



Alfred-Wegener-Institut
Helmholtz-Zentrum für Polar- und Meeresforschung
and
Berliner Hochschule für Technik
Fachbereich III Umweltingenieurwesen - Bau

Classification of a high latitude bog using multispectral drone imagery

Bachelor Thesis

By

Lion Golde

890686

Supervisor: Prof. Dr. Ronny Schomacker

Submission date: 28.09.2023

Abstract

Peatlands store and emit large amounts of greenhouse gases. With the climate changing due to global warming, measuring these emissions helps to get a better understanding of the role of peatlands in the global carbon cycle. Measurements at a bog site of the Siikaneva peatland show that the emissions vary along the different microtopographies shaped by their vegetation and ground water level. To upscale these measurements, a supervised classification of the study area was implemented in this study by testing a method that uses high-resolution multispectral aerial imagery, captured by a UAV (Uncrewed Aerial Vehicle), and a Random Forest classifier. A cohesive orthomosaic of the study area was produced, training data were generated to adjust the Random Forest model, and the study area was classified. The results show that the applied methods were successful in generating a multispectral orthomosaic as well as a classified raster of the study area. A mean classification accuracy of 75.7 % was achieved, which can be considered as a good result. Misclassification rates of neighboring microtopographies with similar vegetation could be mitigated by utilizing a LiDAR (Light Detection and Ranging) sensor in further studies.

Table of Contents

Abstract	i
List of Figures	iii
List of Tables.....	iii
Index of software	v
1. Introduction	1
2. Data & Methods.....	3
2.1 Study Site.....	3
2.2 Field Methods.....	6
2.3 Processing Methods	7
2.3.1 Image Processing	7
2.3.2 Classification	10
3. Results.....	15
3.1 Orthomosaic Raster.....	15
3.3 Classified Raster	22
4. Discussion.....	26
4.1 Image Processing.....	26
4.2 Classification.....	26
5. Outlook.....	27
6. Conclusion.....	28
7. References.....	29

List of Figures

Figure 1: Study site location.....	3
Figure 2: A close-range view of the study area with outlined vegetation classes: 1 = high hummock (HHU) ridge with 1–6-m high pine trees; 2 = hummock (HU); 3 = high lawn (HL) with reddish Sphagnum rubellum; 4 = lawn (L); 5 = hollow (HO); 6 = mud-bottom hollow (MB); 7 = water (W). Figure adapted from Korpela et al., 2020.	4
Figure 3: Operating the UAV during fieldwork.....	6
Figure 4: True color image	9
Figure 5: Near infrared false color image.....	10
Figure 6: Training Points.....	12
Figure 7: Clip - standard deviation map near infrared	13
Figure 8: Clip - ratio map near infrared.....	14
Figure 9: Example - calibrated images flight 1509_108	17
Figure 10: Example - reflectance maps flight 1509_108.....	18
Figure 11: Orthomosaic of study area	19
Figure 12: Mean accuracies with standard deviation and best number of trees	20
Figure 13: Classified raster	22
Figure 14: Clip - classified raster	23
Figure 15: Correlation matrix - actual class on the x-axis, prediction on the y-axis	25

List of Tables

Table 1: Microtopography types at the Siikaneva bog site. Table adapted from Korpela et al., 2020	5
Table 2: Band combinations	9
Table 3: Trained pixels per class	11
Table 4: Image processing performance in Pix4D.....	16
Table 5: Mean decrease in Gini of each input layer	21
Table 6: Area size of land cover classes.....	24

Index of units

mm = millimeter

cm = centimeter

m = meter

ha = hectare

km = kilometer

km² = square kilometer

km/h = kilometers per hour

mb = megabyte

gb = gigabyte

Ghz = Gigahertz

°C = degrees Celsius

% = percent

Index of software

Software	Version Number
Pix4DMapper - Educational	4.6.4
RStudio	2022.07.2 Build 576
QGIS	3.28.0-Firenze
Opanda IExif	2.3

RStudio Package	Version Number
sf	1.0-13
sp	1.6-0
terra	1.7-23
raster	3.6-20
randomForest	4.7-1.1
caret	6.0-94
progress	1.2.2
ggplot2	3.4.2
pheatmap	1.0.12
viridis	0.6.4

1. Introduction

Peatlands are a distinctive wetland type and it is estimated that they store about half as much carbon as present in the Earth's atmosphere in form of CO² (carbon dioxide) (Gorham, 1991; Turunen et al., 2002). Thus, they play a significant role in the global carbon cycle and contribute to the atmosphere's balance of greenhouse gases (Yu, 2011). Peatlands are characterized by high water levels and the accumulation of partially decayed organic matter, forming layers of peat. Approximately 87 % of peatlands are concentrated in the boreal and subarctic regions of North America, Russia, and Fennoscandia (Joosten and Clarke, 2002). Peatlands accumulate organic matter faster than it decomposes due to their waterlogged conditions and nutrient-poor substrate. Consequently, surplus organic matter transforms into peat, forming layers that can reach several meters of depth (Rydin and Jeglum 2013). Since they absorb more CO² than they release, peatlands act as persistent carbon sinks (Yu, 2012). Peatlands can be classified into two primary types: minerotrophic fens, which receive water and nutrients from surrounding runoff and groundwater, and ombrotrophic bogs, relying solely on atmospheric deposition for water and nutrients (Vitt, 2006).

As the effects of global warming increase, these peatlands have gained significant attention due to their potential to influence greenhouse gas emissions into the atmosphere (Bu et al., 2011). Thus, monitoring the greenhouse gas emissions from peatlands is of great importance to better understand the interplay between these ecosystems and the changing climate. The research in this study takes the Siikaneva bog in southern Finland as an example. Chamber systems are being used at the Siikaneva bog to regularly measure its CO² and CH⁴ (Methane) emissions. There are measuring plots for each microtopography. These microtopographies exhibit along a water-level gradient from dry hummock patches to wet water pools and are shaped by their different plant communities. Due to their difference in vegetation and water-level they also emit different quantities of greenhouse gases (Korrensalo et al., 2018).

It is necessary to learn about the spatial distribution of these microtopographies in order to produce an accurate extrapolation of the emission measurements. This supervised classification of the Siikaneva Bog's land cover was conducted by using high-resolution aerial UAV (Uncrewed Aerial Vehicle) imagery and a Random Forest classifier. The objectives of this study were to capture aerial images at Siikaneva using a UAV, create a cohesive orthomosaic of the mapped area, adjust the Random Forest model suitable for the data, and generate a classified raster with the land cover classes and their area size. Finding out if this approach fits the cause of more advanced greenhouse gas emission monitoring at boreal bogs is very helpful for further research.

2. Data & Methods

2.1 Study Site

Siikaneva is a 12 km² boreal peatland in Ruovesi, southern Finland, primarily shaped by minerotrophic fen areas. However, the study site is an ombrotrophic bog within Siikaneva. It is located at 61°50'N, 24°12'E in the southern boreal vegetation zone (**Figure 1**) and surrounded by coniferous forest (Ahti et al., 1968). According to the Juupajoki-Hyytiälä weather station (10 km from the site), the mean annual temperature is 4.2°C. Mean January and July temperatures are -7.2°C and 17.1°C, respectively. The annual rainfall is 707 mm, with one third precipitating as snow (Korrensalo et al., 2020).

