectness of the model that is, how often does the model predict a yes when it is a no

- (false positive +false negative) / total
- equivalent to 1 minus Accuracy
- also known as "Error Ratio"

Precision: When it predicts yes, how often is it correct?

- total positive / predicted yes

Cohen's Kappa: This is essentially a measure of how well the classifier performed as compared to how well it would have performed simply by chance. In other words, a model will have a high Kappa score if there is a big difference between the accuracy and the null error rate that is, the possibility of being wrong each time the model predicts a yes. This analysis was carried out to access the performance of the model and the possibility of having the same or better result if the classification was carried out by chance.

$$K=\frac{2(TP*TN-FN*FP)}{(TP+FP)*(FP+TN)+(TP+FN)*FN+TN)}$$ source: www.dataschool.io

Where TP is the true positive, FP is the false negative, TN is the true negative and FN is the false negative.

Table 1: Confusion matrix binary classifier

Confusion Matrix 1			
n:61993	**Predicted (Yes)**	**Prediction (No)**	**Total**
Actual: Yes	11694	2143	13837
Actual: No	1762	46394	48156
Total	13456	48537	
Confusion Matrix 2			
n:61567	**Predicted (Yes)**	**Predicted (No)**	**Total**
Actual: Yes	12407	1389	13796
Actual: No	1102	46669	47771
Total	13509	48058	
Confusion Matrix 3			
n:66037	**Predicted (Yes)**	**Predicted (No)**	**Total**
Actual: Yes	11487	2973	14460
Actual: No	1764	49813	51577
Total	13251	52786	
Confusion Matrix combined			
n:77795	**Predicted (Yes)**	**Predicted (No)**	**Total**
Actual: Yes	8729	2666	11395
Actual: No	4523	61877	66400
Total	13252	64543	

Table 2: Confusion matrix indices calculated

Indices	**Result**
Confusion Matrix 1	
Accuracy	0.94
Precision	0.87
Misclassification	0.06
Cohen's Kappa	0.82
Confusion Matrix 2	
Accuracy	0.96
Precision	0.92

Misclassification	0.04
Cohen's kappa	0.88
Confusion Matrix 3	
Accuracy	0.93
Precision	0.87
Misclassification	0.07
Cohen's kappa	0.78
Confusion Matrix Combined	
Accuracy	0.91
Precision	0.66
Misclassification	0.09
Cohen's kappa	0.65

Using the confusion matrix, 4 indices were calculated – Accuracy (correctness of the classifier); misclassification rate (how wrong our model was); precision (how correct it is when it predicts a yes); and the most important the Cohen's Kappa (how well the model performed compared to how it would perform if it was random).

According to the computed indices, the accuracy for all the matrices was higher (more than 90%) and precision was lower compared to accuracy. In the model, the accuracy and precision are overrepresented because some deforested areas were calculated as beaver-damaged areas. In addition, the misclassification rate was very low (0.04-0.07) which means that not many samples were classified as beavers while they are forest originally and not many areas were classified as forests when they are beaver-damaged. On the other hand, a higher Cohen's Kappa value denotes how well does the classification work on predefined validation areas of the same class types as training data. For the first three confusion matrix, higher kappa is showing that within the boundary of training and validation samples the model works well.

3.2.1 Analysis of the combined confusion matrix

In the very last stage, 3 different ROI (region of interest) polygon layers were merged into one and then classified again to get better results. On the individual confusion matrix, accuracy and precision were high and misclassification rate was low, and Cohen's kappa value was higher as well. But in the combined matrix, the overrepresentation of beaver damaged areas were reduced compared to the previous individual classification results. That is because the manually

drawn forest polygons were combined to give the machine more training samples to calculate signatures with. The accuracy, precision, and Cohen's kappa value are comparatively lower in the combined matrix. Accuracy was 91%, precision was 66% and Cohen's kappa was 0.65. As mentioned earlier, some forest areas were also included as the beaver damaged areas, so accuracy, precision and Cohen's Kappa don't reflect the real word situation well.

