JNCC Report
No. 667

Developing a framework for using Earth Observation imagery to monitor peatland condition

Becky Trippier, Paul Robinson, Daniel Colson & James Hutchison

October 2020

ISSN 0963-8091

For further information please contact:

Joint Nature Conservation Committee
Monkstone House
City Road
Peterborough PE1 1JY
https://jncc.gov.uk/

This report should be cited as:

Trippier, B., Robinson, P., Colson, D. & Hutchison, J. 2020. Developing a framework for using Earth Observation imagery to monitor peatland condition. *JNCC Report No. 667*, JNCC, Peterborough, ISSN 0963-8091.

JNCC EQA Statement:

This report is compliant with JNCC's Evidence Quality Assurance Policy

https://jncc.gov.uk/about-jncc/corporate-information/evidence-quality-assurance/

Acknowledgments:

We would like to acknowledge the previous work in this field, particularly partners from UKCEH and the James Hutton Institute, which provided a basis for developing this methodology and further exploration. That work was funded by DEFRA and Scottish Government.

Summary

Peatland restoration is a significant target of a number of key policy areas within all four UK administrations. Developing cost-effective methods for monitoring key indicators of peatland condition, such as; vegetation cover of Sphagnum mosses, areas of bare unvegetated peat, soil depth and drainage, can help to inform the targeting of restoration strategies and establish whether efforts are successful in promoting recovery (SNH 2020). This report investigates the use of earth observation data for monitoring peatland condition through exploring whether the amount of bare peat cover can be detected in an automated process from very high-resolution imagery (aerial and satellite) and how this can then be scaled up to lower resolution satellite imagery covering a wider extent. This report outlines a framework for mapping bare peat which includes ways of determining the amount of imagery required to inform models and the best modelling approach for accurate predictions of bare peat cover using machine learning algorithms.

Firstly, we investigated whether Very High Resolution (VHR) satellite imagery or aerial photography can be used to develop a methodology for automating bare peat detection. We compared data from Pleiades-1A satellite imagery and Aerial Photography for Great Britain (APGB) taken over the Peak District, UK in June 2018 and explored differences in spectral band and derived indices signatures between observed vegetated and non-vegetated regions of peat. Thresholding rules were then developed to classify pixels of bare peat in the two sets of imagery.

Secondly, we investigated whether the modelled bare peat from the VHR imagery could be used to predict bare peat cover over wider areas using low resolution Sentinel-2 optical imagery. The classified pixels from VHR data were aggregated to a percentage cover of bare peat at 10m spatial resolution. These data were used to train regression models to predict bare peat cover within Sentinel-2 imagery, allowing prediction over a much wider spatial extent. A variety of different predictor variables and machine learning algorithms were trialled, along with the sampling methodologies for selecting training points and the number of points required for an accurate prediction.

The results of our first investigation showed that the APGB aerial photography displayed distinguishable differences between areas of vegetated and non-vegetated peat, and through thresholding derived indices of NDVI, Brightness, Red/Green and Red/Blue ratios was able to create fine-scale maps of bare peat. However, the methodology was unable to distinguish the bare peat pixels using the Pleiades imagery due to overlap in the spectral and indices signatures when compared with vegetated pixels. In addition, for both data sources, using a pixel-based indices thresholding approach to classify land areas had some limitations, with some neighbouring pixels not meeting threshold criteria as expected, and thresholding rules having limited transferability to different geographical areas. This was likely due to differences in atmospheric conditions, hill shade, blurriness from wind, and sun illumination during different flights impacting the spectral band values. Our research identified several options to address these deficiencies, which could be explored to help improve the fine-scale mapping approach.

The results of our second investigation showed that regression modelling informed by the fine-scale bare peat maps and Sentinel-2 imagery, successfully scaled up these predictions to a wider landscape, producing maps of the amount of bare peat cover per 10m pixel. From our results, we recommend the following workflow (Figure 1) of how best to inform this type of analysis. Training the models with a range of values of bare peat cover per pixel was found to be key in generating accurate results, particularly in extrapolated regions. From this we demonstrate how Sentinel-2 imagery trained on aerial photography derived data, can be used to predict across large areas of the landscape the fraction of bare peat cover below

10m^2. Highest accuracy was achieved where aerial photography was spatially distributed across an area and contained a range of pristine and degraded peatland conditions.

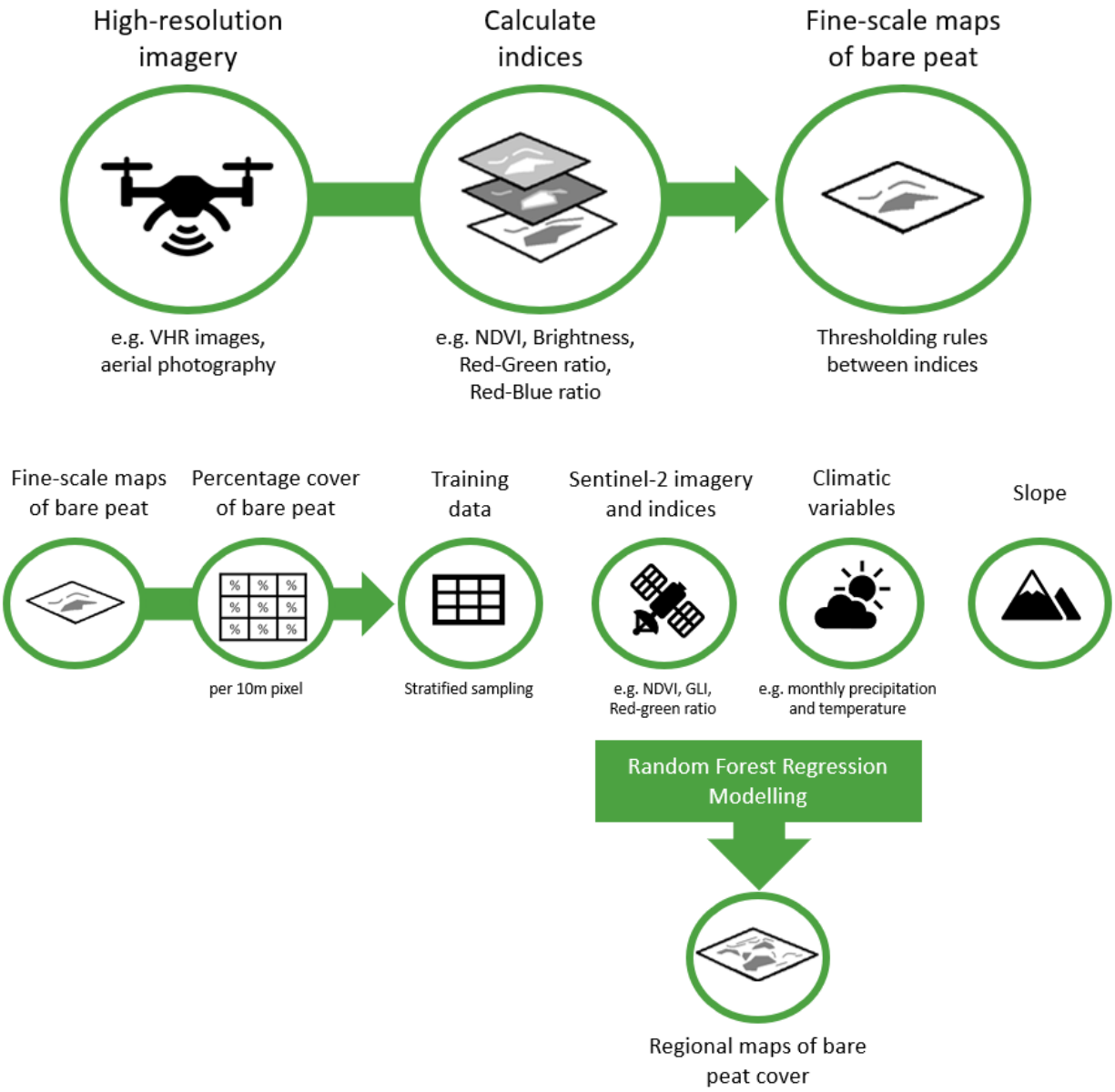


Figure 1: Infographic of the suggested workflow for creating bare peat maps using high-resolution imagery and Sentinel-2 optical imagery.

Contents

1 Introduction

Peatlands occupy over 10% of the UK, estimated to cover an area between 46,000-77,000km^2 (Bain *et al.* 2011). They are areas of naturally accumulated peat formed from layers of decaying plant matter under water-logged conditions, with the three main types in UK being blanket bogs, raised bog and fen (Natural England 2010). Peatlands are important ecosystems for regulating climate and greenhouse gas emissions, regulating drinking water supply, flood prevention, and providing habitat for nationally and internationally important species (FAO 2020). Statistics submitted for Article 17 2013-2018 report to meet the requirements of the Habitats Directive estimate 16.5% of bogs, mires and fens in the UK are not in good condition, with only 1.8% thought to be in good condition and 81.7% being in an unknown state (JNCC 2019). Human activities such as peat extraction, land use change and drainage activities to support agriculture or the forestry industries contribute to the degradation of these natural landscapes (Bain *et al.* 2011). Damaged peatlands can release carbon dioxide back into the atmosphere, adding to greenhouse gas emissions and can have detrimental impacts towards ecosystem function, causing a reduction in the delivery of key ecosystem services (Dickie *et al.* 2015).

Assessing the state of peatlands and restoring these to good condition is therefore a key component for reducing greenhouse gas emissions, preserving biodiversity and habitat resilience. Peatland restoration is significant to a number of key policy areas and targets have been set within all four UK administrations as well as overarching policy targets for the wider UK. The majority of peatland in the UK is in Scotland, where an estimated 70% of blanket bog and 90% of raised bog area is damaged (SNH 2015). Scotland's National Peatland Plan recognises the need for long term monitoring of sites to assess condition and highlights the need for knowledge on the current state of peatlands and the best practices for restoration (SNH 2015). It supports a major programme of peatland restoration funded by the Scottish Government under the Peatland ACTION restoration project and commitment to future support through the Scottish Rural Development programme in efforts to mitigate climate change and create a healthier landscape (SNH 2015). The scheme aims to help meet targets set by the Scottish Biodiversity Strategy Route Map in adhering to the EU biodiversity target of restoring at least 15% of degraded ecosystems and the 2020 Challenge for Scotland's Biodiversity (Scottish Government 2015). In England, restoration is embedded in Defra's 25-year Environment Plan with £10m funding dedicated to peat restoration projects (Defra 2018a, Defra 2018b). In Wales, schemes such as the Welsh Peatlands project under the Sustainable Management Scheme, are helping to meet ministerial targets of bringing Welsh peatlands into sustainable management by 2020 (Welsh Government 2018). In Northern Ireland, peatlands cover around 12% of land and is the only habitat protected under policy to ensure representative areas of habitat are protected for the benefit of future generations (DAERA 2020). The UK Peatland Code is a UK-wide scheme to encourage the private sector to invest in peatland restoration, to aid in the goals of the UK Peatland Strategy set out by the IUCN Peatland Programme (IUCN 2018; SNH 2015) in promoting restoration as an action towards tackling climate change and promoting biodiversity. The UK Peatland Strategy also highlights both national and international policies which recognise their significance in adapting to climate change and supporting rare and threatened species, such as the UN Convention on Biological Diversity Aichi Biodiversity targets, reporting under the EU habitats directive and UN Sustainable Development goals (IUCN 2018). Careful monitoring of the condition of these habitats is therefore essential to understand their impact upon greenhouse gas emissions and maintaining peatland water, biodiversity and carbon (FAO 2020).

Monitoring is essential to assessing the recovery condition of these vital habitats, and the success of restoration projects (Bonnett *et al.* 2009). Current peatland monitoring strategies have been ground based surveys such as the Phase 1 habitat surveys, National Vegetation

Classification (NVC), Heather Trust and Scottish Natural Heritage peat condition surveys and JNCC's Common Standards Monitoring (CSM) (Bonnett *et al.* 2009). These rely on experts to go into the field to survey indicators of condition, such as; vegetation cover of *Sphagnum* mosses, evidence of erosion or burning, areas of bare unvegetated peat, soil depth and drainage (SNH 2020). This can be time consuming, expensive, restricted to accessible sites and data can vary depending upon human error, sampling methods and effort and in their spatial extent. Increasingly, NGOs and peatland partnerships are moving towards using drone-derived data to assist in peatland monitoring strategies, providing imagery that can cover a much wider area than ground observations (Artz *et al*. 2020). Earth observation (EO) data provides a cost-effective rapid methodology for mapping large areas of the UK in real-time. EO can be used to observe types of vegetation and changes in plant cover and growth, as well as other environmental measures such as soil moisture (Copernicus Global Land Service 2019). For peatlands, bare peat is considered a visible indicator of peatlands in poor condition, often resulting from erosion, drying, and damage from overgrazing or trampling (SNH 2020). It can occur in both large expanses and small patches across the landscape, which can be hard to derive from satellite data at lower resolutions. Artz *et al.* (2019) used a time series of Moderate Resolution Imaging Spectroradiometer (MODIS) data to derive peatlands in favourable condition in Scotland, with random forest regression models found to give good predictive capacity of 0.915 ROC accuracy. Work in the region has also been carried out by Blake and Frake (2020), who explored classifying 10m Sentinel-2 imagery into classes of bare peat, rock, water, shadow and other, aggregating these predictions to a national map of area predictions for bare peat cover per 1km. These studies demonstrate how bare peat can be mapped nationally up to 10m spatial resolutions; however, this is relatively coarse when considering the number of small patches of bare peat in the landscape where such an approach may underestimate. Several studies have used high-resolution aerial photography to assess peatland condition, as it has been found successful at detecting fine scale changes in the landscape (Artz *et al.* 2019). Williamson *et al.* (2018) used two high-resolution aerial images to derive the extent of exposed bare peat through using a Random Forest Classification (RFC) to first define areas of bare peat, then used these predictions to train a Random Forest Regression (RFR) model to derive the bare peat coverage from Sentinel-2 imagery, aggregating data to different levels. This methodology was able to accurately estimate the proportion of bare peat within a pixel, with data prepared at 40m resolution data performing the best, however this provided a relatively coarse estimate. Other studies such as Blake *et al.* (2019) used a similar methodology in the Cairngorms, finding this performed best when training the models using imagery from the same year and season as imagery used for the predictor variables. Although these studies have demonstrated that such an approach is possible, these have taken place over a relatively small area with limited suggestions to operationalising such an approach over a wider landscape.

This study aimed to investigate methods of monitoring peatland habitat condition, building on previous studies using EO data and high-resolution aerial photography.This built upon the methodology of Williamson *et al.* (2018)'s study, aiming to derive a framework for deploying this type of habitat analysis rapidly across the landscape. Firstly, we investigated whether we could derive bare peat training pixels from earth observation using high-resolution imagery from two different sources: Pleiades-1A multispectral imagery at 2m spatial resolution captured daily and APGB aerial photography at 25cm and 50cm spatial resolution captured every 3-5 years. The study then focused upon scaling these predictions to 10m Sentinel 2 imagery across a wider extent. This explored various aspects of the methodology to assess how best to sample the training datasets and trialling different combinations of predictor variables, machine learning algorithms and the numbers of training points to use, and their spatial distribution.

The hypotheses tested were:

1. **Is it possible to derive bare peat pixels from high resolution imagery?**
   To investigate whether pixels of bare peat can be determined in the high-resolution imagery using different indices and spectral band data. This aimed to identify whether it is possible to automate a methodology to identify bare peat training points from both the Pleiades-1A and APGB imagery.

2. **What is the best methodology for deriving sample data?**
   This aimed to determine the best sampling methodology to apply to the high-resolution data in order to create the training data for the regression models. Here we used training data from the outputs of hypothesis 1, trialling under four sampling methodologies;
   a. Random Sampling
   b. Equal Interval Sampling
   c. Equal Interval Sampling and filling
   d. Quantile Sampling

3. **Which predictors are important to include in the regression analysis?**
   This explored which predictor variables were important to include in the regression models used to scale up the bare peat cover predictions. This was assessed through rerunning the regression models with different combinations of spectral, indices and environmental layers and evaluating which provided the most accurate predictions.

4. **What is the best performing machine learning technique for extrapolating the data?**
   This tested various types of machine learning algorithms used in the regression analyses; Random Forest Regression (RFR), Quantile Forest Regression (QFR), Support Vector Machines (SVM), Bayesian Regularization of Neural Networks (BRNN) and Boosted Regression Trees (BRT). The performance of each model was assessed to determine which methodology is best for this type of assessment.

5. **What is the minimum number of images/training points required to make accurate predictions across a given extent?**
   This investigated the number of images and training points derived from these required to train accurate regression models. Knowing this can help to determine future data collection, particularly in defining the number of aerial images required to derive bare peatland pixels in Sentinel-2 data. How these images were spatially distributed was also explored.

2 Methodology

Figure 2 summarises the workflow undertaken to detect and scale up bare peat predictions. This was established by building upon previous work commission by DEFRA's Earth Observation Centre of Excellence, conducted by CEH and the James Hutton Institute (Williamson *et al*. 2018). Here, we have further developed R scripts by Ned Horning of the American Museum of Natural History (AMNH, Horning 2018) which were utilised in the Williamson *et al.* (2018) report, by adding in well-established machine learning and sampling approaches to increase the accuracy of these models and evaluate model performance both within the area where our model is trained and outside of this in extrapolated regions.

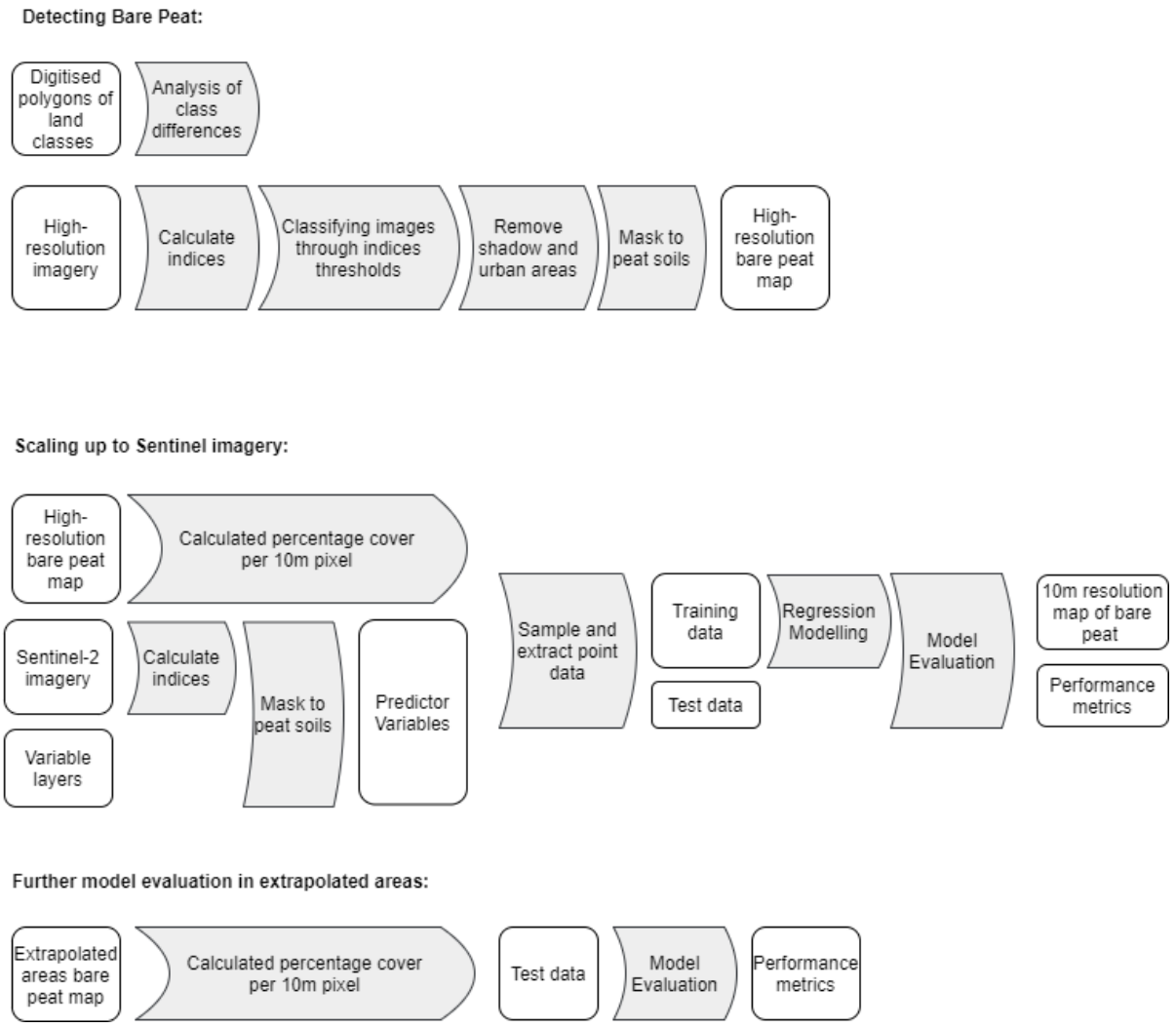


Figure 2: Workflow for the detection of bare peat pixels in high-resolution imagery and the scaling up of these prediction to Sentinel Imagery, with further evaluation in extrapolated regions

The Peak district in the UK was chosen as a study region due to the availability of high- and low-resolution imagery, with previous JNCC projects having focused upon this region. Blanket Bog is estimated to cover 20,838ha of the Peak District and are one of the best studied peatland regions in the UK (Peat District National Park 2011). The AOIs where the imagery was obtained are shown in Figure 3.