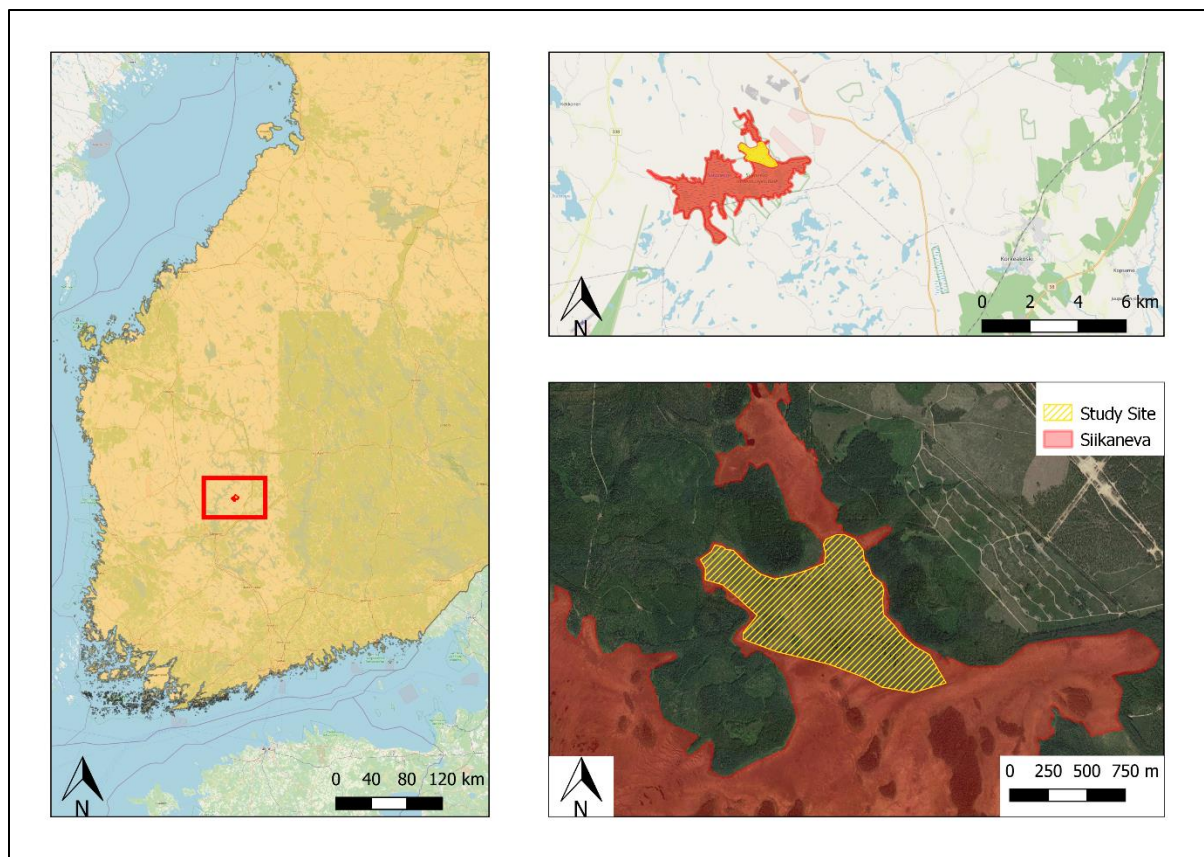
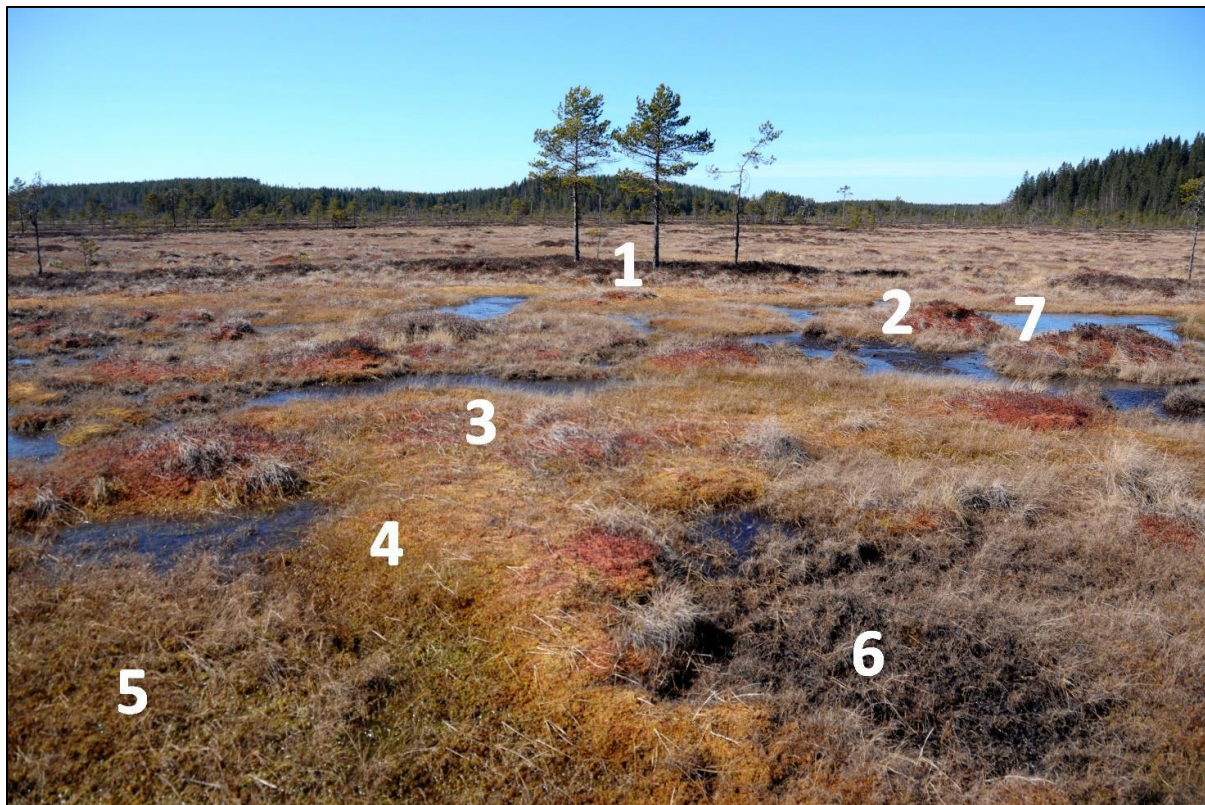


Figure 1: Study site location

Based on the large variety in water table levels, the bog was disaggregated into six vegetated microform classes, high hummock (HHU), hummock (HU), high lawn (HL), lawn (L), hollow (HO), mud bottom (MB), and open water (W)(**Figure 2**). Their species composition varies along the water table gradient from dry hummocks to wet water pools, in the listed order. All microtopographies, except mud bottom, are dominated by Sphagnum mosses (**Table 1**) (Hayward and Clymo, 1982; Rydin, 1993).



*Figure 2: A close-range view of the study area with outlined vegetation classes: 1 = high hummock (HHU) ridge with 1–6-m high pine trees; 2 = hummock (HU); 3 = high lawn (HL) with reddish *Sphagnum rubellum*; 4 = lawn (L); 5 = hollow (HO); 6 = mud-bottom hollow (MB); 7 = water (W). Figure adapted from Korpela et al., 2020.*

For the vegetated microforms, eighteen permanent sample plots were established to represent the different plant community types (Korrensalo et al., 2020). These enable regular greenhouse gas retention and emission measurements for the different microtopographies.

Surface class	Description
High hummock, HHU	High cover of dwarf shrubs (<i>Empetrum nigrum</i> L., <i>Calluna vulgaris</i> (L.) Hull, and <i>Betula nana</i> L.). Bottom layer is dominated by <i>Sphagnum fuscum</i> (Schimp.) H. Klinggr.
Hummock, HU	No shrubs, except for <i>Andromeda polifolia</i> L., which may constitute a significant percentage of the field layer. <i>S. fuscum</i> covers more than 10 % of the bottom layer.
High Lawn, HL	Field layer consists of <i>A. polifolia</i> and <i>Eriophorum vaginatum</i> L. Coverage of <i>S. fuscum</i> is less than 10 % and the dominant <i>Sphagnum</i> species is <i>S. rubellum</i> Wils.
Lawn, L	Field layer may be missing, or scarcely covered by <i>Rhynchospora alba</i> (L.) Vahl., <i>Scheuchzeria palustris</i> L., <i>A. polifolia</i> , <i>E. vaginatum</i> or <i>Carex limosa</i> L. In the bottom layer, the dominant <i>Sphagnum</i> species are <i>S. papillosum</i> Lindb., <i>S. magellanicum</i> Brid. and <i>S. balticum</i> Russ.
Hollow, HO	Field layer may be missing, or has <i>R. alba</i> , <i>S. palustris</i> or <i>C. limosa</i> . Dominant <i>Sphagnum</i> species are <i>S. majus</i> (Russow) C.E.O. Jensen and <i>S. cuspidatum</i> Ehrh. ex Hoffm.
Mud-bottom, MB	Field layer may be missing, or scarcely covered by <i>R. alba</i> , <i>S. palustris</i> or <i>C. limosa</i> . Most of the bottom layer surface is covered by bare peat.
Water, W	Open water surface without ground layer vegetation (pool). A few <i>S. palustris</i> shoots may be found.

Table 1: Microtopography types at the Siikaneva bog site. Table adapted from Korpela et al., 2020

2.2 Field Methods

During a field trip in September 2022, aerial images of the Siikaneva peatland were collected, under mostly overcast sky conditions and low to medium winds. A DJI Phantom 4 Multispectral UAV was used, carrying sensors covering different spectral ranges: four 32 nm wide bands centered in the blue (450 nm), green (560 nm), red (650 nm), and red edge (730 nm) ranges, as well as a 52 nm wide band centered in the near infrared (840 nm) range. Additionally, a sixth camera captured true-color imagery through an RGB-sensor (Red, Green, Blue). The integrated spectral sunlight sensor on top of the UAV captured solar irradiance. With a weight of 1.487 g the Phantom 4 Multispectral falls into the category C2 and requires an operating license of the type A3. The area was checked for NOTAMs (Notice To Air Missions), possible flight restrictions for example due to the nearby airfield or military area.



Figure 3: Operating the UAV during fieldwork

Since there are no clear borders between the bog and fen areas, several AOIs (Areas Of Interest) were defined and prioritized by their distance to the bog site around the measuring plots. Then, flight routes equal to the AOIs were generated in the DJI Pro software on an iPad attached to the RC (Radio Control) the UAV was operated with.

Besides some scattered boardwalks, the bog was rather inaccessible, and it was not feasible to distribute physical ground control points throughout the survey areas. A precise geolocation error correction was therefore not possible during data post-processing.

To fulfill legal regulations requiring the UAV to always be within the operator's visible range, four different takeoff points across the AOIs were determined. The UAV flew the pre-programmed routes at an altitude of 100 m above ground level at 31 km/h speed on average. The camera took a picture every two seconds with a spatial resolution of about 5.3 cm/pixel. This generated an in-track overlap of 75 % and an across-track overlap of 80 %. Due to limited battery power, the UAV's power supply needed to be changed after about 25 minutes of flying. This resulted in a total of 83 individual flights. Overall, an area of 5.1 km² was surveyed.

Additionally, ground truth images were captured with a DJI Mini 2 UAV for inaccessible areas. These were geolocated with the UAV's internal GPS.

2.3 Processing Methods

2.3.1 Image Processing

To create large orthomosaics from the single aerial UAV images, they were processed with Pix4D. Separately for each flight, the images were processed on a high-end Windows 10 workstation with 512 GB RAM (Random Access Memory) and a 2.00 GHz AMD EPYC 7702P CPU (Central Processing Unit). The processing in Pix4D consists of three steps. 1. "Initial processing", 2. "Point cloud and mesh" and 3. "DSM, orthomosaic and index".