3.3 Validation of the model

The confusion matrix overly represented the data. According to the matrix, the model was too accurate and precise, while some forest areas were classified as beaver areas. To validate this result, random polygons were picked to manually check if it was beaver damaged or not. Due to time constraints, 135 random polygons were selected which constituted 1% of the classified areas. After the manual validation using the orthophotos of the beaver damaged areas, 23 out of 135 polygons were truly classified as beaver damaged areas. The rest of the falsely classified areas were mostly deforested areas cut by man as there were clear signs of tractor tyres on the ground, and a few were healthy forest areas as well. According to this validation, the accuracy of the results stands at 17.04%. Although the accuracy is low, it truly represents the condition of the forest in terms of beaver damaged areas. Therefore, although the beaver damaged areas are rare, following this prediction an actual beaver damaged forest area can be found one out of six times instead of guessing randomly. This accuracy can also be increased if more healthy forest areas are added to the training data and iterate the same process again and again. The confusion matrix we got was overestimated which does not truly represent the condition of Estonian forest. So visual validation was done to get the actual accuracy. Therefore, it can be said that, in order to ensure the accuracy of the model, it is essential to conduct independent validation beside the model. This will increase the reliability of the prediction by reducing the possibility of biasness.

3.4 Discussion and Conclusion

According to the study, forest areas close to waterbody are the most vulnerable areas prone to beaver flooding. It was observed that, in some locations that were marked as active beaver locations based on the orthophotos and the hunter beaver points, beavers are no longer active. It was also clearly visible from the orthophotos that, some damaged forest areas are rejuvenating after the beaver's colony has stopped their activities on them. On the other hand, some newly damaged areas were also found to be colonized by beavers recently. Beaver

activities in Estonia are still ongoing and it is important to control these activities to preserve the forests from damage. Also, beavers are not entirely responsible for forest damages in some areas. Some slightly damaged areas were observed to be damaged by humans as seen from the orthophotos of the clear-cut forests in some locations in the study area.

In this study, the final model prediction matrix was closer to the reality of the nature of the forests with the addition of more validation samples. Therefore, a slight improvement of the model with more manual addition of healthy forest polygons would reduce the overrepresented beaver areas and predict the beaver-induced flooding more efficaciously.

In conclusion, beavers in Estonia cause notable damage to the forest ecosystem by creating dams and flooding the surrounding forests. Also, these animals damage large trees which have economic importance. Therefore, measuring beaver-damaged forest areas is crucial to understand the vulnerable sites and undertake necessary action against beaver colonization. In this study, possible beaver locations and active beaver spots were identified using the prediction model in the study area. Beaver-damaged areas have a similarity with the deforested forest, albeit these areas are distinguishable. Due to time limitations, the study was carried out with a limited number of iterations. As a result, the classified model sometimes overrepresented the beaver areas and mixed them up with deforested forests. But it was observed that with several iteration overrepresented areas would decrease and give a true picture of beaver-damaged flooding.

3.5 Key challenges of this study:

The points locations that were collected from the hunters were not truly representing the beaver presence in the forest. These need to be updated to generate good training areas of beaver-induced flooded areas. Also, the unfamiliarity of the researchers on Estonian forest types posed a challenge in mapping healthy forest polygons for training data. In addition, some deforested forest areas close to waterbody raised confusion while generating beaver training polygons, as well as clear-cut forest areas by man. However, the main challenge was running the processing of the workflow. In most of the steps of the workflow, it took a long time to process as 10 m resolution imageries were used with 10 bands and 3 indices together. The most time-consuming step was random forest classification in the SCP plugin which took almost 20 hours to process. After several attempts, the dzetsaka plugin was used as an alternative for random forest classification. Other steps like resampling imageries in SNAP, K-means clustering, generation ROI for training, etc. also took a substantial amount of time, which limited the general processing time for analysis and validation of the generated model. Some multidimensional

errors after the long period of processing also affected the results and a re-processing was carried out for each step. A computer with good processing speed would have produced better results as more trial-and-error would have been carried out for improved results. In addition, the ground-truth exercise would help to validate the actual beaver damaged area and thus improve the result of the classification by adding or subtracting areas based on the actual condition.