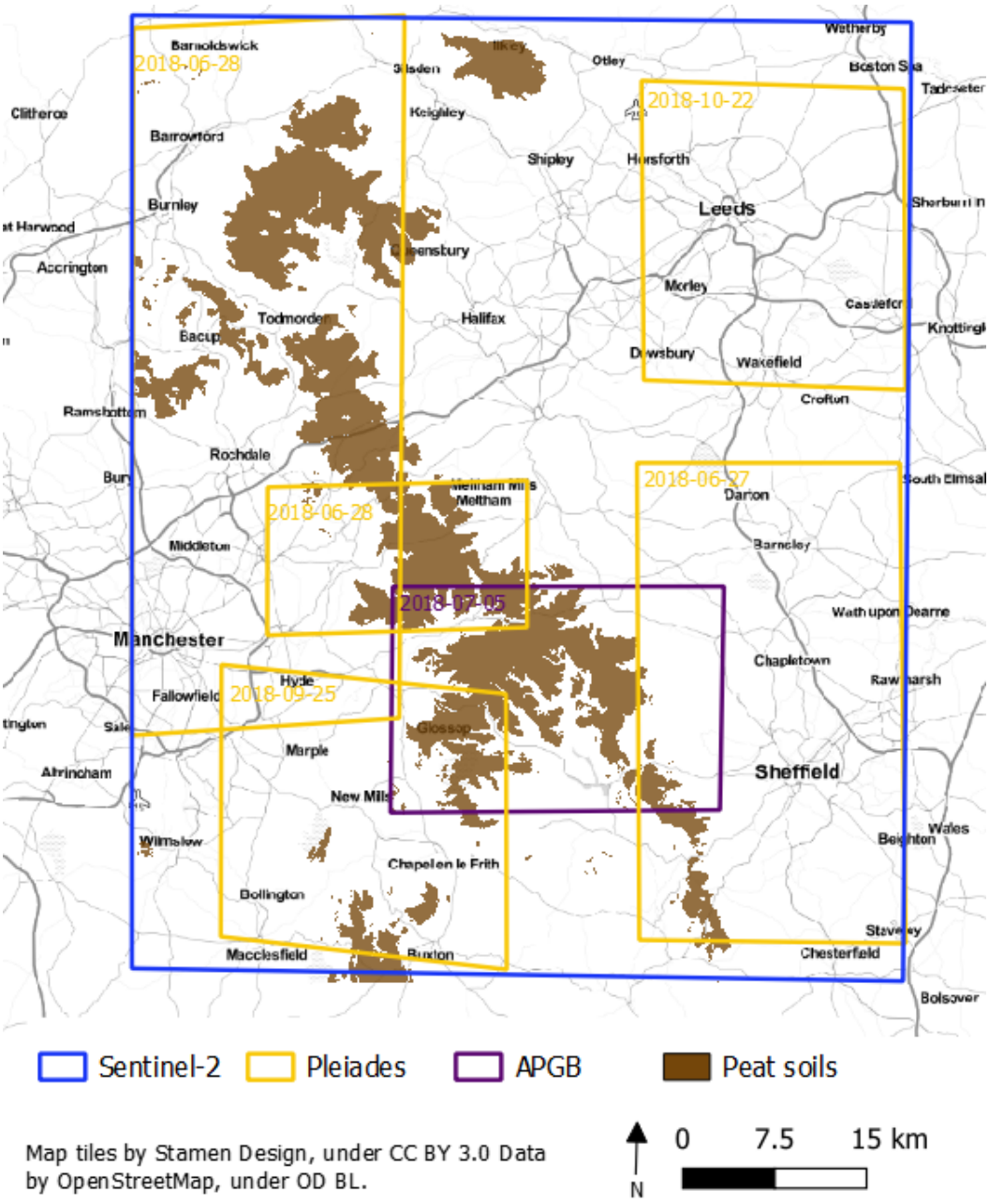


Figure 3: Area of interest in the Peak district, UK, where APGB and Pleiades imagery was obtained from June 2018.

2.1 Data sources

All data were georeferenced to British National Grid (ESPG:27700) and cropped to the area of interest. Data processing and transformation was conducted in R version 3.6.0 (R Core Team 2019), with GDAL functionality integrated through the 'gdalUtils' package (Asher *et al.* 2018). Visual comparisons and some basic manipulation were conducted in QGIS v. 3.4.5-Madeira (QGIS Development Team 2020).

2.1.1 High-resolution imagery

Two data sources were compared for providing high resolution imagery to train areas of bare peatland coverage.

Pleiades-1A
AIRBUS Defence & Space Pleiades-1A multispectral satellite data was acquired at 2m spatial resolution for 2018. The satellite imagery features four spectral bands (blue, green, red and IR) and has an image location accuracy of 3m (CE90) without ground control points (Satellite Imaging Corporation 2019). Five images from the following dates and locations were used in the analyses:

- Pleiades 27th June 2018 - latitude 53, longitude 15
- Pleiades 28th June 2018 - latitude 54 longitude 196
- Pleiades 28th June 2018 - latitude 54 longitude 212
- Pleiades 25th September 2018 - latitude 53 longitude 20
- Pleiades 28th June 2018 - latitude 54 longitude 149

The Pleaides data were processed to an Analysis Ready Data (ARD) product in an Amazon Web Services ec2 virtual machine. The processing steps undertaken produce a topographically correct surface reflectance product, correcting for atmospheric effects. Due to the variations of elevation in the area of interest, this led to inherent distortion in the Pleiades data. To prepare the data for use, an orthorectification step was performed. This enables the topography to be accounted for and attempts to accurately remove image distortions. This step was implemented using the open-source Orfeo Toolbox (CNES 2020) suite of tools. Subsequently, they have been processed to surface reflectance using the Atmospheric and Radiometric Correction of Satellite Imagery (ARCSI) software, v3.1.6 (http://rsgislib.org/arcsi)."

Aerial photography
Aerial photography was provided by DEFRA's APGB service as RGB aerial photography at 25cm and Colour infrared aerial photo (CIR) at 50cm spatial resolution (Bluesky International 2019). 2660 tiled aerial images from 2018 were used, covering a total area of approx. 509km^2. In order to include the Near Infrared band from the CIR photos in our analysis, this image was resampled to 25cm using a bilinear interpolation.

Low-resolution imagery
Sentinel-2 is a multispectral high-resolution imaging mission part of the European Union's Copernicus Programme operated by ESA. Sentinel-2 data were processed to an Analysis Ready Data (ARD) product. The processing steps undertaken produce a topographically corrected surface reflectance product with cloud and topography masks that can be applied to the imagery. The ARCSI software (http://www.rsgislib.org/arcsi) was used to produce ARD. During the processing, the 20m image bands are sharpened to 10m through the application of linear regression models. This produces a final output with 10 by 10m multispectral bands. The 60m bands are used for atmospheric aerosol processes and therefore removed from the final product. This is as per the processing chain developed by JNCC as part of the Sentinel-2 ARD Provision Service (Jones *et al.* 2017). The 2018 Sentinel-2 summer mosaic was generated using the mosaic function within RSGISLIB (https://www.rsgislib.org/rsgislib_imageutils.html). This function is a series of Python bindings that combines multiple raster datasets into a new raster dataset, optimised for working with a large number of raster datasets.

2.1.2 Peatland boundaries

A peatland boundaries layer was obtained from classified peat soil maps of England developed by Evans *et al*. (2014) and a Centre of Excellence commissioned study by Williamson *et al.* (2018). For the purposes of this study, the peatland classifications for 30UWE South Pennines were dissolved to create a single mask for determining the boundary of where peat soils were present. To account for inaccuracies associated with the

peat mapping technique, as visual clues of sub-surface peat are not always reliable, the boundaries layer was buffered by 10m to encompass a wider extent likely to include peat soils.

2.1.3 Variable layers

The CEH 2015 Land cover map (Rowland *et al*. 2017) was obtained at a spatial resolution of 25m. This was used to mask out the urban land class from the derived training points, to ensure no training points were sampled in urban areas. Urban areas were removed as roads and buildings also yield low NDVI values, which could create a source of error in our data when trying to obtain the bare peat pixels from our imagery.

Several variables were trialled in the regression models to assess their importance in predicting the bare peat cover. These included products of Surface Soil Moisture (SSM) and Soil Water Index (SWI) generated by Copernicus' Global Land Service (2019) at 1km spatial resolution obtained for 27- 28 June 2018. A roads layer was also obtained from Ordnance survey's Open Roads open access data layer for the Sentinel imagery AOI (OS 2019). Digital Terrain Model (DTM) data at 2m spatial resolution were collected from Environment Agency's Integrated Height Model (IHM) using Lidar data (Kilcoyne *et al.* 2017). Finally, climatic variables were obtained from the Met Office (2019) HadUK gridded 1km monthly product available through CEDA's archive. Data trialled included monthly precipitation and maximum, minimum and average temperatures for June 2018. All variable layers were resampled to 10m spatial resolution to match that of the Sentinel-2 imagery.

2.2 Deriving bare peat pixels

To assess the differences which could be observed in the imagery between areas of bare and vegetated peat, polygons representing each of the classes were manually digitised in overlapping regions between the two imagery sources. This was performed through visual assessments of the high-resolution imagery, making comparisons with Google Earth® (2019) satellite imagery and Google's publicly available photo library, shown in Figure 4. The photo library provided ground-truthed photographs from sites, although their publishing dates varied greatly and therefore may not be representative of site conditions present during the period our imagery was collected. Points of bare and vegetated peat were randomly sampled from within the digitised polygons.

To observe the different spectral signatures of the bare and vegetated classes, these points were used to extract values from each of the spectral bands of high-resolution imagery. Indices were also calculated from the multispectral data, to see if these could be used to determine between the bare and vegetated peat pixels. The indices analysed included Brightness, Enhanced Vegetation Index (EVI), Green Leaf Index (GLI), Green Normalised Difference Vegetation Index (GNDVI), Normalized difference vegetation index (NDVI), Red-Blue ratio (RB), Renormalized Difference Vegetation Index (RDVI), Red-Green ratio (RG) and Soil Adjusted Vegetation Index (SAVI). Indices such have these have been commonly used in studies mapping vegetation amount and cover, habitat classifications and peat mapping. For instance, NASA (2019) suggests an NDVI less than 0.1 to represent barren rock, snow or sand, with values 0.2-0.3 representing shrub and grassland habitats. Studies, such as Montandon and Small (2008), have also observed that the NDVI values of bare soil are often underestimated, with their observations displaying values averaging 0.2 and is highly variable (SD=0.1). Enhanced Vegetation Index has been widely used in studies as an indicator of vegetation and is thought to be most topographically sensitive than NDVI due to the added soil adjustment factor (Matsushita *et al*. 2007). The indices calculations are summarised in Appendix 1. The spectral and indices profiles of the two peat classes were analysed to assess whether differences in their signatures could be used to identify bare

peat from the APGB and Pleiades imagery sources. The knowledge from this assessment was then used to aid in the development of rules to using threshold values for the indices to separate out the bare peat pixels. The outputs were then validated against digitised polygons of bare and vegetated peat and 100 randomly sampled points throughout the image, to give an indicative accuracy of the predicted results.

Figure 4: Comparisons between the aerial photography (top left) and Google Earth® satellite imagery (top right) and user photographs. Site 1 (bottom left) shows the same site as the top two images, with patches of bare peat, whereas site 2 (bottom right).

2.3 Bare peatland classification

To evaluate how the bare peat layer created through indices thresholding compared against other methodologies, the predicted layer was compared to the results of a random forest classification model. The model was trained on points randomly sampled from the digitised polygons of bare and vegetated peat, taking an equal number of points from each class. The random forest classification model was built using the red, green, blue and near-infrared spectral bands as predictors, with the performance evaluated using 25% of the data. As only a limited number of polygons were digitised, this comparison was only drawn from three classified APGB tiles, however given more time this comparison could be more in depth covering the greater extent.

2.4 Bare peat percentage cover

The bare peat training data derived from the high-resolution APGB imagery at 25cm spatial resolution was scaled up and converted into a percentage cover of bare peat at 10m resolution, to match the Sentinel-2 imagery. This was performed by first aligning the high-resolution training data and sentinel imagery, where one 10m sentinel cell containing 40x40 cells in the high resolution 0.25m. Within each Sentinel-2 cell, cell values were then

extracted from the high-resolution data and summed as a percentage cover per low-resolution pixel, demonstrated in Figure 5.

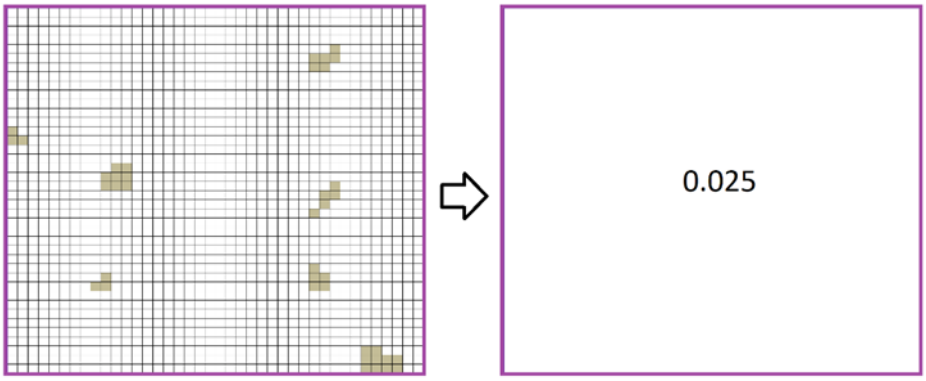


Figure 5: Grid cells in the high-resolution image (black) in comparison with the cell size in the low-resolution image (purple), and how the bare peat pixels (brown) were compiled per cell into a percentage cover.

This generated a 10m raster of bare peat percentage cover per cell. This provided training data for the regression analysis to scale up the bare peat predictions to 10m to match the resolution of the Seninel-2 imagery.

2.5 Regression analysis

The bare peat coverage data were used to derive the training data for the regression modelling. Several methods of sampling these data were trialled:

- **Method 1: Random Sampling** – The desired number of training points were randomly sampled from the data layer.
- **Method 2: Equal interval sampling** – The training data was divided into four categories equally based on the minimum and maximum values, with an additional category for values where bare peat was absent (coverage was 0). These categories were 0, 0-0.25, 0.25-0.5, 0.5-0.75 and 0.75-1 percentage cover of bare peat. The desired number of training points were then equally divided between each category and attempted to be sampled from each category. The actual total points sampled would depend upon the number of points available within each category.
- **Method 2b – Equal interval sampling with categorised filling** – This method was the same as Method 2, however if a category did not have the required number of points available, the remaining points were then divided amongst the other categories and sampled until the actual total points sampled met the desired number of training points.
- **Method 3 – Quantiled sampling of values above 0.1** - The training data was first separated into those where coverage was equal to 0 and those where coverage was greater than 0 but below 0.1%. divided into four categories equally based on the minimum and maximum values

As well as the sampling methodology, the number of training points used to train the regression models were also trialled. The models were run with 2000, 5000, 10000, 20000, 50,000 and 100,000 target points. For each of the model runs, the training data were split; with 75% used to train the models and the remaining 25% used for model evaluation. The models were each run for 5 times and the Root Mean Square Error (RMSE) and R-squared metrics were used to compare their performance. The resulting predicted maps of the five

model runs were averaged for each raster cell, to alleviate variation from the random sampling of training points, to produce the final maps of predicted bare peat cover.

Other aspects of the regression modelling were trialled in order to build a suitable framework for scaling up predictions. The variables used to predict the presence of bare peat in the Sentinel imagery were also assessed through rerunning the models with different combinations of layers and assessing their performance. The variable combinations were chosen through looking at previous studies (Williamson *et al.* 2018; Blake *et al.* 2019) and by producing a correlation matrix to eliminate those highly correlated with each other. This assessment was carried out with the 'Corrval' function in the 'JNCCsdms' package (JNCC 2019) aiding in reducing redundancy and computational demands of the model. Each variable combination was run with 10,000 target training points, to see which had the greater predictive performance. The predictor combinations trialled are shown in Table 1.

Table 1: Predictor Variable Combinations.

Combination name	Layers Included
varstack_sbands	Sentinel bands 1-10: Blue band (1), Green band (2), Red band (3), Red-edge1 band (4), Red-edge2 band (5), Near infrared (narrow) band (6), Near infrared band (8), Short Wave Infrared 1 band (9), Short Wave infrared 2 band (10)
varstack_blake	Blue band (1), Green band (2), Red band (3), Red-edge 1 band (4), Short Wave Infrared 1 band (9), Slope, NDVI
varstack_imp	DTM, Slope, Red/Green ratio, Red/Blue ratio, Brightness, Surface Soil Moisture (SSM)
varstack_all	All variables shown in Figure.
varstack_RD_SWIR	Red-edge 1 band (4), Red-edge 2 band (5), Red-edge 3 band (6), Short Wave Infrared 1 band (9), Short Wave Infrared 2 band (10), Slope, NDVI, Surface Soil Moisture (SSM)
varstack_slope	Red-edge 1 band (4), Red-edge 2 band (5), Red-edge 3 band (6), Short Wave Infrared 1 band (9), Short Wave Infrared 2 band (10), Slope
varstack_NBR	Brightness, Red/Blue ratio, Red/Green ratio, Normalized Burn Ratio (NBR), Surface Soil Moisture (SSM), Slope
varstack_climMax	Brightness, Red/Blue ratio, Red/Green ratio, Normalized Burn Ratio, slope, Monthly rainfall, maximum monthly temperature
varstack_climAv	Brightness, Red/Blue ratio, Red/Green ratio, Normalized Burn Ratio, slope, Monthly rainfall, average monthly temperature
varstack_darkveg	Brightness, NDVI, RG, satimgband7, satimgband2, satimgband4, slope, SSM, SWI, NBR, rainfall

The regression algorithm was also explored, comparing how several machine learning algorithms performed in the models and which produced the most accurate predictions of bare peat cover. Random Forests have been widely used with both classification and regression analyses, particularly with large datasets, with applications including habitat and soil classification mapping (Kilcoyne *et al*. 2017; Veronesi & Schillaci 2019). A random forest is an ensemble machine learning technique using a bagging methodology to train subsets of the data on multiple decision trees (Kuhn & Johnson 2013). In the regression model, the prediction is the mean response across all the trees (Meinshausen 2006). The models were

built in R version 3.6.0 using the 'randomForest' package (Breiman & Cutler 2018). Where training points were sampled randomly under method 1, the random forest predictions were bias corrected, as random forest is known to reduce the variance of regression predictors causing underestimation of larger values and overestimation of smaller values (Zhang & Lu 2012; Xu 2013). Following the methodology set out by Ned Horning (2011), this was performed using a linear regression, adjusting the predictions of the random forest model with the following equation:

$$CorrVal = PredVal * slope + intercept$$

Where *CorrVal* is the bias corrected prediction value, *PredVal* is the predicted value from the model, and *slope* and *intercept* are calculated through a linear model of predicted values compared against the training values. For the stratified sampling methodologies, this correction was not performed as the training data had a suitable variance where correction was deemed unnecessary.

Quantile regression forests (QRF) are based on the random forest methodology but differ where they keep the value of all the observations for each node in each tree, as opposed to just the mean, and evaluates the empirical quantile estimates (Meinshausen 2006). Estimates incorporating conditional quantiles have proven to provide accurate prediction, however, can create bias in feature selection and prediction bias (Nguyen *et al*. 2014). The models were built in R using the 'QRT' R package (Meinhausen 2017) called through the 'caret' wrapper package (Kuhn *et al.* 2020).

As well as random and quantile regression forests, three other machine learning algorithms popular with regression analysis were trialled. These included Boosted Regression Trees (BRTs), Support Vector Machines (SVM) and Bayesian Regularization of Neural Networks (BRNN) algorithms (Veronesi & Schillaci 2019; Räsänen *et al*. 2019; Fukuda *et al.* 2013). Many of these or similar algorithms were similarly trialled in Rudiyanto's (2018) analysis of mapping peat thickness in tropical peatlands, with quantile regression forests and random forests resulting in two of the three best modelling approaches outperforming continuous parametric models, such as Neural Networks and SVMs. These models were built using the 'brnn' (Rodriguez & Gianola 2020), 'brt' (Zheng & Yu 2018) and 'kernlab' (Karatzoglou *et al*. 2019) packages respectively, called through using the 'caret' wrapper package (Kuhn *et al*. 2020).

2.6 Model performance in extrapolated areas

The evaluation statistics RMSE and R-squared were used to assess at how well the models performed against the test dataset, within our initial training data region. To assess how well they predicted bare peat cover in regions outside of the training region, APGB imagery from the same time period were also obtained for two additional areas, shown in Figure 6. This assessment aimed to inform the practicalities of the scaling approach, to see if one area of aerial photography was sufficient, or whether additional areas would be required in order to produce an accurate map of bare peat coverage.

The performance in the extrapolated areas was assessed by further digitising bare and vegetated peat polygons in the areas 'Eval1 and 'Eval2, given a value of 1 was where bare peat was present and 0 where peat soils were vegetated. The polygons were then used to extract values from the predicted bare peat cover output by the models. A threshold of 5% cover was used to classify the predicted values, with values above 5% representing the presence of bare peat. A confusion matrix comparing the polygon dataset and the predicted bare peat absence/presence was then generated. This showed how well the model was able to predict cells where bare peat was present.

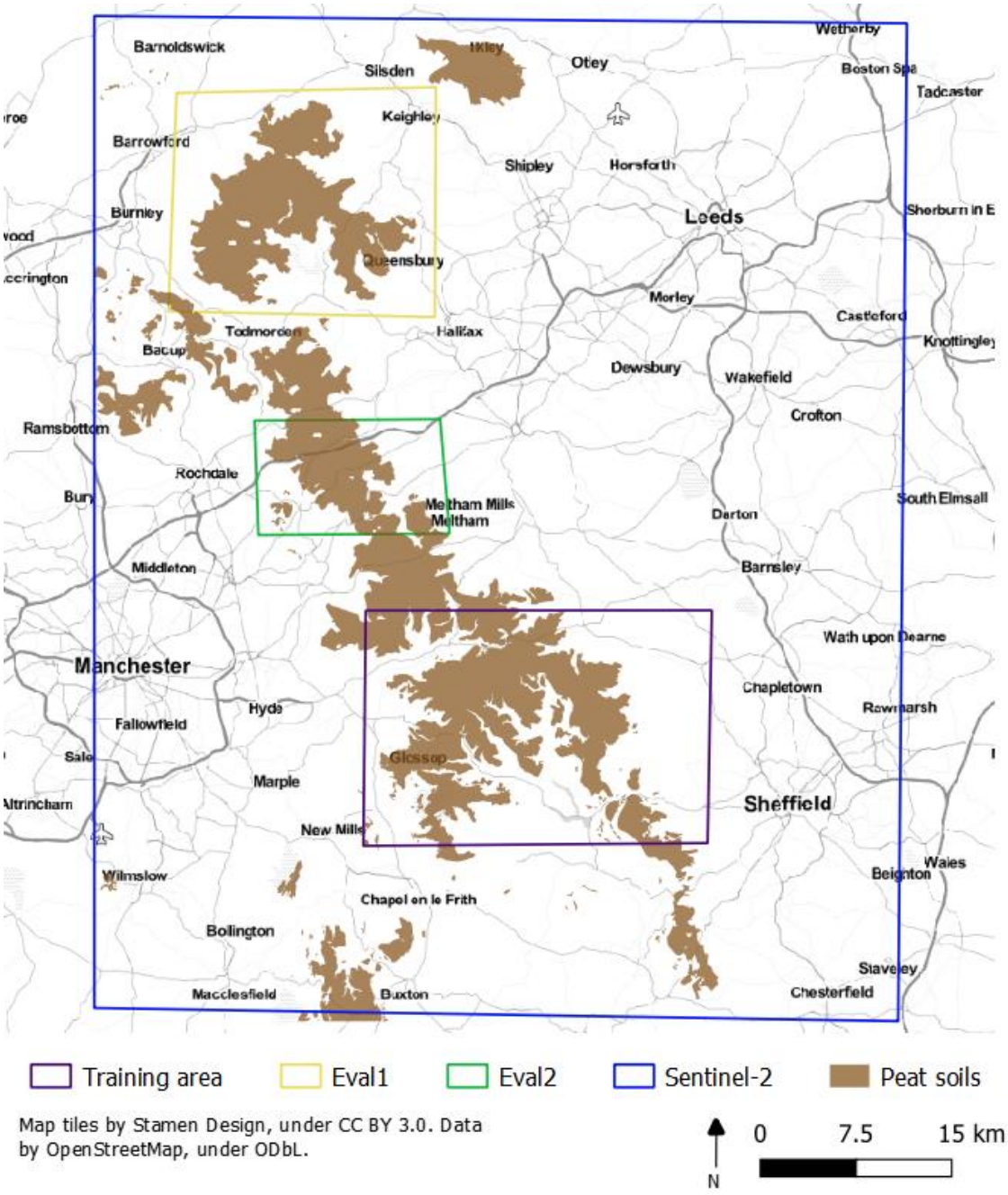


Figure 6: The areas displaying where APGB imagery tiles were obtained to assess the extrapolated model predictions across the Sentinel-2 image, derived from training points within the APGB training area.

