The potential for modelling peatland habitat condition in Scotland using long-term MODIS data

Rebekka R.E. Artz[a], Sally Johnson[b], Patricia Bruneau[b], Andrea J. Britton[a], Ruth J. Mitchell[a], Louise Ross[c], Gillian Donaldson-Selby[a], David Donnelly[a], Matt Aitkenhead[a], Alessandro Gimona[a] and Laura Poggio[a]

a The James Hutton Institute, Craigiebuckler, Aberdeen, AB15 8QH, UK.

b Scottish Natural Heritage, Silvan House, 231 Corstorphine Rd, Edinburgh EH12 7AT, UK.

c School of Biological Sciences, University of Aberdeen, Cruickshank Building, St Machar Drive, Aberdeen, AB24 3UL, UK.

Corresponding author: rebekka.artz@hutton.ac.uk

Abstract

Globally, peatlands provide an important sink of carbon in their near natural state but potentially act as a source of gaseous and dissolved carbon emission if not in good condition. There is a pressing need to remotely identify peatland sites requiring improvement and to monitor progress following restoration. A medium resolution model was developed based on a training dataset of peatland habitat condition and environmental covariates, such as morphological features, against information derived from the Moderate Resolution Imaging Spectroradiometer (MODIS), covering Scotland (UK). The initial, unrestricted, model provided the probability of a site being in favourable condition. Receiver operator characteristics (ROC) curves for restricted training data, limited to those located on a peat soil map, resulted in an accuracy of 0.916. The kappa statistic was 0.8151, suggesting good model fit. The derived map of predicted peatland condition at the suggested 0.56 threshold was corroborated by data from other sources, including known restoration sites, areas under known non-peatland land cover and previous vegetation survey data mapped onto inferred condition categories. The resulting locations of the areas of peatland modelled to be in favourable ecological condition were largely confined to the North and West of the country, which not only coincides with prior land use intensity but with published predictions of future retraction of the bioclimatic space for peatlands. The model is limited by a lack of spatially appropriate ground observations, and a lack of verification of peat depth at training site locations, hence future efforts to remotely assess peatland condition will require more appropriate ground-based monitoring. If appropriate ground-based observations could be collected, using remote sensing could be considered a cost-efficient means to provide data on changes in peatland habitat condition.

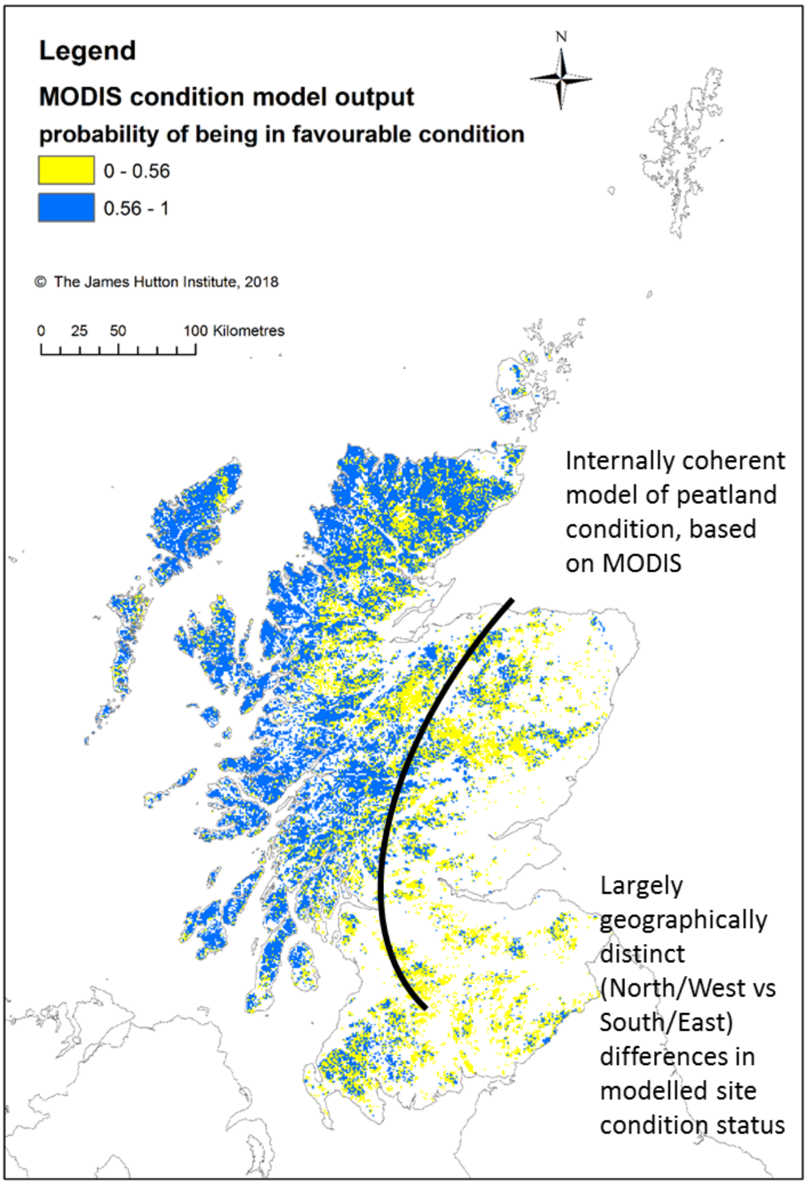
Keywords: peatland, habitat condition, remote sensing, MODIS, modelling, mapping

HIGHLIGHTS

- A MODIS-based model of peatland condition was constructed across the land area of Scotland. Restricting the spatial extent to peat locations provided a kappa statistic of 0.8151, suggesting good model fit.
- Comparison with various other spatial datasets containing information about partial aspects of peatland condition further suggested that the model returned appropriate condition classification outputs.
- The resulting spatial model of peatland condition across Scotland suggest a strong geographical divide, in line with historical land use intensity, but also with published predictions in the reduction of bioclimatic space for peatland

- The approach is suitable for larger areas of peatland with low fragmentation but could be revisited as higher resolution satellite data sources become realistically available.

GRAPHICAL ABSTRACT



1. Introduction

Peatlands are one of the largest terrestrial carbon stores and can continually sequester carbon over millennia if they are in near natural condition (Yu et al., 2010). The term 'peatland' refers to peat soils combined with the plant communities that occur on their surfaces. However, peatlands in the temperate zones of the Northern Hemisphere (blanket bogs, raised bogs and fens) have been altered for centuries for various human purposes. Whilst the peat soil that originally supported the peatland habitat remains to varying degrees, the habitats have been affected by an estimated decline in ecological condition over 50% of the blanket bog and more than 90% of the raised bog area in the UK (JNCC, 2011). However, these estimates are generally based on figures for sites under nature conservation protection. In Scotland, conservation status only applies to a small proportion of the peatland habitats; and only a subset of these are designated for their peatland habitat (other sites may be designated for breeding birds, whose population the habitat underpins). In Scotland, only a small proportion of the total peatland area is designated for its blanket and raised bog habitat: a conservative estimate based on 20-year old land cover mapping data from the Land Cover of Scotland 1988 surveys and the boundaries for extent of designated sites suggests a figure of about 14% (A. Coupar, SNH, pers. Comm). The condition of peatland habitat in designated sites in the UK is currently monitored using a rolling 6-year Common Standards Monitoring programme of ground-based assessments (JNCC, 2004, 2009). This programme classifies the current habitat condition of designated habitats into favourable or unfavourable condition categories (see below for details). However, these surveys are labour-intensive and therefore costly to implement (£14 million over 6 years for the first phase; Anon, 2006). In addition, field-based surveys are by necessity based on observations at a relatively small number of point locations, rather than a comprehensive survey of the entire land area. Therefore, remote observations that could inform policy makers of the current state of peatland habitat condition over large areas are of potentially high value.

Peatlands in Scotland cover an estimated 19,000 square kilometres, nearly a quarter of the land mass (Chapman et al., 2009). They can occur as large continuous areas of blanket bog or small areas of lowland raised bogs and fens that are interspersed amongst other habitats. Plant species distribution and soil moisture in undisturbed Northern European peatlands reflect gradients in site hydrology and chemistry as well as climatic gradients, and result in a highly complex repeating mosaic (Harris et al., 2015; Lindsay, 2010; Harris and Bryant, 2009a,b; Belyea and Malmer, 2004) at scales that can be less than 5 m (Lees et al., 2018). The challenge therefore is to find mapping solutions that can measure peatland ecological condition at the appropriate scale. Northern European peatlands can appear visually relatively homogeneous at the 500 m spatial resolution of the moderate resolution image spectrometer (MODIS) aboard the Terra/Aqua satellites, yet display high complexity at spatial resolutions finer than that of even most modern high resolution satellite data sources (e.g. Landsat and Sentinel series). Decline in peatland habitat condition can take the shape of relatively minor damage to the vegetation composition through, for example, excessive grazing, or, at the other extreme, can be caused by full land use conversion. Detecting damage in peatlands produces further challenges for mapping efforts in terms of the spatial extent and complexity, as well as the effects on the vegetation and hydrological components of the system. Most attempts at mapping habitat condition via remote sensing to date have utilised the visible, near infrared and shortwave infrared spectral ranges of satellite data sources (referred to hereafter as optical signals, to distinguish from radar-based approaches). Large scale, full land use conversion, such as afforestation or conversion to agricultural land, results in a very easily recognised change in optical signals at typically 50 - >500 m resolution. Other damage types may be similarly large in scale, but relatively transient (e.g. burning to alter the vegetation specifically for sports shooting purposes). Finally, some damage types can be relatively minor in terms of the changes observed in the optical signals. For example, where displacement of peatland-specific vegetation towards proportionally more grasses occurs, due to overgrazing, atmospheric pollution or fertilisation, this may result in only a minor shift in the visible and infrared range within a satellite image. At the other end of damage types are those that necessitate use of high spatial resolution data; for example, erosion gullies range from <1m to >10 m in width. Peatland drainage channels are typically only 0.5 metre across (the width of the Cuthbertson plough used for most older drains) and they can be as far as 20 to 100 metres apart even in areas that are targeted for drainage. The drain spacing in areas targeted for forestry plantations can be as small as 3 m but typically are approximately 10 m for most upland drainage (Robinson, 1980). Nevertheless, such damage features, whilst relatively small in their individual areal extent, are often densely repeated across the landscape and therefore can cause changes in site hydrology across extensive areas of peatland (e.g. Holden et al., 2017, 2011; Luscombe et al., 2016). These types of damage not only cause decline in habitat condition, but also lead to habitat fragmentation, which, ultimately, can lead to negative effects on genetic diversity of the species inhabiting peatlands (e.g. Wilson and Provam, 2003).

Conversely, to reverse peatland habitat decline and fragmentation of peatland habitat by historical damage, there have been significant efforts to restore peatlands in Scotland since the late 1990s. These efforts have been stepped up as the climate abatement potential of peatland restoration efforts have been recognised within the Scottish Government and UK Governments strategic plans (e.g. Climate Change Plan, 2018). Internationally, the publication of the 2013 Supplement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories: Wetlands (Wetlands Supplement) has focused the attention of Nation states on the potential to mitigate against carbon emissions from peatlands through restoration, and several successful global landscape scale projects have been highlighted by the IUCN UK Peatland Programme (Cris et al., 2014). At present, the UK Government is assessing the feasibility of implementing the Wetlands Supplement into its annual reporting on GHG emissions under UNFCCC and Kyoto Protocol obligations. The Peatland Action programme is Scottish Government's policy instrument to achieve their target to restore 50,000 hectares of peatland by 2020 (Climate Change Plan, 2018). There are now at least 150 completed restoration projects all over Scotland that have been carried out under the Peatland ACTION programme since 2012, covering approximately 15,000 hectares that are 'on the road to recovery' (Artz and McBride, 2016). Monitoring the progress of so many widely dispersed sites is challenging and cost-effective measures to remotely assess progress are therefore in the public interest.