References

Ceccato, P., et al. "Detecting Vegetation Leaf Water Content Using Reflectance in the Optical Domain." *Remote Sensing of Environment* 77 (2001): 22-33.

Gao, B.-C. 1996. NDWI - A normalized difference water index for remote sensing of vegetation liquid water from space. Remote Sensing of Environment 58: 257-266.

Tape, D. K; Benjamin, M.J; Christopher, D.A; Ingmar. N; and Guido, G. (2018) Tundra be Dammed: Beaver colonization in the Arctic. Global Change Biology DOI: 10.1111/gcb.14332

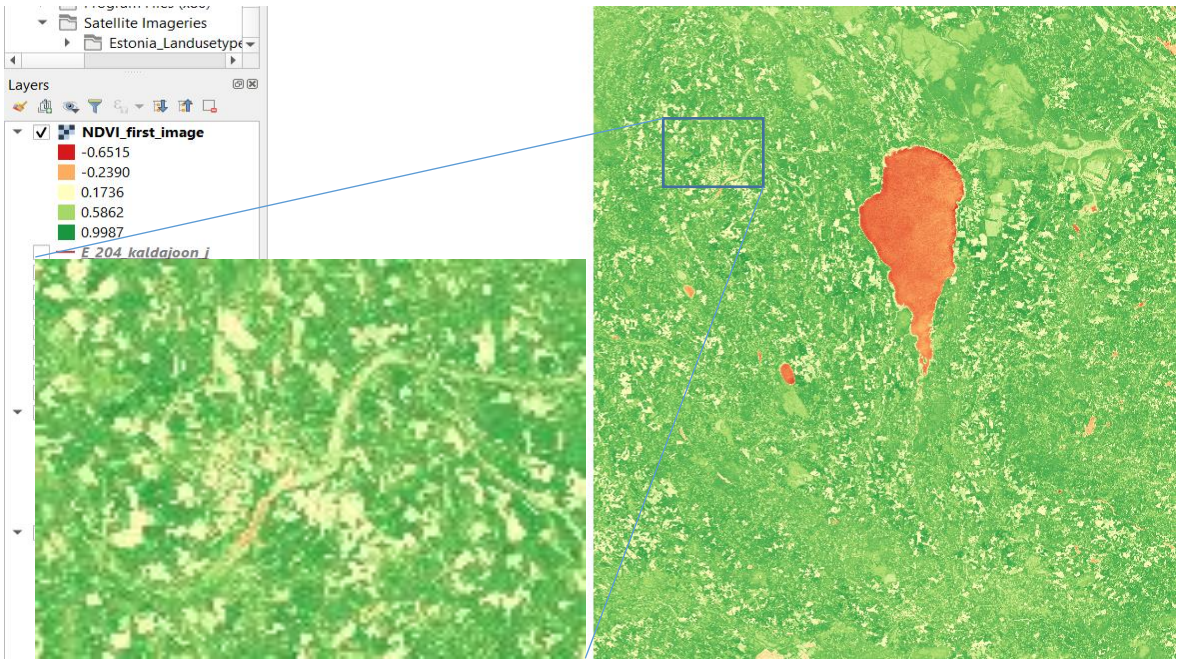
Nick, P. (2017) Beaver the Disruptor: Tolerating Disorder in the heart of Vancouver. Lanscapes/Paysages Summer2017, Vol. 19 Issue 2, p51-53p

Imad, D Github blog-post 2017 – K-means clustering
blog-posts/Kmeans-Clustering.ipynb at master · ImadDabbura/blog-posts · GitHub

Confusion matrix formula Simple guide to confusion matrix terminology (dataschool.io)