As well as looking at the extrapolated predictions for presence and absence of bare peat, we wanted to assess the values of bare peat cover which were being generated. For this, several tiles where bare peat was clearly present from the aerial photography, were thresholded as per the training area methodology outlined in Section 2.2, to identify the bare peat pixels. These were then scaled up to give a percentage cover of bare peat at 10m resolution, which were then used to further evaluate the model predictions in the extrapolated regions.

2.7 Further experimenting of improving predictions

Further testing was carried out to see whether the predictions improved locally when trained on a greater number of tiles from within a spatial range, and overall if predictions improved across the whole area of interest when trained on data covering a range of spatial distributions. Several zones from within each of the Training, Eval1 and Eval2 areas were selected where bare peat was clearly seen in the aerial photography, with each zone made up of between one and four 1km^2 tiles.

The zones are shown in Figure 7, and regression models were trained with data from:

- Eval 1 Zone 1 – 4 x 1km^2 tiles
- Eval 1 Zone 3 – 1 km^2 tile
- Eval 1 Zone 1 & 3 – 5 x 1km^2 tiles
- Eval 1 Zone 1 & 3 and Eval 2 Zone 2 – 7 x 1km^2 tiles
- Eval 1 Zone 1 & 3 and Eval 2 Zone 2 & 3 – 9 x 1km^2 tiles
- Eval 1 Zone 1 & 3 and Eval 2 Zone 2, 3 & 4 – 10 x 1km^2 tiles
- Eval 1 Zone 1 & 3 and Eval 2 Zone 2, 3 & 4 and Training Zone 1 – 12 x 1km^2 tiles
- Eval 1, Eval 2 and Training areas -18 tiles in total: 8 x 1km^2 tiles from Training area, 5 x1km^2 tiles from Eval 1 and 5 x 1km^2 tiles from Eval2.

The results were assessed using 2000, 5000 and 10,000 training points sampled by equal intervals will category filling (method 2b). The models were also run with all available training data created from the training area and two additional areas Eval 1 and Eval 2 to see if overall this produced a higher performing model than that of a smaller sample across the three areas.

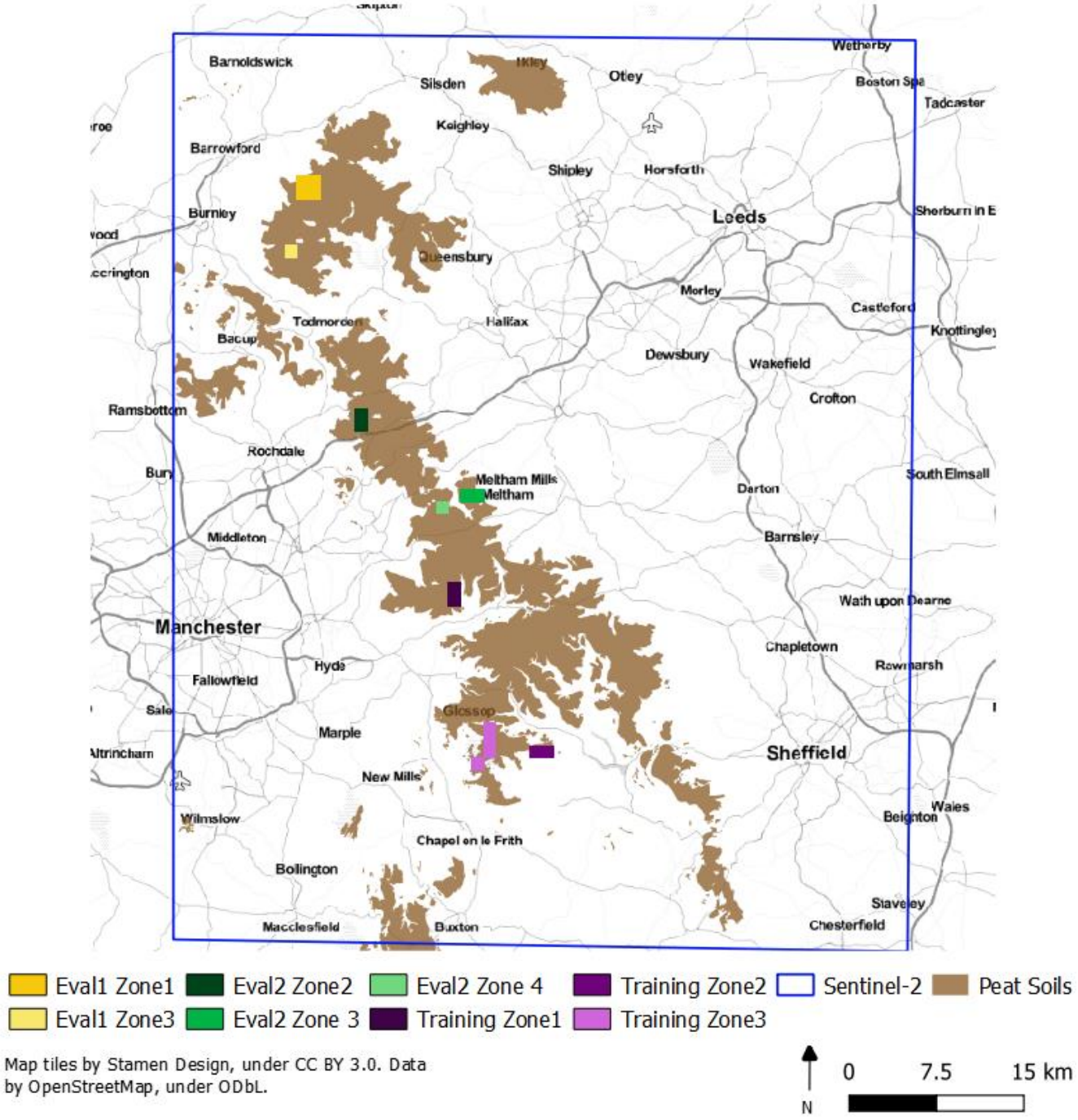


Figure 7: The model extents used to examine reach of the predictions using training data from Eval1.

3 Results and Discussion

3.1 Deriving bare peat pixels from the high-resolution imagery

The bare and vegetated peat classes demonstrated different responses between the APGB aerial photography and the Pleiades imagery. Figure 8 displays significant difference between the imagery sources, with a significant difference between the means of 54.3 in their red bands (t=199.1, p=<2.2e-16) and -56.6 in their near-infrared bands (t=-74.9, p=<2.2e-16). The reflectance values in the APGB imagery had higher mean values and greater variance within the red, blue and green bands in comparison to the Pleiades imagery. Only the near-infrared band demonstrated a higher mean value in the Pleiades data. Some of the differences in spectral values may be due to where the sensors are capturing the data, with the Pleiades data captured further from the earth's surface than the aerial photography requiring pre-processing steps to atmospherically correct the imagery. However, the APGB imagery being captured within the atmosphere means variation may arise from weather conditions during flights. These results suggest that values associated with these classes may not be comparable between the aerial and satellite derived data.

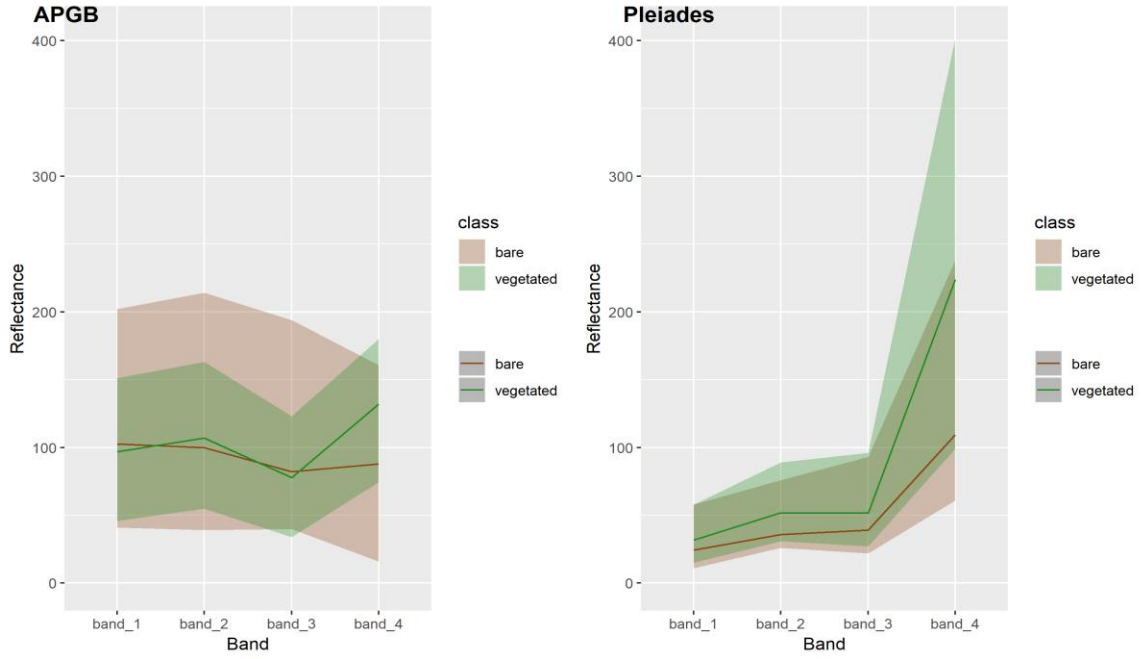


Figure 8: Spectral profiling of bare and vegetated peat training points from the APGB Aerial Imagery (left) and Pleiades Imagery (right).

The boxplots shown in Figures 9 to 12 display graphically the differences between the areas of bare and vegetated peat from the two sources of imagery, with the bold line within the box denoting the mean, the upper and lower edges of the box display the upper and lower quartile and the dots displaying outliers. In the APGB imagery, the spectral signature shown in Figure 9 demonstrates greatest differentiation between bare and vegetated peat pixels in the near-infrared band, significant increasing where vegetation was present (t=78.6, p=<2.2e-16). A significant increase with vegetation is also apparent in the green band (t=12.8, p=<2.2e-16).

In the Pleiades imagery, the vegetated pixels were significantly higher values in all the bands compared with the bare peat pixels (p=<2.2e-16). Like the APGB imagery, the near-infrared band had the greatest difference between the vegetated and bare peat classes, with a mean

difference of 114.4 (t=93.8, p=<2.2e-16), highlighting its importance in both data sources to differentiating between the two classes.

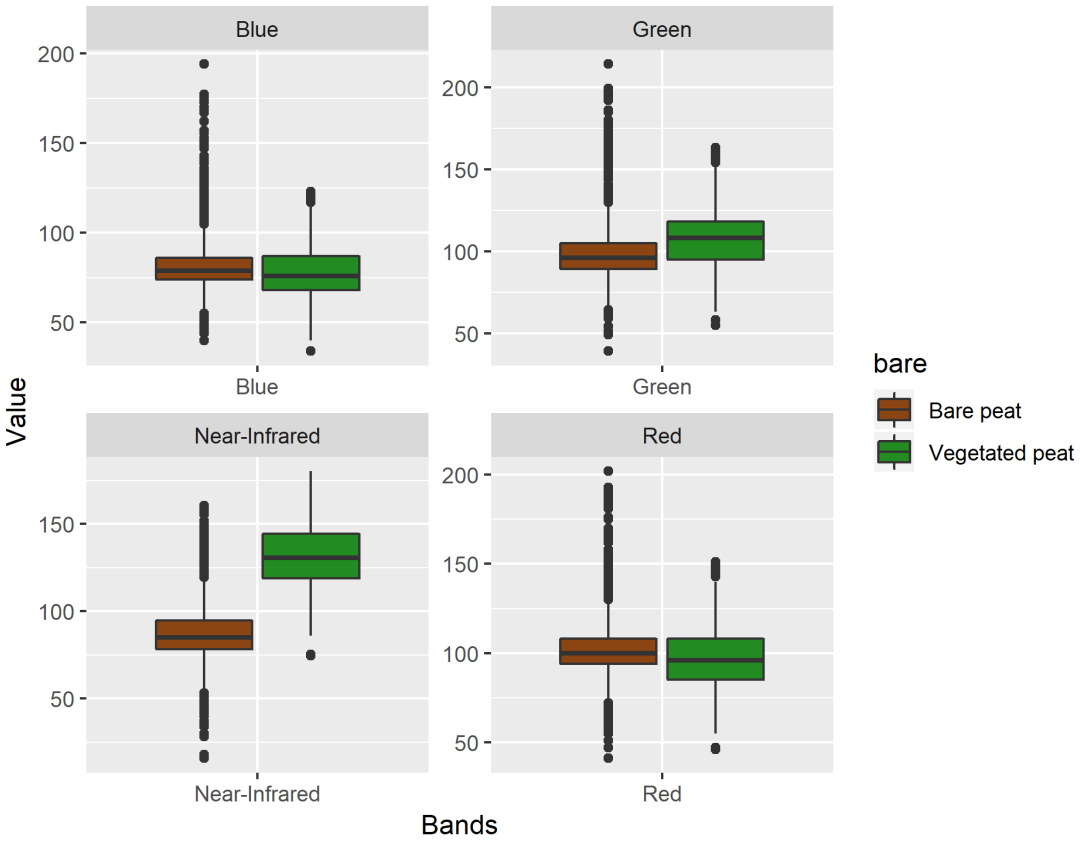


Figure 9: The spectral profile of the APGB imagery for bare and vegetated peat soils.

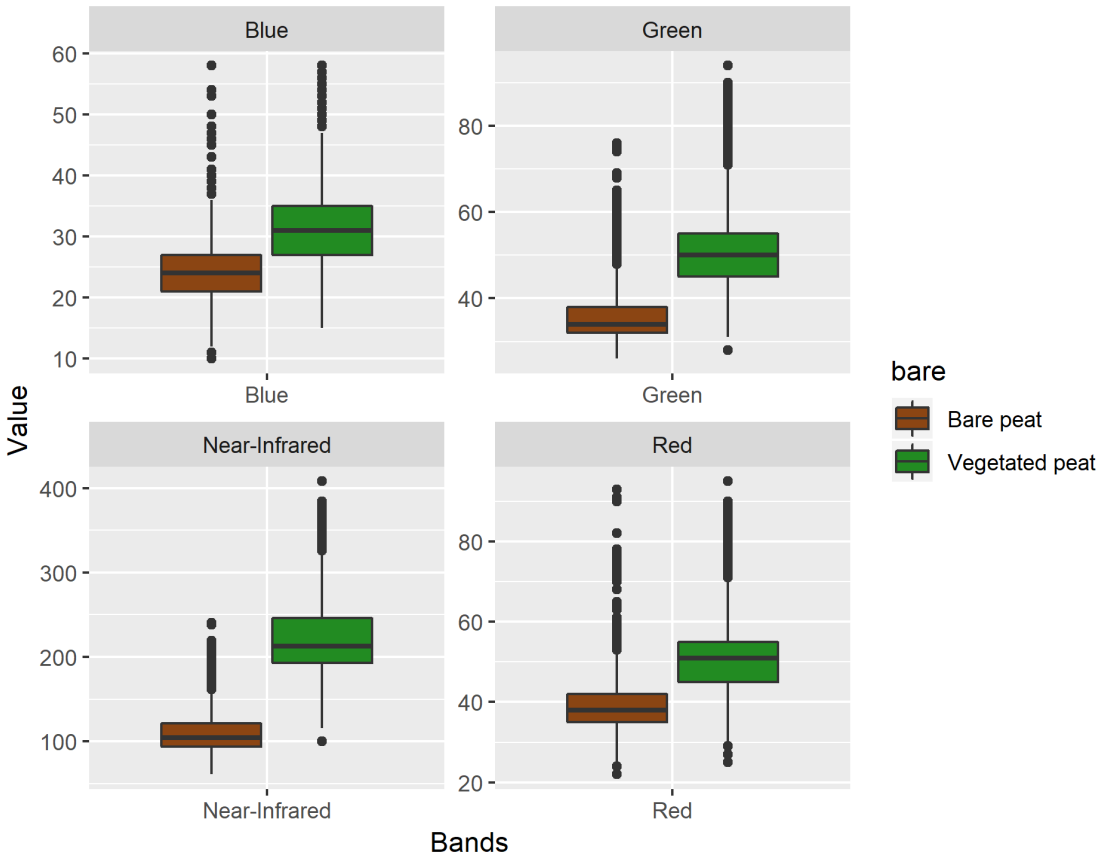


Figure 10: The spectral profiles of the Pleiades imagery for bare and vegetated peat soils.

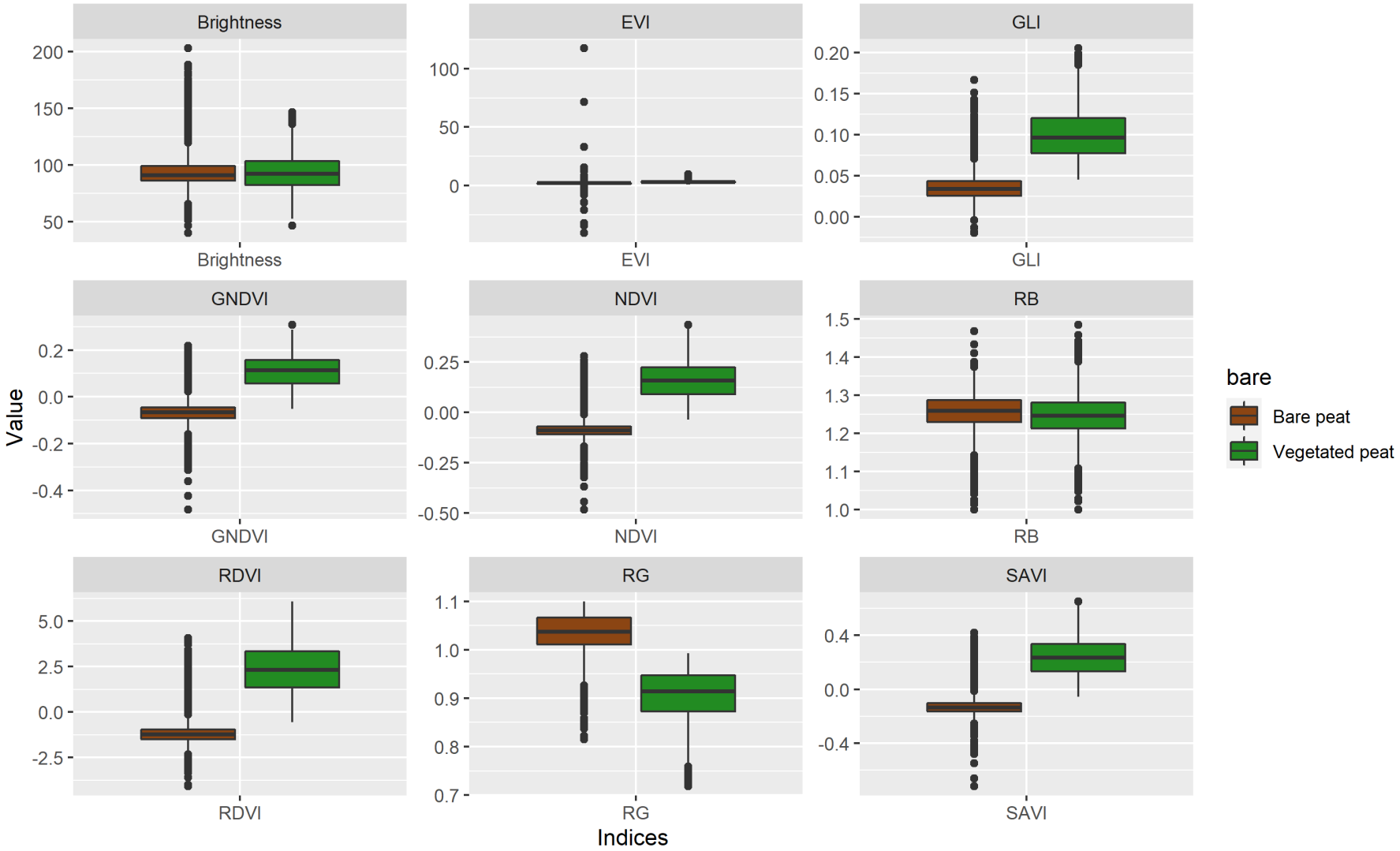


Figure 11: The indices profile of the APGB imagery for bare and vegetated peat soils.