Remote sensing methods have been successfully utilised to map vegetation in peatlands, using Landsat, Sentinel and other high resolution series satellite data sources (e.g. Harris and Bryant, 2009a,b). Very high resolution satellite imagery (e.g. GeoEye-1, IKONOS, etc.) has been successfully used to detect fine scale changes in vegetation (e.g. Mehner et al., 2004), or features such as drains (e.g. Connolly and Holden, 2017) in smaller scale, site-level or regional studies. In addition, even higher resolution visible range data from unmanned aerial vehicles (UAV), manned helicopter flights or airborne hyperspectral monitoring flights have been successfully used to monitor restoration progress in peatlands (e.g. Knoth et al., 2013) or to distinguish floristically discrete peatland biotopes (e.g. Harris et al., 2015, Middleton et al., 2012). However, there are often great costs in acquiring such very high resolution images and also in the subsequent image classification analysis (e.g. Harris and Bryant, 2009a,b). Techniques using remote sensing data sources must be able to detect not only short-term disturbances (e.g. burning) but also long-term changes in peatlands as peatland vegetation is relatively slow-growing. In this study, we investigated the potential of Moderate Resolution Imaging Spectroradiometer (MODIS) data to model peatland condition as defined by the Common Standards Monitoring protocol (CSM, JNCC 2004, 2009). Ground-based data that are required to build national scale models are generally scarce for peatland environments. The CSM monitoring programme is probably the best source of UK peatland condition ground observations that have been collected with a standard protocol, however, the number of observations for any given year is often relatively low and spatially poorly distributed. The training data available to us within the currently complete CSM dataset spanned the period of 2002-2012 and we therefore sought satellite imagery within this period. Although there are data with higher spatial resolution optical data that are freely available in these time slices (e.g. Landsat), there can be challenges in acquiring temporally matching images with low cloud cover across large spatial areas from these data sources. MODIS has a much higher pass frequency (1-2 days) than Landsat (8 days), and in addition, the long-term MODIS archive does not suffer from missing data, such as strips missing due to e.g. the scan line corrector failure issues that affected Landsat 7. Spatio-temporal modelling generally requires some form of gap filling for missing pixels due to cloud cover. Due to the oceanic location, Scotland has a moist temperate climate, which means its landmass is frequently cloud covered in a semi-consistent spatial pattern with greater persistence along coasts and in the mountain areas (Perry and Hollis, 2005). A higher pass frequency increases the chance of finding space-time neighbour images that can be used to gap fill across missing pixel values due to cloud cover (Poggio et al., 2012). For these reasons, MODIS data was selected as being one of the most likely cost-effective source of long term data for a first attempt at modelling national scale peatland condition.

2. Data and methodology

2.1. MODIS data preparation and modelling

A set of indices (below) were acquired as 8-day composite products from the MODIS satellite for mainland Scotland, the Western Isles and Orkney for the period 2002-2011. Composites contain the best possible observations obtained during the period, based on parameters such as view angle, absence of clouds, cloud shadows or aerosols. We used time series of indices describing vegetation greenness (Enhanced Vegetation Index, EVI and Soil Adjusted Vegetation Index, SAVI), water availability (Normalised Water Difference Index, NDWI, Gao et al, 1996, using the index based on NIR and the short-wave IR band at 2130), land surface temperature (LST, Wan, 1999) and vegetation productivity (Gross Primary Productivity, Running et al., 2004). The median of MODIS data for the 12 years were used, with cloud gaps filled using the method described in Poggio et al. (2012). Briefly, this method is an example of a hybrid Generalised Additive Model (GAM)-geostatistical space-time model and included the fitting of a spatio-temporal smoother with related covariables and a further spatial component through kriging of GAM residuals. Depending on the type of cloud data loss (e.g. at extremes, highly temporal but widespread cloud cover, versus localised but persistent cloud cover), this method is very competent at restoring missing data due to clouds. A simulation with data from the year 2005 provided competent reproduced spatial patterns and local features of the MODIS EVI product, even with substantial amounts of missing pixels (i.e. up to 80% of missing values, see Poggio et al., 2012 for further details). The spatio-temporal interpolation for missing areas due to cloud was performed using only MODIS images on dates that were close together, as per Poggio et al. (2012). Data for the Shetland Islands were excluded because the high cloud coverage in both spatial and temporal terms, i.e. highly persistent and extensive cloud coverage across the Shetland Islands from most of the year across the entire 2002-2011 time slice) did not allow for the implementation of the method.

The available point information about peat condition from the CSM rolling programme (see training data below) was used as training data for a Random Forest model (Hengl et al., 2004; see Figure 1 for a workflow diagram) that included the statistical relationships between morphological features such as elevation, slope and topographic wetness index (Sorensen et al., 2006) and the MODIS time series of EVI, SAVI, LST, NDWI and GPP (as above). Other covariates included average snow cover (Poggio and Gimona, 2015), elevation and interpolated percentage of organic matter in the soil (Poggio and Gimona, 2014). An additional covariate was a scale-invariant tensor product smooth of space-time dimensions. This surface, relating the coordinates (x and y) of the points, was created using a Generalized Additive Model (GAM, Wood, 2006). The random forest method is a modification of the regression kriging approach (Hengl et al, 2004), an established technique for digital soil mapping (McBratney et al, 2003). The validation statistics were calculated on an out-of-sample set obtained by randomly sampling 30% of the locations from the dataset. This split was repeated 100 times and the statistics averaged over the iterations, following a traditional bootstrapping aggregating approach to obtain more representative samples of the points given the relatively low number of locations available.

The resulting model returned probabilities of a given pixel to be in favourable condition. Model fit, and the optimum threshold for classification, were assessed using Receiver Operator Characteristics (ROC) curves (ROCR package, Sing et al., 2005), using the model output that was truncated to locations on peat soil (see below). ROC curves graphically illustrate the diagnostic ability of a model that is using a binary classifier, as its discrimination threshold is varied. Confusion matrix statistics of the binary classifiers were calculated using the caret package (Kuhn et al., 2016), using the training dataset truncated to locations on peat soil (see below).

2.2. Site condition monitoring training and internal validation data

Scottish Natural Heritage staff provided data from the rolling six-year Site Condition Monitoring programme, which is based on Common Standards Monitoring (CSM) guidance (JNCC, 2004, 2009), for designated upland and lowland peatland sites in the period of interest. The CSM method assesses site condition based on several criteria for each habitat type. These include i) feature extent, ii) vegetation composition, iii) vegetation structure, and iv) physical structure. Vegetation composition assessments include frequency of taxa which are indicators of favourable condition, cover of taxa which are indicators of favourable condition, and others which are indicators of unfavourable condition. Vegetation structure assessments include vegetation height, removal or destruction of plant parts by grazing, browsing, burning and trampling, accumulation of plant litter in the sward, and dieback of typical species. The spatial scale of these assessments is generally 4 m^2 for peatland habitats, except in cases where features may be small or fragmented in area, such as transition mires, ladder fens and quaking bogs. In such cases, an assessment might be made on individual, smaller, stands or patches. Finally, the assessment of physical structure, includes attributes for levels of ground disturbance, burning, drainage or drying, indicating damage to the habitat. Some attributes (e.g. burning, erosion) are assessed while travelling between sample locations or as line of sight from the sampling location, i.e. over a much wider area than the 4 m^2

vegetation quadrats, as such features generally occur more sporadically. For each of the selected attributes, one or more target(s) are set, as specified in the relevant guidance (JNCC, 2004, 2009). For a given location, all attributes must pass the stated target(s) at the sample point; if one attribute fails, then that particular sample point is considered to fail the CSM assessment. For the purposes of national level summaries, CSM reporting normally combines individual point pass/fail rates for the entire (peatland) habitat within a designated site. For a given designated site to meet favourable condition, each habitat type within the site must have met the pass status in at least 90% of the sampled locations.

The designated site condition by Scottish Natural Heritage is generally reported for the site centroid location, which can be several kilometres away from the nearest individual point locations visited for the assessment. In addition, it reports the condition of the whole site rather than just the peatland habitat as peatland habitat sometimes only covers a fraction of the designated site alongside other habitat types. Therefore, and to more effectively upscale from the usually several 4 m^2 individual point observations taken per MODIS pixel at individual designated sites, we used the original point location CSM data (see workflow, Figure 1). The point location data were summarised to favourable or unfavourable status, by assuming that a single fail for any given category would equate to unfavourable ecological status at that point location (i.e. analogous to the CSM methodology). These were distributed as per Fig. 2. Out of the original 951 points, 7 were located on the Shetland Mainland, and a further 2 points had co-ordinates outside of the Scottish land area. Excluding these points resulted in a final training dataset of 943 training point locations. All point location data were combined across years of observations, as data for individual years were either too low in number or spatially poorly dispersed (data not shown).

2.3. Peat soil mask

The initial model output represented the entire land area for Scotland (see 2.1), which includes other soil types beyond peat. In Scotland, peat soils are defined as soils with an organic horizon of >50 cm (the Scottish Soil Survey definition of peat), although blanket bog habitat can occur on <50 cm of peat depth. To limit the model output to areas with peat soil, we employed two potential masks of peat extent: a) the modelled peat extent by Aitkenhead (2016) and b) a mask created by a simpler model than that of Aitkenhead, by combining data from three spatial peat mapping data sources and limiting the locations of peat within the mixed polygons to areas defined with a slope threshold. The spatial data sources contained data on peat-containing soil polygons from the National Soil Map of Scotland (full national cover, 1:250,000), which was GIS intersected with peat polygons from the Soil Map of Scotland (partial national cover, 1:25,000). A further GIS intersection was made with the peat polygons from the UK DigiMap version 6 (British Geological Survey). In areas where the 1:250,000 maps specified 100% peatland, or both the 1:25,000 maps and the UK DigiMap datasets agreed that peat was present, these were attributed to be 100% peat soil. The remaining polygons where peat was a proportion of the area of the polygon (varying between 30 and 75%) rather than a spatially discrete area were spatially limited to areas using a slope threshold based on a 5 m DTM (Terrain 5, Ordnance Survey, UK), in order to spatially allocate the peat to the shallower slopes. This slope threshold was chosen based on National Soil Inventory of Scotland (NSIS1, 1978-88, Lilly et al., 2010) soil profiles, which included statistics on the slope at the location of each soil pit. Averages and standard errors of the slope data from the soil pit locations on peat and non-peat soils were calculated and a slope of 15% was found to be the upper 84th percentile for peat soil extent whereas non-peat soils were found to be distributed to steeper slopes. The predictive capacity of the resulting spatially explicit model of peat distribution was tested against the 10 km grid point location data in the NSIS1 database, which contains 728 peat and 2457 non-peat soil locations across Scotland.

2.4. Independent additional model validation I: Assessment of classification threshold, based on areas under known landcover (proxy for condition I)

The distribution of the training data points (Figure 2) was clearly not fully representative of the peatland condition across the whole of Scotland. This is due to the distribution of designated sites, which is not random across the Scottish land area. In addition, nature conservation protection tends to apply to sites that were examples of good condition at time of designation, rather than sites in need of management. We therefore tried to find additional datasets to validate the model outputs, especially to test the model in areas where training data were lacking. Unfortunately there are no other long-term national scale monitoring programmes in existence, and therefore we were forced to use other datasets of land cover and vegetation community composition, that indicate condition by proxy, instead.

The UK National Forest Inventory produces an annual update of forest cover for the UK. These data, when GIS intersected with a peat extent map as above, produce a layer of peatland sites currently under forestry, which would be classified as being in unfavourable condition in a CSM-based assessment based on the vegetation criteria alone. We also used a previously existing dataset of digitised areas of peat erosion (Cummins et al., 2011), which would similarly fail to meet the CSM criteria for favourable status due to the presence of bare peat. A third independent habitat condition dataset was obtained from the Royal Society for Protection of Birds (RSPB) Scotland for the Forsinard reserve (England, 2008), which is a reserve that includes extensive areas of peatland in good habitat condition as well as large areas undergoing restoration after former afforestation. Here, we assumed that sites in good habitat condition as per RSPB's methodology would be equal to favourable condition under CSM methodology as many criteria are similar, and that restoration sites have not yet fully recovered to favourable condition as the vegetation criteria of the CSM assessment would not be yet met.

2.5. Independent additional model validation II: Assessment of model classification threshold against manually assessed drainage status (proxy for condition II)

As the preceding three validation datasets were very small, we also assessed 500 m blocks, aligned with the MODIS pixels and occurring on peat soils, for evidence of drainage to produce a further external validation dataset. Any sites affected by drainage would also be classified as being in unfavourable condition under CSM methodology. High-resolution aerial photography was provided under licence by GetMapping©. This imagery provides full coverage of Scotland at a spatial resolution of 0.25 m, with a rolling programme of flights ensuring imagery is no more than five years old (and usually less than three years old). Only the RGB imagery was used for this project. A total of 400 georeferenced points across Scotland's peat soils were randomly generated using conditioned Latin hypercube sampling (Minasny and McBratney, 2006) and used as centroids for 500 m blocks within the MODIS pixels. These sites were selected using a stratification approach designed to ensure that there was equal representation across different elevation, easting, northing and climate ranges. The corresponding 500 m images were extracted from the GetMapping© imagery. The images were overlaid onto the peat mask and only images with > 50% cover on peat were selected (i.e. only those that did not include edge effects due to the conversion from points to 500 m blocks, n=221). Peatland drainage classes (1-6, Table 2) in the remaining blocks were assigned based on a visual classification that considered the density of drains in each image block. The drains were digitised for a subset of 49 blocks and assigned a 0.25 m buffer either side of the drain, to estimate the density of pixels assigned to drains within a 500 m block. The drain pixel density was assessed using the resolution of Getmapping (0.25 m). A second attribute included any additional features that could contribute to drainage effects such as erosion gullies or complete land cover change to crop/forestry cover as these would necessitate drainage before planting. We assumed sites in class 1 would be in favourable condition, whereas all other classes would be in unfavourable, and increasingly worse, condition. The visual examination process was carried out iteratively and by two people working independently at first. Disagreements were subsequently solved by consensus through a second review involving both assessors.