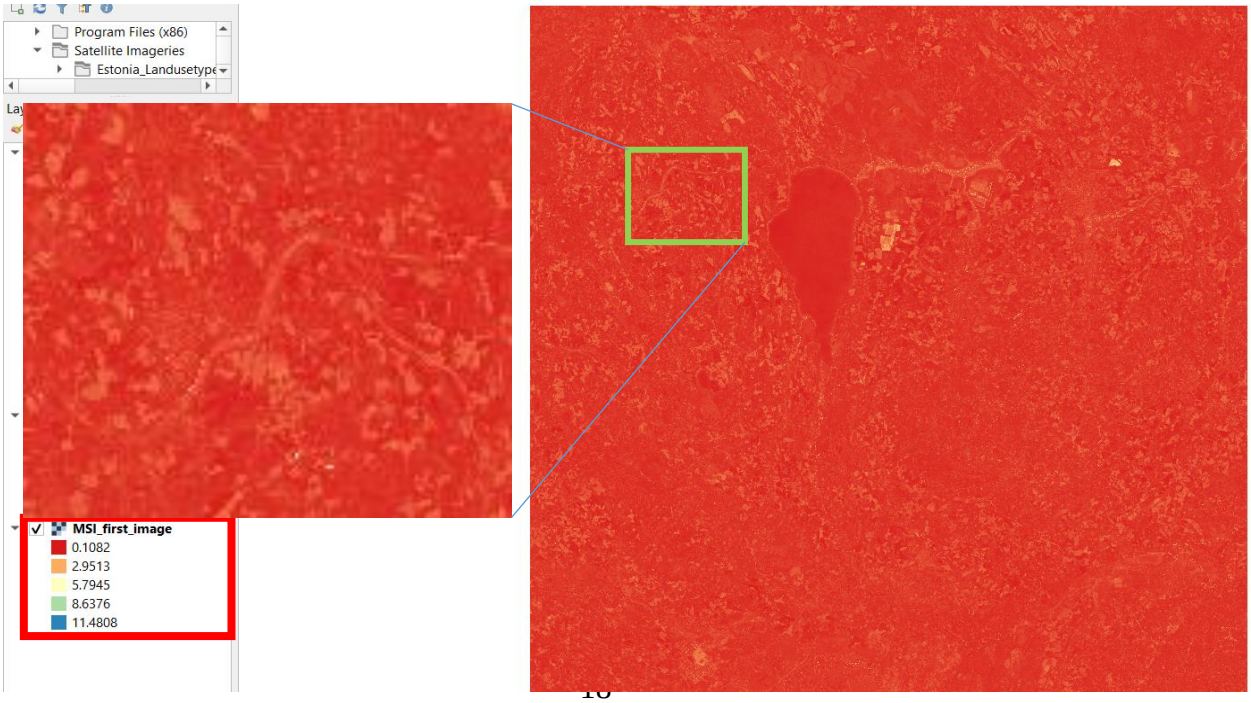
Annexes

Annex 1 Figure 1



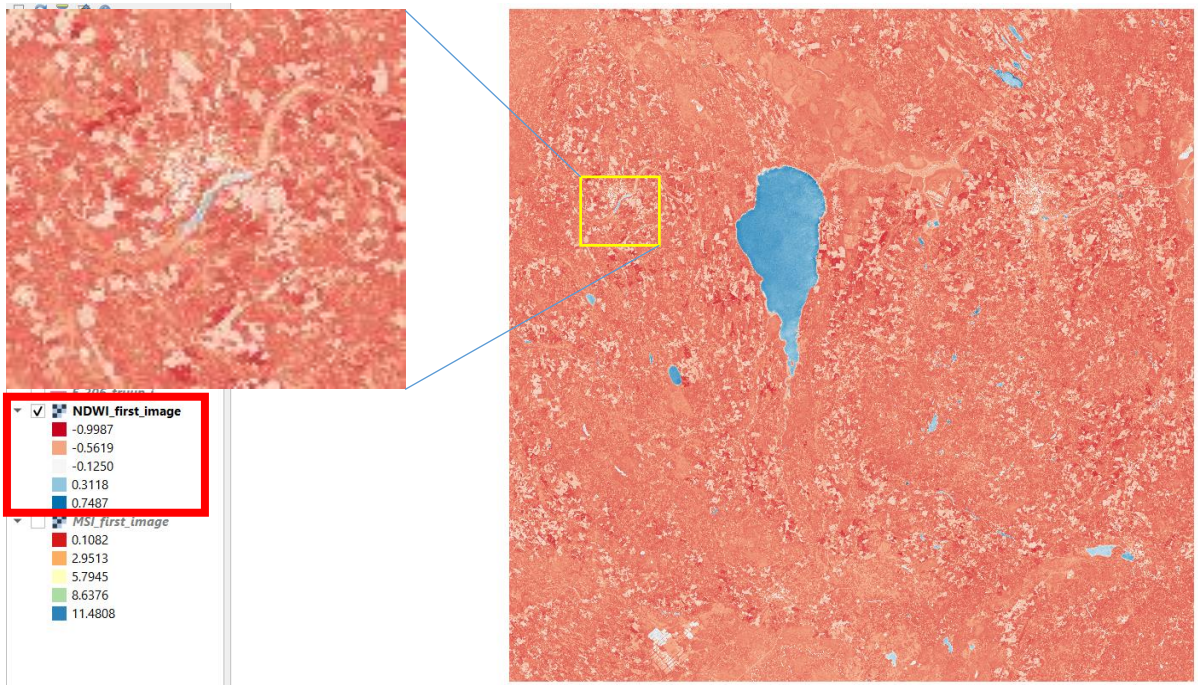
Annex 1 Figure 1 – NDVI calculated for one spectral image. The zoom layer shows the absence of forest cover around water bodies as explained in the text