In the first step, the images were calibrated. The Pix4D algorithm identified specific features of each image and extracted them as keypoints. Based on these keypoints, matching images were recognized. Default settings were used, except for the calibration method. With the default calibration method, some results in the blue band came out blurry. These deficiencies were enormously mitigated by choosing the alternative calibration method which is “optimized for aerial nadir images with accurate geolocation, low texture content and relatively flat terrain” (Pix4D online user manual).

As the area is characterized by relatively flat terrain, the second step did not perform very well. Since the point cloud and the resulting digital surface model was not of interest for this study, this step was omitted.

In the third step the single tiles were merged for each multispectral band as well as for the RGB-images. The sun irradiance data were considered in the process. To get the raw-value-images without color correction, the reflectance maps were generated in addition to the default orthomosaics. Because of slight resolution differences between the flights, all scenes were downsampled to 6 cm/pixel. The processing results were exported as GeoTIFFs with a reflectance map of each band for every flight.

After visually assessing the bog area, the single bands of all 12 flights within that area were combined into raster stacks in RStudio. Corrupted pixels with values < 0 and > 1 were masked out. The raster stacks were then merged into one large raster, after being georeferenced manually with the QGIS Georeferencer. Eventually a final area of 0.75 km² was determined, removing remained forest and fen areas on the edges. The final raster was cut and used to generate a true-color (**Figure 4**) and an infrared-false-color (**Figure 5**) version in QGIS (**Table 2**). Therefore, the display value range of each applied band were adjusted manually to achieve an accurate result.

The coordinates of the aerial ground truth images, captured with the mini UAV, were extracted with Opanda IExif.

Channel	True color image	Near infrared false color image
3 (Red)	Red	Near Infrared
2 (Green)	Green	Red
1 (Blue)	Blue	Green

Table 2: Band combinations

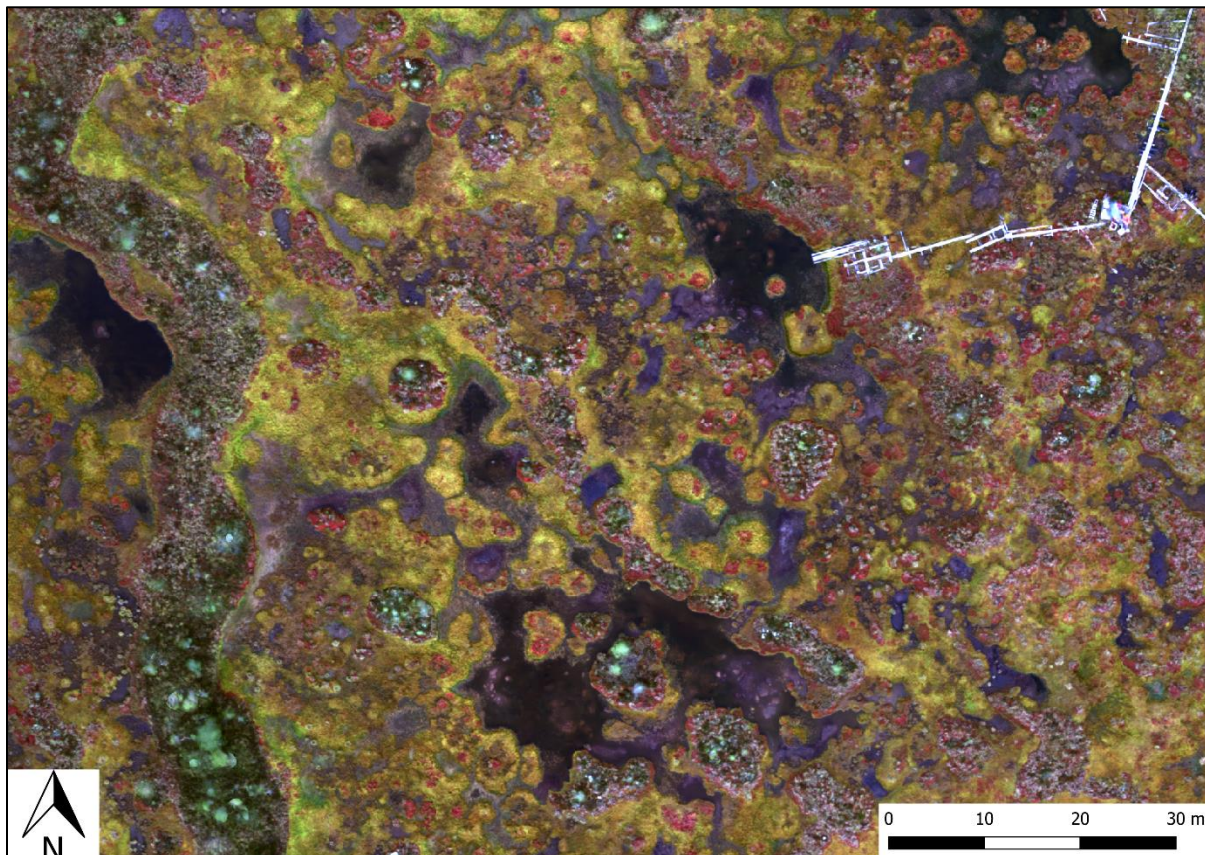


Figure 4: True color image

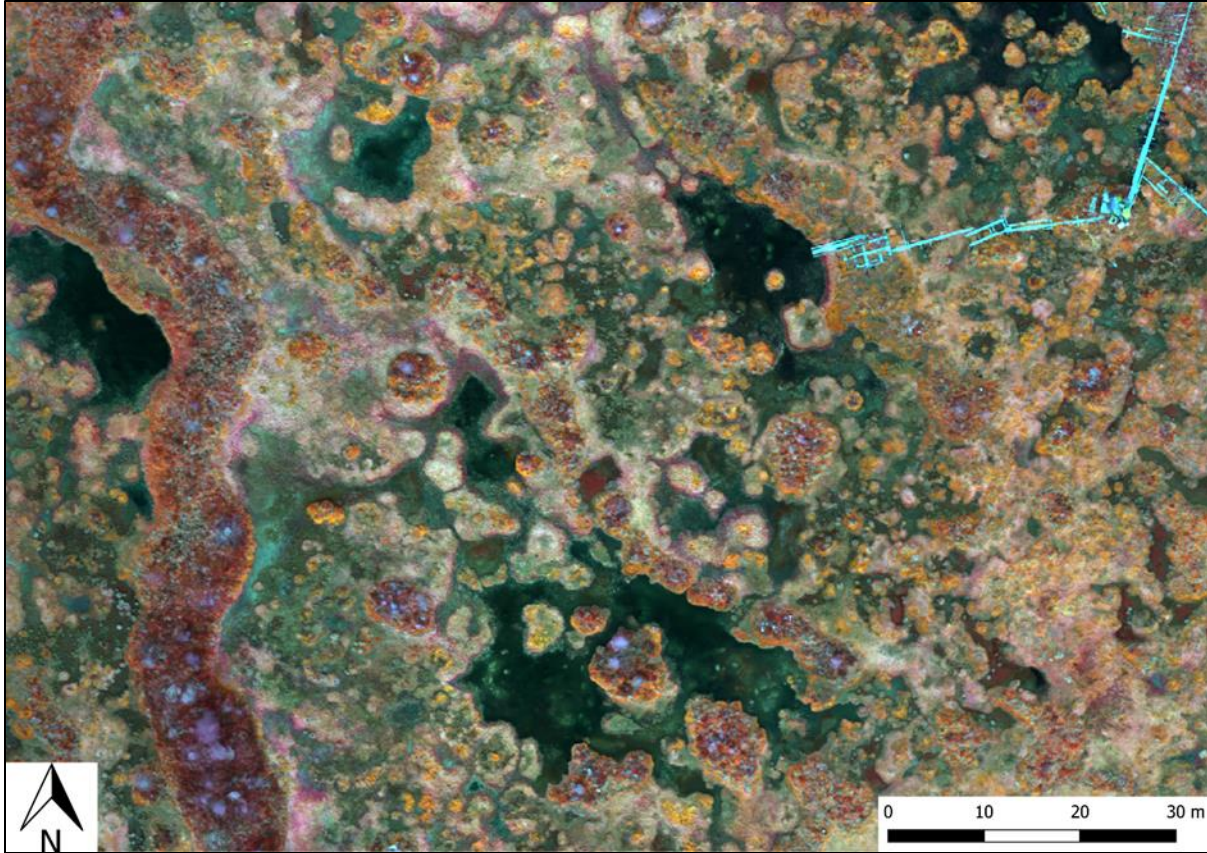


Figure 5: Near infrared false color image

2.3.2 Classification

To find out the area size of each land cover class, representing the different microtopographies as well as boardwalks and mineral soil islands, a supervised classification utilizing a Random Forest model was realized. The Random Forest algorithm is a powerful ensemble learning technique used in machine learning. It requires a set of trained data to learn from and has been successfully used to map land cover classes (Colditz, 2015; Haas and Ban, 2014; Tsutsumida and Comber, 2015). It operates by creating a multitude of decision trees during training, each with a unique subset of the data and features. Thus, the Random Forest classifier is very effective in dealing with complex datasets such as multi-band rasters. Combining the predictions of multiple decision trees reduces overfitting and increases the model's accuracy (Breiman, 1984; Özcan et al., 2021).

The training and validation data for the Random Forest classifier must be statistically independent. The training samples must be balanced and representative of the target classes (Belgiu et al., 2016). Therefore, a set of 500 randomly distributed training points was generated in RStudio (**Figure 6**).