2.6. Independent model validation III: Assessment of model classification threshold against site condition proxies based on published vegetation composition data (proxy for condition III)

We were aware that the assessment datasets under external model validation I and II suffered from a lack of detail on the components of the specific vegetation composition criteria that were assessed at the same spatial resolution as the training dataset (4 m^2). To overcome this limitation, we assessed the feasibility of inferring habitat condition from vegetation survey data from previously published studies. In the mid-2000s, Ross et al. (2012) resurveyed the plant communities of Scottish uplands that had been previously described in detail by McVean and Ratcliffe (1962). Similarly, Britton et. al. (2009, 2017a, b) reported on resurveys of vegetation composition plots first surveyed between 1963 and 1987 by Birse & Robertson (1976) and Birse (1980, 1984). Here, we re-interpreted the datasets from the recently resurveyed locations with the JNCC (2004, 2009) CSM approach as previously stated. Not all JNCC CSM qualifying criteria could be assessed based on the published vegetation cover data of Ross et al (2012) and Britton et al (2009; 2017a) alone, as, for example, the CSM methodology also assesses browsing, burning and extent and activity of erosion. These CSM criteria, except for browsing, were instead assessed visually across the relevant 500 m block for each data point from the resurveyed locations, by overlay with aerial photography (GetMapping©) as above. We assumed that the single 4 m^2 vegetation quadrat observation was representative of the vegetation community across the whole 500 m pixel but excluded sites where this was clearly not the case based on the aerial photography assessment (where there was visibly mixed vegetation within a 500 m pixel). We also excluded sites where the relevant pixel was on less than 50% peat as per model assessment II.

3. Results

3.1. Unconstrained MODIS-based prediction of peatland condition

The predicted probability of a peatland site being in favourable condition across Scotland was modelled using MODIS data against a training dataset of peatland condition status from 943 point locations from the latest available round of the CSM programme. The unconstrained MODIS-based model returned data for all of Scotland's land area, based on the 100-fold validation of an out-of-sample split of 30%. This unconstrained model was predominantly driven by site elevation and NDWI. Sites above 755 m returned a very low probability of being in favourable condition, presumably as peat depths at such altitudes would be shallower and the growing season short, thereby magnifying the effects of any damage done to such sites. Another significant discriminating factor was the NDWI of vegetation water content. Sites in unfavourable condition would be expected to have lower and more variable water tables, thus placing constraints on water availability in peatlands reliant on rainfall as water inputs. This unconstrained model, however, was for the entire Scottish land area which includes areas that are not on peat soil. Therefore, this model output was further constrained with the masks of spatial peatland extent.

3.2. Peat mask validity

The peat mask we devised by slope limiting a GIS intersected map originating from three data sources of peat soil information was 74.6% accurate in detecting peat and 82.4% accurate for non-peat (Fig. 3). There was no distinct geographical pattern for any of these incorrectly identified locations. There was also no correlation of any locations that were incorrectly predicted with polygons that contained less than 100% peat in the 1:250,000 soils map or with steeper slopes.

The peat extent mask by Aitkenhead (2016) was based on a neural network built using a mosaic of 2013 Landsat 8 summer image data, using all 11 30-m bands, and various covariates including elevation, slope, slope curvature, aspect, rainfall and temperature as well as land cover mapping information (please refer to Aitkenhead, 2016 for the methodology). This produced a model output of peat soil distribution with an overall accuracy of 86.4%. The two models of peat extents were largely similar, although the Aitkenhead (2016) model suggests overall lower peat coverage and smaller sizes for individual peat areas (Supplementary Figure 1). In addition, a significant proportion of the non-peat training points that were incorrectly classified as peat in the model presented here were not predicted to be peat by the Aitkenhead (2016) model. This raises the distinct possibility that the 1:250,000 Soils of Scotland map overestimated peat, both in 100% peat polygons and in mixed soil polygons and may be due to the partial extrapolation from land cover at the time.

3.3. Constrained model

Using only the training data points that co-located on the peat mask (716 points), we assessed Receiver Operator Characteristics. ROC curves for restricted training data, limited to those located on a peat soil map, suggested a threshold of 0.56 of the probability to be in favourable condition for classification of a site as being in favourable status (Table 1, Supplementary Fig 2). The model was assessed as having an accuracy of 0.916, and the kappa statistic was 0.8151, suggesting good model fit (Table 1). Constraining the MODIS model outputs with the Aitkenhead (2016) peat mask suggested a similar threshold value of 0.562 of the probability to be in favourable condition, despite some spatial differences in the predicted peat areas (Table 1, Supplementary Figures 2 and 4). This is an encouraging result, suggesting that the model is spatially consistent. One of the limitations of the model was that the training dataset was not a fully representative sample of the peat biogeophysical space across Scotland; Figure 2 shows that the input data were strongly clustered. We therefore attempted to find additional data sources that could test the model outputs for verification of the threshold value for the classification.

3.4. Model assessment I: based on areas with known site condition or drainage status

Areas with known site condition from the various GIS maps provided by the UK National Forest Inventory, previous peat erosion surveys (Cummins et al., 2009) and the RSPB Forsinard Habitat Condition Monitoring Programme (England, 2008), were assessed visually against GetMapping© aerial imagery and a grid of the 500 m MODIS pixels. Only 70 locations could be identified from these three data sources where a peatland area in known condition occupied at least 70% of a 500 m MODIS pixel, and where this pixel was located on an area with more than 50% peat (Figure 4). An assessment of the model fit to these 70 locations in known condition showed that the model was able to distinguish areas in assumed favourable condition (near natural, average probability significantly above 0.56) from those in unfavourable condition due to complete land cover conversion (afforested) or severe erosion (Fig 4). However, although other areas in unfavourable condition such as drained and restored areas had a significantly lower average probability of being in favourable condition than natural areas, such areas could not be distinguished from near natural peatland based on the model threshold of 0.56 (Fig. 4). Hence, the model threshold of 0.56 would have correctly placed near natural areas into the favourable condition category and eroded and afforested areas correctly into the unfavourable category. However, the drained and restored areas would have been classified as being in favourable condition on the basis of the returned probabilities for the tested areas (Fig. 4). Although the aim of peatland restoration is to restore the habitat to its former functionality, inclusive of its vegetation complement, the restoration sites in the RSPB Forsinard reserve have only recently been restored from former afforestation, and even in the oldest restoration sites, vegetation has not yet fully recovered to that of a near natural community (Hancock et al., 2018. In addition, there was no trend in the predicted probability of a site being in favourable condition that was dependent on the year when restoration had been carried out (Fig. 4). This may, however, have been due to the low number of restoration sites assessed (n=14).

3.5. Model validation using manually assessed peatland drainage status

The small validation dataset above (section 3.4) suggested that decline in peatland condition due to drainage may not be detected with our model. However, this may have been due to the low number of observations for this category (n=9). Therefore, we created a larger dataset based on digitisation of high resolution aerial photography (Fig. 5). In addition, at low and medium drain density categories, only part of each 500 m block was affected by drainage. Calculation of the pixel density proportion that the drains occupied within 500 m blocks in the digitised subsample returned averages of 0.15% for Class 1-2; 0.25% for Classes 3-4; and 1.19% for Classes 5 and 6. This may seem low but is due to the small width of these drains (0.5 m) coupled with drain intervals that range from 3 m (forestry) to >20 m (upland drainage). The ranges of the proportion that drain pixels occupied, however, was quite large, with the maxima almost overlapping with the minima of the next category.

We extracted the modelled probability of being in favourable condition at each of the drainage assessment site from the constrained model outputs. This showed a decline in the average probability of being in favourable condition across the drainage class gradient (Table 2). Sites without any drainage features (Class 1) had an average predicted probability of being in favourable condition of 0.67 +/- 0.3, whereas all other drainage classes had significantly lower average predicted probabilities (Table 2). However, only sites in drainage classes 5 and 6, and some sites in Class 4, would be classified as unfavourable based on the model threshold of 0.56. It is possible that the resolution of MODIS is insufficient to detect potentially localised effects of drainage, especially where drainage is not applied across the entire area occupied by a MODIS pixel. Conversely, however, drains are not always effective at draining the landscape and may thus not lead to significant effects on vegetation and/or site hydrology.

3.6. Model assessment II: Site condition proxy based on published vegetation composition data

A final external validation attempt of the model was made that included a similar hybrid in terms of resolution to our training dataset. The starting dataset from the McVean and Ratcliffe resurvey (Ross et al., 2012) included 107 moorland and wetland vegetation quadrats. Of these, 63 sites met the condition of being located on over 50% peat. Inferring condition status from these resulted in 25 pixels of favourable condition and 38 pixels of unfavourable condition (Table 3, Fig 6a). The majority of sites inferred to be in unfavourable condition failed on the basis of a) greater than 50% ericaceous species cover, b) cover of non-peatland ruderal species such as *Holcus lanatus* exceeding 1%, or, c) in a relatively small number of cases, cover of tree species that exceeded the 10% threshold. Sites at higher altitude more frequently fell into the inferred unfavourable condition class due to site erosion (Fig. 6a). The average predicted probability of being in favourable condition was significantly different (Table 3) for the resurveyed McVean and Ratcliffe sites inferred to be in favourable or unfavourable condition, although there was substantial overlap between the two categories.

The Birse and Robertson resurvey (Britton et al., 2009) included 132 moorland and peatland vegetation sites, all of which met the condition of being on more than 50% peat as per our peat extent model. The inferred condition status from this dataset resulted in 49 pixels in inferred favourable condition and 83 pixels in inferred unfavourable condition (Table 3, Fig 6b). Within this dataset, the most common reason for sites to be inferred to be in unfavourable condition was a failure to meet the threshold for the required number of indicator species, followed by a few sites exceeding the threshold for ericaceous cover (i.e. > 50%). The average predicted probability of being in favourable condition for the resurveyed Birse and Robertson sites inferred to be in favourable or unfavourable condition were also statistically significantly different (Table 3), however, the group averages were also substantially lower than for the two groups from the McVean and Ratcliffe data. The boxplots of the distributions of MODIS probabilities to be in favourable condition for each group are shown in Supplementary Fig. 2, which demonstrates that both datasets have 'tails' into low predicted probabilities of favourable condition for the sites inferred to be in favourable condition as well as those inferred to be in unfavourable condition.

3.7. Predicted condition from the constrained model for the entire peatland resource

The unconstrained model was built using 943 training points. Following constraining of the spatial output with a peat mask, the model statistics suggested that a threshold value of 0.56 could be used to successfully predict condition (Table 3). The external validation procedures above (sections 3.4-3.6) provided an additional 486 data points (n=70 for validation I, n=221 for validation II and n=195 for validation III), which contributed spatial locations that were not, or only sparsely, covered by the CSM training dataset. While the results from the additional validation cannot directly compared as they were largely based on proxies of the CSM methodology, the data nevertheless suggested that the threshold value was not unrealistic for differentiating clear examples of favourable and unfavourable condition (e.g. near natural sites versus those with erosion, full land cover conversion or a compromised vegetation community). The secondary validation did, however, show that there were limitations in the detection of unfavourable site condition due to drainage, and that the condition of restoration sites may be estimated as more favourable than it might be on the ground.

Based on the observed threshold of 0.56 of the predicted probability to be in favourable condition, we created a map of predicted condition status (Fig. 7) by allocating pixels with a probability <0.56 to the unfavourable category and those pixels with a probability >0.56 to the favourable condition category. The resulting map of peatland condition suggested significant geographical differences in the spatial distribution of peatland in predicted favourable or unfavourable condition.

4. Discussion

The model of peatland condition, constrained to peat soil locations, was of good predictive capacity (Table 1). There are, of course, caveats with this approach of first modelling at full national scale, and then constraining to peat soil extent later: Firstly, the accuracy of the constrained condition model is critically dependent on the accuracy of the peat extent model(s). The simple peat extent model we created in this study was only moderate in its ability to predict where peat existed (Fig.2), so therefore our approach may have simply been serendipitous in improving the accuracy of the final, constrained, condition model. However, we also tested our approach by constraining the condition model with a previously published peat extent model that suggested a slightly different spatial distribution (Aitkenhead, 2016), with no significant differences to the model statistics (Table 1). The reason for the improved accuracy after constraining with a peat extent mask, we believe, lies in the distribution of blanket bog habitat. As stated earlier, blanket bog habitat in Scotland can occur on organic soils of less than 50 cm (the definition applied by the survey teams who created the (National) Soil Map of Scotland and the National Soil Inventory of Scotland). We believe that blanket bog on such shallower organic soils may be more susceptible to drought phases due to limitations in the water storage potential of such soils and that such occurrences would have resulted in a different signal in the MODIS NDWI to sites on peat more than 50 cm deep. NDWI and similar water indices have been previously tested by others (e.g. Meingast et al., 2014, Kalacska et al, 2018, and references in both) and found to have a strong relationship with surface volumetric water content in northern peatlands. Hence, excluding such shallower site would have correctly improved model accuracy for peatland habitat, and we believe our approach to be valid given the limitations of the various data sources and the nature of blanket bog habitat occurrence.