Annex 1 Figure 2



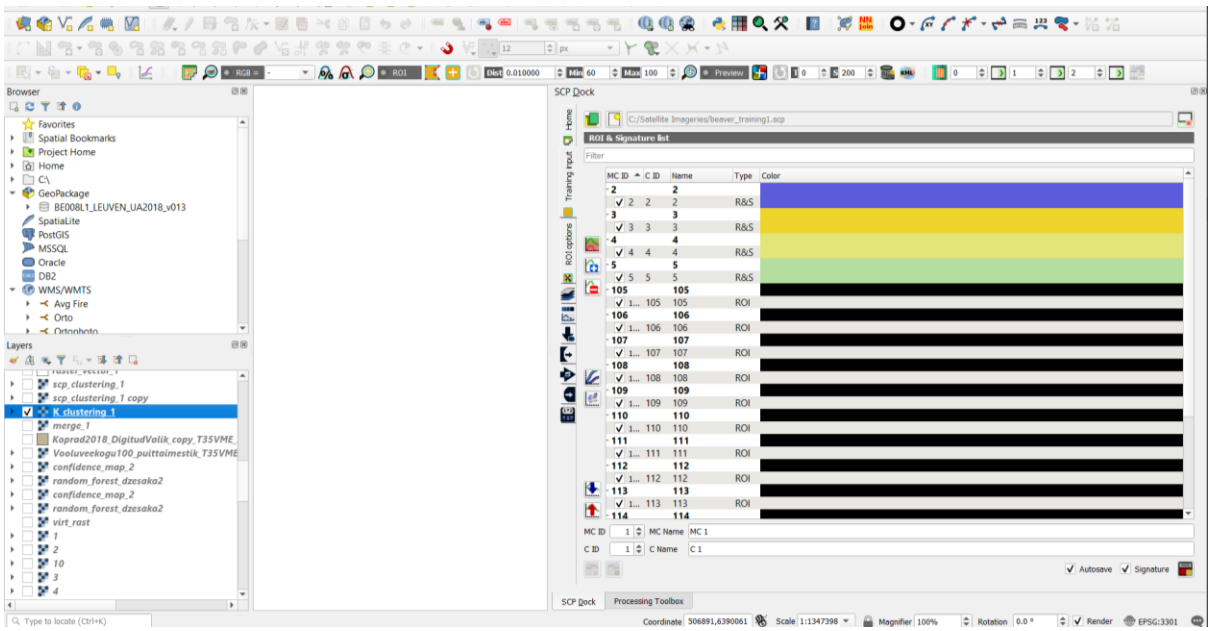
Annex 1 Figure 2 – MSI calculated for one spectral image. The zoom layer shows water stress of some vegetations around flowing waterbodies