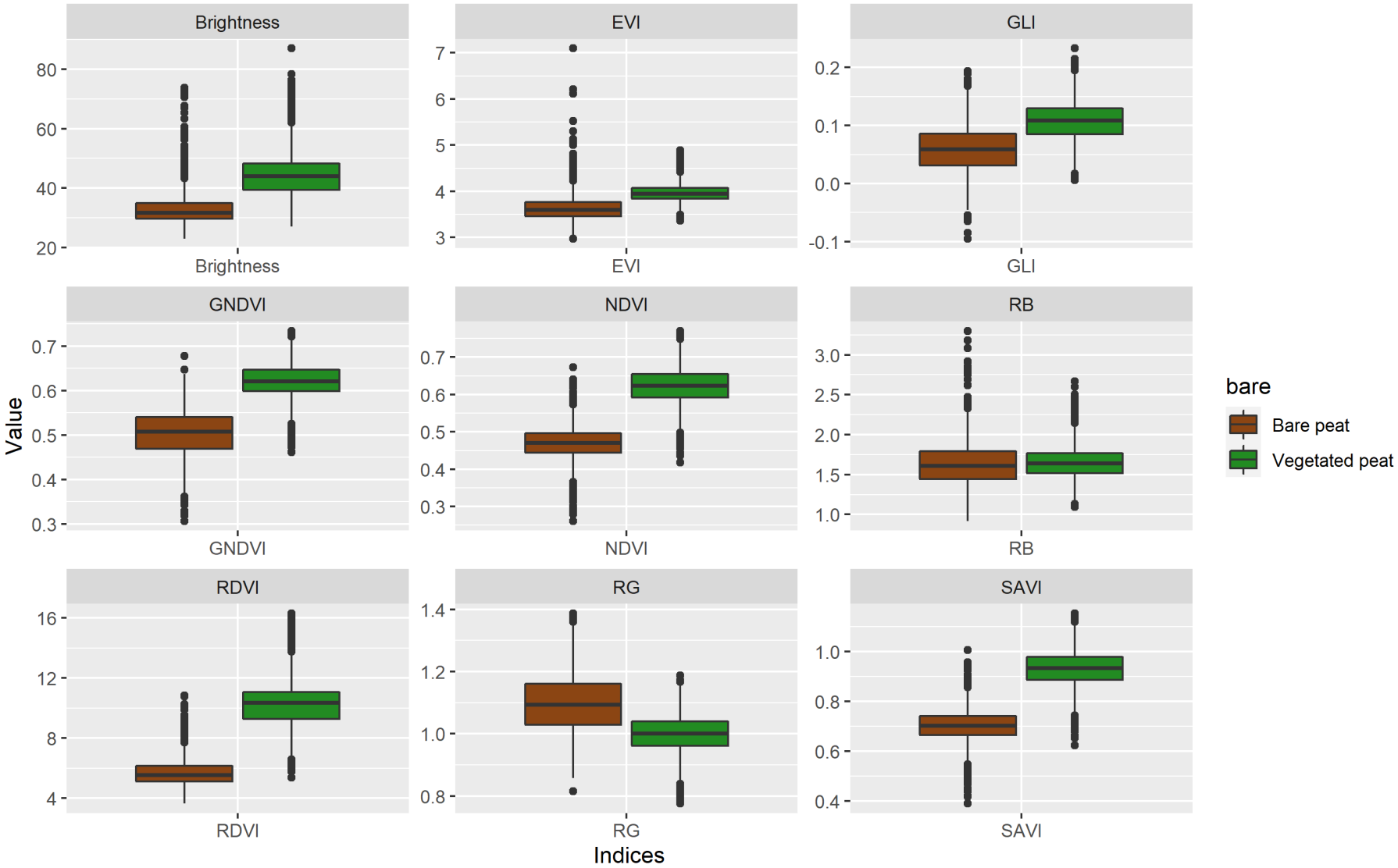


Figure 12: The indices profile of the Pleiades imagery for bare and vegetated peat soils.

Similarly, there were significant differences between the indices profiles shown with the different imagery sources, described in Figure 11 and Figure 12. Overall, the indices were significantly correlated between the two imagery sources demonstrating similar trends in their distributions (SAVI: Pearson's correlation 0.9, t=118.2, p=<2.2e-16). However, NDVI displayed significantly greater values in the Pleiades imagery when compared with the APGB imagery (mean diff=0.51, t=424.9, p=<2.2e-16).

In the Pleiades imagery, a paired t-test of NDVI values showed a significant difference of 0.15 between the vegetated and bare peat classes (t=103.8, p=<2.2e-16). However, examination of the NDVI layer did not show any discernible pattern to distinctly separate out the regions of bare peat visible in the imagery, demonstrated in Figure 13.This would suggest that although this has shown to help to determine bare soil in previous studies from remote sensing imagery, patterns with bare peat may be harder to distinguish in this type of imagery. EVI, RNDVI and GNDVI also displayed similar differences between the bare and vegetated classes, although similarly these layers failed to display a distinguishable threshold between bare peat and vegetated peat pixels. The brightness index also demonstrated a significant difference between the two classes with a mean difference of 12.0 (t=52.2, p=<2.2e-16). The narrower data ranges and no overlap in the interquartile ranges of the bare and vegetated classes may indicate its suitability to aid in distinguishing between the two classes in the Pleiades data.

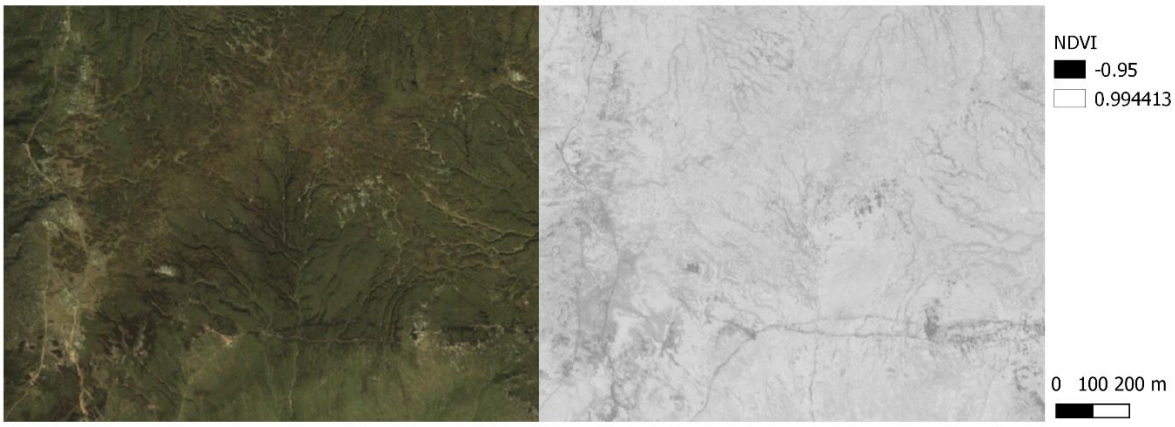


Figure 13: Eroded peatland seen in the Pleiades imagery (left) and compared to the calculated NDVI value (right).

The Pleiades imagery did not demonstrate a discernible difference in NDVI between vegetated and bare peat pixels. There were some significant differences displayed between the classes, however none of the indices alone seemed able to separate out the bare and vegetated peat in the imagery, showing overlap between class values which would be difficult to separate through thresholding. Combinations of these and the spectral bands could be further explored to determine if thresholding rules for bare peat pixels could be established with the Pleiades imagery.

The mean difference in NDVI between the vegetated and bare classes was greater and more distinct in the APGB data in comparison to the Pleiades data, with less overlapping values observed between the different peat classes. The APGB imagery displayed greater correlation visually between NDVI values and areas of bare peatland, shown in Figure 14. Figure 11 demonstrates that most bare peat pixels tended to have low NDVI values below 0. The paired t-test showed a significant difference between the means of the vegetated and bare peat classes of 0.24 (t=98.7, p=<2.2e-16). As with the Pleiades imagery, the APGB imagery also displayed significantly higher Red-Green values with the bare peat class (t-test=-83.5, p=<2.2e-16). The Red-green values displayed no overlap in their interquartile ranges with the majority of the vegetated class having values below 1.0 and the bare class

with values greater than 1.0. Similarly, the brightness index again showed little difference in the APGB data with only a slight significant difference between the means of the vegetated and bare pixels of -1.4 (t=-2.8, p=0.06).

Figure 14: NDVI threshold comparison with the APGB aerial photography.

The significant differences displayed with the Red/Green ratio and NDVI in the APGB imagery were examined for their use in automating a method of identifying the bare peat pixels. The NDVI values however displayed similar values in both bare peat pixels and hill shaded areas in the imagery. By extracting points from two tiles where the shading was particularly dominant, brightness could further help to separate the classes. Therefore, a thresholding rule proved possible to determine between the shadow pixels (where brightness values were below 0) and rocky, non-peat pixels (where brightness values were above 1), such as paths, which have similar NDVI values to bare peat. A combination of these layers could therefore be used to derive training data from the APGB imagery for the classification model using the following threshold rules:

1. Bare peat: Brightness below 100, NDVI less than 0, RB greater than 1.25, RG greater than 1
2. Vegetated peat: Brightness below 120, NDVI greater than 0, RB greater than 1.25, RG less than 1.1
3. Rock: Brightness greater than 120, RB greater than 1 and RG greater than 0.99
4. Shade: RG less than 0.9, Brightness less than 60 and RB less than 1

The class distinction present with the APGB indices but not present in those indices derived from Pleiades imagery is evident by there being notably less overlap between class ranges. This could be due to the relatively higher variance in the near-infrared band of the Pleiades data, and greater differences seen between the red and near-infrared bands, both of which were used in several of the indices calculations.

Validation against the digitised polygons demonstrated an accuracy of 83.81% (P=<2.2e-16) with 95% confidence intervals of 83.73% and 83.9%. The validation method against points created through visual assessments of the data is inherently bias towards larger patches of bare peat clearly visible in the imagery and spatially biased in its distribution. However, this does provide an indicative accuracy rating which suggests this method to be effective at detecting bare peat. Validation against 100 randomly sampled points also provided similar results with 89.8% accuracy assessed. However, the majority of these points fell within vegetated regions, with only 8% found within small patches of bare peat and of these none gave a positive prediction of bare peat.

3.1.2 Bare peat thresholding results

The results of the indices thresholding to derive bare peat is shown in Figure 15.

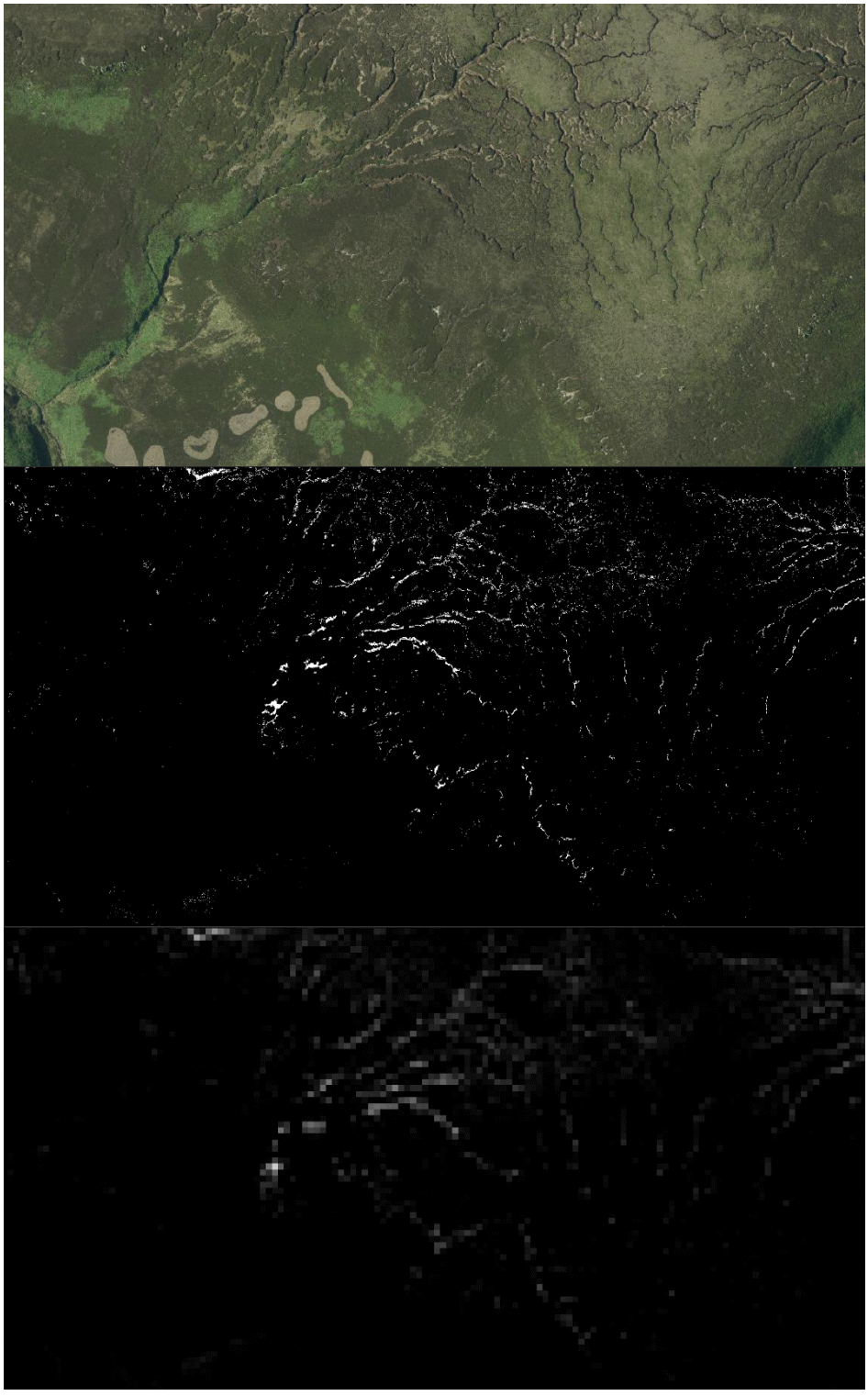


Figure 15: The images show how the bare peat training data was derived from the APGB aerial photography (top), classified into bare peat pixels using indices thresholds (middle), and then scaled to a percentage cover of bare peat at 10m pixels (bottom).

Visual comparisons show close alignment between the areas of bare peat in the aerial photography and the determined bare peat layer created through the thresholding rule, shown in Figure 15. Some uncertainty is introduced where patches of bare peat are not fully picked up in the thresholding processes and so training data may be underestimating the amount of bare peat. Similarly, as this is a pixel-based approach, slight variations between neighbouring pixel values may mean within patches some pixels of bare peat are missed due to their values falling outside of the thresholding range. Further methodologies of object-based classification alongside indices thresholding may be able to overcome this and provide a more accurate training classifier. Uncertainty is also introduced when scaling up to 10m resolution, as this assumed that the bare peat coverage in each 25cm cell of the training data layer is 100%, where, in actuality, this may not be the case.

The performance of the derived bare peat layer through thresholding was examined against a random forest classification, to see how the automated methodology compared with other common methods. The random forest classification informed by digitised bare peat polygons, demonstrated a model accuracy of 97.55% with a user accuracy of 92.96% for the bare peat class. Out of the four spectral bands tested the NIR band was found to have the greatest importance in determining the model predictions (Gini imp. 147732.82). The classified tiles showed visual similarity with the classification derived from our thresholding methodology, shown in Figure 16, drawing out similar bare peat pixels.

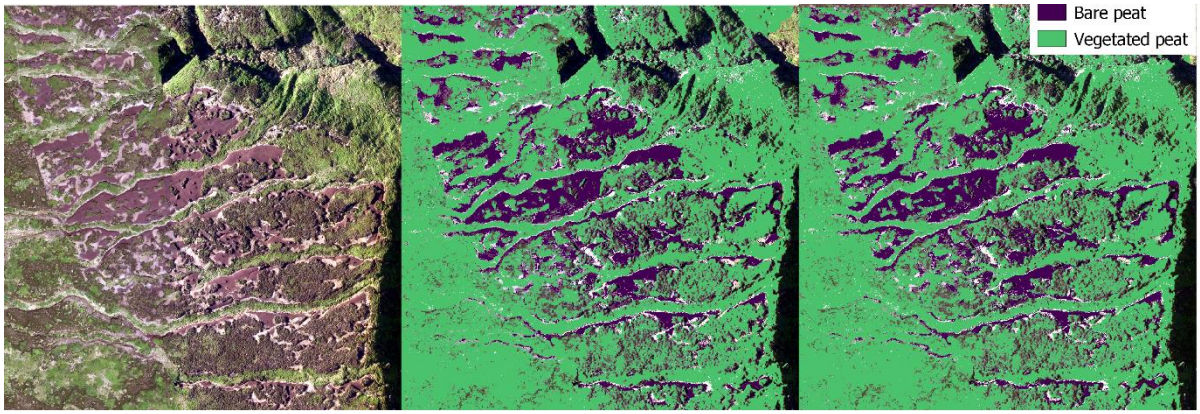


Figure 16: Comparison of the APGB aerial photography tile SE0500 (left) with the thresholded image (middle) and random forest classification prediction (right).

3.2 Regression model methodology to upscale bare peat coverage to Sentinel-2

3.2.1 What is the best methodology for deriving sample data?

The bare peat training data were scaled up from the high-resolution imagery to a percentage cover per sentinel-2 pixel, shown in Figure 15. Values ranged from 0% to 100% bare cover, the distribution of which is shown in Figure 17. The training area covered 494km^2 with the majority of data values falling below 5% of bare peat cover. The methodology for sampling the training data greatly influenced the range of values being used to train the models and the resulting range of predicted cover values.

Where points were sampled randomly under method 1, this led to underprediction where the training data were trained and tested against very low cover values, resulting in predictions with a vastly lower range of 0 – 40% cover compared with the training data. When the output predictions were corrected, the range of values still suggested underestimation of the bare peat, with a maximum cover value of 42%.

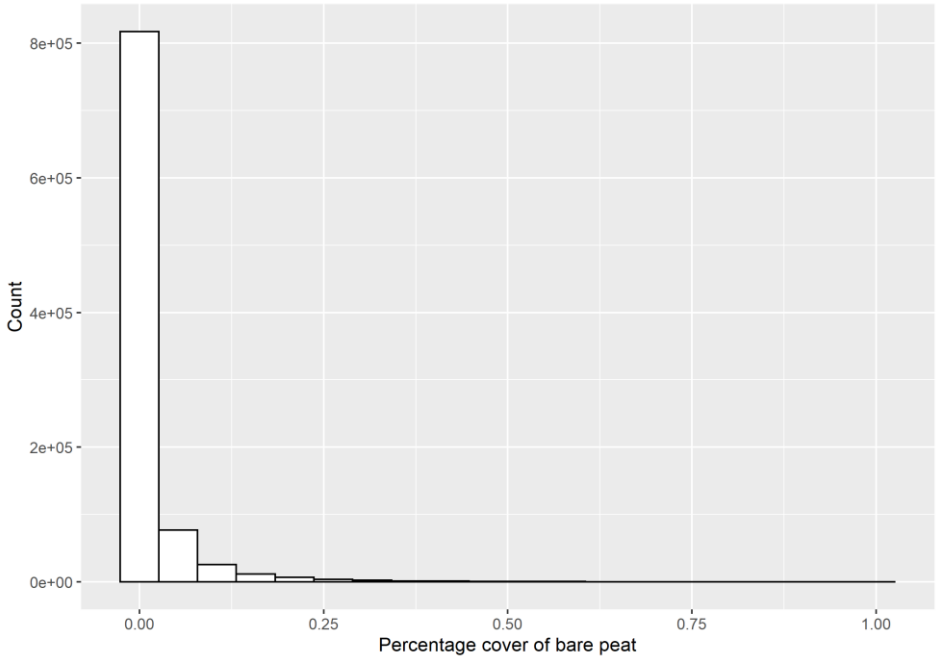


Figure 17: A histogram of the distribution of bare peat coverage training data

Methodologies 2, 2b and 3 aimed to sample a greater distribution of values within the training data, reducing the overinclusion of 0-5% coverage. The results of all four methodologies are shown below in Table 2, modelled with 10,000 target training points using all the variables and a Random Forest (RF) algorithm, with RFC denoting where randomised results have been corrected.

Table 2: A comparison of the different sampling methodologies.

Method	**1 – Random**		**2 – Equal interval**	**2b – Equal interval with category filling**	**3 – Quantiled sampling of values above 0.1**
Max tries	5	-	5	5	
Target training points	10000	-	10000	10000	10000
Training points	7500	-	5263	7500	7500
Test points	2500	-	1756	2500	2500
Best RMSE model	RF4	-	RF1	RF1	RF5
Best RMSE value	0.035	-	0.116	0.109	0.088
Best R-squared model	RF3	-	RF1	RF1	RF3
Best R-squared value	0.408	-	0.75	0.747	0.534
Model results	RF	RFC	RF	RF	RF
Mean RF RMSE	0.037	-	0.117	0.111	0.089
Eval1 RMSE	0.054	-	0.15	0.135	0.116
Eval1 R-squared	0.152	-	0.057	0.094	0.072
Eval1 P/A Accuracy	0.497	0.512	0.610	0.631	0.626
Eval2 RMSE	0.04	-	0.147	0.125	0.106
Eval2 R-squared	0.01	-	0.016	0.028	0.028
Eval2 P/A Accuracy	0.525	0.468	0.563	0.570	0.563
Mean P/A Accuracy	0.511	0.490	0.587	0.600	0.594

The results demonstrated that, overall, the RMSE performance metrics were small suggesting good fit to the observed test data. The random sampling methodology had the lowest RMSE performance scores against the test datasets both within the training and extrapolated areas (Eval1 and Eval2), which may be due to the large number of low values being assessed in the test datasets in all three areas. The R-squared results would suggest that within the training area, the equal interval stratified sampling methods 2 and 2b perform much better than the other two methodologies at accounting for the variance in the observed data and fit the regression line more closely. We also see these methods showing greater accuracy within the extrapolated areas when looking at the presence and absence measure of bare peat. This can also be seen in the predicted bare peat cover layers where the predictions from methodologies 2 and 2b show closer alignment with the range of values of the training data. Figure 18 demonstrates the underpredictions seen with methodologies 1 and 3, despite all the predictions displaying a similar pattern of cover.