In our view, the model carries some potential to detect differences in site condition at national scale, although it should not be used to infer actual condition at site level given the moderate resolution of the model input and hence output. Many Scottish and UK peatlands can show significant fragmentation at smaller scales than this model can predict. In addition, our external validation using manually assessed drainage suggested that the model overestimated the site condition for drained sites (Table 2). A limiting factor here may have been error terms introduced by the visual assessment. The methodology was successful, as the average pixel densities were statistically different between drainage categories in the subset of blocks where drains were fully digitised. However, the ranges of drain pixel density per category were quite wide and hence a more stringent approach would have been to fully digitise the drains in all 500 m blocks and form drainage categories based on the statistics of these (i.e. relate percent cover of drains to the modelled probability of favourable condition). However, this was not feasible within the constraints of this project. In addition, not all functional drains may be visible on aerial imagery (e.g. if they are overgrown) or conversely, not all visible drain features may function (equally) as active drains in the landscape and some peat piping and drains may not be visible from aerial photographs. Connolly and Holden (2017) used automated image analysis tools to identify drains in peatlands, however their test area was relatively small and the drains more organised (for peat cutting) than in a typical UK upland. Nevertheless, our results are encouraging in that there was some distinction between undrained and heavily drained sites. although complete CSM assessments at locations with different drain densities would be required to validate this further.

There was also no observed relationship of the probability to be in favourable condition with time passed since restoration activities. We assume that the lack of an observed restoration effect is at least in part due to the use of median annual images spanning 2000-2011, during which most of the restoration work on the ground on the sites we identified had been carried out, thus obscuring potential year-on-year changes by interpolating between pre-restoration and post-restoration condition. It is, however, curious, that the model predicts most of these restoration sites to be in favourable condition as this is not the case on the ground. Many of the restoration sites included do not yet have the required vegetation community to pass the CSM assessment, with keystone species still lacking (e.g. Hancock et al., 2018).

The final attempt to assess the model using inferred site condition from previous vegetation surveys augmented with visual assessment of erosion and burning produced similar, but even less robust results. Although there were statistical differences between the site groups classed as being in inferred favourable or unfavourable condition and the threshold for these datasets was similar to that obtained earlier, there were large ‘tails’ in the distributions of these observations that included low probabilities to be in favourable condition even in sites inferred to be in favourable condition. This may be due to a discrepancy in the resolution of the vegetation data, as these originated from single 4 m^2 surveys and hence are less likely to be representative of the condition across the 500 m MODIS pixels than our training data, which consisted of multiple observations of site condition per 500 m MODIS pixel. We believe that our visual assessment for these survey sites across the wider 500 m block for the CSM criteria that were not captured by the vegetation community composition did produce a marginally better validation dataset, however the results further highlight the need for spatially more representative ground observations if remote assessments are to be developed further (see also below). Again, full CSM assessments would be required to validate our model outputs.

To our knowledge, this is the first attempt to directly model peatland habitat condition using remotely sensed data at national level. There have been several other studies that classified peatland vegetation types, rather than condition. Generally, these attempted to build high resolution models of vegetation types in relatively small areas (e.g. Mehner et al., 2004, Knoth et al., 2013; Harris et al., 2015, Middleton et al., 2012), however Pflugmacher et al. (2007) attempted mapping across a larger geographical region for the St Petersburg region in Russia using a sub pixel proportional cover approach. They trained a MODIS-based model on mapped peatland sites of different site nutrition types that were either mined for peat or not and were able to build a reasonably accurate model. Connolly et al (2011) were able to detect various disturbance factors, such as burning, that could result to decreased peatland condition, using a MODIS EVI-based model for the Wicklow area in Ireland. Krankina et al. (2008) further noted the usefulness of moderate resolution remote sensing data in mapping peatlands across larger geographical regions in Russia. We believe that our approach is a potentially cost-effective method to detect peatland condition across large continuous areas (range of several km^2) where there is a low degree of internal fragmentation. Others have noted the potential for remotely sensing greenhouse gas exchange (Lees et al., 2018; Gatis et al., 2017) using MODIS data. As such, there are limitations due to the MODIS image data resolution in a UK context of significant areas of small, heavily fragmented peatlands. Although computationally more intensive, Landsat 7 images may provide an alternative over the time frame used in this study, although the data loss due to the scan line corrector failure needs to be addressed via appropriate gap filling techniques. This was not feasible within the constraints of this project. Going forward, Landsat 8 data may be useful if this approach is to be revisited with the data from the next tranche of data from the CSM programme. Other optical alternatives such as using Sentinel-2 image data are not (yet) viable at present due to the relatively low data availability of images with low levels of cloud across higher altitudes and coastal areas due to the time lag between the launch of Sentinel-2A in 2015 and Sentinel-2B in 2017. However, we are currently working on a spatially more restricted model of peatland condition using Sentinel-2 data.

A potentially highly policy relevant observation in our study is the observation that peatlands more likely to be in favourable condition were predominantly located in the North and West of Scotland. Not only does this observation parallel the historical land use intensity across Scotland, but it could also suggest that climate change impacts already add to existing pressures. Gallego-Sala and colleagues and Clark et al. (both 2010) used bioclimatic envelope models to predict the likely geographical distribution of blanket bogs in the UK under UKCIP02 climate projections. Their findings suggested that the blanket bog bioclimatic space would decline dramatically under a high emissions scenario, with predominately western and northern coastal areas of Scotland remaining inside suitable bioclimatic space by the 2080s. Our model suggests that the distribution of peatland habitat in ecologically favourable condition may already be skewed towards western and northern areas. There are confounding issues of course, as our model is designed to detect human impacts as part of the overall condition, however the distribution of sites predicted to be more likely to be in favourable may have a climatic component related to rainfall. The findings of Ross et al. (2012) and Britton et al. (2017a, b) that there was some evidence of degradation of the peatland plant communities through pollution and climate change (e.g. increase in graminoid cover) also adds weight. Therefore, the climate sensitivity of blanket peatlands may be higher than predicted by current bioclimatic envelope models, especially given that these used the then available maps of spatial extent of peat soils as training data (i.e. not a map of currently active peatland which would be smaller in spatial extent and more fragmented). Conversely, as discussed, our model appeared to be too optimistic at predicting the condition of drained and restored peatlands. To date, there is no map in existence of peat drainage across Scotland. Robinson (1990) is the only source we were able to find that compiled the percentage of land drained, but this did not distinguish peat soils from other soil types and only reported averages for regions that were roughly analogous to the modern-day Local Authority boundaries. More work is required to fully ascertain the current condition of peatlands remotely, and although this is only a first, and moderate scale, attempt, maps of peatland condition could perhaps be used as a more appropriate input dataset for bioclimatic envelope modelling to predict future climate sensitivity.

Acknowledgements

We gratefully acknowledge comments on an earlier version of this manuscript by an anonymous reviewer and discussions with Andrew Nolan about the model outputs.

Funding: All James Hutton Institute authors are supported by the Scottish Government's Rural and Environment Research and Analysis Directorate under the current Strategic Research Programme (2016-2021). Sally Johnson, Patricia Bruneau and Louise Ross did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors for this project. The peat spatial extent model was created in part within a UK Government – Department for Business, Energy and Industrial Strategy-funded project (TRN860/07/2014, Scoping the use of the methodology set out in Chapters 2 and 3 of the '2013 Supplement to the 2006 IPCC Guidelines for National Greenhouse Gas Inventories: Wetlands in the UK GHG Inventory: Land Use, Land Use Change and Forestry (LULUCF)), with further updates created within the Strategic Research Programme (2016-2021) funding.

References

Aitkenhead, M.J. 2016. Mapping peat in Scotland with remote sensing and site characteristics. European J. Soil Sci. DOI: 10.1111/ejss.12393.

Anon. 2006. Common Standards Monitoring – Lessons learned from the first 6-year report. Joint Nature Conservation Committee. December 2006. http://jncc.defra.gov.uk/pdf/comm06P19.pdf (Last accessed: 25/9/2018).

Artz, R.R.E.; McBride, A. 2016. Data from the Peatland Action Programme and their use for evaluations of ecosystem benefits. Report for ClimateXChange Scotland (https://www.climatexchange.org.uk/media/1485/cxc_peatland_action_data_uses.pdf)

Belyea, L. R., Malmer, N. 2004. Carbon sequestration in peatland: Patterns and mechanisms of response to climate change. Global Change Biol., 10, 1043–1052.

Birse, E.L. Robertson, J.S. 1976. Plant communities and soils of the lowland and southern upland regions of Scotland. The Macaulay Institute for Soil Research, Aberdeen, UK.

Birse, E.L. 1980. Plant communities of Scotland: a preliminary phytocoenonia. The Macaulay Institute for Soil Research, Aberdeen, UK.

Birse, E.L. 1984. The phytocoenonia of Scotland: additions and revision. Macaulay Institute for Soil Research, Aberdeen, UK.

Breiman, L. 2001. Random Forests Machine Learning, 45, 5-32.

Britton, A.J., Beale, C.M., Towers, W. Hewison, R.L. 2009. Biodiversity gains and losses: evidence for homogenisation of Scottish alpine vegetation. Biolog. Cons. 142, 1728–1739.

Britton, A.J., Hester, A.J., Hewison, R.J., Potts, J.M., Ross, L.C. 2017a. Climate, pollution and grazing drive long-term change in moorland habitats, Appl. Veg. Sci. 20, 194-203.

Britton, A.J., Hewison, R.L., Mitchell, R.J., Riach, D. 2017b, Pollution and climate change drive long-term change in Scottish wetland vegetation composition, Biol. Con. 210, 72.

Chapman, S., Bell, J., Donnelly, D., Lilly, A. 2009. Carbon stocks in Scottish peatlands. Soil Use Manage. 25, 105–112.

Clark, J., Gallego-Sala, A.V., Allott, T.E.H., Chapman, S.J., Farewell, T., Freeman, C., House, J.I., Orr, H.G., Prentice, I.C., Smith, P. 2010. Assessing the vulnerability of blanket peat to climate change using an ensemble of statistical bioclimatic envelope models. Cli. Res. doi:10.3354/cr00929.

Climate Change Plan, 2018. The Scottish Government's Climate Change Plan, Third Report on Proposals and Policies 2018-2032 (RPP3). http://www.gov.scot/Publications/2018/02/8867

Connolly, J., Holden, N.M. 2017. Detecting peatland drains with Object based Image Analysis and Geoeye-1 imagery. Carbon Balance Manag. 12, 7.

Connolly, J., Holden, N.M., Seaquist, J., Ward, S. 2011. Detecting recent disturbance on Montane blanket bogs in the Wicklow mountains, Ireland using the MODIS enhanced vegetation index. Int J Remote Sens 32, 2377-2393.

Cris, R. Buckmaster, S. Bain, C. Reed, M. (Eds). 2014. Global Peatland Restoration demonstrating SUCCESS. IUCN UK National Committee Peatland Programme, Edinburgh.

Cummins, R., Donnelly, D., Nolan, A., Towers, W., Chapman, S., Grieve, I., Birnie, R.V. 2011. Peat erosion and the management of peatland habitats. Scottish Natural Heritage Commissioned Report No. 410.

England, B. 2008. Habitat condition monitoring methodology, Repeat Survey 2008, Forsinard Flows Nature Reserve, Sutherland. RSPB Scotland.

Gallego-Sala, A.V., Clark, J.M., House, J.I., Orr, H.G., Prentice, I.C., Smith, P., Farewell, T., Chapman, S.J. 2010. Bioclimatic envelope model of climate change impacts on blanket peatland distribution in Great Britain. Clim. Res. 2010, doi: 10.3354/cr00911.

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

Gatis, N., Anderson, K., Grand-Clement, E., Luscombe, D., Hartley, I.P., Smith, D., Brazier, R.E. 2017. Evaluating MODIS vegetation products using digital images for quantifying local peatland CO_2 gas fluxes. Remote Sensing in Ecol. Con. 3,217–231.

Hancock, M.H., Klein,D. Andersen, R., Cowie, N.R. 2018. Vegetation response to restoration management of a blanket bog damaged by drainage and afforestation. Appl Veg Sci doi 10.1111/avsc.12367.

Harris, A., Bryant, R.G. 2009a. A multi-scale remote sensing approach for monitoring northern peatland hydrology: present possibilities and future challenges. J. Env. Manag. 90, 2178-2188.