Annex 1 Figure 3



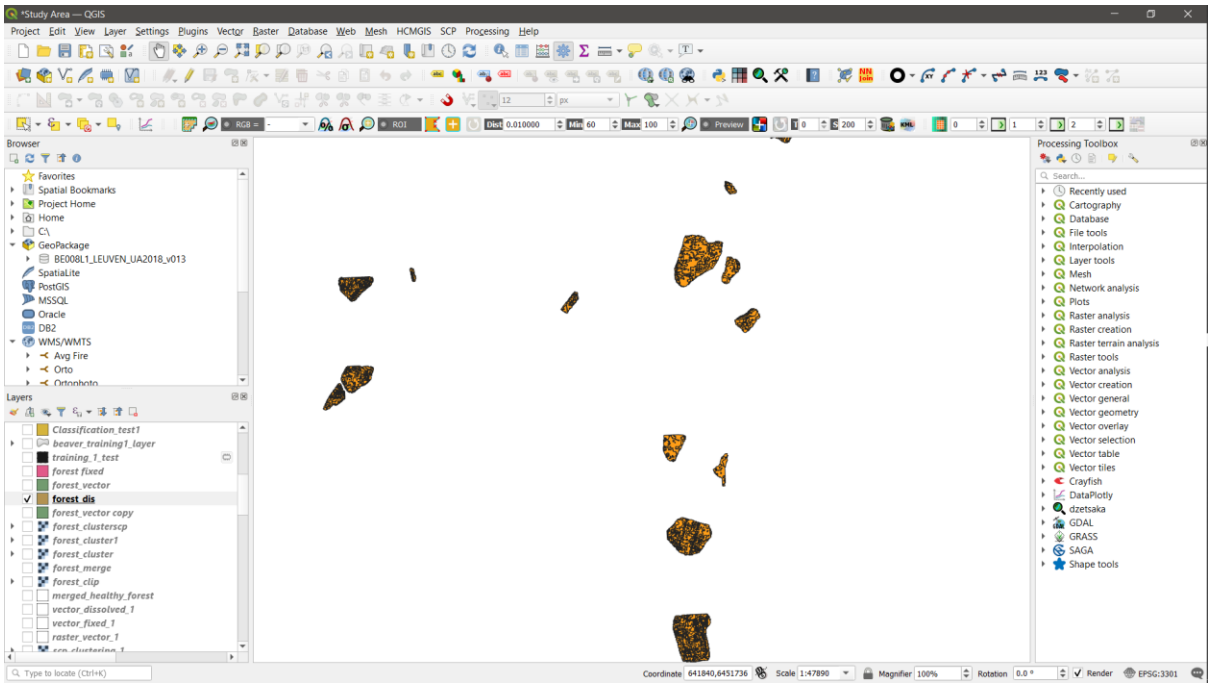
Annex 1 Figure 3 – NDWI calculated for one spectral image. The zoom layer shows the absence of water in some vegetation and the presence of water in the soil in some areas around rivers