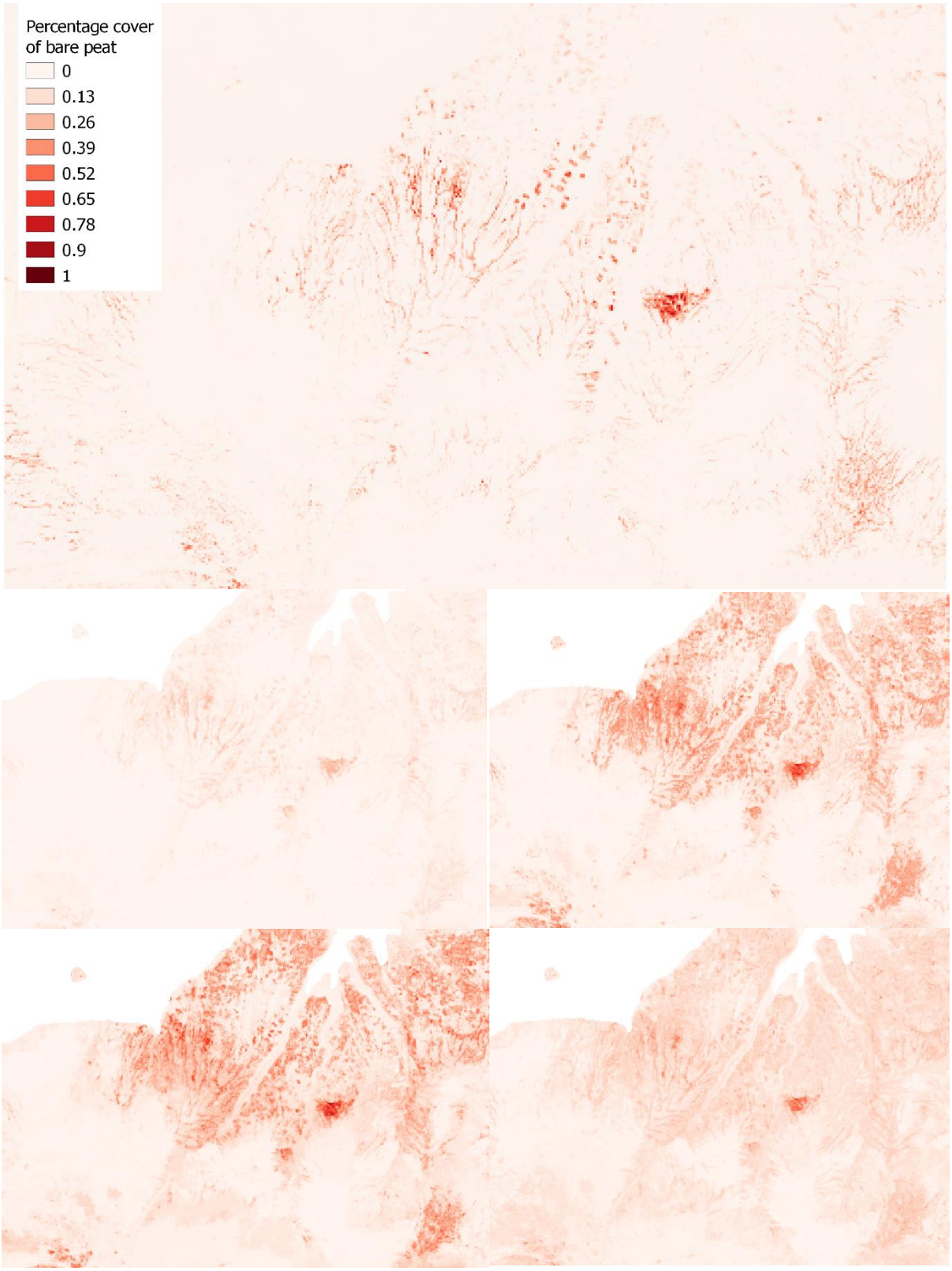


Figure 18: A comparison of the predictions resulting from the four sampling methodologies mapped to the same scale; the training data (top), and the predictions from method 1 (middle left), method 2 (middle right), method 2b (bottom left) and method 3 (bottom right).

3.2.2 Which predictors are important to include in the regression analysis?

The correlation matrix of the predictor variables, Figure, highlighted the variable layers from Sentinel spectral bands, derived indices from these bands and other variable data sources which were highly correlated. Unsurprisingly the sentinel bands and indices were highly correlated as the indices were derived from these data, having high negative and positive relationships.

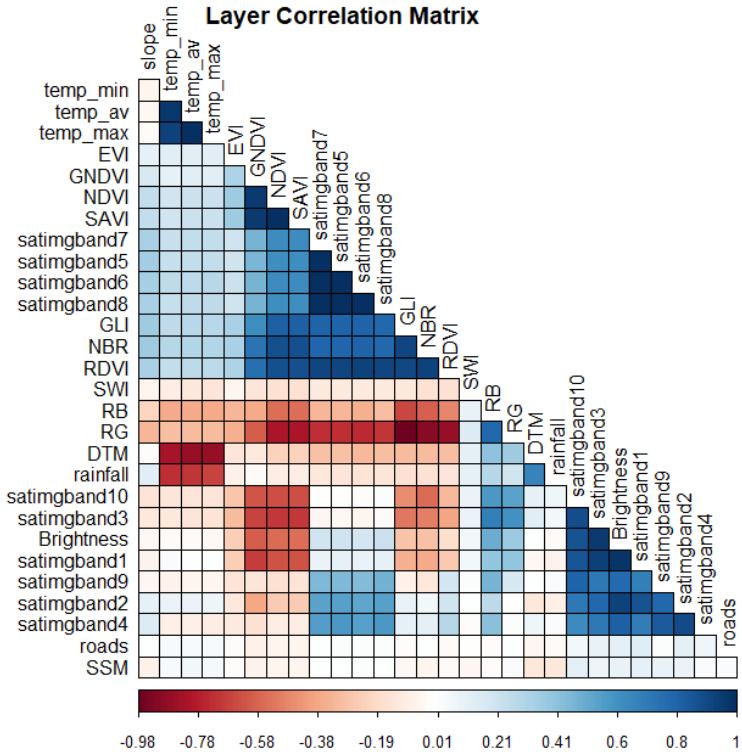


Figure 19: Comparison matrix of the correlation between variables trialled in the regression analysis, with 0 representing no correlation, -1 being highly negatively correlated and 1 being highly positively correlated.

The most important predictors when all the variables were included in the model, were the Red/Green Ratio (63), with Green Leaf Index (GLI) (50) and the Near infrared (narrow) band of the Sentinel Imagery. This is unsurprising as all three have been found to be indicators relating to detecting vegetation. The red and green optical bands are known to be sensitive to canopy colouration and ratios have been explored for their use in phenological vegetation discrimination (Motohka *et al*. 2010). Similarly, the near infrared and short-wave infrared bands are commonly used in indices calculations such as for NDVI, which is highly associated to plant productivity and canopy chlorophyll density (Zhao *et al*. 2007; Rajah *et al*. 2019). GLI is a widely used spectral vegetation indices, derived from the red, green and blue optical bands, known to relate to vegetation cover and chlorophyll content (Hunt Jr *et al.* 2013)

Other variables which were highlighted as having some importance in distinguishing bare peat cover were rainfall, DTM, slope and the Short-wave Infrared 2 band. These factors are likely to influence where areas of bare peat is located as a lack of precipitation would lead to less water availability for plants to utilise, which could result in less vegetated peatlands. The DTM and slope showed visual correlation with where the peat soil boundaries were located. Greater slope has been shown to influence erosion, and so this could also lead to patches of bare unvegetated peat, where soils regularly experience erosion from wind and precipitation

(Li *et al.* 2018). The short-wave infrared band in Sentinel-2 sensors are also sensitive to cellulose and lignin absorption, which may help explain its importance in detecting bare peat, with obvious differences in vegetation and dry matter expected (Pasqualotto *et al.* 2019).

Table 3: A comparison of the different variable combinations used in the regression models.

Combination name	***Varstack_all***		***Varstack_ climAv***		***Varstack_darkveg***	
Variables	All variables		Brightness, RB, RG, NBR, slope, monthly rainfall, average monthly temperature		Brightness, NDVI, RG, Near infrared (narrow), Green, Red-edge 1, slope, NBR, monthly rainfall	
Method	2	2b	2	2b	2	2b
Best RMSE model	RF1	RF1	RF1	RF5	RF1	RF1
Best RMSE value	0.116	0.109	0.122	0.115	0.12	0.114
Best R-squared model	RF1	RF1	RF1	RF5	RF1	RF1
Best R-squared value	0.750	0.747	0.717	0.716	0.737	0.718
Most important variable	RG	RG	RG	RG	RG	RG
Eval1 RMSE	0.150	0.135	0.078	0.077	0.114	0.105
Eval1 R-squared	0.057	0.094	0.074	0.078	0.142	0.150
Eval1 P/A Accuracy	0.610	0.631	0.636	0.626	0.641	0.662
Eval2 RMSE	0.147	0.125	0.062	0.06	0.079	0.072
Eval2 R-squared	0.016	0.028	0.024	0.027	0.044	0.042
Eval2 P/A Accuracy	0.563	0.570	0.524	0.542	0.577	0.587
Mean P/A Accuracy	0.587	0.600	0.580	0.584	0.609	0.624

The variable combinations which demonstrated the best predictive accuracy with training points sampled by equal interval methodologies are summarised in Table 3. The combination named ‘varstack_darkveg’ demonstrated the best fit with the lowest RMSE values and highest accuracy at predicting presence and absence of bare peat in the extrapolated areas. This included the Red/Green ratio as highlighted earlier, but also NDVI and Near-Burn Ratio indices which have been used in studies to determine the amount of vegetation productivity and highlight the removal of vegetation, reliant on information from the Red-edge and Near infrared spectral sensors (Amos *et al.* 2018). This combination also included the slope and monthly rainfall which similarly displayed high importance in identifying bare peat pixels.

3.2.3 What is the best performing machine learning technique for extrapolating the data?

Random Forest consistently demonstrated to be the best performing model having more accurate predictions of bare peat in the resulting maps in comparison to the other machine learning algorithms trialled. Table 4 displays the results from the different machine learning algorithms, training each model on 100,000 training points using the variables Brightness, RB, RG, NBR, SSM and Slope. Overall, RF and QRF displayed greater predictive performance, with RF performing slightly more consistently across the models with lower RMSE values and higher R-squared values. This can be seen in the results in Table 4 and further evident in the results found in Appendix 2. The mapped outputs from each model are shown in Figure 20. Interestingly the range of values differed vastly between the predicted results, despite being trained of the same training data. The results for SVM suggests this method greatly underpredicted the bare peat coverage. The comparison between RF and QRF interestingly show less variability in the QRF results, with crisper definitions of the locations of high bare peat cover.

Table 4: A comparison of the machine learning algorithms used in the regression models.

Model	*Mean RMSE*	*Eval1 RMSE*	*Eval1 R-squared*	*Eval1 P/A Accuracy*	*Eval2 RMSE*	*Eval2 R-squared*	*Eval2 P/A Accuracy*
RF	0.094	0.055	0.142	0.533	0.026	0.031	0.531
QRF	0.101	0.055	0.143	0.472	0.026	0.030	0.503
BRNN	0.108	0.154	0.005	0.513	0.197	0.000	0.514
SVM	0.144	0.047	0.096	0.451	0.027	0.299	0.503
BRT	0.113	0.055	0.096	0.528	0.032	0.025	0.539

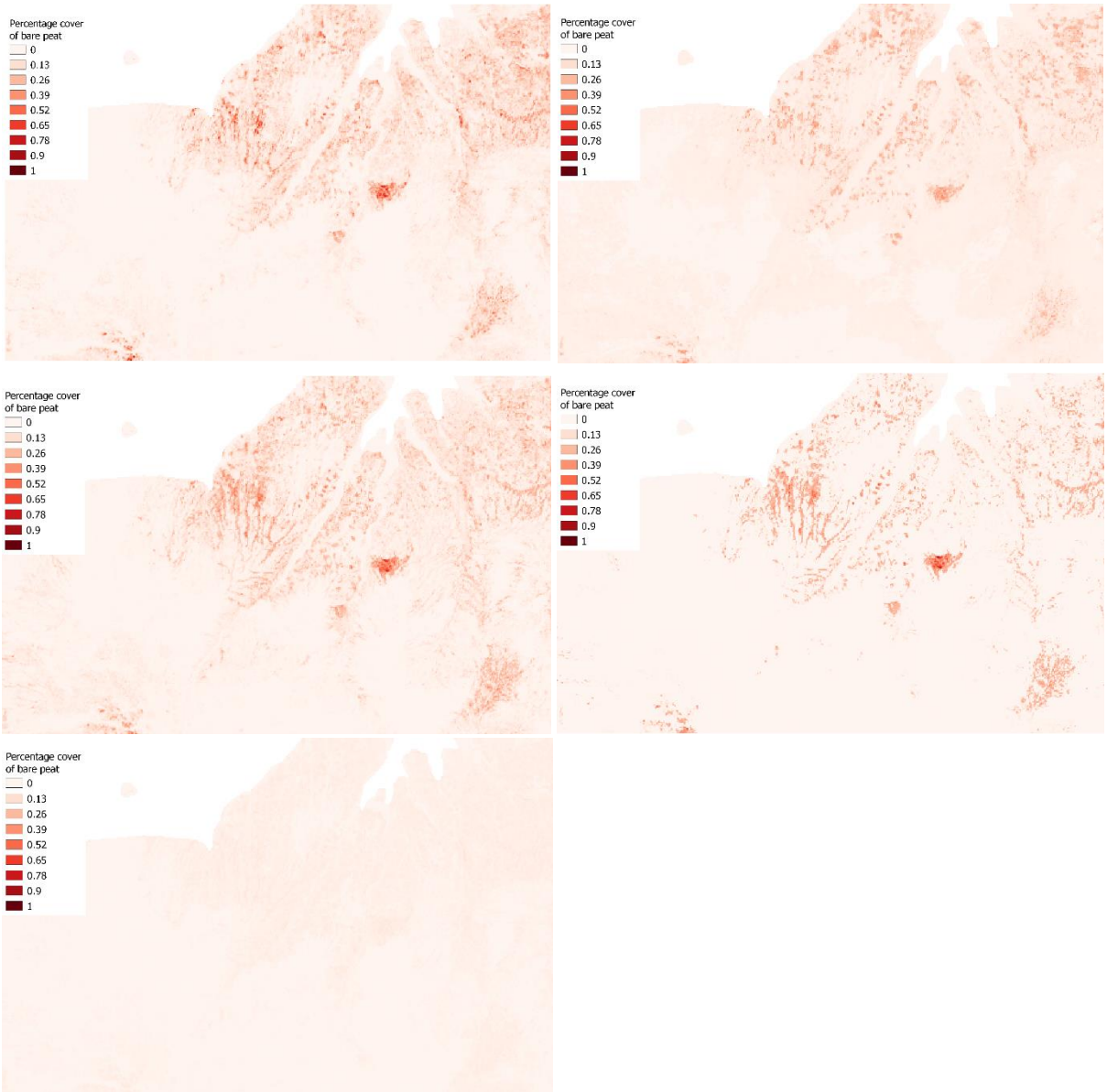
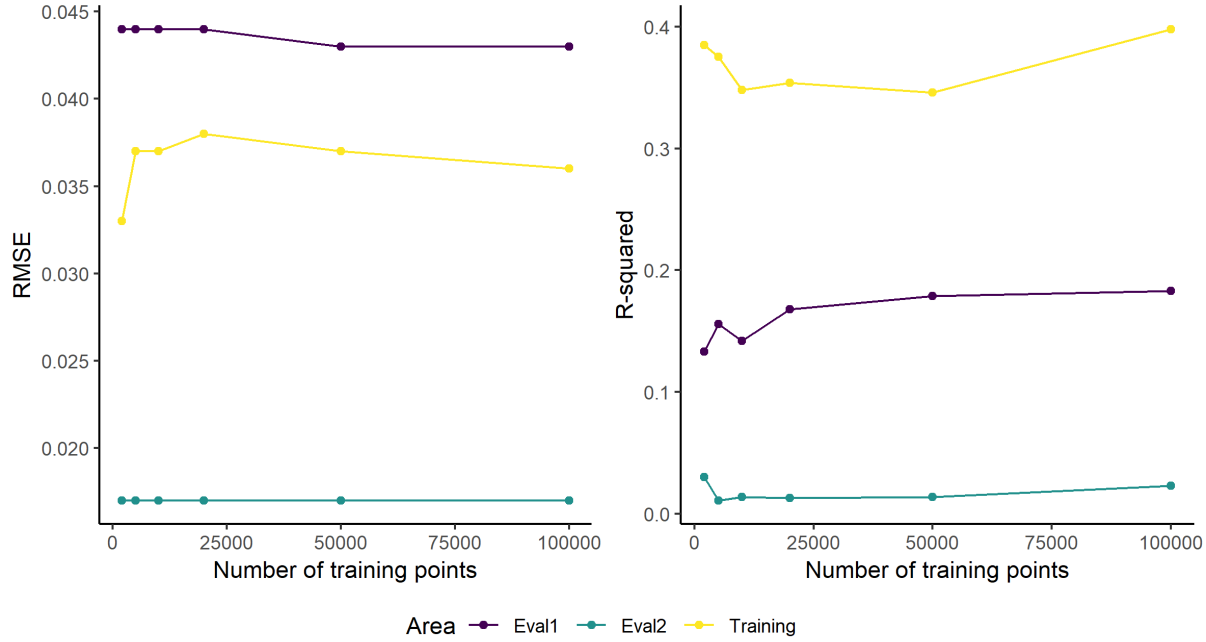


Figure 20: A comparison of the predicted maps from the different machine learning algorithms; BRNN (top left), BRT (top right), RF (middle left), QRF (middle right), and SVM (bottom).

3.2.4 What is the minimum number of images/training points required to make accurate predictions across a given extent?

The models trained on data from the training area were trialled with different numbers of training points, to see if including a greater number of training points produced a more accurate prediction when extrapolating predictions across the whole sentinel extent. The results were highly variable across the different sampling methodologies modelled with differing numbers of training points, as seen in Figure 21.

Method 1:



Method 2:

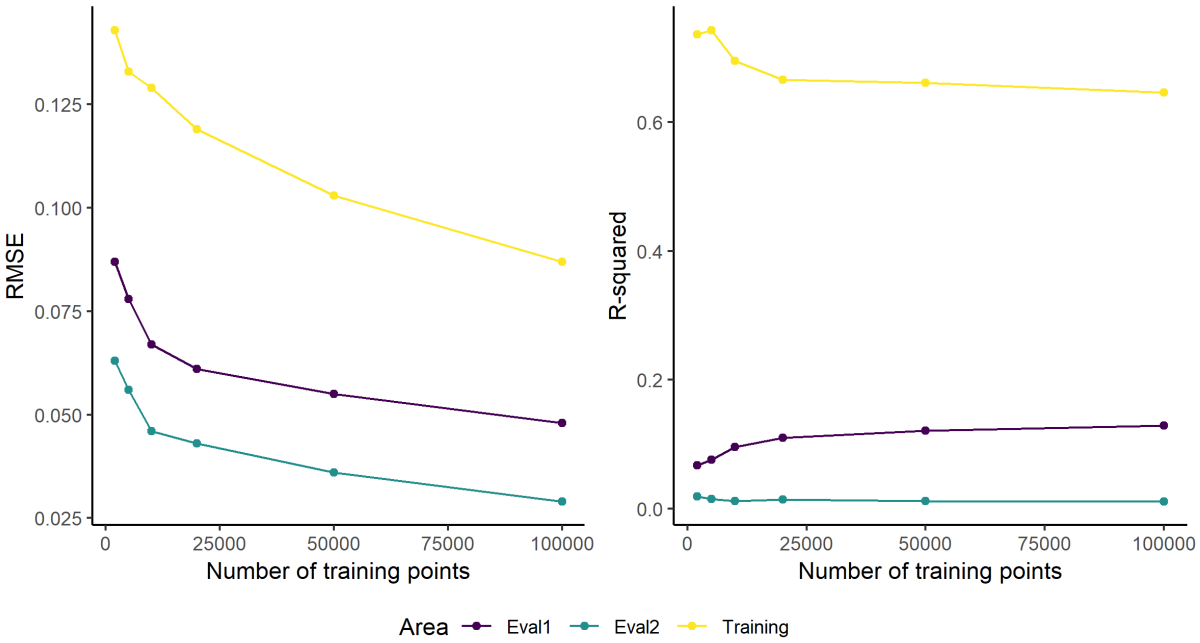


Figure 21: A comparison of model performance with different numbers of training points when the bare peat cover training data were randomly sampled (Method 1) and sampled using categorised equal intervals (Method 2).

The Random sampling methodology (Method 1) demonstrated greater model accuracy in predicting the presence and absence of bare peat in the extrapolated areas with 20000 training points (0.472% Accuracy). However, when the cover values were assessed with the RMSE and R-squared metrics, these results were variable with 100000 points having greater predictive strength in Eval 1 and 2000 points in Eval 2. With the equal interval sampling methodology (Method 2), overall, the RMSE scores decreased with a greater number of training points, suggesting a better fit of the models. The results varied however, with 2000 and 5000 training points seeming to provide a greater predictive fit in Eval2 with higher R-squared scores (RMSE 0.063,0.056) and increased accuracy in predicting the presence and absence of bare peat pixels in both the extrapolated areas (0.578%, 0.584%). However, training the models with 100,000 points seemed to provide a better fitting model in Eval 1. The results of the point measures were inconclusive, suggesting that model accuracy does not necessarily increase with a greater number of training points and the distribution of the training data may play a greater role in determining the predictive strength of the models.

The spatial distribution and number of imagery tiles included in the training data was explored through using different combinations of data from the three of the APGB acquired regions; training, Eval 1 and Eval 2. Training and predicting the models first in the Eval 1 area, towards the north west of the area of interest, it was found that the models had performed better when trained on four 1km^2 tiles in zone 1 compared with on a single 1km^2 tile of zone 3 (Z1 r^2 0.615, Z3 r^2 0.444). Interestingly, when these tiles were combined to train the models on all 5 tiles, the RMSE value increased to 0.063 but the R-squared value decreased to 0.583, suggesting better fit to the test data despite capturing less variation compared with zone 1. This may have been due to the bare peat cover in the zone 3 having a low range of values, with lower levels of bare peat present.

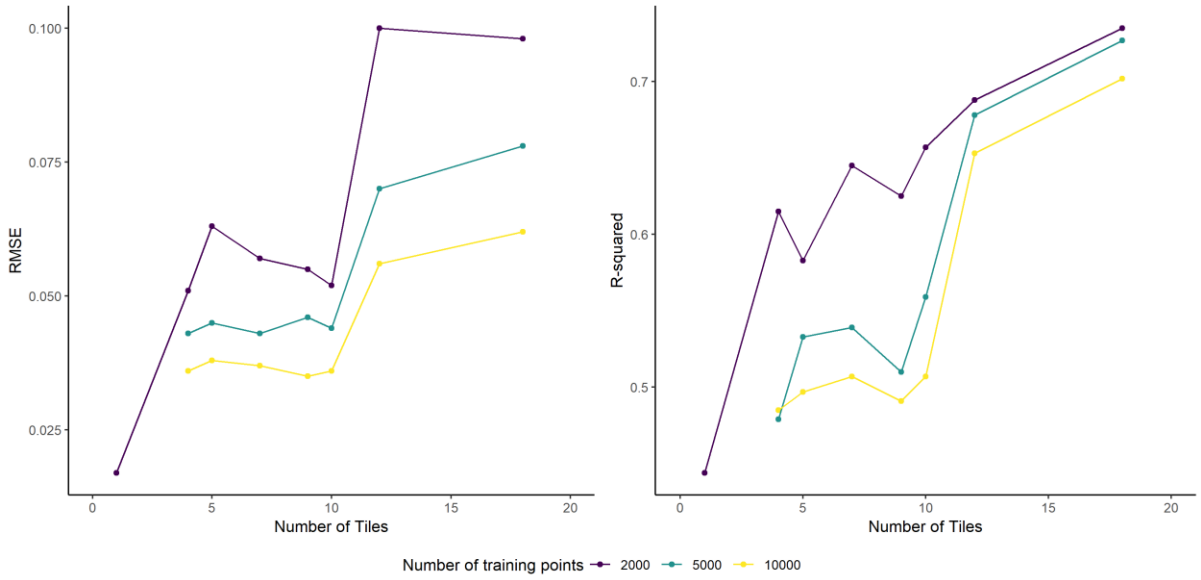


Figure 22: A comparison of model performance with training data derived from different numbers of 1km tiles of the APGB high-resolution imagery.