Harris, A., Bryant, R.G. 2009b. Northern Peatland Vegetation and the Carbon Cycle: A Remote Sensing Approach. In: Baird, A.J., Belyea, L.R., Comas, X., Reeve, A., and Slater, L. , editor(s). Carbon Cycling in Northern Peatlands: Geophysical Monograph Series. Washington D.C., USA: American Geophysical Union; 2009. p. 79-98.

Harris, A., Charnock, R., Lucam R.M. 2015. Hyperspectral remote sensing of peatland floristic gradients. Remote Sensing of Environ. 162,99-111.

Hengl, T., Heuvelink, G.B.M., Stein., A. 2004. A generic framework for spatial prediction of soil variables based on regression-kriging. Geoderma, 120, 75-93.

Holden, J., Green, S.M., Baird, A.J., Grayson, R.P., Dooling, G.P., Chapman, P.J., Evans, C.D., Peacock, M., Swindles, G. 2017. The impact of ditch blocking on the hydrological functioning of blanket peatlands. Hyd. Proc. 31, 525-539.

Holden, J., Wallage, Z.E., Lane, S.N. and McDonald, A.T. 2011. Water table dynamics in undisturbed, drained and restored blanket peat. J.Hydrol. 402, 103–114.

Joint Nature Conservation Committee, 2004. Common standards monitoring guidance for lowland wetland habitats. ISSN 1743-8160 (online).

Joint Nature Conservation Committee, 2009. Common standards monitoring guidance for upland habitats. ISSN 1743-8160 (online).

Joint Nature Conservation Committee 2011. Towards an assessment of the state of UK Peatlands, JNCC report No. 445.ISSN 0963 8901 (online).

Kalacska, M., Arroyo-Mora, J.P., Soffer, R.J., Roulet, N.T., Moore, T.R., Humphreys, E., Leblanc, G., Lucanus, O., Inamdar, D. 2018. Estimating peatland water table depth and net ecosystem exchange: a comparison between satellite and airborne imagery. Remote Sens. 10,687. Doi: 10.3390/rs10050687.

Knoth, C., Klein, B., Prinz, T., Kleinebecker, T. 2013. Unmanned aerial vehicles as innovative remote sensing platforms for high-resolution infrared imagery to support restoration monitoring in cut-over bogs. Appl. Veg. Sci. 16, 509–517.

Krankina, O.N., Pflugmacher, D., Friedl, M., Cohen, W.B., Nelson, P., Baccini, A. 2008. Meeting the challenge of mapping peatlands with remotely sensed data. Biogeosciences 5, 1809-1820.

Kuhn, M., Wing, J., Weston, S., Williams, A., Keefer, C., Engelhardt, A., Cooper, T., Mayer, Z., Kenkel, B., the R Core Team, Benesty,M., Lescarbeau, R., Ziem, A., Scrucca, L., Tang Y., Candan, C. 2016.

caret: Classification and Regression Training. R package version 6.0-71. https://CRAN.R-project.org/package=caret

Lees, K.J., Quaife, T., Artz, R.R.E., Khomik M., Clark, J.M. 2018. Potential for using remote sensing to estimate carbon fluxes across northern peatlands – A review. Sci. Tot. Environ. 615, 857–874.

Lilly, A., Bell, J.S., Hudson, G., Nolan, A.J. & Towers. W. (Compilers). 2010. National soil inventory of Scotland (NSIS_1); site location, sampling and profile description protocols. (1978-1988). Technical Bulletin. Macaulay Institute, Aberdeen.

Lindsay R. 2010. Peatbogs and Carbon: A Critical Synthesis. University of East London: London.

Luscombe, D.J., Anderson, K., Grand-Clement, E., Gatis, N., Ashe, J., Benaud, P., Smith, D., Brazier, R.E. 2016. How does drainage alter the hydrology of shallow degraded peatlands across multiple spatial scales? J. Hydrol. 541, 1329-1339.

McBratney, A., Santos, M., Minasny, B. 2003. On digital soil mapping. Geoderma 117,3-52.

McVean, D. N., Ratcliffe, D. A., 1962. Plant communities of the Scottish Highlands. Monograph no. 1 of the Nature Conservancy. HMSO, London.

Meingast, K.M., Falkowski, M.J., Kane, E.S., Potvin, L.R., Benscoter, B.W., Smith, A.M.S., Bourgeau-Chaves, L.L., Miller, M.E. 2014. Spectral detection of near-surface moisture content and water-table position in northern peatland ecosystems. Remote Sens Environ 152, 536-546.

Mehner, H., Cutler, M., Fairbairn, D., Thompson, G. 2004. Remote sensing of upland vegetation: the potential of high spatial resolution satellite sensors. Global Ecol. Biogeog. 13, 359–369.

Middleton, M., Närhi, P., Arkimaa, H., Hyvönen, E., Kuosmanen, V., Treitz, P., Sutinen, R. 2012. Ordination and hyperspectral remote sensing approach to classify peatland biotopes along soil moisture and fertility gradients. Remote Sens Environ 124, 596-609.

Minasny, B., McBratney, A.B. 2006. A conditioned Latin hypercube method for sampling in the presence of ancillary information. Computers & Geosci. 32, 1378-1388.

Perry, D., Hollis, D. 2005. The development of a new set of long-term climate averages for the UK. Int. J. Climatol. 25,1023-1039.

Pflugmacher, D., Krankina, O.N., Cohen, W.B. 2007. Satellite-based peatland mapping: Potential of the MODIS sensor. Global Planetary Change 56, 248–257.

Poggio, L.; Gimona, A. 2015. Sequence-based mapping approach to spatio-temporal snow patterns from MODIS time-series applied to Scotland., Int. J. App. Earth Obs. Geoinf. 34, 122-135.

Poggio, L., Gimona, A., Brown, I., 2012. Spatio-temporal MODIS EVI gap filling under cloud cover: an example in Scotland. ISPRS J.Photogrammetry Remote Sens. 72, 56-72.

Poggio, L.; Gimona, A., 2014. National scale 3D modelling of soil organic carbon stocks with uncertainty propagation - an example from Scotland. Geoderma 232-234, 284-299.

Robinson. M 1990. Impact of improved land drainage on river flows. Report No. 113 Institute of Hydrology, Crowmarsh, Gifford, Wallingford, Oxon OX10 8BB, UK

Ross, L.C., Woodin, S.J., Hester, A.J., Thompson, D.B.A., Birks, H.J.B. 2012. Biotic homogenization of upland vegetation: patterns and drivers at multiple spatial scales over five decades. J. Veg. Sci. 23, 755-770.

Running S., R. R.; Nemani, F. A.; Heinsch, M.; Zhao, M.; Reeves & Hashimoto., H. 2004. A continuous satellite-derived measure of global terrestrial primary production BioScience, 54, 47-560.

Sing, T., Sander, O., Beerenwinkel, N., Lengauer, T. 2005. ROCR: visualizing classifier performance in R. Bioinformatics 21(20), 3940-3941.

Sorensen, R., Zinko, U., Seibert, J. 2006. On the calculation of the topographic wetness index: evaluation of different methods based on field observations. Hydrol Earth System Sci 10, 101–112.

Wan, Z., 1999. MODIS Land-Surface Temperature Algorithm Theoretical Basis Document (LST ATBD) NASA, NASA.

Wilson, P.J., Provan, J. 2003. Effect of habitat fragmentation on levels and patterns of genetic diversity in natural populations of the peat moss Polytrichum commune. Proc Biol Sci. 270, 881-6.

Wood, S. 2006. Generalized Additive Models: An Introduction With R, Chapman and Hall/CRC Press.

Yu, Z., Loisel, J., Brosseau, D.P., Beilman, D.W., Hunt, S.J. 2010. Global peatland dynamics since the Last Glacial Maximum. Geophys. Res. Lett. 37, L13402, doi. 10.1029/2010GL043584, 2010.

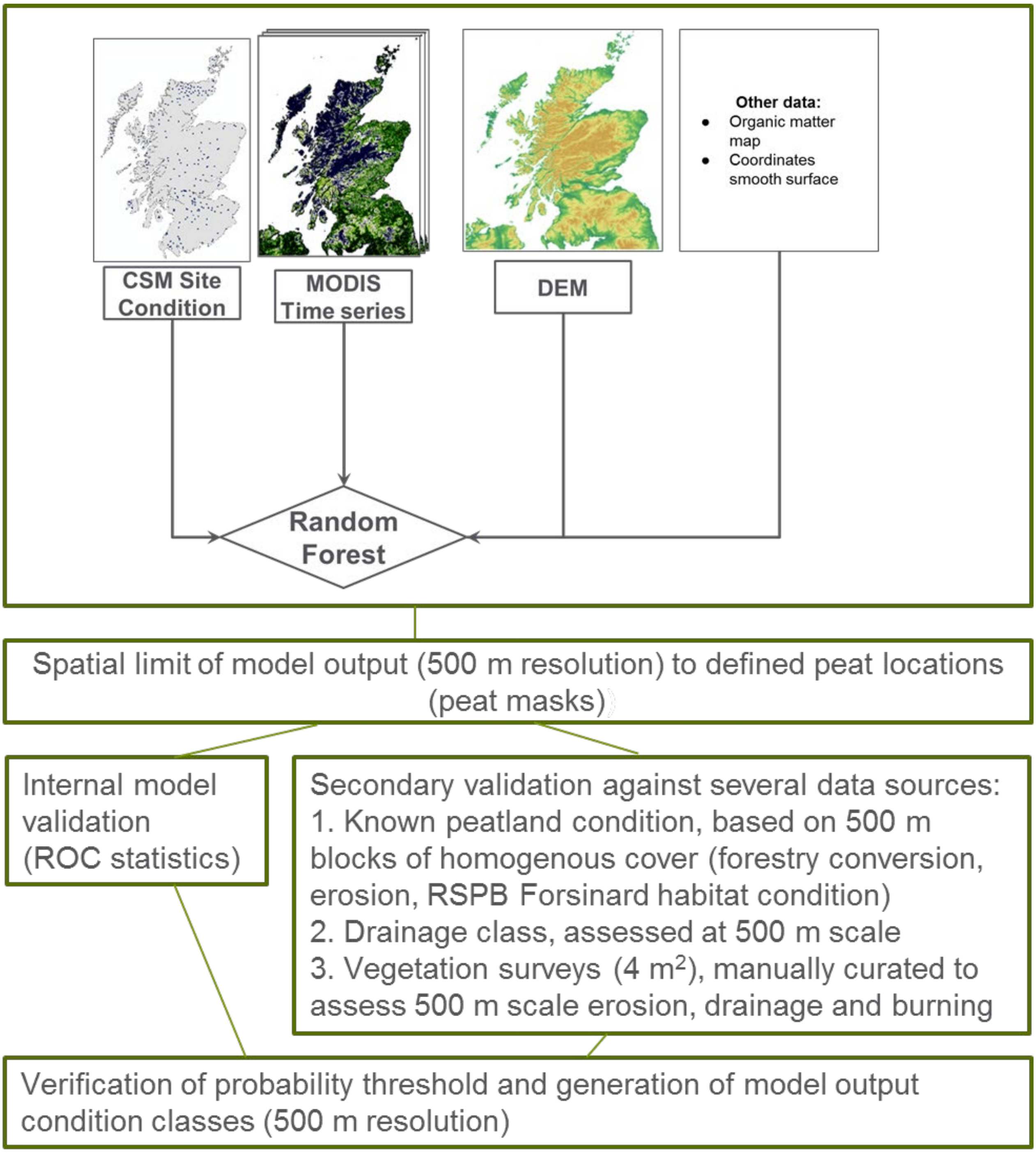


Figure 1. Workflow.