The training was conducted in QGIS by going through the training points and visually determining which class each trained pixel belongs to. For this, the true color image, an infrared false color composition as well as the geolocated ground truth imagery were taken into consideration to make the most accurate decisions. Due to the imbalance of represented classes after training 500 points, additional training points were created. Therefore, distinct areas with high representation of certain classes were defined and then used to generate extra randomized points within these extents (**Figure 6**). The training was then continued until a minimum suitable number of trained pixels per class and a total of 722 trained points was reached (**Table 3**). Since the classes HU and HHU were hard to distinguish during training, they were combined into one class labeled as HU.

Class	BW	HL	HO	HU	L	MB	MSI	OW
Trained pixels	32	149	122	104	75	75	80	85

Table 3: Trained pixels per class

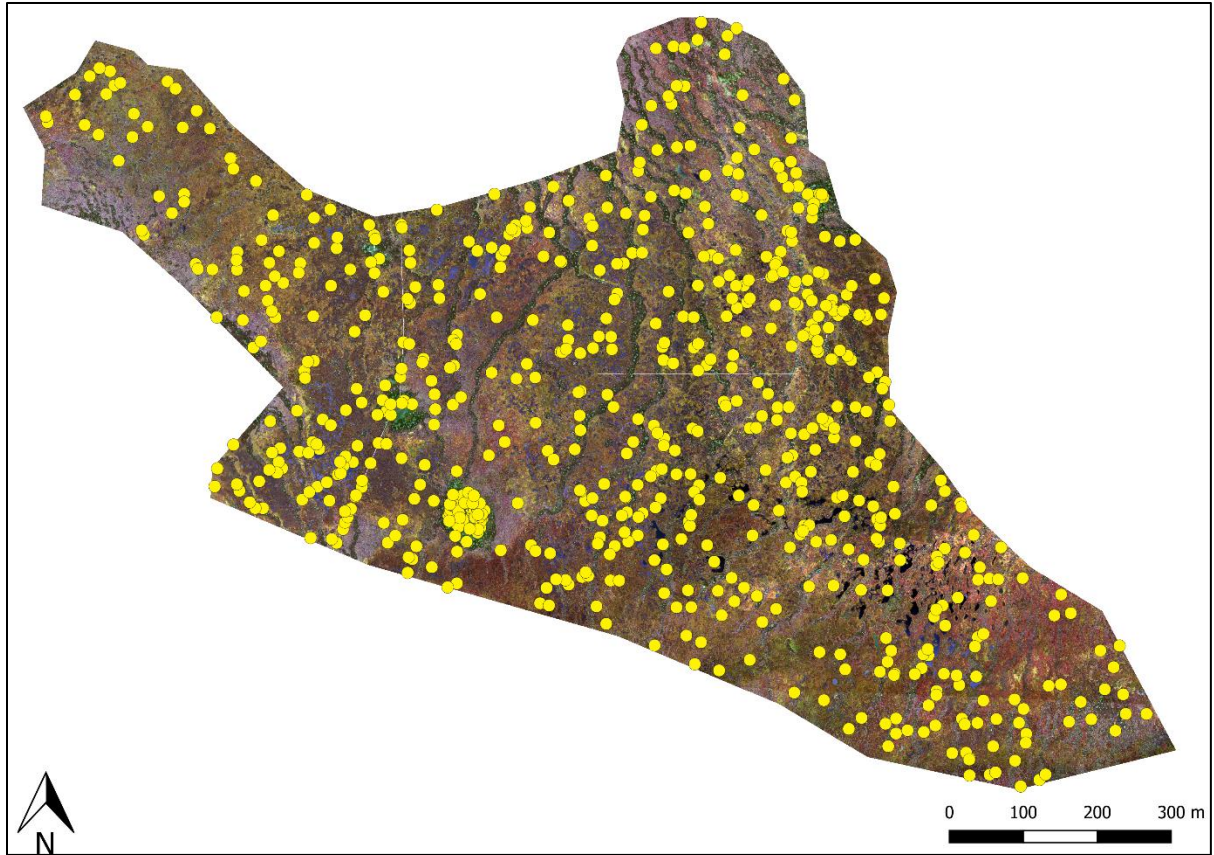


Figure 6: Training Points

To consider the surrounding area of each pixel in the classification, a standard deviation map (Figure 7) was created for each multispectral layer. Additionally, a ratio map (Figure 8) was created to mitigate the effects of different sun irradiance. For the standard deviation maps, the standard deviation of the surrounding 9x9 pixels was calculated for each pixel. The ratio maps were composed by dividing each pixel value by the sum of pixel values from all bands.