When training data were sampled from both Eval 1 and Eval 2 regions, we see a slight increase in the R-goodness of fit of the models, as shown in Figure 22. This suggests using training data with greater spatial variation can help the model to better account for variation in the data and lead to more accurate predictions of bare peat cover. The amount of bare peat visible in the Eval 1 and Eval 2 areas were fairly low, with mainly smaller patches of bare cover present, with a small range of training data values. This can be seen from the distributions of training data extracted from individual tiles, displayed in Figure 23, in graphs A-D. The vast difference seen when tiles from the training area are added (Figure 23, graphs

E-F) are evident by increases in both the RMSE and R-squared metrics when data were sampled with between 12 and 18 tiles. The R squared values notably increased, particularly when sampling 5000 and 10000 training points, likely due to more points being sampled with higher bare peat cover values. Graphs E and F, in Figure 23, demonstrate the distribution of data in one of the 1km tiles in the training area, showing data with 50-100% peat cover. The improvement in model performance with data from these areas suggests that the distribution of the training data greatly influences the modelled predictions. These findings highlight that the accuracy of our modelled predictions increase with the inclusion of data across a wider spatial distribution and of representative values across the range of high and low percentage cover of bare peat.

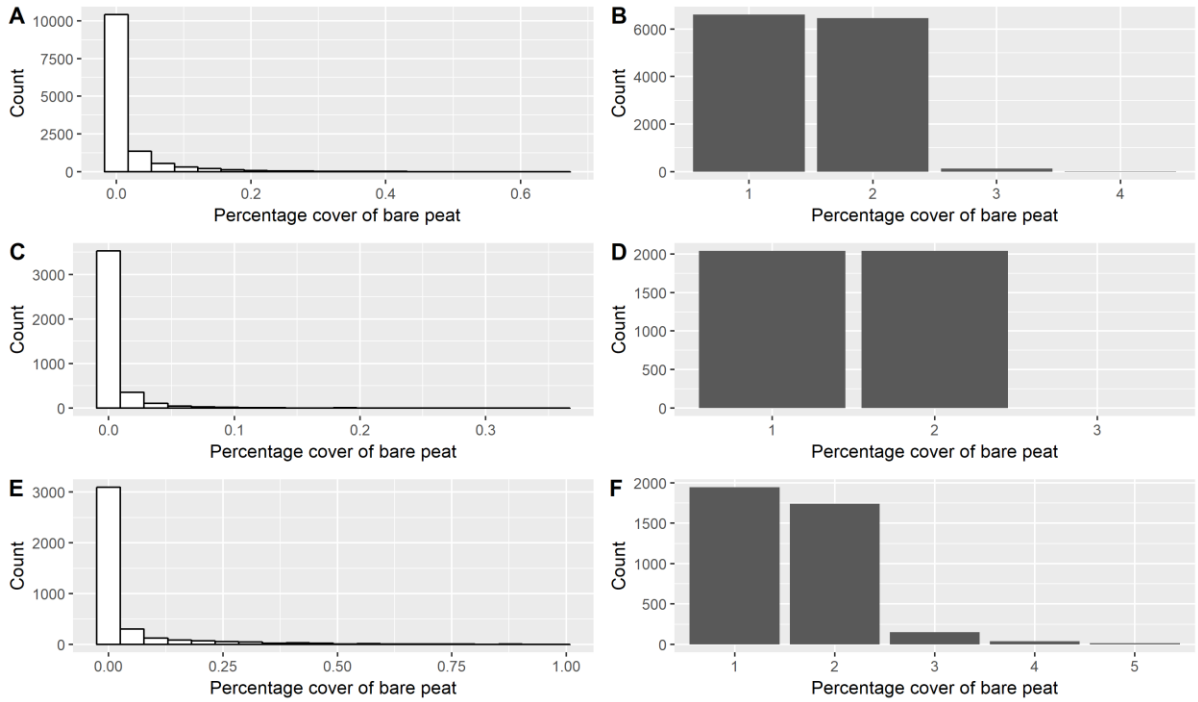


Figure 23: Plots of the distribution of training data points from within single 1km tiles in Eval1 (A&B), Eval2 (C&D) and the Training (E&F) areas, shown as histograms of the raw (A&C&E) and categorised (B&D&F) data.

The regression models run with all of the available training data from the three areas, had a wide distribution of predicted values ranging from 0 to 100% coverage and a variance of 0.0022. Compared with the model developed from only within the Eval 1 area, the model trained on all available training data produced a better fit to the regression, overall capturing more variation in the data with R^2 values of 0.718 compared to 0.497. Differences in the importance values were also notable, with the influence of variable factors of rainfall and slope having a greater importance locally with the Eval 1 model than compared to training on all three areas, where these factors are relatively less important than the influence of near-infrared band. The limited range of values found in the test data derived from Eval1 may explain these trends and although it shows to accurately predict in the area it is trained on, in areas these predictions are extrapolated to it performs poorly. This is evident when looking at the bare peat cover predictions in an extrapolated region roughly 13km East of Eval1, where the local model seems to underpredict in comparison to the model trained on all three areas; Eval1, Eval2 and the Training area. Interestingly, the bare peat predictions seem more accurate when trained on data from only the Training area, an approximate distance of 40km south from this site. Figure 24 shows these results, suggesting that when extrapolating predictions, the distance from the site is less influential than the distribution of values provided in the training data. This is also apparent from the accuracy of bare peat presence

in Eval 2, which is located roughly midway between Eval 1 and the training area. Accuracy tended to be higher with 59% accuracy in predicted presence of bare peat when modelled with the training area dataset as opposed to 55% accuracy when modelled with the Eval 1 training dataset.

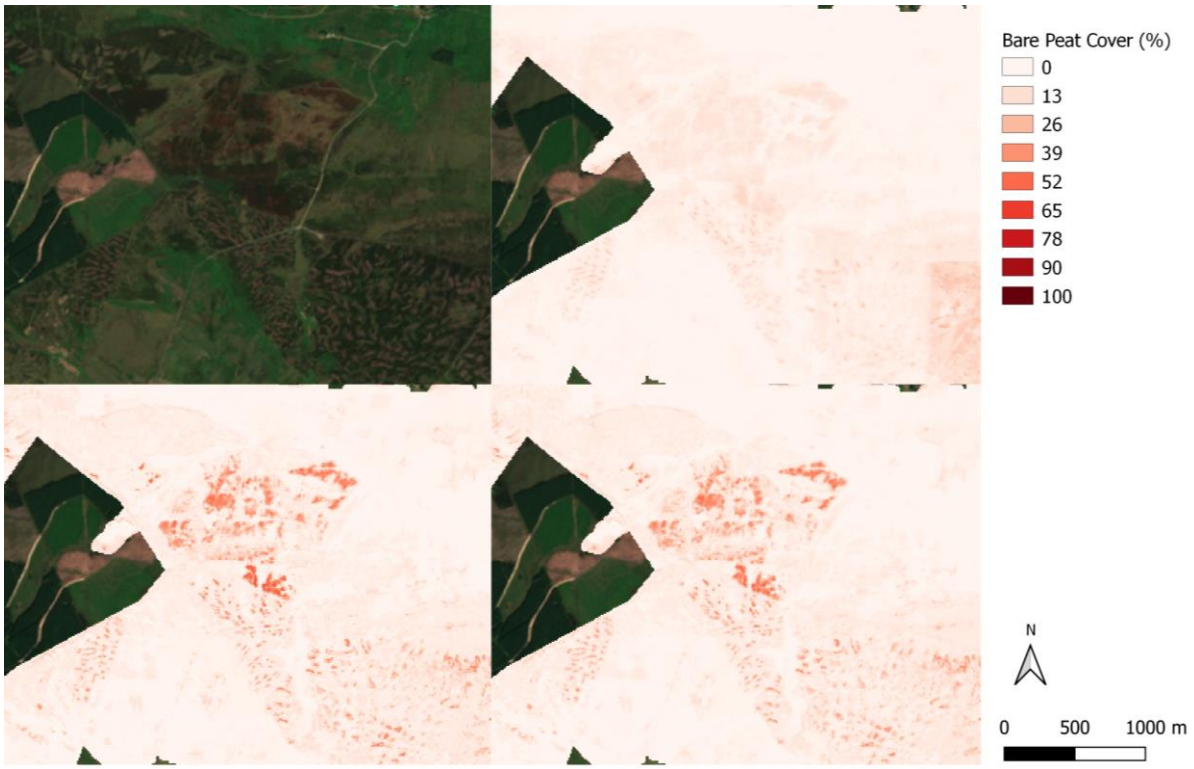


Figure 24: A comparison of predicted bare peat cover for Bare Hill, Ilkley Moor, displaying the Sentinel optical imagery (top left), modelled predictions when trained on data only from Eval 1 (top right), only from the training area (bottom left) and when trained on data from across all the three regions (bottom right).

Model performance was shown to be greater when data were sampled from all available tiles as opposed to the 18 spatially distributed tiles used in the above analysis, with R^2 values of 0.836 compared to 0.702, as well as higher accuracy scores when predicting presence and absence from observed polygons in the Eval 1 and 2 areas. The models were less influenced by factors of rainfall and slope with higher relative importance of the spectral band and indices predictors. The output prediction maps visibility aligned better to areas of bare peat within the training area, highlighting the need to include more training tiles in areas of high bare peat cover. This reiterates the importance of including imagery tiles from a range of high and low peat cover, with the inclusion of more tiles in areas known to have high amounts of bare peat helping to improve the detection of bare peat overall, particularly in extrapolated areas.

4 Conclusions

The workflow outlined in this study further developed upon the work of Williamson *et al*. (2018), exploring the practicalities of using earth observation data to identify bare peat pixels and scaling predictions across a wider landscape.

We have presented that it is possible to partially automate a process of deriving bare peat pixels from high-resolution aerial photography, through thresholding indices derived from the imagery. However, this process was informed by a manual procedure of digitising areas of vegetated and non-vegetated peat, and so would require some expertise to assess these in the imagery and develop thresholding rules which could be applied to multiple tiles. It was found that threshold values which would identify bare peat pixels in one image, would not necessarily be applicable across multiple images due to variation in brightness and reflectance levels, hill-shading and variates of vegetation present. This meant that thresholding rules were case specific and would not necessarily be transferable between sites requiring manual input to assess individual images. Differences may also be explained through flights capturing the imagery being conducted over different days, atmospheric conditions or blurriness from wind and sun illumination. Further work could help to explore these differences within the imagery derived from aerial photography and look to identify if there is a more consistent methodology with which bare peat can be identified. The sensor capture of the APGB imagery also caused some uncertainty in our results where the near infrared band was captured at a different resolution from the RGB imagery, and was consequently resampled to match the 25cm resolution of the red, green and blue bands. Ideally these would all be captured at the same resolution to inform this type of analysis and create accurate indices layers. The detection of bare peat using a pixel-based methodology provided a fast analysis and image classification, however it was found to miss some visible pixels which fell outside of the thresholding limits causing some gaps between neighbouring cells of bare peat. Using an object-based methodology may help to alleviate this and generate more accurate classifications of bare peat, by segmenting the imagery then classifying based on our established thresholds. Furthermore, our study found that it was not possible to develop a thresholding rule for the Pleiades satellite imagery with the indices we trialled, due to the overlap of indices signatures between areas of vegetated and bare peat. This could be further explored to develop a methodology for bare peat identification, trialling different indices and combinations to see if this could be distinguished from the Pleiades data.

The classified maps of where bare peat were present in the high-resolution imagery, provide a detailed account of areas in poor condition which can be used to target management practices for peat restoration. Aggregating this to a percentage cover of bare peat per 10m pixel as opposed to a simple presence-absence approach, provides a more descriptive and representative way of displaying the data informed by the aerial photography. However, the carries over uncertainty where pixels of bare peat were missed in the thresholding process, resulting in an underprediction of cover estimates when scaled up to the larger cell size. This highlights the importance of the initial bare peat identification stage in order to obtain accurate cover data to inform the regression models. The regression analysis workflow has been fully automated and can be deployed to predict the amount of bare peat across a wider extent than covered by the training data. This can help to provide cost effective assessments of vast areas of peatland, inform ground teams as to where to target their monitoring efforts. This study found random forest to be the best performing machine learning algorithm for the regression modelling, however, the quantile regression forest algorithm performed highly with predicted maps providing a more defined bare peat map, reducing the noise resulting with the random forest predictions. The variables found most important in predicting the presence of bare peat included the Red-Green ratio, the Green-Leaf Index and the Near infrared (narrow) band of the Sentinel Imagery. Environmental variables of monthly rainfall,

soil wetness index and slope were also predictably important to inform the models, with drier steeper conditions relating to higher bare peat coverage.

The most influential step in the workflow was the methodology used to sample the training points created in the thresholding step. Stratified random sampling of the data based on equal interval categories from 0 to 100% bare peat cover and randomly sampling an equal number of points from each category seemed to provide the better fitting models with the highest predictive accuracy both within the training areas and in extrapolated areas, when compared with random sampling methods. As the distribution of training points was skewed to lower cover values with fewer records of high coverage, “filling” with values from the lower categories was used to include more training points, which was able to account for more variability in the resulting predictions. Inclusion of greater numbers of training data with high coverage of bare peat could be further explored to see how this would influence the models, and whether this would still produce realistic predictions of bare peat cover across the landscape or tend to overpredict cover. Predictions locally improved when sampling from a greater number of image tiles, likely due to the increased range of values present. Predictions of cover in extrapolated regions improved when training data were present from a range of high to low cover and were spatially varied, with extrapolation distance from the training area demonstrating little importance to predictions. This was due to spectral band predictors and derived indices having a higher influence in determining bare peat cover, which may raise difficulties when comparing between imagery swaths captured on different dates or transferability of predictions between years. Further exploration is needed to assess as to whether change over time in bare cover can then be determined using time series imagery.

This study has demonstrated that high-resolution imagery has the potential to inform bare peat mapping in lower resolution data and indicating the condition of peatland habitats. Observations from representative sites undergoing a range of poor and recovering conditions are key for capturing the variation between different sites and improving predictions of bare peat cover, as well as being able to detect when a site is beginning to degrade. Observations from sites where large patches of bare peat are located are particularly helpful in providing a basis on which to train the models, allowing predictions of areas of bare peat outside of a region which has been examined with the aerial photography. This analysis has highlighted several considerations when using such an approach and areas for improvement in the methodology, with a need for further ground-observations for validating the accuracy of predictions and assessing their limitations. Maps of bare peat could help to flag peatland areas which are in poor condition, target sites in need of further ground monitoring and provide an indication of the success of peatland restoration measures. This assessment provides novel techniques of applying machine learning to earth observation data to inform practical management measures and a cost-effective way to aid the design of monitoring schemes.

5 References

Amos, C., Petropoulos, G.P. & Ferentinos, K.P. 2018. Determining the use of Sentinel-2A MSI for wildfire burning & severity detection. *International Journal of Remote Sensing* **40**(3):905-930. doi: 10.1080/01431161.2018.1519284

Artz, R. Evans, C., Crosher, I., Hancock, M., Scott-Cambell, M., Pilkinton, M., Jones, P., Chandler, D., McBride, A., Ross, K. & Weyl, R. 2020. Update: The State of UK Peatlands. IUCN UK Peatland Programme's Commission of Inquiry on Peatlands. Available from: https://www.iucn-uk-peatlandprogramme.org/sites/default/files/2020-01/IUCN_S~1.PDF [Accessed 16/04/2020].

Artz, R.R.E., Johnson, S., Bruneau, P., Britton, A.J., Mitchell, R.J., Ross, L., Donaldson-Selby, G., Donnelly, D., Aitkenhead, M.J., Gimona, A. & Poggio, L. 2019. The potential for modelling peatland habitat condition in Scotland using long-term MODIS data. *Science and the Total Environment* **660**: 428-442.

Asher Greenberg, J. & Mattiuzzi, M. 2018. Package 'gdalUtils'. Wrapper for the Geospatial Data Abstraction Library (GDAL). Version 2.0.1.14. Available from: https://cran.r-project.org/web/packages/gdalUtils/gdalUtils.pdf [Accessed 15/07/2019].

Bain, C.G., Bonn, A., Stoneman, R., Chapman, S., Coupar, A., Evans, M., Gearey, B., Howat, M., Joosten, H., Keenleyside, C., Labadz, J., Lindsay, R., Littlewood, N., Lunt, P., Miller, C.J., Moxey, A., Orr, H., Reed, M., Smith, P., Swales, V., Thompson, D.B.A., Thompson, P.S., Van de Noort, R., Wilson, J.D. & Worrall, F. 2011. IUCN UK Commission of Inquiry on Peatlands. IUCN UK Peatland Programme, Edinburgh.

Blake, D. & Frake, K. 2020. GIG analysis report – 92279 – Scottish bare peat map. SNH. A2812417

Blake, D. 2019. Random forest regression for assessing extent of bare peat. A2891120. Unpublished manuscript.

Bluesky International. 2019. APGB: Aerial Photography for Great Britain. Available from: https://www.apgb.co.uk/ [Accessed on 08/07/2019].

Bonnett, S.A.F., Ross, S., Linstead, C. & Maltby, E. 2009. A review of techniques for monitoring the success of peatland restoration. University of Liverpool. *Natural England Commissioned Reports, Number 086*.

Breiman, L. & Cutler, A. 2018. Package 'randomForest': Breiman and Cutler's Random Forests for Classification and Regression. Version 4.6-14. Available from: https://cran.r-project.org/web/packages/randomForest/randomForest.pdf [Accessed 30/09/2019].

Copernicus Global Land Service. 2019. Copernicus Global Land Service. Vegetation products. Available at: https://land.copernicus.eu/global/themes/Vegetation [Accessed 31/07/2019].

CNES. 2020. OrfeoToolbox: Open Source processing of remote sensing images. Available from: https://www.orfeo-toolbox.org/ [Accessed 02/07/2020].

DAERA. 2020. Peatlands. Available from: https://www.daera-ni.gov.uk/articles/peatlands [Accessed 10/07/2020].

Defra. 2018a. A Green Future: Our 25 Year Plan to Improve the Environment. Defra, London.

Defra. 2018b. Grants for Peatlands Restoration. Press Release. Available from: https://www.gov.uk/government/news/grants-for-peatlands-restoration [Accessed 10/07/2020].

Dickie, I., Evans, C., Smyth, M.A. & Artz, R. 2015. Scoping the Natural Capital Accounts for Peatland, *work package 3 of Report NR0165 for Defra*.

Evans, C., Artz, R., Moxley, J., Smyth, M-A., Taylor, E., Archer, N., Burden, A., Williamson, J., Donnelly, D., Thomson, A., Buys, G., Malcolm, H., Wilson, D., Renou-Wilson, F. & Potts J. 2017. Implementation of an emission inventory for UK peatlands. *Report to the Department for Business, Energy and Industrial Strategy*. Centre for Ecology and Hydrology, Bangor. 88pp.

FAO. 2020. Peatlands mapping and monitoring – Recommendations and technical overview. Rome. https://doi.org/10.4060/ca8200en

Freeman, E.A., Frescino, T.S. & Moisen, G.G. 2018. Pick your flavour of Random Forest. Available from: https://cran.r-project.org/web/packages/ModelMap/vignettes/Vquantile.pdf [Accessed 23/10/2019].

Fukuda, S., De Baets, B., Waegeman, W., Verwwaeren, J. & Mouton, A.M. 2013. Habitat prediction and knowledge extraction for spawning European grayling (*Thymallus thymallus* L.) using a broad range of species distribution models. *Environmental Modelling & Software* **47**: 1-6.

Google Earth. 2019. V 7.1.8.3036(32-bit). Peak D, French Polynesia. 53° 27’ 0.285”N, 1°49’23.31”W, Eye alt 43.36km. DigitalGlobe 2012. Available from: http://www.earth.google.com [Accessed September 15, 2019].

Horning, N. 2018. Random Forest for Remote Sensing. Available from: https://github.com/nedhorning/RandomForestForRemoteSensing [Accessed 29/07/2019].

Horning, N. 2011. Training Guide for adjusting random forests regression images. American Museum of Natural History - v1, Center for Biodiversity and Conservation. Available from http://biodiversityinformatics.amnh.org/ [Accessed 15/10/2019].

Hunt Jr, E.R., Doraiswamy, P.C., McMurtrey, J.E., Daughtry, C.S.T., Perry, E.M. & Akhmedov, B. 2013. A visible band index for remote sensing leaf chlorophyll content at the canopy scale. *International Journal of Applied Earth Observation and Geoinformation* **21**:103–112. http://dx.doi.org/10.1016/j.jag.2012.07.020

Index DataBase (IDB). 2020. Index DataBase: A database for remote sensing indices. Available from: https://www.indexdatabase.de/ [Accessed 13/04/2020].

The International Union for the Conservation of Nature (IUCN). 2018. UK Peatland Strategy 2018-2040. IUCN UK Peatland Programme, Edinburgh. Available from: https://portals.iucn.org/library/sites/library/files/documents/2018-015-En.pdf [Accessed 07/07/2020].

JNCC. 2019. ‘JNCCSDMs’ package: Functions to Generate Species Distribution Models using the JNCC SDM Framework. Available from: https://github.com/jncc/sdms [Accessed 01/10/2019].

JNCC. 2018. The Fourth Article 17 UK Habitats Directive Report reporting for 2013-2018 under the EU Habitats Directive. JNCC, Peterborough.

JNCC. 2011. Towards an assessment of the state of UK peatlands. JNCC Report No.445. JNCC, Peterborough, ISSN 0963-8091.

Karatzoglou, A., Smola, A., Hornik, K., National ICT Australia, Maniscalco, M.A. & Teo, C.H. 2019. Package 'kernlab' : Kernel-based Machine Learning Lab. Version 0.9-29. Available from: https://cran.r-project.org/web/packages/kernlab/kernlab.pdf [Accessed 10/04/2020]

Keyworth, S., Jarman, M. & Medcalf, K. 2009. Assessing the Extent and Severity of Erosion on the Upland Organic Soils of Scotland using Earth Observation: A GIFTSS Implementation Test. Final Report. Available from: https://www.envsys.co.uk/wp-content/uploads/2015/02/Mapping_Peat_Erosion.pdf [Accessed 29/07/2019].