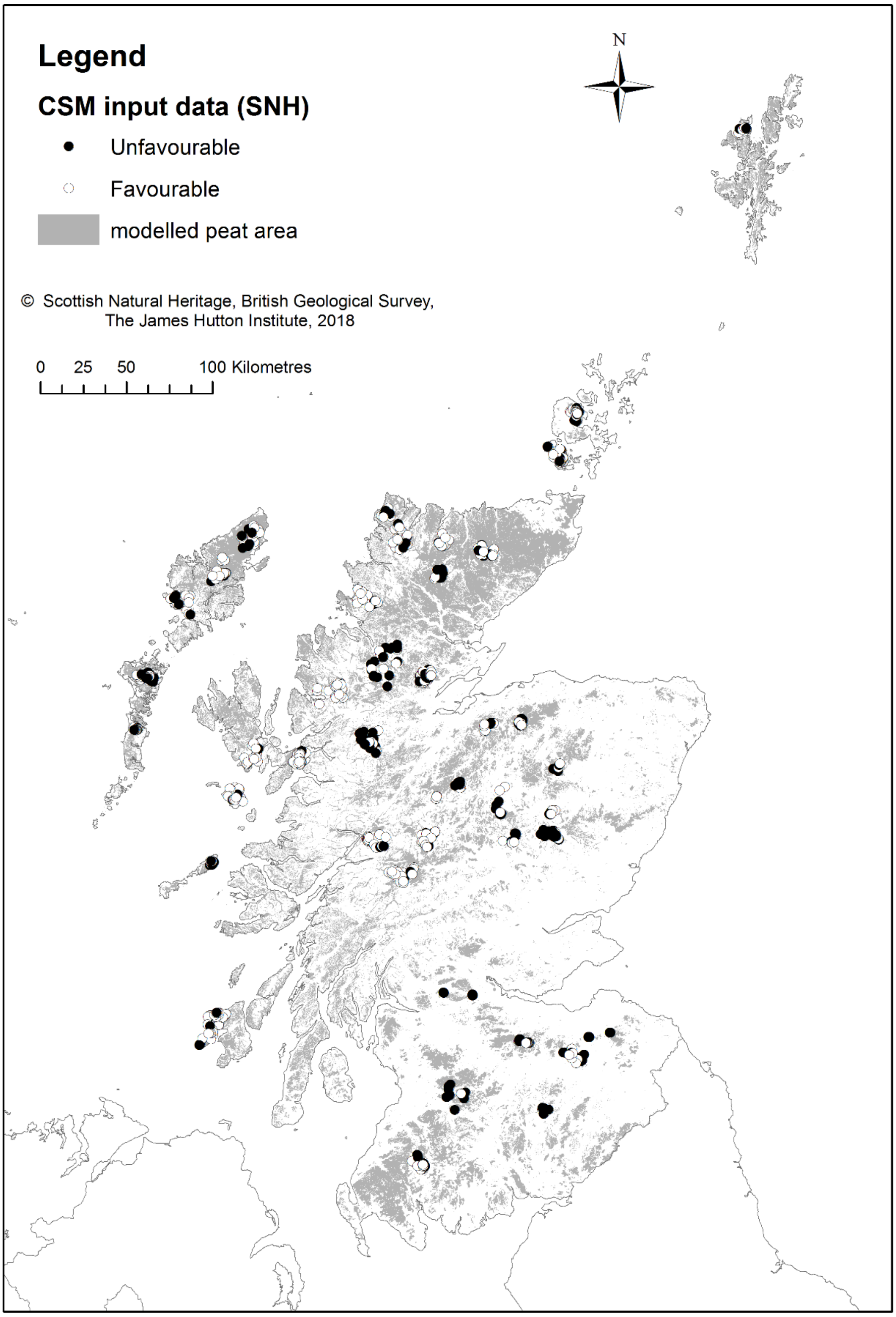
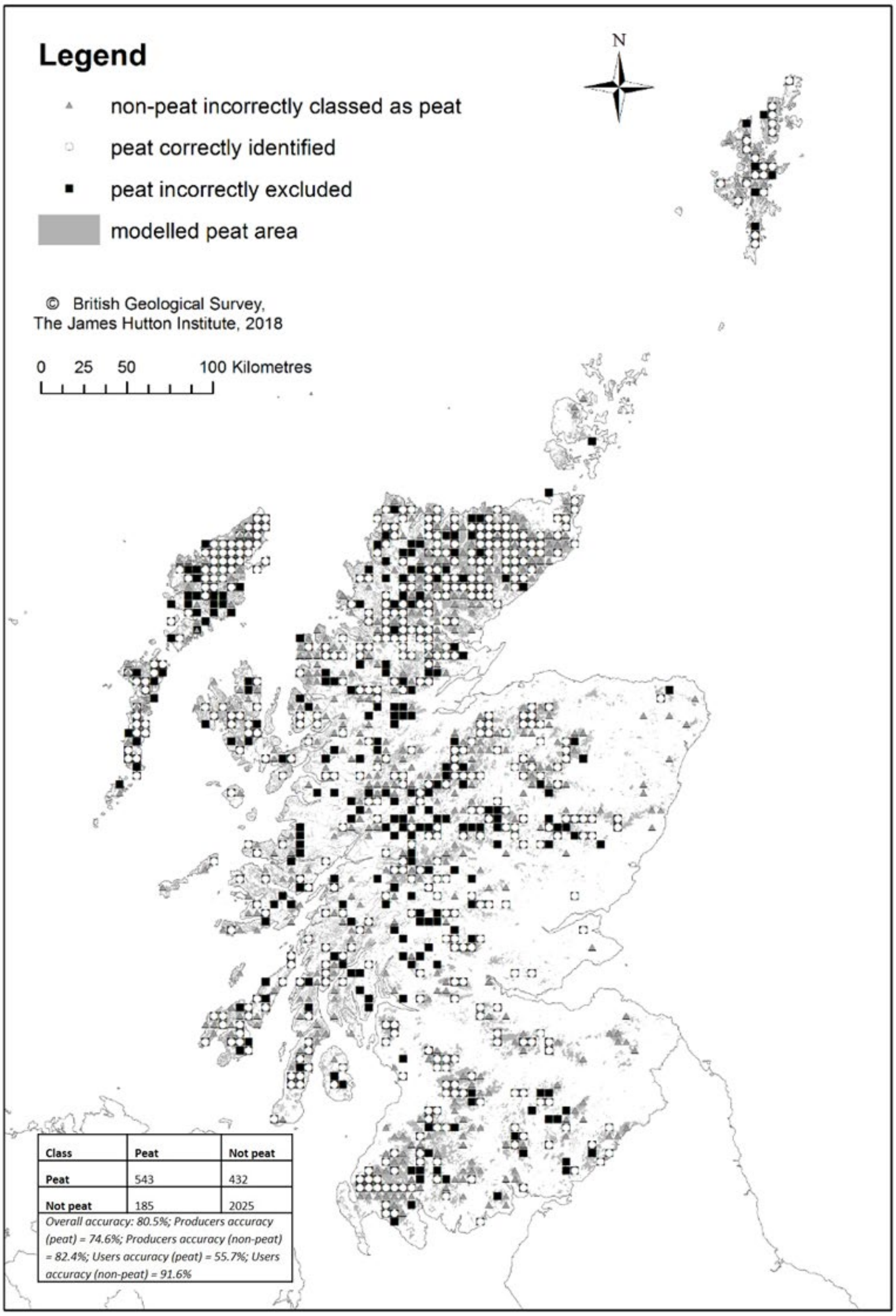


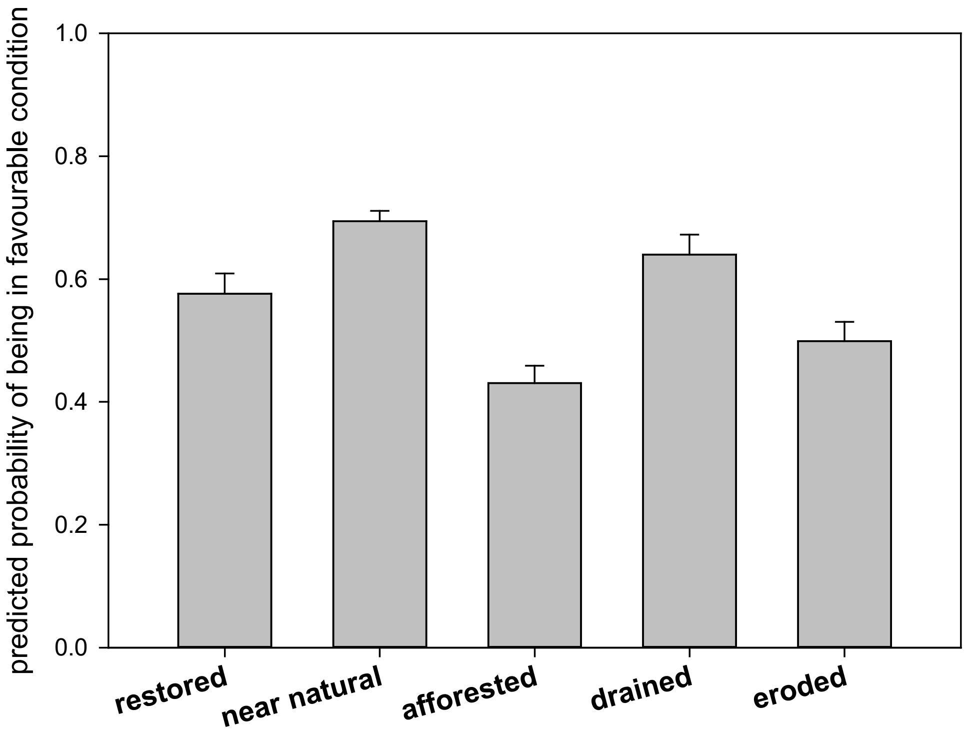
Figure 2. Common Standards Monitoring (CSM) training data for the model, with point locations that meet the criteria for site favourable condition in white (n= 602) and unfavourable condition in black circles (n= 349). The peat soil area (peat mask) as modelled by our approach is shown in grey. The training points in Shetland, as well as two points that didn't locate in Scotland, were ignored during model construction.



Class	Peat	Not peat
Peat	543	432
Not peat	185	2025

Overall accuracy: 80.5%; Producers accuracy (peat) = 74.6%; Producers accuracy (non-peat) = 82.4%; Users accuracy (peat) = 55.7%; Users accuracy (non-peat) = 91.6%

Figure 3. Validation of the peat mask (underlying grey areas) against the NSIS (1978-88) point location dataset. Point locations in white circles are correctly identified peat locations (n=543), locations in black squares (n= 185) are peat that the model incorrectly excludes and locations in grey triangles (n= 432) are non-peat soils the model incorrectly assumes to be peat locations. Table in inset shows the error matrix and model statistics.



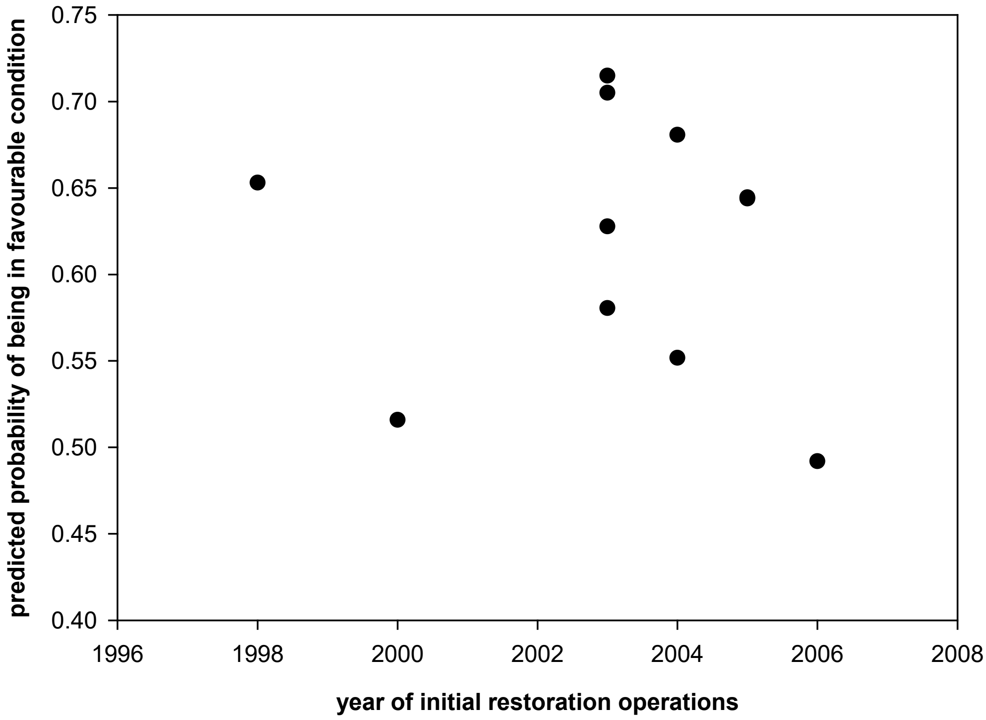


Figure 4. Additional external validation of the model using extracted probabilities at 70 locations where condition is known. Upper figure shows the predicted probability of being in favourable condition for (left to right) restored, near natural, afforested, drained and eroded sites (n=14, 23, 13, 9 and 16, respectively). Lower figure: Correlation of predicted probability of a site being in favourable condition against the year of restoration.