Annex 1 Figure 4



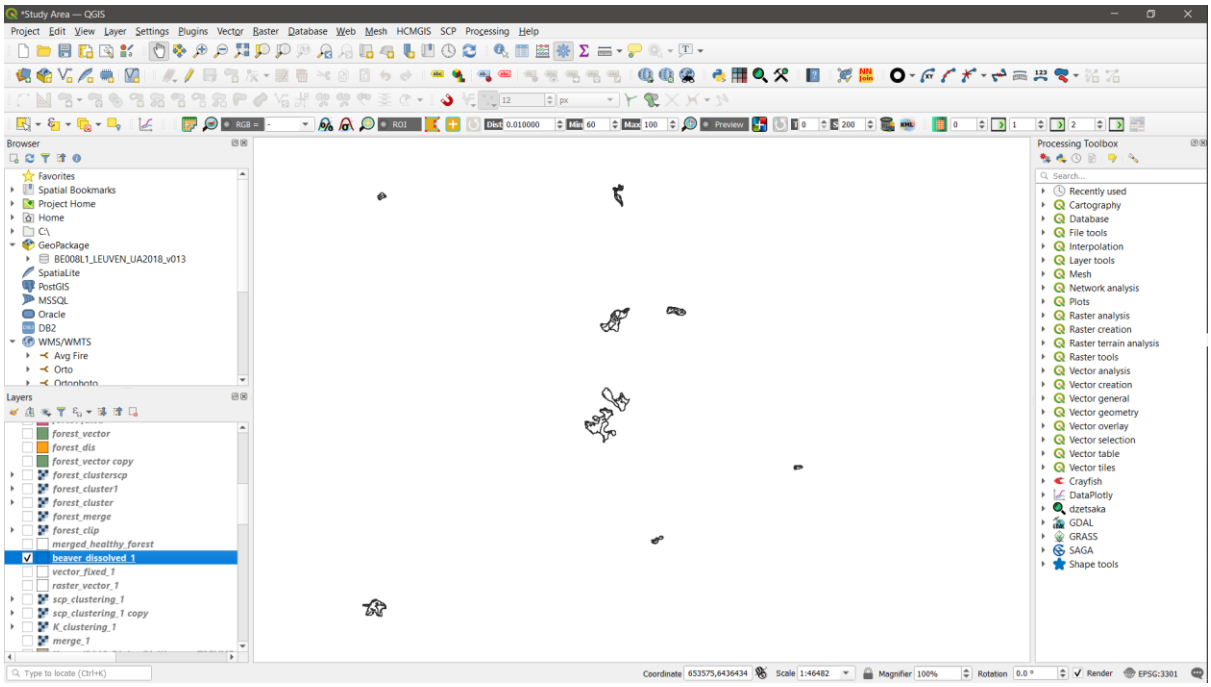
Annex 1 Figure 4 – Clustering results for beaver and forest polygons into classes. The first 5 classes represent beaver and the next 20 classes represent forest polygons

Annex 1 Figure 5



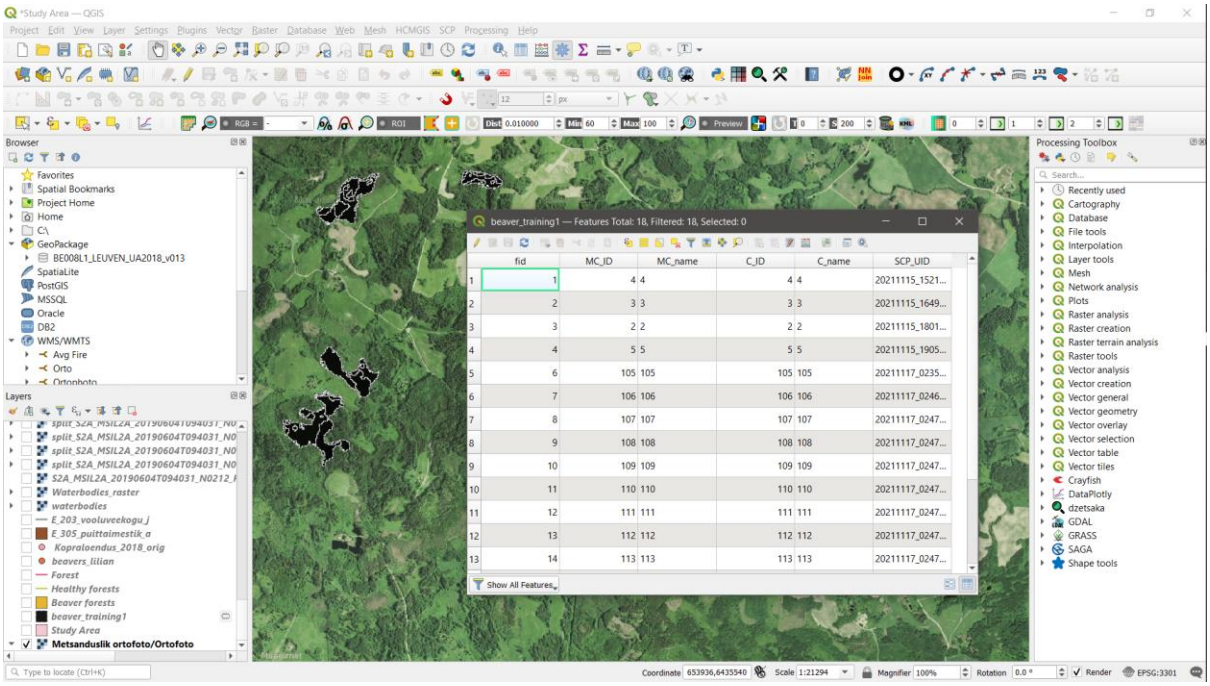
Annex 1 Figure 5 – Vectorising and dissolving results for forest sample polygons