Kilcoyne, A.M., Alexander, R., Cox, P. & Brownett, J. 2017. Living Maps: Satellite based Habitat Classification. Evidence Project SD1705.

Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., Benestry, M & R Core Team. 2020. Package 'caret', version 6.0-85. Classification and Regression Training. Available from: https://cran.r-project.org/web/packages/caret/caret.pdf [Accessed 24/01/2020].

Kuhn, M. & Johnson, K. 2013. Applied Predictive Modeling. Springer; New York.

Li, C., Grayson, R., Holden, J. & Li, P. 2018. Erosion in peatlands: Recent research progress and future directions. *Earth-Science Reviews* **185**: 870-886.

Matsushita, B., Yang, W., Chen, J., Onda, Y. & Qui, G. 2007. Sensitivity of the Enhanced Vegetation Index (EVI) and Normalized Difference Vegetation Index (NDVI) to Topographic Effects: A Cast Study in High Density Cypress Forest. *Sensors (Basel)* **7**(11):2636-2651.

Meinshausen, N. 2017. Package 'quanregForest': Quantile Regression Forests. Version 1.3.7. Available from: https://cran.r-project.org/web/packages/quantregForest/-quantregForest.pdf [Accessed 02/10/2019].

Meinshausen, N. 2006. Quantile Regression Forests. *Journal of Machine Learning Research* **7**: 983-999.

Met Office, Hollis, D., McCarthy, M., Kendon, M., Legg, T. & Simpson, I. (2019): HadUK-Grid Gridded Climate Observations on a 1km grid over the UK, v1.0.1.0 (1862-2018). Centre for Environmental Data Analysis, 14 November 2019. http://dx.doi.org/10.5285/d134335808894b2bb249e9f222e2eca8

Minayeva, T., Bragg, O. & Sirin, A. (2016). Peatland biodiversity and its restoration. In A. Bonn, T. Allott, M. Evans, H. Joosten, & R. Stoneman (Eds.), Peatland Restoration and Ecosystem Services: Science, Policy and Practice. 44-62: Cambridge University Press.

Montandon, L.M. & Small, E.E. 2008. The impact of soil reflectance on the quantification of the green vegetation fraction from NDVI. *Remote Sensing of Environment* **112**: 1835-1845.

Motohka, T., Nasahara, K.N., Oguma, H. & Tsuchida, S. (2010) Applicability of green-red vegetation index for remote sensing of vegetation phenology. *Remote Sensing* **2**(10): 2369–2387. https://doi.org/10.3390/rs2102369

Natural England. 2010. England's peatlands: carbon storage and greenhouse gases (NE257). Available from: http://publications.naturalengland.org.uk/publication/30021 [Accessed 16/04/2020].

Ordnance Survey. 2019. Order OS Open Data. Available from: https://www.ordnancesurvey.co.uk/opendatadownload/products.html [Accessed 01/08/2019].

Pasqualotto, N., Delegido, J., Van Wittenberghe, S.V., Rinaldi, M. & Moreno, J. 2019. Multi-Crop Green LAI Estimation with a New Simple Sentinel-2 LAI Index (SeLi). *Sensors (Basel)* **19**(4):904.

Peak District National Park. 2011. Blanket Bog. V.1. April 2011. Available from: https://www.peakdistrict.gov.uk/__data/assets/pdf_file/0016/134053/Blanket-Bog.pdf. [Accessed 29/07/2019].

QGIS Development Team (2020). QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://qgis.osgeo.org" .

R Core Team. 2019. R: A Language and Environment for Statistical Computing. Vienna, Austria. Available from: https://www.R-project.org/

Rajah, P., Odindi, J., Mutanga, O. & Kiala, Z. (2019) The utility of Sentinel-2 Vegetation Indices (VIs) and Sentinel-1 Synthetic Aperture Radar (SAR) for invasive alien species detection and mapping. *Nature Conservation* **35**: 41-61.

Räsänen, A., Elsakov, V. & Virtanen, T. 2018. Usability of one-class classification in mapping and detecting changes in bare peat surfaces in the tundra. *International Journal of Remote Sensing* **40** (11): 4083-4103.

Rodriguez, P.P. & Gianola, D. 2020. Package 'brnn': Bayesian Regularization for Feed-Forward Neural Networks. Version 0.8. Available from: https://cran.r-project.org/web/packages/brnn/brnn.pdf [Accessed 10/04/2020].

Rowland, C.S., Morton, R.D., Carrasco, L., McShane, G., O'Neil, A.W. & Wood, C.M. 2017. Land Cover Map 2015 (25m raster, GB). NERC Environmental Information Data Centre. https://doi.org/10.5285/bb15e200-9349-403c-bda9-b430093807c7

Satellite Imaging Corporation. 2019. Pleiades-1A Satellite Sensor. Available from: https://www.satimagingcorp.com/satellite-sensors/pleiades-1/ [Accessed 29/07/2019].

Scottish Government. 2015. Scotland's Biodiversity – a Route Map to 2020. The Scottish Government, Edinburgh. June 2015.

Scottish National Heritage (SNH). 2020. Peatland Condition Assessment. Available from: https://www.nature.scot/sites/default/files/2017-10/Guidance-Peatland-Action-Peatland-Condition-Assessment-Guide-A1916874.pdf [Accessed 01/07/2020].

Scottish Natural Heritage (SNH). 2015. Scotland's National Peatland Plan: Working for our future. Available from: https://www.nature.scot/scotlands-national-peatland-plan-working-our-future [Accessed 10/07/2020].

The IDB Project. 2019. Index Database: A database for remote sensing indices. Available from: https://www.indexdatabase.de/ [Accessed 28/08/2019].

United Nations. 2015. Transforming Our World: The 2030 Agenda For Sustainable Development. New York: UN Publishing.

Veronesi, F & Schillaci, C. 2019. Comparison between geostatistical and machine learning models as predictors of topsoil organic carbon with a focus on local uncertainty estimation. *Ecological Indicators* **101**: 1032-1044.

Welsh Government. 2018. Welsh Government Rural Communities – Rural Development Programme 2014-2020: Sustainable Management Scheme Collaborative Projects. Welsh Government. Available from: https://gov.wales/sites/default/files/publications/2019-10/sustainable-mananagement-scheme-selected-projects-and-map.pdf [Accessed 10/07/2020].

Williamson, J., Morton, D., Artz, R., Burden, A., Rowland, C., Tornero, L., O'Neill, A., Poggio, L., Khomik, M., Donnelly, D. & Evans, C.D. 2018. The role of earth observation in an integrated framework for assessing peatland habitat condition and its impact on greenhouse gas accounting. Final report to Defra, Project number MI07. Centre for Ecology & Hydrology, Bangor. Unpublished.

Xu, R. 2013. Improvements to random forest methodology. Graduate Theses and Dissertations. 13052. https://lib.dr/iastate.edu/etd/13052

Zhao, D., Huang, L., Li, J. & Qi, J. 2007. A comparative analysis of broadband and narrowband derived vegetation indices in predicting LAI and CCD of a cotton canopy. *ISPRS Journal of Photogrammetry and Remote Sensing* **62**(1): 25–33.

Zhang, G. & Lu, Y. 2012. Bias corrected random forests in regression. Journal of Applied Statistics **39**(1): 151-160.

Zheng L. & Yu, P. 2018. Package 'brt': Biological Relevance Testing. Version 11.3.0. Available from: https://cran.r-project.org/web/packages/brt/brt.pdf [Accessed 10/04/2020].

Appendix 1

JNCC ARD band numbers with Sentinel-2 optical imagery:

JNCC ARD bands	Spectral band
1	Blue
2	Green
3	Red
4	Red-edge 1
5	Red-edge 2
6	Red-edge 3
7	Near Infrared (narrow)
8	Near Infrared
9	Short Wave Infrared 1
10	Short Wave Infrared 2

Indices calculations from multispectral imagery bands, derived from IDB Project (2019) alongside expert knowledge from JNCCs EO specialists.

Indices	Indices name	Indices Equation
NDVI	Normalized difference vegetation index	$NDVI = \frac{(NIR - Red)}{(NIR + Red)}$
EVI	Enhanced Vegetation Index	$EVI = G\ x \frac{(NIR - Red)}{(NIR + (C1\ x\ Red - C2\ x\ BLUE) + L}$ Where L is a soil adjustment factor, C1 and C2 are coefficients for correcting aerosol scattering in the red band by use of the blue band. General values used are L = 1, G = 2.5, C1=6.0 and C2=7.5 (Matsushita, et al. 2007).
GLI	Green Leaf Index	$GLI = \frac{2 * Green - Red - Blue}{2 * Green + Red + Blue}$
GNDVI	Green Normalised Difference Vegetation Index	$GNDVI = \frac{NIR - Green}{NIR + Green}$
RDVI	Renormalized Difference Vegetation Index	$RDVI = \frac{NIR - Red}{(NIR + Red)^{0.5}}$
SAVI	Soil Adjusted Vegetation Index	$SAVI = \frac{NIR - Red}{NIR + Red + L}\ (1 + L)$ Where L = 0.5
R/B	Red-Blue ratio	$RB = \frac{Red}{Blue}$
R/G	Red-Green ratio	$RG = \frac{Red}{Green}$
Brightness	Brightness	$Brightness = \frac{(Red + Blue + Green)}{3}$
NBR	Normalized Difference NIR/SWIR Normalized Burn Ratio	$NBR = \frac{\mathrm{NIR} - \mathrm{SWIR}}{\mathrm{NIR} + \mathrm{SWIR}}$

Appendix 2

2.1 Method 1: Random Sampling Results

Testing	Combination name	training points	best model	best RMSE	best R-sq model	best R-sq	Model	mean RMSE	most important variable	Eval 1 best RMSE	Eval 1 mean RMSE	Eval1 best R-sq	Eval1 mean R-sq	Eval1 Accuracy	Eval 2 best RMSE	Eval 2 mean RMSE	Eval2 best R-sq	Eval2 mean R-sq	Eval2 Accuracy	Mean accuracy
variables	varstack_sbands	10000	RF1	0.038	RF4	0.318	RF	0.041	NIR (narrow)	0.043	0.043	0.192	0.181	0.492	0.021	0.022	0.034	0.029	0.531	0.512
							RFcorrect							0.493					0.487	0.490
							QRF	0.045	NIR (narrow)		0.043		0.179	0.462		0.021		0.029	0.521	0.491
	varstack_blake	10000	RF3	0.037	RF5	0.242	RF	0.040	slope	0.044	0.045	0.134	0.123	0.456	0.017	0.018	0.037	0.037	0.507	0.482
							RFcorrect													0.521
							QRF	0.042	NDVI		0.045		0.120	0.446		0.018		0.036	0.521	0.484
	varstack_ imp	10000	RF5	0.037	RF5	0.348	RF	0.039	RG	0.044	0.045	0.142	0.045	0.477	0.017	0.018	0.014	0.018	0.521	0.499
							RFcorrect							0.498					0.445	0.471
							QRF	0.420	RG		0.446		0.045	0.456		0.017		0.017	0.521	0.489
	varstack_all	10000	RF4	0.035	RF3	0.408	RF	0.037	NIR (narrow)	0.054	0.101	0.152	0.060	0.497	0.040	0.090	0.010	0.011	0.524	0.511
							RFcorrect							0.512					0.468	0.490
							QRF	0.041	RG		0.084		0.081	0.446		0.075		0.014	0.521	0.484
	varstack_RD_SWIR	10000	RF3	0.038	RF3	0.279	RF	0.041	SWIR 2	0.043	0.043	0.193	0.184	0.478	0.018	0.019	0.042	0.034	0.531	0.505
							RFcorrect							0.487					0.468	0.477
							QRF	0.044	SWIR 2		0.044		0.183	0.446		0.019		0.034	0.524	0.485
	varstack_slope	10000	RF1	0.04	RF2	0.247	RF	0.042		0.043	0.044	0.170	0.148	0.462	0.018	0.019	0.023	0.024	0.542	0.502
							RFcorrect							0.458					0.468	0.463
							QRF	0.416			0.441		0.146	0.446		0.194		0.023	0.528	0.487
	varstack_NBR	10000	RF2	0.037	RF2	0.332	RF	0.041	RG	0.043	0.043	0.176	0.043	0.456	0.016	0.017	0.050	0.017	0.510	0.483
							RFcorrect							0.478					0.445	0.461
							QRF	0.045	RG		0.043		0.043	0.446		0.017		0.017	0.521	0.484
	varstack_climMax	10000	RF1	0.037	QRF1	0.453	RF	0.040	RG	0.045	0.045	0.118	0.045	0.456	0.017	0.017	0.028	0.017	0.503	0.480
							RFcorrect							0.478					0.441	0.459
							QRF	0.040	RG		0.045		0.045	0.451		0.017		0.017	0.521	0.486
	varstack_ climAv	10000	RF1	0.034	RF4	0.322	RF	0.040	RG	0.045	0.045	0.108	0.045	0.462	0.017	0.017	0.040	0.017	0.503	0.483
							RFcorrect							0.478					0.433	0.456
							QRF	0.043	RG		0.045		0.045	0.451		0.017		0.017	0.521	0.486
	varstack_darkveg	10000	RF1	0.036	RF4	0.331	RF	0.395	RG	0.043	0.045	0.184	0.188	0.497	0.019	0.020	0.036	0.039	0.524	0.511
							RFcorrect							0.507					0.483	0.495
							QRF	0.423	RG		0.045		0.191	0.446		0.021		0.039	0.524	0.485
	varstack_import	10000	RF5	0.038	QRF1	0.321	RF	0.040	RG	0.064	0.085	0.051	0.126	0.456	0.047	0.074	0.023	0.022	0.507	0.482
							RFcorrect							0.458					0.468	0.463
							QRF	0.043	RG		0.088		0.126	0.446		0.076		0.019	0.521	0.484
Points	varstack_imp	2000	RF5	0.033	BRNN2	0.385	RF	0.040	RG	0.044	0.046	0.133	0.046	0.487	0.017	0.018	0.030	0.018	0.514	0.501
							RFcorrect							0.502					0.441	0.472
							QRF	0.043	RG		0.046		0.046	0.451		0.018		0.018	0.521	0.486
							BRNN	0.041	RG		0.058		0.576	0.472		0.582		0.582	0.503	0.488
							SVM	0.046	RG		0.049		0.049	0.446		0.017		0.017	0.521	0.484

Testing	Combination name	training points	best model	best RMSE	best R-sq model	best R-sq	Model	mean RMSE	most important variable	Eval 1 best RMSE	Eval 1 mean RMSE	Eval1 best R-sq	Eval1 mean R-sq	Eval1 Accuracy	Eval 2 best RMSE	Eval 2 mean RMSE	Eval2 best R-sq	Eval2 mean R-sq	Eval2 Accuracy	Mean accuracy
Points	varstack_imp	5000	RF4	0.037	RF1	0.375	RF	0.400	RG	0.044	0.045	0.156	0.045	0.472	0.017	0.017	0.011	0.017	0.507	0.489
							RFcorrect							0.493					0.433	0.463
							QRF	0.430	RG		0.045		0.045	0.451		0.018		0.018	0.521	0.486
	varstack_imp	10000	RF5	0.037	RF5	0.348	RF	0.039	RG	0.044	0.045	0.142	0.045	0.477	0.017	0.018	0.014	0.018	0.521	0.499
							RFcorrect							0.498					0.445	0.471
							QRF	0.420	RG		0.446		0.045	0.456		0.017		0.017	0.521	0.489
	varstack_imp	20000	RF3	0.038	RF4	0.354	RF	0.039	RG	0.044	0.044	0.168	0.044	0.477	0.017	0.018	0.013	0.018	0.521	0.499
							RFcorrect							0.507					0.449	0.478
							QRF	0.042	RG		0.044		0.044	0.456		0.018		0.018	0.521	0.489
	varstack_imp	50000	RF5	0.037	RF3	0.346	RF	0.038	RG	0.043	0.044	0.179	0.044	0.472	0.017	0.018	0.014	0.018	0.521	0.496
							RFcorrect							0.493					0.445	0.469
							QRF	0.041	RG		0.044		0.044	0.456		0.018		0.018	0.521	0.489
	varstack_imp	100000	RF2	0.036	RF3	0.398	RF	0.037	RG	0.043	0.044	0.183	0.171	0.472	0.017	0.018	0.023	0.022	0.517	0.495
							RFcorrect							0.498					0.449	0.473
							QRF	0.040	RG		0.044		0.170	0.456		0.018		0.023	0.521	0.489
							BRNN	0.042	RG		0.918		0.001	0.451		0.929		0.001	0.503	0.477
							SVM	0.048	RG		0.049		0.099	0.446		0.017		0.223	0.521	0.484
	varstack_NBR	2000	RF2	0.036	RF5	0.340	RF	0.040	RG	0.043	0.044	0.026	0.017	0.456	0.017	0.017	0.198	0.043	0.503	0.480
							RFcorrect							0.473					0.437	0.455
							QRF	0.042	RG		0.045		0.017	0.446		0.017		0.044	0.521	0.484
	varstack_NBR	5000	RF1	0.041	QRF1	0.380	RF	0.043	RG	0.044	0.044	0.187	0.174	0.456	0.016	0.017	0.060	0.052	0.510	0.483
							RFcorrect							0.468					0.445	0.456
							QRF	0.045	RG		0.044		0.171	0.446		0.017		0.052	0.521	0.484
	varstack_NBR	10000	RF2	0.037	RF2	0.332	RF	0.041	RG	0.043	0.043	0.176	0.043	0.456	0.016	0.017	0.050	0.017	0.510	0.483
							RFcorrect							0.478					0.445	0.461
							QRF	0.045	RG		0.043		0.043	0.446		0.017		0.017	0.521	0.484
	varstack_NBR	20000	RF4	0.04	QRF1	0.313	RF	0.041	RG	0.043	0.043	0.198	0.167	0.467	0.016	0.017	0.051	0.050	0.510	0.489
							RFcorrect							0.478					0.452	0.465
							QRF	0.044	RG		0.044		0.165	0.446		0.017		0.049	0.521	0.484
							BRNN	0.042			0.131		0.004	0.456		0.018		0.000	0.517	0.487
							SVM	0.048			0.049		0.096	0.446		0.017		0.022	0.521	0.484
	varstack_NBR	50000	RF2	0.039	RF3	0.346	RF	0.039	RG	0.043	0.043	0.187	0.043	0.472	0.016	0.017	0.060	0.017	0.507	0.489
							RFcorrect							0.478					0.452	0.465
							QRF	0.042	RG		0.043		0.043	0.446		0.017		0.017	0.521	0.484
	varstack_NBR	100000	RF5	0.042824	RF5	0.344	RF	0.040	RG	0.043	0.043	0.193	0.173	0.467	0.016	0.016	0.049	0.052	0.514	0.490
							RFcorrect							0.478					0.449	0.463
							QRF	0.043	RG		0.043		0.175	0.446		0.016		0.051	0.521	0.484

Testing	Combination name	training points	best model	best RMSE	best R-sq model	best R-sq	Model	mean RMSE	most important variable	Eval 1 best RMSE	Eval 1 mean RMSE	Eval1 best R-sq	Eval1 mean R-sq	Eval1 Accuracy	Eval 2 best RMSE	Eval 2 mean RMSE	Eval2 best R-sq	Eval2 mean R-sq	Eval2 Accuracy	Mean accuracy
methods	varstack_NBR	20000	RF4	0.04	QRF1	0.313	RF	0.041	RG	0.043	0.043	0.198	0.167	0.467	0.016	0.017	0.051	0.050	0.510	0.489
							RFcorrect							0.478					0.452	0.465
							QRF	0.044	RG		0.044		0.165	0.446		0.017		0.049	0.521	0.484
							BRNN	0.042			0.131		0.004	0.456		0.018		0.000	0.517	0.487
							SVM	0.048			0.049		0.096	0.446		0.017		0.022	0.521	0.484
	varstack_imp	100000	RF2	0.036	RF3	0.398	RF	0.037	RG	0.043	0.044	0.183	0.171	0.472	0.017	0.018	0.017	0.018	0.517	0.495
							RFcorrect							0.498					0.449	0.473
							QRF	0.040	RG		0.044		0.170	0.456		0.018		0.018	0.521	0.489
							BRNN	0.042			0.918		0.001	0.451		0.929		0.929	0.503	0.477
							SVM	0.048			0.049		0.099	0.446		0.017		0.017	0.521	0.484
	varstack_imp	2000	RF5	0.033	BRNN2	0.385	RF	0.040	RG	0.044	0.046	0.133	0.046	0.487	0.017	0.018	0.030	0.018	0.514	0.501
							RFcorrect							0.502					0.441	0.472
							QRF	0.043	RG		0.046		0.046	0.451		0.018		0.018	0.521	0.486
							BRNN	0.041			0.058		0.576	0.472		0.582		0.582	0.503	0.488
							SVM	0.046			0.049		0.049	0.446		0.017		0.017	0.521	0.484