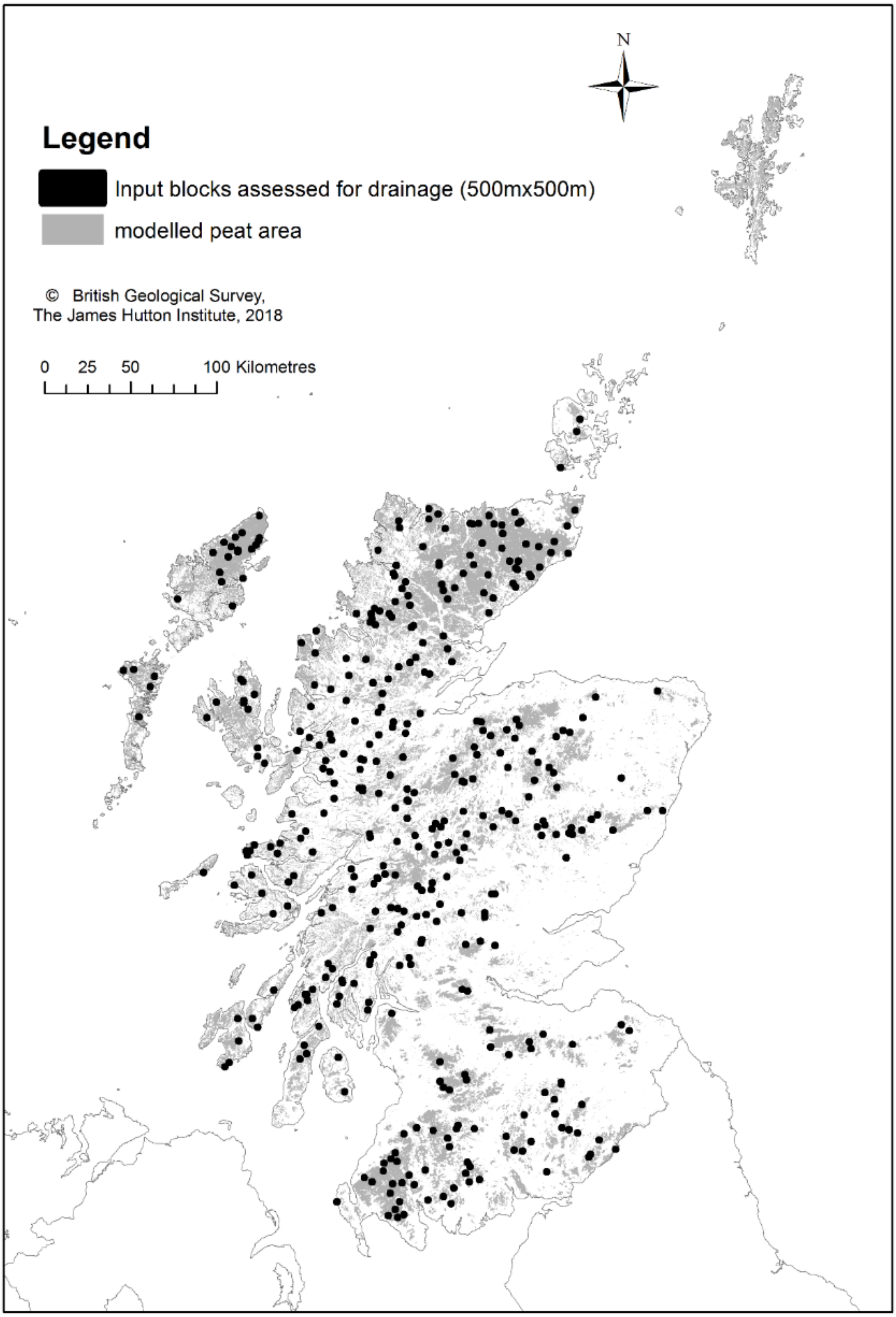


Figure 5. Distribution of the 500 m input blocks (n=400) assessed for drainage features (black circles). The peat soil area (peat mask) as modelled by our approach is shown in grey.

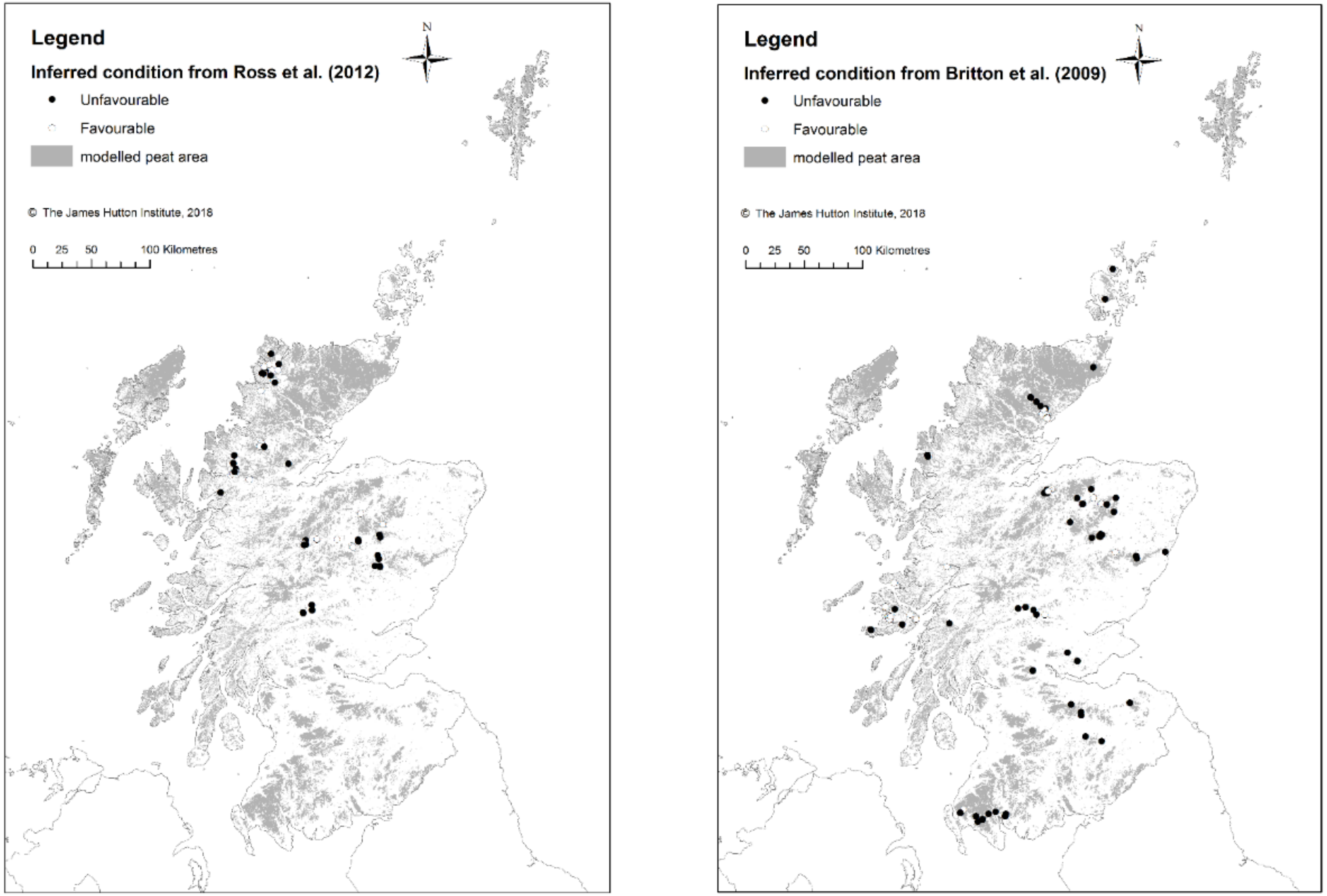


Figure 6. Inferred condition from McVean and Ratcliffe resurveys (Ross et al., 2012, left figure) and Birse and Robertson resurveys (Britton et al., 2009, right figure). Base map shown is modelled peat extent (grey). Locations (circles) in white (n= 25 in left figure and 49 in right figure) denote sites with inferred favourable condition, whereas sites in black (n= 38 in left figure and 83 in right figure) denote sites with inferred unfavourable condition.

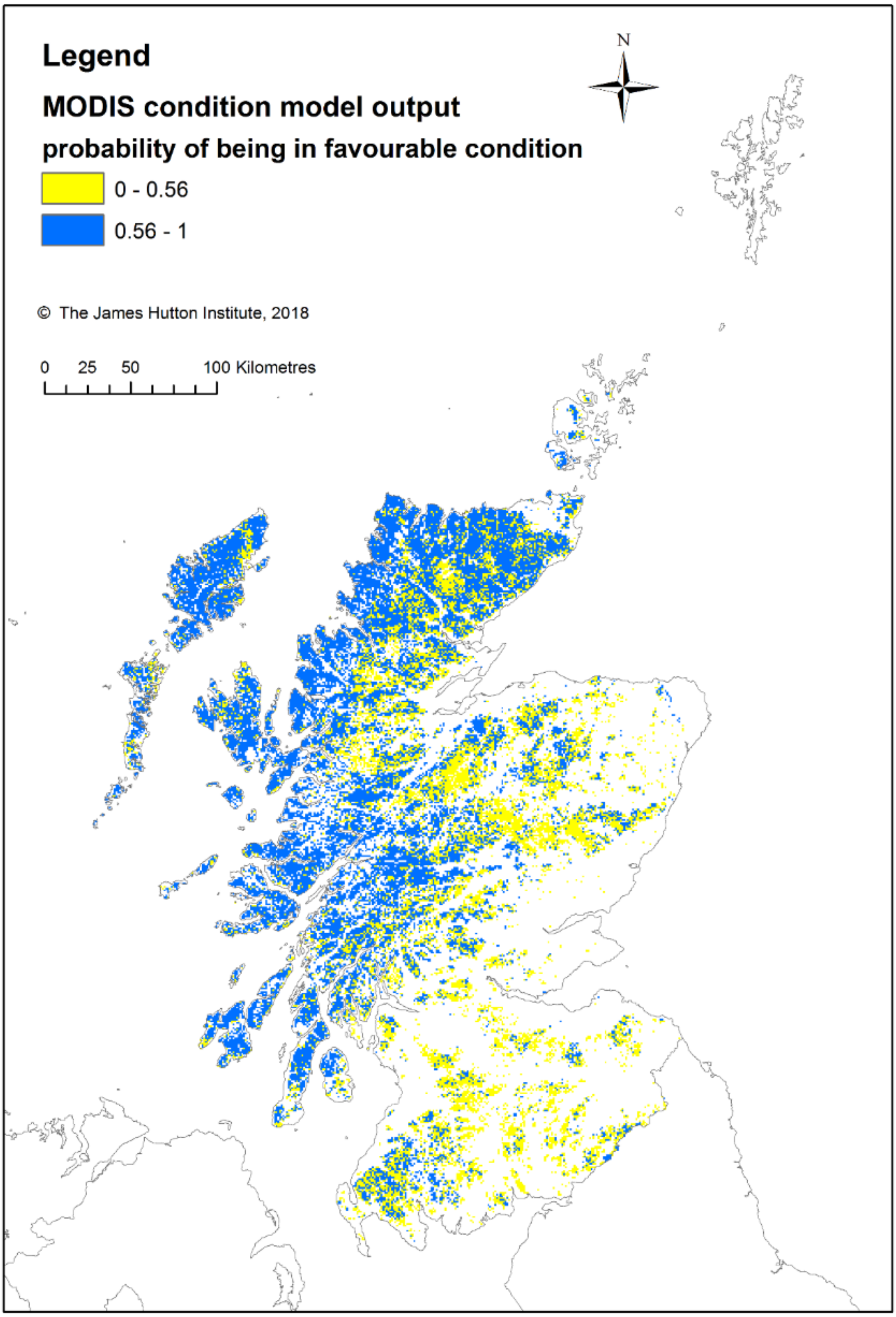


Figure 7. Predicted areas with favourable (blue) or unfavourable (yellow) peatland condition, based on a 56% probability threshold of the MODIS-based model limited to the peat mask developed in this study, as per model evaluation (Suppl. Fig. 2).

Table 1. Model statistics for the constrained model outputs classified to favourable/unfavourable categories, as per the threshold suggested by the ROC analysis.