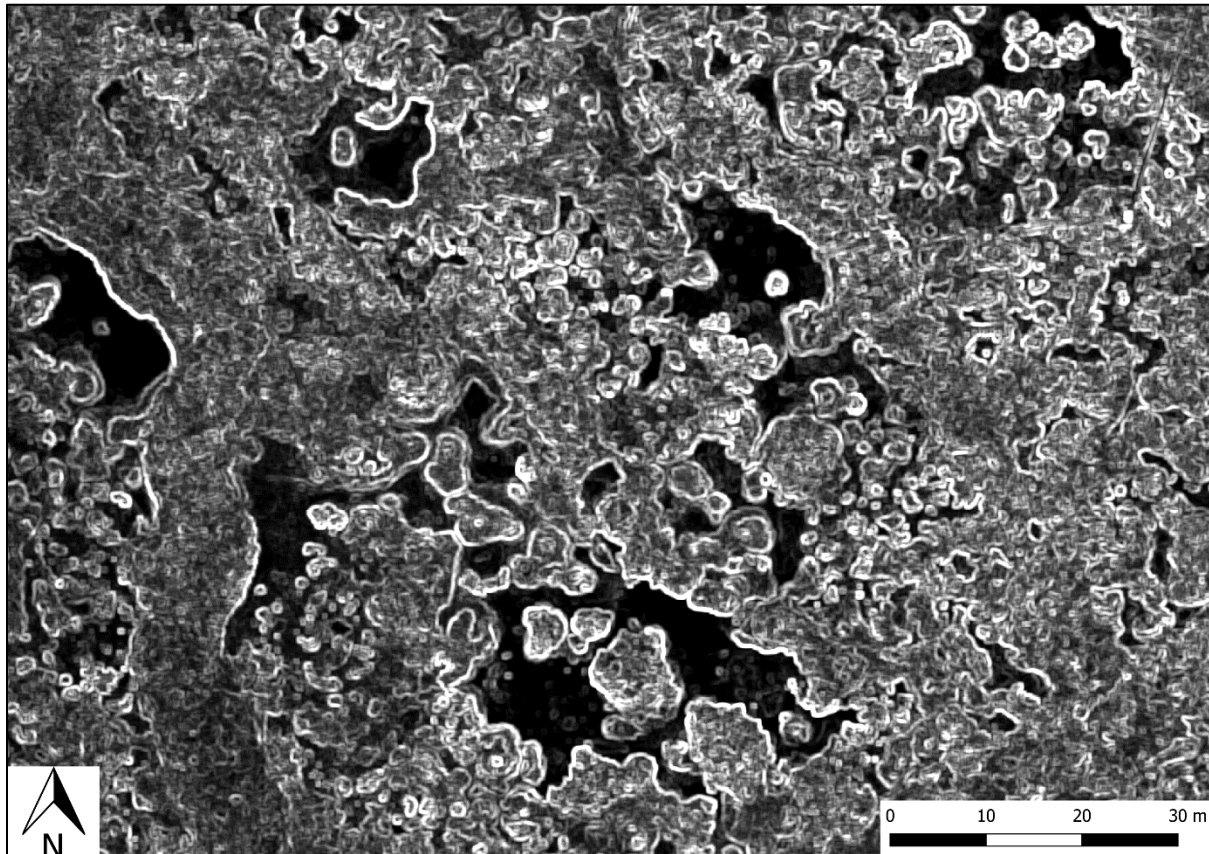


Figure 7: Clip - standard deviation map near infrared

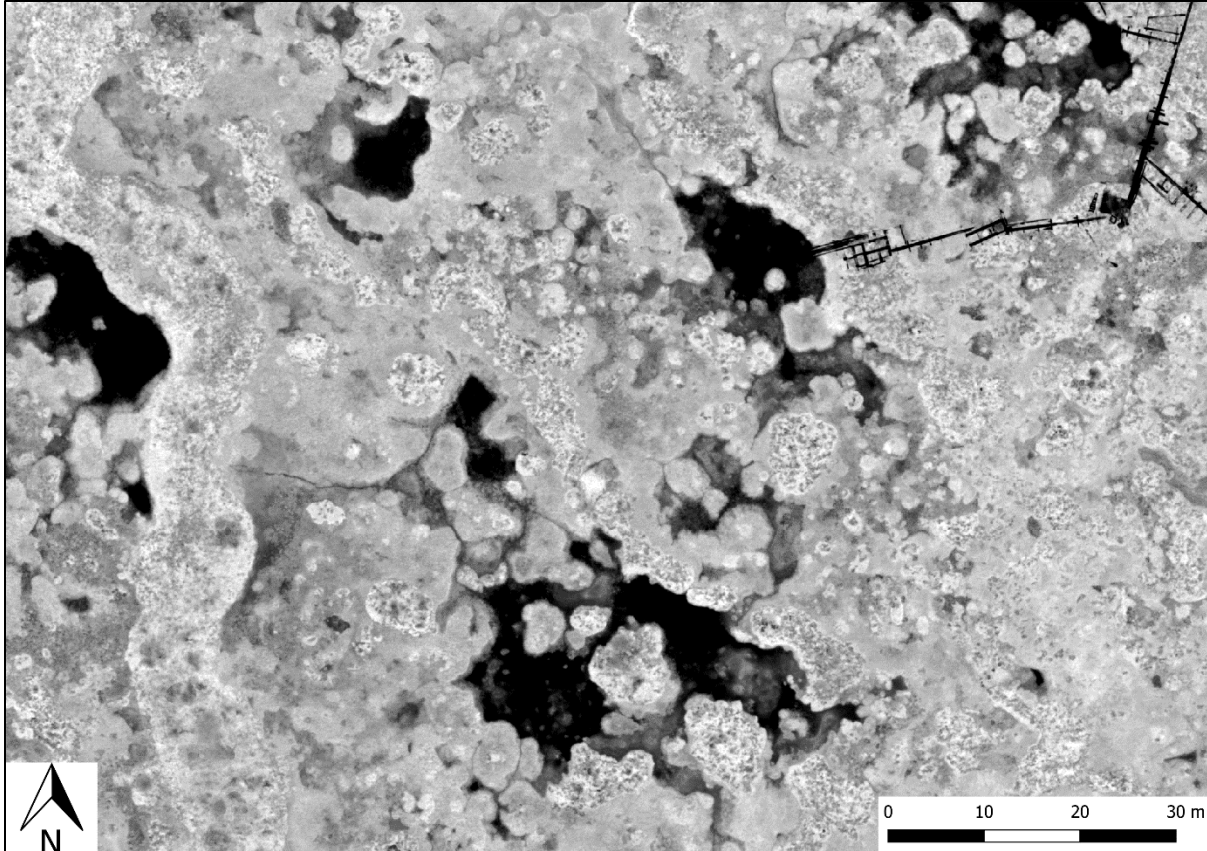


Figure 8: Clip - ratio map near infrared

Then, all 15 raster layers (five bands plus their standard deviation and ratio maps) were combined into one raster stack. From that raster stack and the coordinates of the training points, the training values and their class labels were extracted into a data frame and then split into 70 % training and 30 % test data.

The Random Forest model was created and set to loop different numbers of decision trees, ranging from 5 to 1500 in increments of 5. For every number of decision trees, 100 iterations were looped. This provided a reliable mean accuracy value for each number of trees which then enabled a calculation for the best number of trees. Therefore, the moving average with a window size of 10 was created. This was done to identify the point where the classification performance stabilized, and thus further increasing the number of trees no longer significantly improved the mean accuracy at the set threshold of 0.1 %.

Eventually the Random Forest model was designed for classifying the area of interest using the raster stack of all 15 input layers. To predict the land cover classes, the best number of decision trees and the entire set of trained data points were employed. Also, the mean decrease in Gini was calculated for each input layer. The Gini impurity index reflects the likelihood of achieving minimal misclassification probability (Wang et al., 2020). Last, the classified raster was exported and used to calculate the size of each class area.

3. Results

3.1 Orthomosaic Raster

As outlined in the Processing Methods section, the aerial images of all single flights were merged in Pix4D, resulting in a reflectance map for each multispectral band (**Figure 10**). Images of the bog area were calibrated properly, whereas images of the surrounding forest could not be calibrated (**Table 5, Figure 9**). The geolocation of the processed flights was accurate to about 2 meters. Using the Georeferencer in QGIS, the geolocation error was mitigated to a few centimeters. The raster of the 12 combined flights (**Figure 11**) covered an area of 74.64 ha around the bog site of Siikaneva, featuring a resolution of 6 cm/pixel. With a file size of 8.36 gb it was stored as a GeoTIFF which contained all five multispectral layers, as well as its geolocation.

Flight	Median of keypoints per image	Images calibrated	Median of matches per calibrated image
1509_107	10548	1366 out of 1926 (70%)	2108.24
1509_108	10821	2851 out of 3023 (94%)	6293.06
1509_110	11489	3048 out of 3060 (99%)	4550.02
1509_112	11409	1170 out of 1170 (100%)	5427.42
1709_105	11211	2917 out of 2922 (99%)	7075.3
1709_106	11369	2862 out of 2862 (100%)	3888.17
1709_109	11067	1134 out of 1152 (98%)	7638.15
1709_111	10645	2929 out of 3102 (94%)	4036.54
2009_108	10890	2433 out of 2460 (98%)	5493.31
2009_109	10783	2986 out of 3017 (98%)	6976.36
2009_110	10689	2820 out of 2880 (97%)	6377.79
2009_111	10561	1513 out of 1764 (85%)	4557.95