2.2 Method 2: Equal Interval Sampling Results

Testing	Combination name	target training points	training points	test points	best model	best RMSE	best R-sq model	best R-sq	Model	mean RMSE	most important variable	Eval 1 best RMSE	Eval 1 mean RMSE	Eval1 best R-sq	Eval1 mean R-sq	Eval1 Accuracy	Eval 2 best RMSE	Eval 2 mean RMSE	Eval2 best R-sq	Eval2 mean R-sq	Eval2 Accuracy	Mean accuracy
variables	varstack_sbands	10000	5263	1756	RF5	0.131	RF4	0.689	RF	0.133	NIR (narrow	0.090	0.093	0.178	0.093	0.554	0.072	0.076	0.027	0.076	0.552	0.553
									QRF	0.141	NIR (narrow		0.093		0.093	0.508		0.075		0.075	0.535	0.521
	varstack_blake	10000	5263	1756	RF5	0.138	RF5	0.645	RF	0.143	NDVI	0.082	0.084	0.104	0.084	0.574	0.060	0.064	0.022	0.064	0.570	0.572
									QRF	0.151	NDVI		0.084		0.084	0.492		0.064		0.064	0.524	0.508
	varstack_imp	10000	5263	1756	RF2	0.129	RF3	0.695	RF	0.132	RG	0.067	0.070	0.096	0.070	0.590	0.046	0.049	0.012	0.049	0.538	0.564
	varstack_all	10000	5263	1756	RF1	0.116	RF1	0.750	RF	0.118	RG	0.150	0.159	0.057	0.041	0.610	0.147	0.152	0.016	0.017	0.563	0.587
									QRF	0.124	RG		0.163		0.040	0.518		0.156		0.018	0.535	0.526
	varstack_RD_SWIR	10000	5263	1756	RF3	0.130	RF3	0.679	RF	0.136	Red-edge 2	0.104	0.110	0.123	0.110	0.621	0.063	0.068	0.027	0.068	0.580	0.600
									QRF	0.143	Red-edge 2		0.110		0.110	0.508		0.069		0.069	0.552	0.530
	varstack_slope	10000	5263	1756	RF1	0.141	RF1	0.636	RF	0.145	Red-edge 2	0.104	0.110	0.128	0.110	0.621	0.074	0.076	0.032	0.076	0.598	0.609
									QRF	0.154	Red-edge 2		0.110		0.110	0.533		0.774		0.077	0.573	0.553
	varstack_NBR	10000	5263	1756	RF5	0.132	RF5	0.680	RF	0.136	RG	0.077	0.081	0.123	0.129	0.615	0.044	0.048	0.033	0.033	0.559	0.587
									QRF	0.143	RG		0.081		0.127	0.503		0.048		0.032	0.531	0.517
									BRNN	0.153			0.119		0.005	0.585		0.133		0.000	0.538	0.562
									SVM	0.164			0.103		0.104	0.554		0.100		0.041	0.535	0.544
	varstack_climMax	10000	5263	1756	RF2	0.123	RF2	0.724	RF	0.126	RG	0.077	0.083	0.093	0.083	0.631	0.059	0.064	0.025	0.064	0.535	0.583
									QRF	0.132	RG		0.083		0.083	0.518		0.065		0.065	0.528	0.523
	varstack_climAv	10000	5263	1756	RF1	0.122	RF1	0.717	RF	0.125	RG	0.078	0.084	0.074	0.084	0.636	0.062	0.065	0.024	0.065	0.524	0.580
									QRF	0.131	RG		0.084		0.084	0.492		0.064		0.064	0.535	0.514
	varstack_darkveg	10000	5263	1756	RF1	0.120	RF1	0.737	RF	0.123	RG	0.114	0.120	0.142	0.120	0.641	0.079	0.082	0.044	0.082	0.577	0.609
									QRF	0.130	RG		0.119		0.119	0.492		0.081		0.081	0.528	0.510
	varstack_import	10000	5263	1756	RF1	0.122	RF1	0.717	RF	0.125	GLI	0.078	0.084	0.074	0.084	0.636	0.062	0.065	0.024	0.065	0.524	0.580
									QRF	0.131	GLI		0.084		0.084	0.492		0.064		0.064	0.535	0.514
	varstack_import	5000			RF1	0.120	RF1	0.737	RF	0.123	RG	0.114	0.120	0.142	0.120	0.641	0.079	0.082	0.044	0.082	0.577	0.609
									QRF	0.130	RG		0.119		0.119	0.492		0.081		0.081	0.528	0.510
points	varstack_imp	545	409	136	RF1	0.148	RF1	0.803	RF	0.158	RG	0.106	0.114	0.028	0.020	0.559	0.071	0.087	0.015	0.010	0.486	0.522
	varstack_imp	2000	1281	428	RF2	0.143	RF2	0.736	RF	0.153	RG	0.087	0.094	0.067	0.047	0.636	0.063	0.067	0.019	0.018	0.521	0.578
									QRF	0.160	RG		0.092		0.050	0.518		0.065		0.016	0.510	0.514
									BRNN	0.174			1.069		0.005	0.590		1.080		0.001	0.550	0.570
									BRT	0.179			0.116		0.089	0.564		0.085		0.024	0.500	0.532
	varstack_imp	5000	3013	1006	RF2	0.133	RF2	0.743	RF	0.143	RG	0.078	0.083	0.076	0.059	0.641	0.056	0.059	0.015	0.015	0.528	0.584
	varstack_imp	10000	5263	1756	RF2	0.129	RF3	0.695	RF	0.132	RG	0.067	0.070	0.096	0.082	0.590	0.046	0.049	0.012	0.015	0.538	0.564
	varstack_imp	20000	9763	3256	RF1	0.119	RF3	0.666	RF	0.119	RG	0.061	0.064	0.110	0.090	0.554	0.043	0.045	0.014	0.013	0.535	0.544
	varstack_imp	50000	21087	7031	RF2	0.103	RF4	0.661	RF	0.104	RG	0.055	0.055	0.121	0.117	0.569	0.036	0.037	0.012	0.016	0.542	0.556
	varstack_imp	100000	36087	12031	RF2	0.087	RF2	0.646	RF	0.088	RG	0.048	0.048	0.129	0.127	0.544	0.029	0.029	0.011	0.015	0.500	0.522
	varstack_NBR	2000	1281	428	RF5	0.153	RF5	0.711	RF	0.159	RG	0.097	0.101	0.113	0.110	0.610	0.059	0.065	0.025	0.027	0.528	0.569

Testing	Combination name	target training points	training points	test points	best model	best RMSE	best R-sq model	best R-sq	Model	mean RMSE	most important variable	Eval 1 best RMSE	Eval 1 mean RMSE	Eval1 best R-sq	Eval1 mean R-sq	Eval1 Accuracy	Eval 2 best RMSE	Eval 2 mean RMSE	Eval2 best R-sq	Eval2 mean R-sq	Eval2 Accuracy	Mean accuracy
points	varstack_NBR	10000	5263	1756	RF5	0.132	RF5	0.680	RF	0.136	RG	0.077	0.081	0.123	0.129	0.615	0.044	0.048	0.033	0.033	0.559	0.587
									QRF	0.143	RG		0.081		0.127	0.503		0.048		0.032	0.531	0.517
									BRNN	0.153	RG		0.119		0.005	0.585		0.133		0.000	0.538	0.562
									SVM	0.164	RG		0.103		0.104	0.554		0.100		0.041	0.535	0.544
	varstack_NBR	20000	9763	3256	RF1	0.123	RF1	0.639	RF	0.125	RG	0.073	0.075	0.125	0.129	0.605	0.040	0.042	0.035	0.037	0.545	0.575
	varstack_NBR	50000	21087	7031	RF5	0.109	RF3	0.619	RF	0.110	RG	0.062	0.064	0.139	0.139	0.549	0.032	0.033	0.034	0.031	0.549	0.549
	varstack_NBR	100000			RF1	0.093	RF1	0.588	RF	0.094	RG	0.047	0.055	0.097	0.142	0.533	0.024	0.026	0.028	0.031	0.531	0.532
									QRF	0.101	RG		0.055		0.143	0.472		0.026		0.030	0.503	0.488
									BRNN	0.108			0.154		0.005	0.513		0.197		0.000	0.514	0.513
									SVM	0.144			0.047		0.096	0.451		0.027		0.299	0.503	0.477
									BRT	0.113			0.055		0.096	0.528		0.032		0.025	0.538	0.533
methods	varstack_NBR	10000	5263	1756	RF5	0.132	RF5	0.680	RF	0.136	RG	0.077	0.081	0.123	0.129	0.615	0.044	0.048	0.033	0.033	0.559	0.587
									QRF	0.143	RG		0.081		0.127	0.503		0.048		0.032	0.531	0.517
									BRNN	0.153	RG		0.119		0.005	0.585		0.133		0.000	0.538	0.562
									SVM	0.164	RG		0.103		0.104	0.554		0.100		0.041	0.535	0.544
	varstack_imp	2000	1281	428	RF2	0.143	RF2	0.736	RF	0.153	RG	0.087	0.094	0.067	0.047	0.636	0.063	0.067	0.019	0.018	0.521	0.578
									QRF	0.160	RG		0.092		0.050	0.518		0.065		0.016	0.510	0.514
									BRNN	0.174			1.069		0.005	0.590		1.080		0.001	0.550	0.570
									BRT	0.179			0.116		0.089	0.564		0.085		0.024	0.500	0.532
	varstack_NBR	100000	36087	12031	RF1	0.093	RF1	0.588	RF	0.094	RG	0.047	0.055	0.097	0.142	0.533	0.024	0.026	0.028	0.031	0.531	0.532
									QRF	0.101	RG		0.055		0.143	0.472		0.026		0.030	0.503	0.488
									BRNN	0.108			0.154		0.005	0.513		0.197		0.000	0.514	0.513
									SVM	0.144			0.047		0.096	0.451		0.027		0.299	0.503	0.477
									BRT	0.113			0.055		0.096	0.528		0.032		0.025	0.538	0.533

2.3 Method 2b: Equal Interval Sampling with categorised filling Results

Testing	Combination name	target training points	training points	test points	best model	best RMSE	best R-sq model	best R-sq	Model	mean RMSE	Eval 1 best RMSE	Eval 1 mean RMSE	Eval1 best R-sq	Eval1 mean R-sq	Eval1 Accuracy	Eval 2 best RMSE	Eval 2 mean RMSE	Eval2 best R-sq	Eval2 mean R-sq	Eval2 Accuracy	Mean accuracy
variables	varstack_NBR	10000	7500	2500	RF4	0.130	RF4	0.633	RF	0.133	0.071	0.074	0.137	0.074	0.610	0.039	0.042	0.032	0.042	0.545	0.578
									QRF	0.141		0.074		0.074	0.487		0.042		0.042	0.503	0.495
	varstack_climMax	10000	7500	2500	RF5	0.114	RF4	0.722	RF	0.116	0.072	0.075	0.080	0.075	0.626	0.055	0.058	0.023	0.058	0.528	0.577
									QRF	0.122		0.075		0.075	0.482		0.058		0.058	0.528	0.505
	varstack_climAv	10000	7500	2500	RF5	0.115	RF5	0.716	RF	0.118	0.077	0.079	0.077	0.079	0.626	0.060	0.062	0.027	0.062	0.542	0.584
									QRF	0.125		0.079		0.079	0.492		0.061		0.061	0.524	0.508
	varstack_all	10000	7500	2500	RF1	0.109	RF1	0.747	RF	0.111	0.135	0.155	0.094	0.042	0.631	0.125	0.146	0.028	0.021	0.570	0.600
									QRF	0.118		0.157		0.043	0.492		0.150		0.020	0.524	0.508
	varstack_import	10000	7500	2500	RF4	0.125	RF4	0.666	RF	0.127	0.141	0.165	0.115	0.116	0.631	0.132	0.150	0.055	0.042	0.570	0.600
									QRF	0.133		0.161		0.114	0.533		0.146		0.042	0.524	0.529
	varstack_darkveg	10000	7500	2500	RF1	0.114	RF1	0.718	RF	0.118	0.105	0.111	0.150	0.111	0.662	0.072	0.077	0.042	0.077	0.587	0.624
									QRF	0.125		0.111		0.111	0.477		0.076		0.076	0.528	0.502

2.4 Method 3: Quantiled sampling of values above 0.1 Results

Testing	Combination name	target training points	training points	test points	best model	best RMSE	best R-sq model	best R-sq	Model	mean RMSE	Eval 1 best RMSE	Eval 1 mean RMSE	Eval1 best R-sq	Eval1 mean R-sq	Eval1 Accuracy	Eval 2 best RMSE	Eval 2 mean RMSE	Eval2 best R-sq	Eval2 mean R-sq	Eval2 Accuracy	Mean accuracy
variables	varstack_sbands	10000	7500	2500	RF2	0.094	RF2	0.460	RF	0.097	0.072	0.074	0.161	0.074	0.585	0.065	0.066	0.048	0.066	0.552	0.569
									QRF	0.100		0.074		0.074	0.579		0.067		0.067	0.545	0.562
	varstack_climMax	10000	7500	2500	RF5	0.090	RF1	0.490	RF	0.092	0.071	0.074	0.082	0.070	0.595	0.057	0.061	0.051	0.056	0.577	0.586
									QRF	0.095		0.073		0.071	0.600		0.060		0.056	0.528	0.564
	varstack_NBR	10000	7500	2500	RF1	0.093	RF1	0.443	RF	0.097	0.067	0.069	0.139	0.069	0.579	0.048	0.050	0.058	0.050	0.594	0.587
									QRF	0.100		0.068		0.068	0.528		0.050		0.050	0.549	0.539
	varstack_climAv	10000	7500	2500	RF5	0.091	RF3	0.491	RF	0.092	0.058	0.060	0.120	0.071	0.621	0.069	0.071	0.049	0.060	0.570	0.595
									QRF	0.096		0.060		0.071	0.605		0.071		0.060	0.545	0.575
	varstack_all	10000	7500	2500	RF5	0.088	RF3	0.534	RF	0.089	0.116	0.132	0.072	0.064	0.626	0.106	0.125	0.028	0.028	0.563	0.594
									QRF	0.093		0.126		0.074	0.590		0.120		0.030	0.563	0.576

2.5 Results from sampling from all the training data: Training area, Eval 1 and Eval 2

Sampling method	Combination name	target training points	training points	test points	best model	best RMSE	best R-sq model	best R-sq	Model	mean RMSE	most important variable	Eval1 Accuracy	Eval2 Accuracy	Mean Accuracy
1	varstack_NBR	20000	15000	5000	RF1	0.038	RF5	0.333	RF	0.040	RG	0.467	0.521	0.494
									RFcorrect			0.488	0.449	0.468
									QRF	0.043	RG	0.446	0.521	0.484
									BRNN	0.042	RG	0.456	0.517	0.487
									SVM	0.048	RG	0.446	0.521	0.484
									BRT	0.044	RG	0.456	0.521	0.489
	varstack_climMax	10000	7500	2500	RF2	0.035	RF1	0.337	RF	0.038	RG	0.462	0.503	0.483
									RFcorrect			0.483	0.433	0.458
									QRF	0.041	RG	0.451	0.521	0.486
2	varstack_NBR	1000	681	228	QRF4	0.144	QRF4	0.812	RF	0.159	RG	0.559	0.528	0.543
									QRF	0.165	RG	0.528	0.514	0.521
	varstack_NBR	5000	3018	1008	RF2	0.135	QRF5	0.765	RF	0.139	RG	0.600	0.552	0.576
									QRF	0.141	RG	0.508	0.535	0.521
	varstack_NBR	10000	5268	1758	QRF1	0.122	QRF1	0.789	RF	0.127	RG	0.626	0.559	0.593
									QRF	0.127	RG	0.518	0.535	0.526
	varstack_NBR	20000	9768	3258	RF3	0.114	QRF3	0.760	RF	0.117	RG	0.636	0.556	0.596
									QRF	0.118	RG	0.513	0.510	0.512
	varstack_NBR	50000	21196	7068	QRF4	0.099	QRF2	0.755	RF	0.100	RG	0.600	0.570	0.585
									QRF	0.993	RG	0.523	0.517	0.520
	varstack_NBR	100000	36196	12068	QRF4	0.079	QRF4	0.801	RF	0.085	RG	0.579	0.563	0.571
									QRF	0.080	RG	0.518	0.507	0.512

Sampling method	Combination name	target training points	training points	test points	best model	best RMSE	best R-sq model	best R-sq	Model	mean RMSE	most important variable	Eval1 Accuracy	Eval2 Accuracy	Mean Accuracy
2	varstack_climMax	10000	5263	1756	QRF2	0.107	QRF2	0.838	RF	0.111	RG	0.631	0.552	0.592
									QRF	0.111	RG	0.559	0.531	0.545
	varstack_darkveg	10000	5263	1756	QRF1	0.106	QRF1	0.841	RF	0.111	RG	0.626	0.580	0.603
									QRF	0.110	RG	0.523	0.510	0.517
2b	varstack_NBR	10000	7500	2501	QRF3	0.117	QRF3	0.786	RF	0.122	RG	0.600	0.545	0.573
									QRF	0.123	RG	0.523	0.517	0.520
	varstack_climMax	10000	7500	2501	QRF5	0.105	QRF5	0.825	RF	0.110	RG	0.651	0.556	0.604
									QRF	0.110	RG	0.538	0.528	0.533
	varstack_darkveg	10000	7500	2501	QRF1	0.102	QRF1	0.836	RF	0.110	RG	0.651	0.580	0.616
									QRF	0.110	RG	0.523	0.524	0.524
	varstack_darkveg	2000	1497	504	RF5	0.132	RF5	0.810	RF	0.138	RG	0.636	0.542	0.589
									QRF	0.144	RG	0.538	0.542	0.540
	varstack_darkveg	5000	3748	1252	RF5	0.115	RF5	0.825	RF	0.122	RG	0.656	0.577	0.617
									QRF	0.125	RG	0.538	0.521	0.530
3	varstack_NBR	10000	7500	2500	RF3	0.092	RF3	0.447	RF	0.096	RG	0.605	0.573	0.589
									QRF	0.099	RG	0.533	0.552	0.543
	varstack_climMax	10000	7500	2500	RF3	0.087	RF3	0.512	RF	0.090	RG	0.621	0.584	0.602
									QRF	0.093	RG	0.651	0.542	0.597

2.6 Results from patch sampling analysis

Tiles	Combination name	target training points	best model	best RMSE	best R-sq model	best R-sq	Model	mean RMSE	mean R-sq	Eval1 Accuracy	Eval2 Accuracy	Mean Accuracy	TZ2&3 best RMSE	TZ2&3 mean RMSE	TZ2&3 best R-sq	TZ2&3 mean R-sq
ES1Z1	varstack_darkveg	2000	RF5	0.051	RF5	0.615	RF	0.057	0.518	0.615	0.559	0.587	0.061	0.067	0.033	0.028
							QRF	0.061	0.468	0.523	0.524	0.524		0.065		0.030
		5000	RF3	0.043	RF3	0.479	RF	0.044	0.448	0.631	0.538	0.585	0.058	0.061	0.039	0.034
							QRF	0.047	0.416	0.528	0.524	0.526		0.061		0.036
		10000	RF1	0.036	RF4	0.485	RF	0.037	0.467	0.641	0.542	0.591	0.056	0.057	0.090	0.076
							QRF	0.039	0.429	0.533	0.528	0.531		0.057		0.071
ES1Z3	varstack_darkveg	2000	RF1	0.017	QRF1	0.444	RF	0.020	0.291	0.446	0.535	0.491	0.052	0.052	0.239	0.219
							QRF	0.021	0.316	0.446	0.524	0.485		0.053		0.200
ES1Z1 ES1Z3	varstack_darkveg	2000	RF3	0.063	RF1	0.583	RF	0.066	0.520	0.533	0.538	0.536	0.051	0.051	0.264	0.231
							QRF	0.071	0.466	0.523	0.524	0.524		0.052		0.222
		5000	RF3	0.045	RF3	0.533	RF	0.048	0.495	0.533	0.531	0.532	0.051	0.052	0.257	0.242
							QRF	0.052	0.438	0.523	0.524	0.524		0.052		0.241
		10000	RF4	0.038	RF5	0.497	RF	0.039	0.463	0.538	0.545	0.542	0.051	0.051	0.273	0.265
							QRF	0.042	0.414	0.523	0.524	0.524		0.052		0.266
ES1Z1 ESZ3 ES2Z2	varstack_darkveg	2000	RF4	0.057	RF4	0.645	RF	0.063	0.576	0.528	0.531	0.530	0.051	0.052	0.247	0.203
							QRF	0.068	0.518	0.523	0.524	0.524		0.052		0.198
		5000	RF4	0.043	QRF4	0.539	RF	0.047	0.505	0.538	0.531	0.535	0.051	0.051	0.252	0.240
							QRF	0.050	0.445	0.523	0.524	0.524		0.051		0.239
		10000	RF4	0.037	RF4	0.507	RF	0.037	0.485	0.538	0.531	0.535	0.051	0.051	0.285	0.265
							QRF	0.041	0.422	0.518	0.521	0.519		0.051		0.265

Tiles	Combination name	target training points	best model	best RMSE	best R-sq model	best R-sq	Model	mean RMSE	mean R-sq	Eval1 Accuracy	Eval2 Accuracy	Mean Accuracy	TZ2&3 best RMSE	TZ2&3 mean RMSE	TZ2&3 best R-sq	TZ2&3 mean R-sq
ES1Z1	varstack_darkveg	2000	RF1	0.055	RF1	0.625	RF	0.061	0.565	0.528	0.545	0.537	0.051	0.052	0.304	0.265
ESZ3							QRF	0.065	0.517	0.523	0.521	0.522		0.052		0.257
ES2Z2		5000	RF4	0.046	RF4	0.510	RF	0.048	0.461	0.538	0.542	0.540	0.052	0.053	0.288	0.244
ES2Z3							QRF	0.052	0.405	0.523	0.528	0.526		0.053		0.242
		10000	RF2	0.035	RF5	0.491	RF	0.038	0.452	0.533	0.542	0.538	0.053	0.053	0.261	0.255
							QRF	0.041	0.378	0.523	0.521	0.522		0.053		0.262
ES1Z1	varstack_darkveg	2000	RF3	0.052	RF3	0.657	RF	0.059	0.589	0.538	0.535	0.537	0.052	0.052	0.207	0.217
ESZ3							QRF	0.062	0.560	0.523	0.521	0.522		0.053		0.193
ES2Z2		5000	RF5	0.044	QRF5	0.559	RF	0.046	0.503	0.544	0.531	0.538	0.051	0.052	0.232	0.185
ES2Z3							QRF	0.046	0.454	0.523	0.521	0.522		0.053		0.180
ES2Z4		10000	RF5	0.036	RF4	0.507	RF	0.037	0.481	0.538	0.531	0.535	0.052	0.053	0.237	0.179
							QRF	0.044	0.427	0.518	0.524	0.521		0.053		0.180
ES1Z1	varstack_darkveg	2000	RF5	0.100	RF5	0.688	RF	0.102	0.674	0.641	0.542	0.591	0.092	0.102	0.195	0.166
ESZ3							QRF	0.107	0.653	0.523	0.542	0.533		0.101		0.166
ES2Z2		5000	RF5	0.070	RF5	0.678	RF	0.072	0.669	0.579	0.542	0.561	0.075	0.082	0.201	0.202
ES2Z3							QRF	0.078	0.626	0.518	0.528	0.523		0.082		0.195
ES2Z4		10000	RF3	0.056	RF3	0.653	RF	0.058	0.611	0.554	0.535	0.544	0.067	0.072	0.210	0.200
TZ1							QRF	0.062	0.565	0.523	0.531	0.527		0.074		0.202
ES1Z1	varstack_darkveg	2000	RF5	0.098	RF5	0.735	RF	0.107	0.706	0.646	0.573	0.610	0.056	0.059	0.508	0.499
ESZ3							QRF	0.103	0.689	0.538	0.538	0.538		0.059		0.498
ES2Z2		5000	RF2	0.078	RF2	0.727	RF	0.081	0.703	0.600	0.577	0.588	0.044	0.046	0.588	0.566
ES2Z3							QRF	0.087	0.671	0.538	0.524	0.531		0.047		0.559
ES2Z4 TZ1,2&3		10000	RF3	0.062	RF3	0.702	RF	0.064	0.669	0.579	0.563	0.571	0.037	0.038	0.636	0.630
							QRF	0.069	0.631	0.528	0.524	0.526		0.038		0.629