Model based on peat soil extent as described in this work		
Reference/Prediction	0 (unfavourable)	1 (favourable)
0 (unfavourable)	227	22
1 (favourable)	39	428
Accuracy : 0.9148; 95% CI : (0.8919, 0.9342); No Information Rate : 0.6285; P-Value [Acc > NIR] : <2e-16; Kappa : 0.8151; Mcnemar's Test P-Value : 0.0405; Sensitivity : 0.8534; Specificity : 0.9511; Pos Pred Value : 0.9116; Neg Pred Value : 0.9165; Prevalence : 0.3715; Detection Rate : 0.3170; Detection Prevalence : 0.3478; Producers accuracy (0) = 0.853; Producers accuracy (1) = 0.951; Users accuracy (0) = 0.912; Users accuracy (1) = 0.916; Balanced Accuracy : 0.9022		
Model based on peat soil extent as per Aitkenhead (2016)		
Reference/Prediction	0 (unfavourable)	1 (favourable)
0 (unfavourable)	174	20
1 (favourable)	30	341
Accuracy : 0.9115; 95% CI : (0.885, 0.9336); No Information Rate : 0.6389; P-Value [Acc > NIR] : <2e-16; Kappa : 0.8061; Mcnemar's Test P-Value : 0.2031; Sensitivity : 0.8529; Specificity : 0.9446; Pos Pred Value : 0.8969; Neg Pred Value : 0.9191; Prevalence : 0.3611; Detection Rate : 0.3080; Detection Prevalence : 0.3434; Producers accuracy (0) = 0.853; Producers accuracy (1) = 0.944; Users accuracy (0) = 0.897; Users accuracy (1) = 0.895; Balanced Accuracy : 0.8988		

Table 2. Predicted probability of being in favourable condition for drainage sites

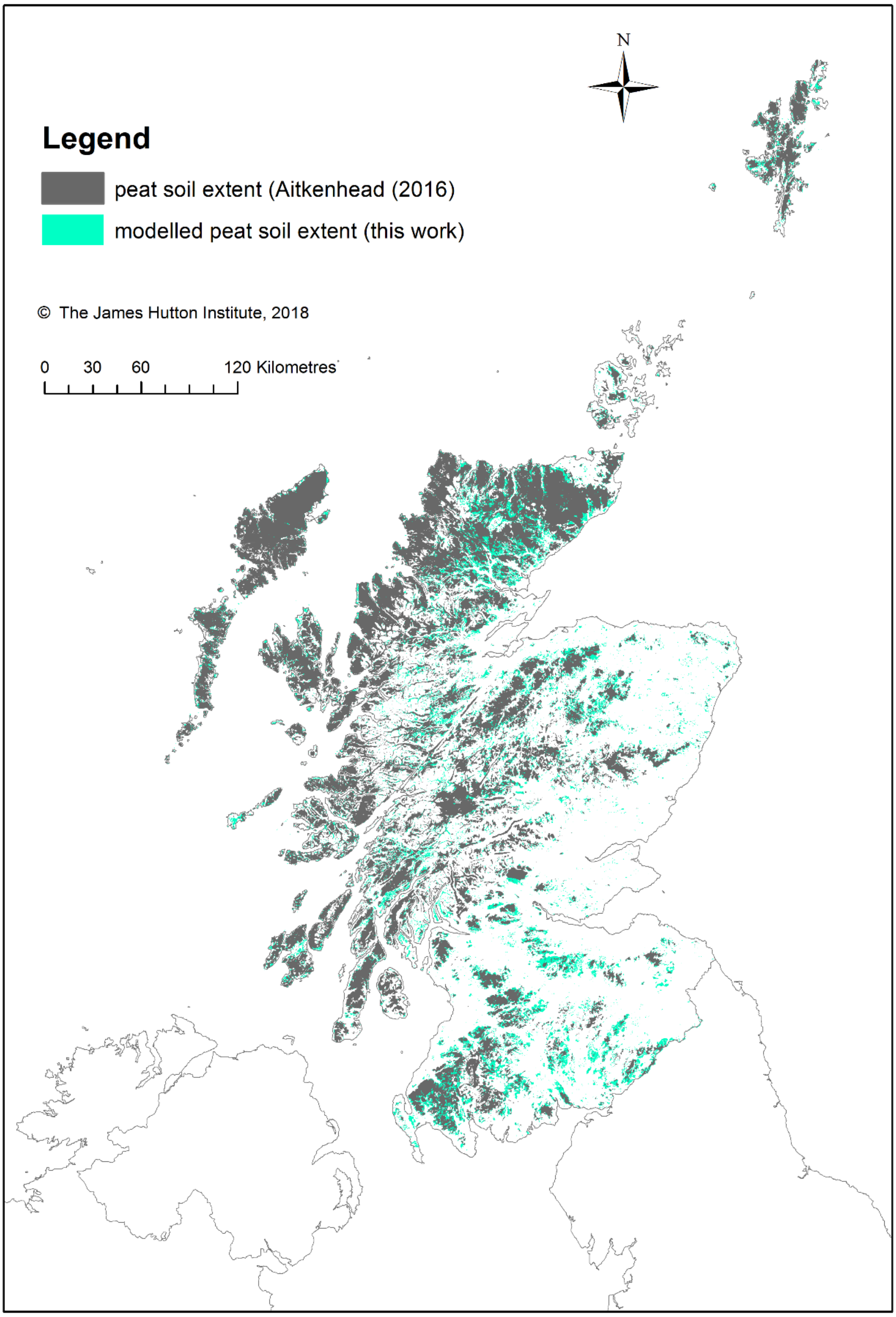
Drainage class	Class description	Number of observations	Predicted probability (average)	Standard error
1	no drains, no other features contributing to drainage	26	0.67	0.03
2	no drains but low numbers of other features contributing to drainage present	113	0.59	0.02
3	low number/density of drains, low number of other features contributing to drainage	29	0.61	0.02
4	low or medium number/density of drains, but a medium-large proportion of other features contributing to drainage	12	0.57	0.04
5	medium number/density of drains and medium other features contributing to drainage	29	0.53	0.03
6	high density of drainage channels (intervals of <20 m between drains), no or only sporadic other features contributing to drainage	12	0.50	0.05

Table 3. Predicted probability of being in favourable condition for sites with inferred condition status from previously published vegetation surveys.

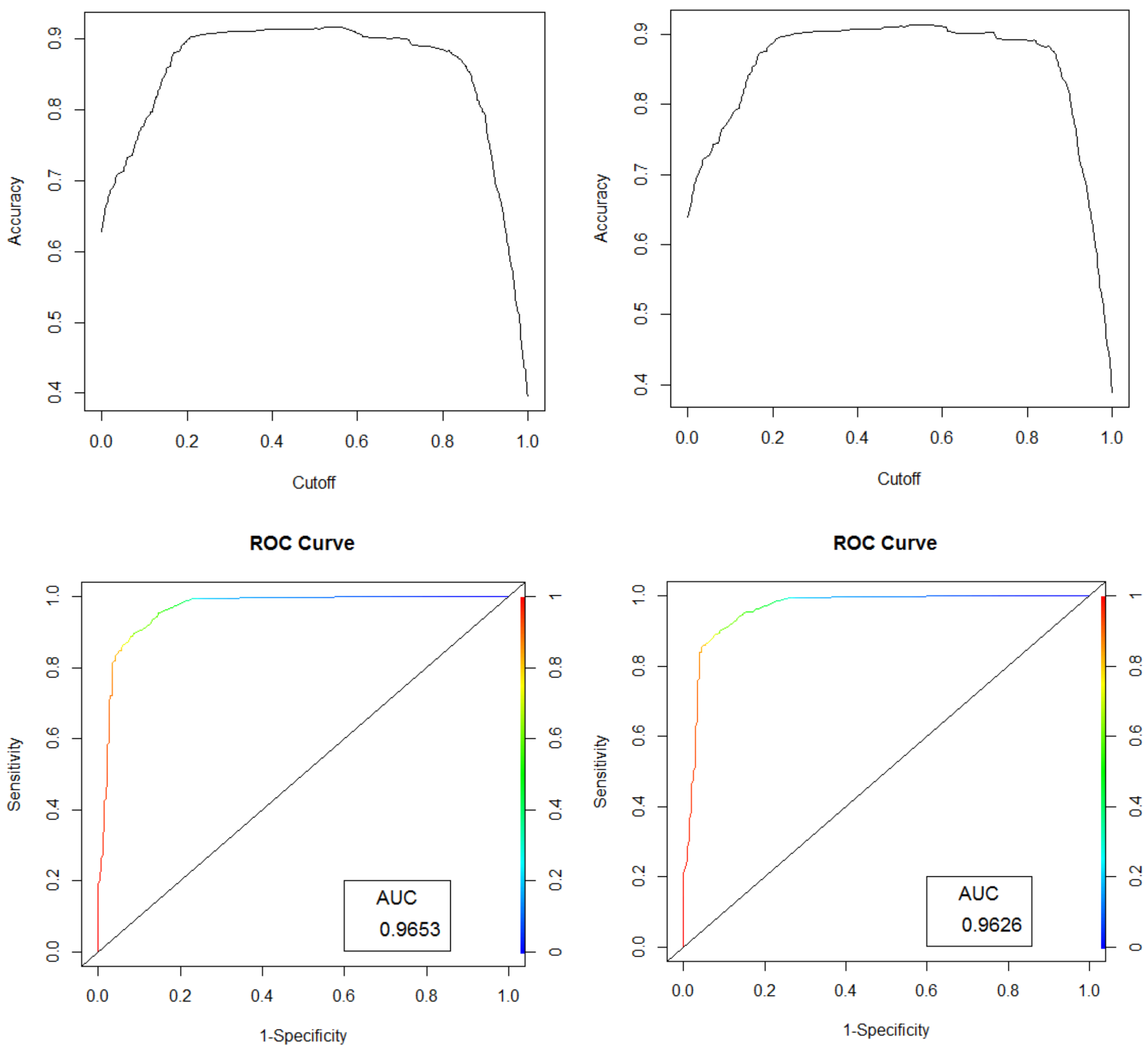
Data source	Inferred condition status (n)	Predicted average probability to be in favourable condition (average +/- SEM) $
Ross et al. (2012)	Favourable (25)	0.71 (0.03) a
Ross et al. (2012)	Unfavourable (38)	0.57 (0.04) b
Britton et al. (2017)	Favourable (49)	0.63 (0.02) c
Britton et al. (2017)	Unfavourable (83)	0.48 (0.03) d

$ significant differences between group tested with 2-way ANOVA within each data set (i.e. datasets derived from Ross and Britton et al. tested separately), different letters denote significantly different groups.

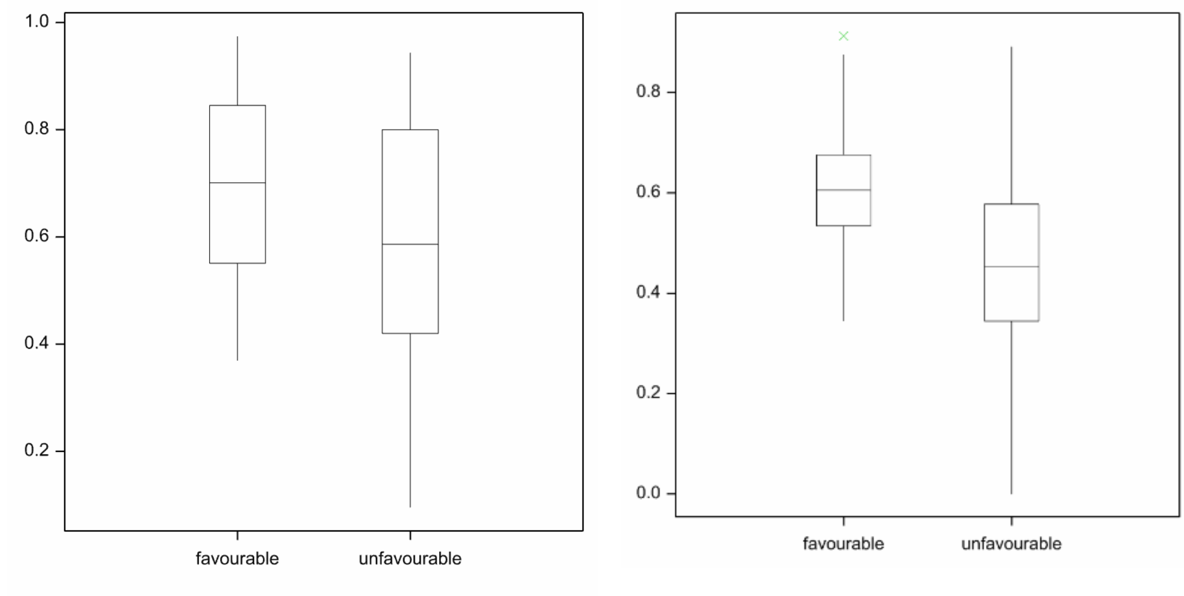
Supplementary Figure 1. Comparison of the peat extent model outputs.



Supplementary Figure 2. Model evaluation plots (top) and ROC curves (bottom). The model evaluation plot shows the accuracy of the predictions if a threshold is set (cutoff, x axis) to define the binary classes. The ROC curve graphically plots the false positive fraction (1-specificity) against the true positive fraction (sensitivity) for the threshold chosen by the model evaluation. Left graphs show the outputs for the condition model constrained to the peat extent model described in this work, while the graphs on the right show the outputs for the model constrained to the peat extent modelled by Aitkenhead (2016).



Supplementary Figure 3. Box plots of the probability ranges for the inferred condition on the McVean and Ratcliffe (left) and Birse and Robertson (right) resurveyed plots. Y-axis denotes the probability of being in favourable condition that was returned for each category (see Table 3).



Supplementary Figure 4. Predicted areas with favourable (blue) or unfavourable (yellow) peatland condition, based on a 56.2% probability threshold of the MODIS-based model limited to the peat mask of Aitkenhead (2016), as per model evaluation (Suppl. Fig. 2).