Table 4: Image processing performance in Pix4D

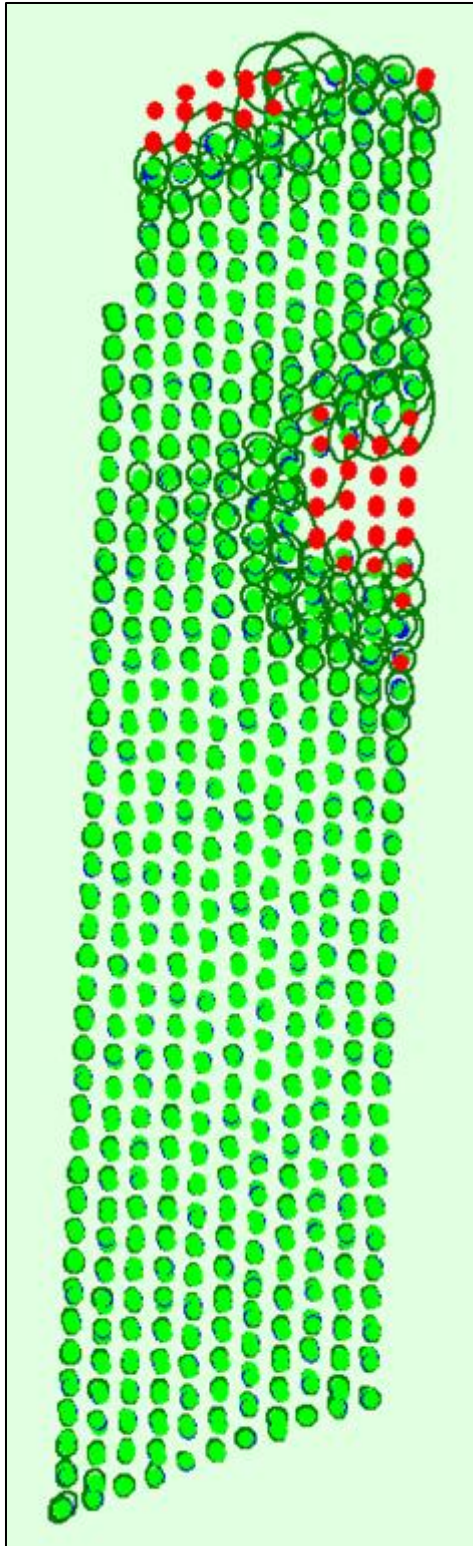


Figure 9: Example - calibrated images flight
1509_108

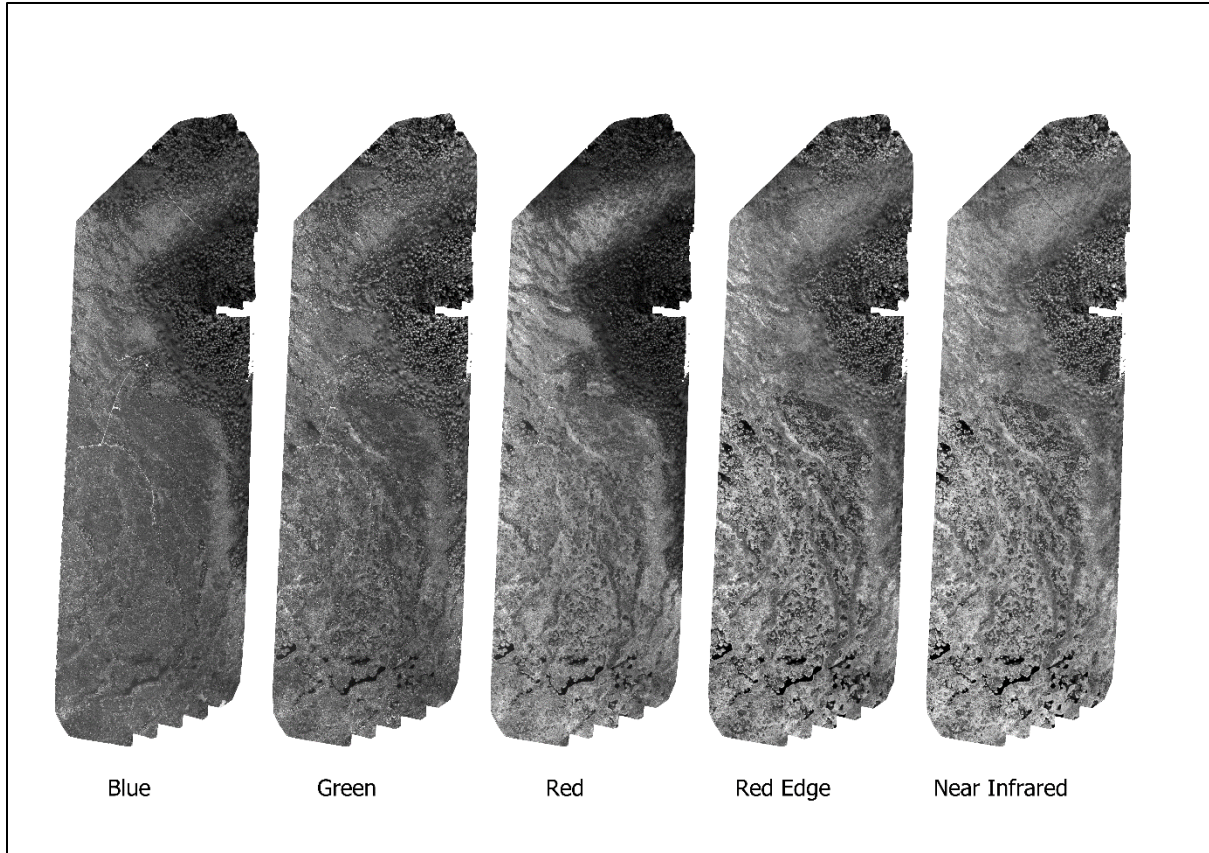


Figure 10: Example - reflectance maps flight 1509_108

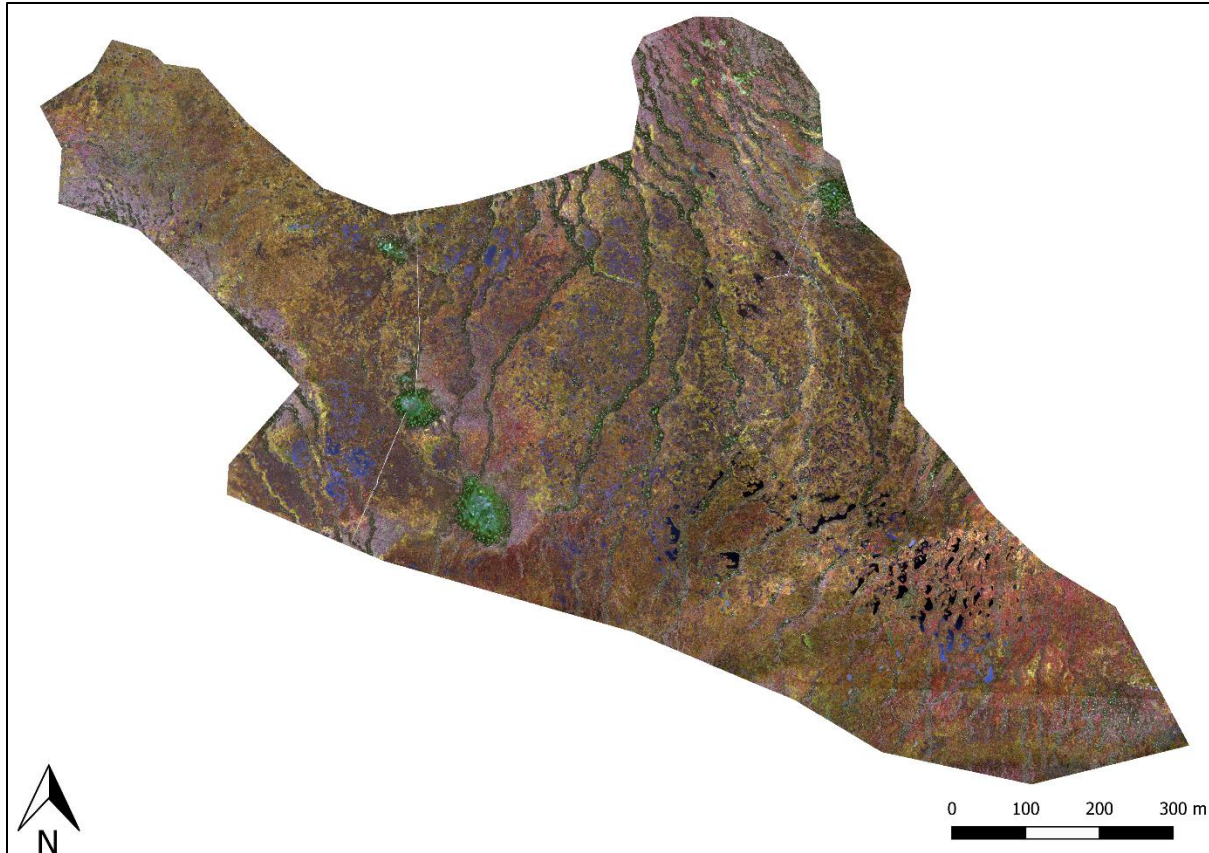


Figure 11: Orthomosaic of study area

3.2 Classification Model

To achieve the best classification results, the Random Forest Model was adjusted to strike the ideal balance between using the fewest trees possible and maintaining a high level of accuracy. Therefore, different numbers of trees were tested, as outlined in the classification section. The best number of trees showed to be 95. More than 95 trees did not improve the mean accuracy and standard deviation of accuracies anymore (**Figure 12**).

The mean decrease in Gini values indicated that the ratio map of the green band was the most important layer in the classification. The standard deviation map of the red edge band was the least important. Overall, the standard deviation maps had the lowest impact on the classification (**Table 6**).

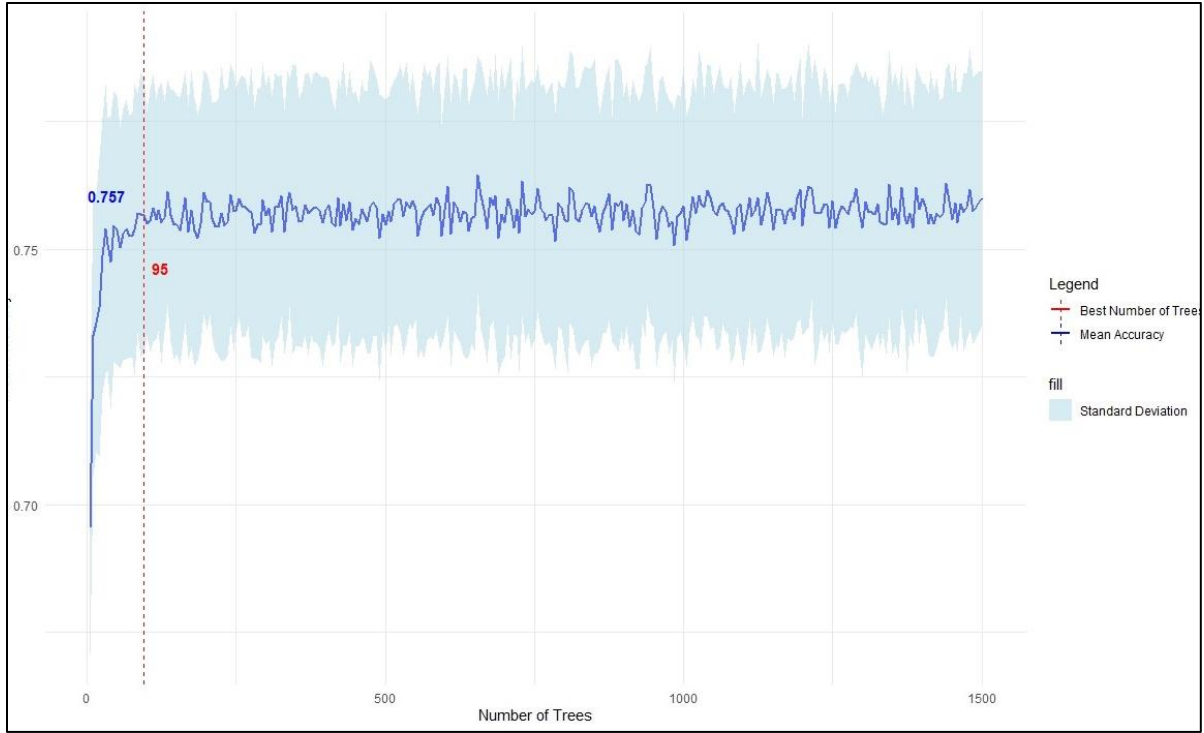


Figure 12: Mean accuracies with standard deviation and best number of trees

Raster Layer	Mean decrease in Gini
Blues Band	33.695
Green Band	46.089
Red Band	50.314
Red Edge Band	49.243
Near Infrared Band	47.401
Blue Ratio	53.346
Green Ratio	74.467
Red Ratio	51.16
Red Edge Ratio	42.061
Near Infrared Ratio	40.929
Blue Standard Deviation	35.479
Green Standard Deviation	22.606
Red Standard Deviation	22.726
Red Edge Standard Deviation	20.09
Near Infrared Standard Deviation	29.118

Table 5: Mean decrease in Gini of each input layer

3.3 Classified Raster

The resulting classified raster is a 83.2 mb GeoTIFF with a spatial resolution of 6 cm/pixel and an extent of 74.64 ha, just like the orthomosaic (**Figure 14**). The raster displays the areas of each land cover class. It is characterized by strips of hummock, mostly adjoined by fields of high lawn, lawn and hollow (**Figure 13**). Within these field, mud bottom areas with bare peat occur next to open water pools. The boardwalk is displayed very accurately and there are a few mineral soil islands that can be visually recognized very clearly. From the six microtopographies, hollow is the most frequent one with an area coverage of 21.578 ha, the least frequent one is open water with an area of 2.935 ha (**Table 7**).