Annex 1 Figure 6



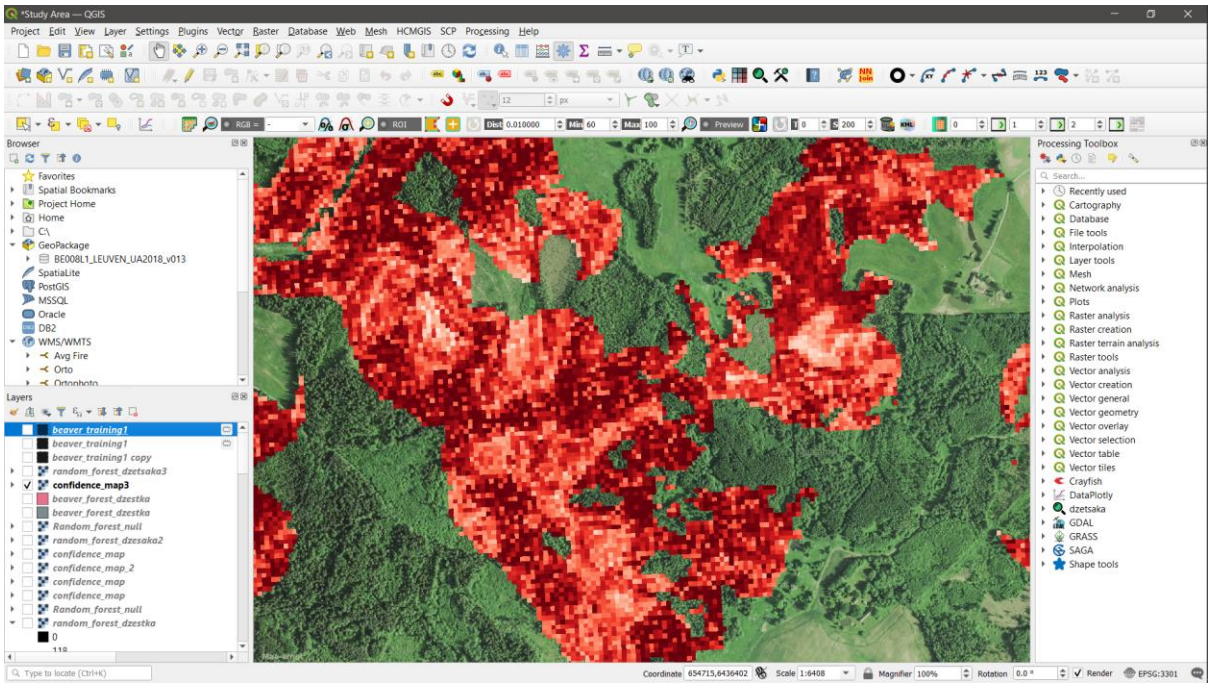
Annex 1 Figure 6 – Vectorising and dissolving results for beaver-damaged area sample polygons

Annex 1 Figure 7



Annex 1 Figure 7 – Classification of forest and beaver polygons with unique ID to prepare the samples as input for random forest classification

Annex 1 Figure 8



Annex 1 Figure 8 – Confidence map of forest classification. Darker areas show more confidence in the model in the classification of that area as a healthy forest.