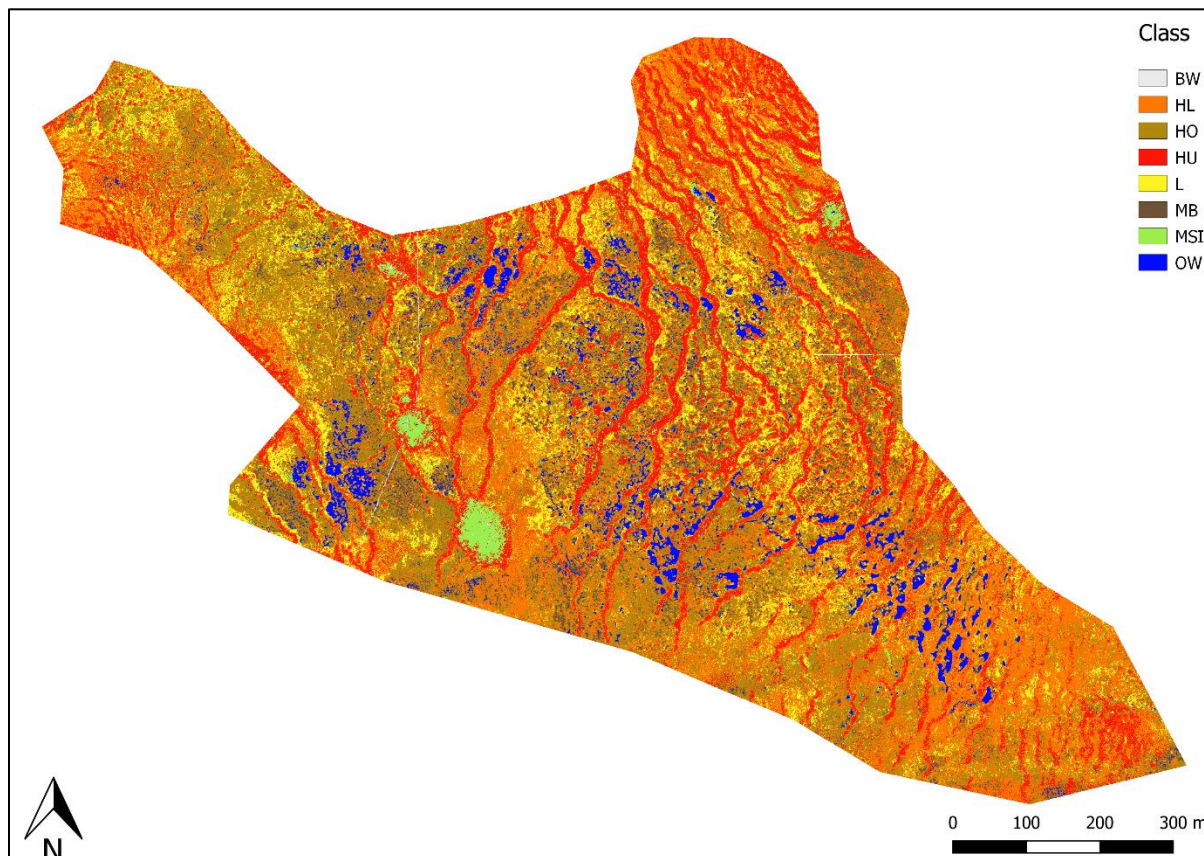


Figure 13: Classified raster

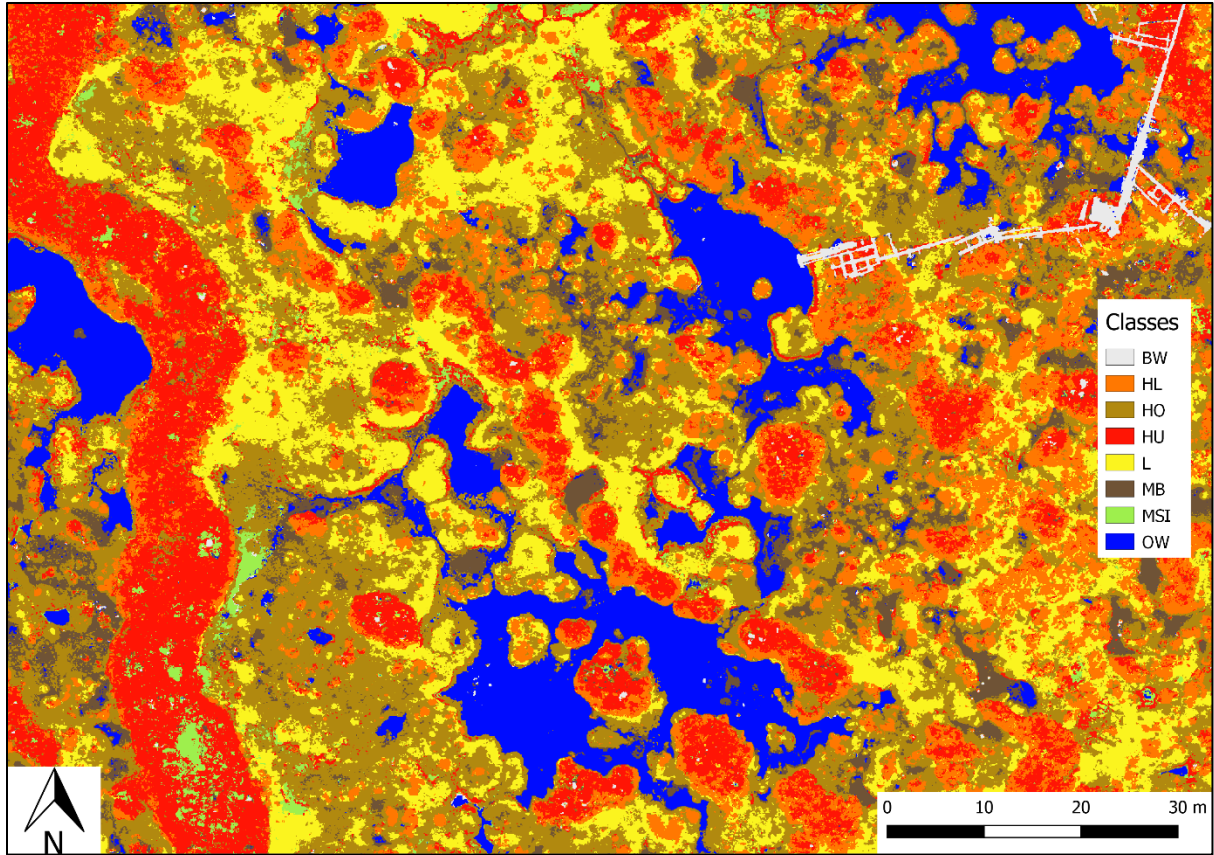


Figure 14: Clip - classified raster

Class	Area [ha]
Boardwalk	0.201
High Lawn	20.972
Hollow	21.578
Hummock	14.189
Lawn	9.527
Mud Bottom	3.587
Mineral Soil Island	1.656
Open Water	2.935

Table 6: Area size of land cover classes

The correlation matrix shows how well the Random Forest model performed at predicting each land cover class (**Figure 15**). The Random Forest model performed best at predicting the open water class with a correctness of 92.24 %. Significantly misclassified were 4.87 % as mud bottom and 2.37 % as boardwalk. The worst performing class was hummock. 62.11 % of actual hummock pixels were classified correctly, significantly misclassified were 16.42 % as high lawn, 9 % as hollow, 8.87 % as mineral soil islands and 3.05 % as lawn. In general, spatially neighboring microtopographies commonly had the highest misclassification rates, e.g. 14.92 % of actual lawn was classified as hollow, 10.15 % of mud bottom as open water.

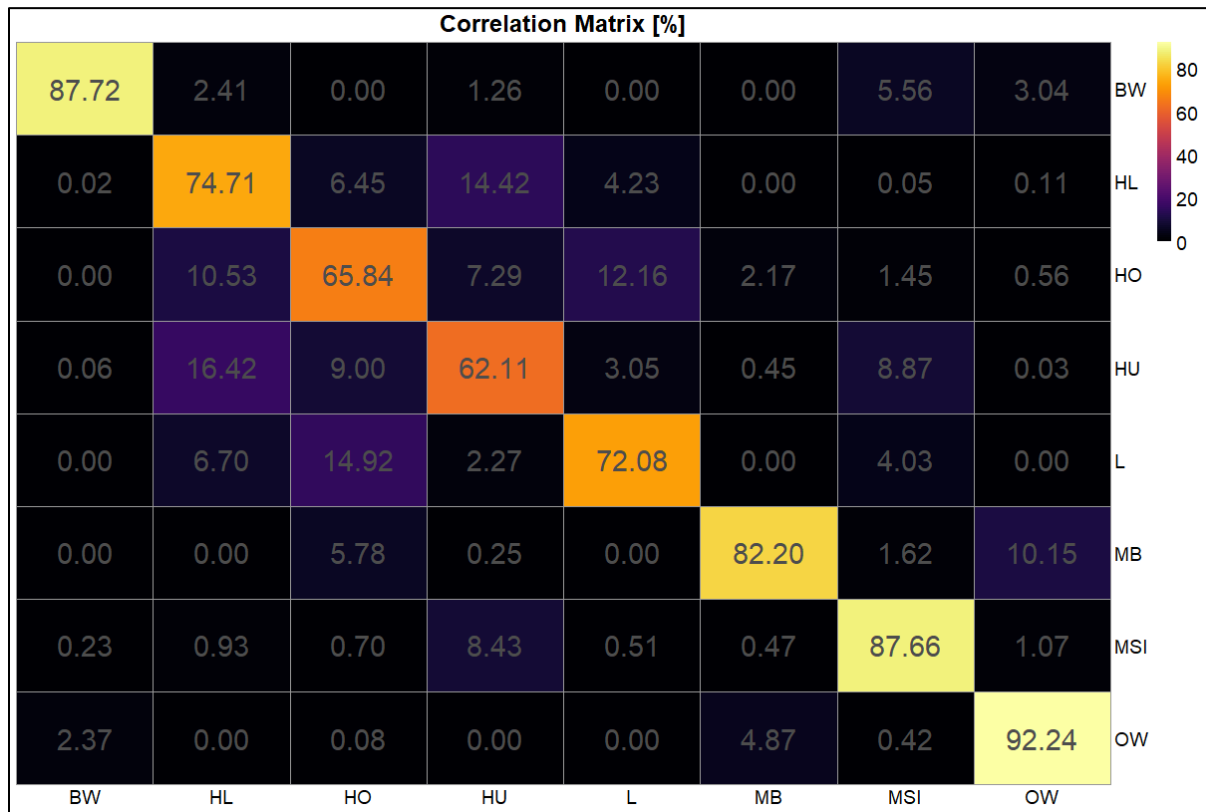


Figure 15: Correlation matrix - actual class on the x-axis, prediction on the y-axis

4. Discussion

4.1 Image Processing

The image processing in Pix4D needed a lot of memory and was very time consuming. Processing each flight used up to 30 gb of RAM and took several hours to complete. Thus, testing out the right processing options took its time. However, nearly all images of the targeted area could be calibrated and matched, so the results turned out very well. Images of forest area could not be calibrated due to the shorter distance between tree crowns and the UAV, which is why some flights show a lower rate of calibrated images. These forest areas were not of interest for this study though. Since the microtopographies are characterized by different elevation levels, a 3D (three dimensional) model of the surface could have been valuable, but was not possible to generate at such a small scale with the two-dimensional aerial imagery. To generate a 3D surface model, the use of a LiDAR (Light Detection and Ranging) sensor would be helpful, such as done by Korpela (2020).

4.2 Classification

Generating training data from completely randomized training points did not work out very well. Due to the uneven occurrence of classes in the classification area, the initial 500 training points did not result in a balanced training set. Adding points within specific areas diminished the level of randomness, but balanced the training data for a more accurate model. A way to work with a fully randomized training data set would be to increase the number of training points until all classes would be sufficiently represented, and then use the ensemble margin (Mellor et al., 2015). The training takes a lot of time though. Overall, 75.7 % can be considered as a good accuracy. A similar study by Korpela 2020 tested different classification methods using RGB imagery and LiDAR data. In that study, the best results, with an accuracy of 71%, were achieved by using a Random Forest model and a combination of RGB imagery and LiDAR-generated index maps. The infrared false color composition helped to visually determine the classes. Especially at distinguishing mud bottom from open water, due to the reflectance of vegetation in the near infrared band. However, distinguishing neighboring classes with similar vegetation during training turned out to be difficult sometimes without a 3D elevation model.

Thus, misclassification of neighboring areas could be the result of inaccurately trained data. Another inaccuracy-factor could be the different capture-times of the aerial imagery. The UAV recorded the data within several hours at three days. Different lighting and weather conditions could have affected the reflection in each image. The relatively high Gini values of the red edge and near infrared band indicate that multispectral data are very useful for such a classification. Even though the quality relies on many factors, achieving a higher accuracy with multispectral imagery than another study did with RGB imagery and LiDAR, leads to the conclusion that multispectral data are very valuable for classifying boreal bog areas. A combination of LiDAR data and multispectral imagery could result in a more accurate classification.

Further analysis of the data with a deeper focus on the images processing as well as the classification goes beyond the scope of this study.

5. Outlook

With the resulting areas of each microtopographie and the greenhouse gas emissions, measured at the Siikaneva bog site, emissions can be estimated on a larger scale. The use of a UAV to capture multispectral imagery for mapping the Siikaneva bog area turned out very successful. The ground truth images taken by the second UAV showed to be very helpful for determining the land cover classes at inaccessible areas during training. Overall, the method developed in this study fits the cause of greenhouse gas emission monitoring very well and it can be applied at any peatland area within visual range of a possible spot to operate a UAV. As outlined in the discussion section, the additional use of a LiDAR sensor could enhance the accuracy of this method.

6. Conclusion

Within this study the following objectives, as outlined in the introduction, were successfully achieved:

- Generating a georeferenced orthomosaic from multispectral UAV imagery
- Designing a Random Forest model with randomly generated training points
- Perform a classification of the study area
- Evaluate the quality of the generated orthomosaic and the performance of the classification model

7. References

- Ahti, T., Hämet-Ahti, L., & Jalas, J. (1968). Vegetation zones and their sections in northwestern Europe. *Annales Botanici Fennici*, 5(3), 169–211.
- Belgiu, M., & Drăguț, L. (2016). Random forest in remote sensing: A review of applications and future directions. *ISPRS Journal of Photogrammetry and Remote Sensing*, 114, 24–31. <https://doi.org/10.1016/j.isprsjprs.2016.01.011>
- Breiman, L., Friedman, J. H., Olshen, R. A., & Stone, C. J. (2017). *Classification And Regression Trees* (1. Aufl.). Routledge. <https://doi.org/10.1201/9781315139470>
- Bu, Z., Hans, J., Li, H., Zhao, G., Zheng, X., Ma, J., & Zeng, J. (2011). The response of peatlands to climate warming: A review. *Acta Ecologica Sinica*, 31(3), 157–162. <https://doi.org/10.1016/j.chnaes.2011.03.006>
- Clymo, R. S., & Hayward, P. M. (1982). The Ecology of Sphagnum. In A. J. E. Smith (Hrsg.), *Bryophyte Ecology* (S. 229–289). Springer Netherlands. https://doi.org/10.1007/978-94-009-5891-3_8
- Colditz, R. (2015). An Evaluation of Different Training Sample Allocation Schemes for Discrete and Continuous Land Cover Classification Using Decision Tree-Based Algorithms. *Remote Sensing*, 7(8), 9655–9681. <https://doi.org/10.3390/rs70809655>
- Dolapci, B., & Özcan, C. (2021). Automatic Ship Detection and Classification using Machine Learning from Remote Sensing Images on Apache Spark. *Journal of Intelligent Systems: Theory and Applications*, 94–102. <https://doi.org/10.38016/jjista.772145>
- Gorham, E. (1991). Northern Peatlands: Role in the Carbon Cycle and Probable Responses to Climatic Warming. *Ecological Applications*, 1(2), 182–195. <https://doi.org/10.2307/1941811>
- Haas, J., & Ban, Y. (2014). Urban growth and environmental impacts in Jing-Jin-Ji, the Yangtze, River Delta and the Pearl River Delta. *International Journal of Applied Earth Observation and Geoinformation*, 30, 42–55. <https://doi.org/10.1016/j.jag.2013.12.012>
- Joosten, H., & Clarke, D. (2002). Wise use of mires and peatlands: Background and principles including a framework for decision-making. International Peat Society ; International Mire Conservation Group.
- Korrensalo, A., Männistö, E., Alekseychik, P., Mammarella, I., Rinne, J., Vesala, T., & Tuittila, E.-S. (2018). Small spatial variability in methane emission measured from a wet patterned boreal bog. *Biogeosciences*, 15(6), 1749–1761. <https://doi.org/10.5194/bg-15-1749-2018>
- Korrensalo, A., Mehtätalo, L., Alekseychik, P., Uljas, S., Mammarella, I., Vesala, T., & Tuittila, E.-S. (2020). Varying Vegetation Composition, Respiration and Photosynthesis Decrease Temporal Variability of the CO₂ Sink in a Boreal Bog. *Ecosystems*, 23(4), 842–858. <https://doi.org/10.1007/s10021-019-00434-1>
- Mellor, A., Boukir, S., Haywood, A., & Jones, S. (2015). Exploring issues of training data imbalance and mislabelling on random forest performance for large area land cover classification using the ensemble margin. *ISPRS Journal of Photogrammetry and Remote Sensing*, 105, 155–168. <https://doi.org/10.1016/j.isprsjprs.2015.03.014>

- Rydin, H., & Jeglum, J. K. (2013). *The Biology of Peatlands*. Oxford University Press.
<https://doi.org/10.1093/acprof:osobl/9780199602995.001.0001>
- Rydin, H., & Rydin, H. (1993). Interspecific Competition between Sphagnum Mosses on a Raised Bog. *Oikos*, 66(3), 413. <https://doi.org/10.2307/3544935>
- Tsutsumida, N., & Comber, A. J. (2015). Measures of spatio-temporal accuracy for time series land cover data. *International Journal of Applied Earth Observation and Geoinformation*, 41, 46–55.
<https://doi.org/10.1016/j.jag.2015.04.018>
- Turunen, J., Tomppo, E., Tolonen, K., & Reinikainen, A. (2002). Estimating carbon accumulation rates of undrained mires in Finland—application to boreal and subarctic regions. *The Holocene*, 12(1), 69–80.
<https://doi.org/10.1191/0959683602hl522rp>
- Vitt, D. H. (2006). Functional Characteristics and Indicators of Boreal Peatlands. In R. K. Wieder & D. H. Vitt (Hrsg.), *Boreal Peatland Ecosystems* (Bd. 188, S. 9–24). Springer Berlin Heidelberg.
https://doi.org/10.1007/978-3-540-31913-9_2
- Wang, Z., Zuo, R., & Jing, L. (2021). Fusion of Geochemical and Remote-Sensing Data for Lithological Mapping Using Random Forest Metric Learning. *Mathematical Geosciences*, 53(6), 1125–1145.
<https://doi.org/10.1007/s11004-020-09897-8>
- Yu, Z. (2011). Holocene carbon flux histories of the world’s peatlands: Global carbon-cycle implications. *The Holocene*, 21(5), 761–774. <https://doi.org/10.1177/0959683610386982>
- Yu, Z. C. (2012). Northern peatland carbon stocks and dynamics: A review. *Biogeosciences*, 9(10), 4071–4085. <https://doi.org/10.5194/bg-9-4071-2012>
- Pix4D online user manual.
<https://support.pix4d.com/hc/en-us/sections/360003718992-Manual>