

Article

## Empirical Modelling of Vegetation Abundance from Airborne Hyperspectral Data for Upland Peatland Restoration Monitoring

Beth Cole <sup>1,\*</sup>, Julia McMorrow <sup>2</sup> and Martin Evans <sup>2</sup>

<sup>1</sup> Centre for Landscape and Climate Research, University of Leicester, Bennett Building, University Road, Leicester LE1 7RH, UK

<sup>2</sup> School of Environment Education and Development, University of Manchester, Oxford Road, Manchester M13 9PL, UK; E-Mails: Julia.mcmorrow@manchester.ac.uk (J.M.); Martin.Evans@manchester.ac.uk (M.E.)

\* Author to whom correspondence should be addressed; E-Mail: bc132@le.ac.uk; Tel.: +44-116-252-3796.

Received: 22 November 2013; in revised form: 20 December 2013 / Accepted: 31 December 2013 / Published: 9 January 2014

---

**Abstract:** Peatlands are important terrestrial carbon stores. Restoration of degraded peatlands to restore ecosystem services is a major area of conservation effort. Monitoring is crucial to judge the success of this restoration. Remote sensing is a potential tool to provide landscape-scale information on the habitat condition. Using an empirical modelling approach, this paper aims to use airborne hyperspectral image data with ground vegetation survey data to model vegetation abundance for a degraded upland blanket bog in the United Kingdom (UK), which is undergoing restoration. A predictive model for vegetation abundance of Plant Functional Types (PFT) was produced using a Partial Least Squares Regression (PLSR) and applied to the whole restoration site. A sensitivity test on the relationships between spectral data and vegetation abundance at PFT and single species level confirmed that PFT was the correct scale for analysis. The PLSR modelling allows selection of variables based upon the weighted regression coefficient of the individual spectral bands, showing which bands have the most influence on the model. These results suggest that the SWIR has less value for monitoring peatland vegetation from hyperspectral images than initially predicted. RMSE values for the validation data range between 10% and 16% cover, indicating that the models can be used as an operational tool, considering the subjective nature of existing vegetation survey results. These predicted coverage images are the first quantitative landscape scale monitoring results to be produced

for the site. High resolution hyperspectral mapping of PFTs has the potential to assess recovery of peatland systems at landscape scale for the first time.

**Keywords:** hyperspectral; peatland; restoration; partial least squares regression (PLSR); plant functional type (PFT); vegetation abundance

---

## 1. Introduction

Upland landscapes are an important environmental, conservation, social and economic resource on national and global levels [1]. They are often places of scenic natural beauty and biodiversity hotspots. The ecology of peatlands are unique and the biodiversity value is high due to the national and international importance of rare or threatened plant and animal species which they support [2,3]. As well as providing a habitat for important bird and insect species [4,5] and ecologically-important plants [6], uplands also have a rich cultural heritage, having been shaped by human activity over centuries [7]. The importance of peatland ecology is recognised under legislation. Blanket bogs are a UK Biodiversity Action Plan (BAP) priority habitat (UKBAP, 2008), and a priority for conservation under the EC habitats directive. The importance of peatlands has become widely recognised due to the diversity of ecosystem services they provide [8–10]. Examples include agricultural services, air and water regulation and carbon storage and sequestration. Peatlands are the single largest carbon reserve in the UK, storing around 3 billion tonnes of carbon [11].

All of these features and ecosystem services require peatlands to be in a favourable condition with vegetation cover. The potential to sequester carbon is greatly reduced and carbon stores are rapidly lost when peat becomes exposed [12]. A range of pressures including overgrazing, burning, pollution and climate change have led to large areas of degradation and exposed peat. Restoration projects aim to reduce peat erosion by restoring vegetation coverage and to return peatlands to functioning ecosystems. Monitoring is crucial, given the investment in restoration. Remote sensing of peatlands offers the potential to monitor restoration success by charting the succession of vegetation communities as an alternative to the time-consuming and expensive traditional field survey methods.

Vegetation is the most obvious physical representation of an ecosystem, so monitoring vegetation can be one of the most important tools for understanding ecosystem properties, functions and condition. Functional composition, therefore, can be an important indicator of underlying ecosystem properties and ecosystem health [13]. Functional classifications, although generally derived to describe static conditions of a landscape, have also been used to predict response to disturbance [14]. Plant species reflect habitat changes in a detailed way by their presence, disappearance or absence [15], so monitoring species composition can be used to assess change in the peatland environment.

The discipline of landscape ecology has benefited from the partnership with remote sensing and the provision of high value spatially-explicit data [16–18]. The opportunities this presents for environmental monitoring have recently been recognised by national conservation bodies [19,20]. However, presently, peatland applications are concerned with the immediate priority of basic habitat inventories. Ecologically, peatlands are characterised by a wide diversity and complexity in their composition and interaction between different vegetation types [21]. Peatland restoration monitoring,

therefore, requires monitoring of vegetation composition on a fine spatial scale, to match the spatial scales of the processes under observation [22]. A pixel size of less than 3.45 m is recommended, in an intact upland peatland, to determine patch size in the vegetation [23]. Advances in the spatial and spectral resolutions of remote sensing data, and of modelling techniques using this data, have made the remote sensing of such ecological variables more feasible [18].

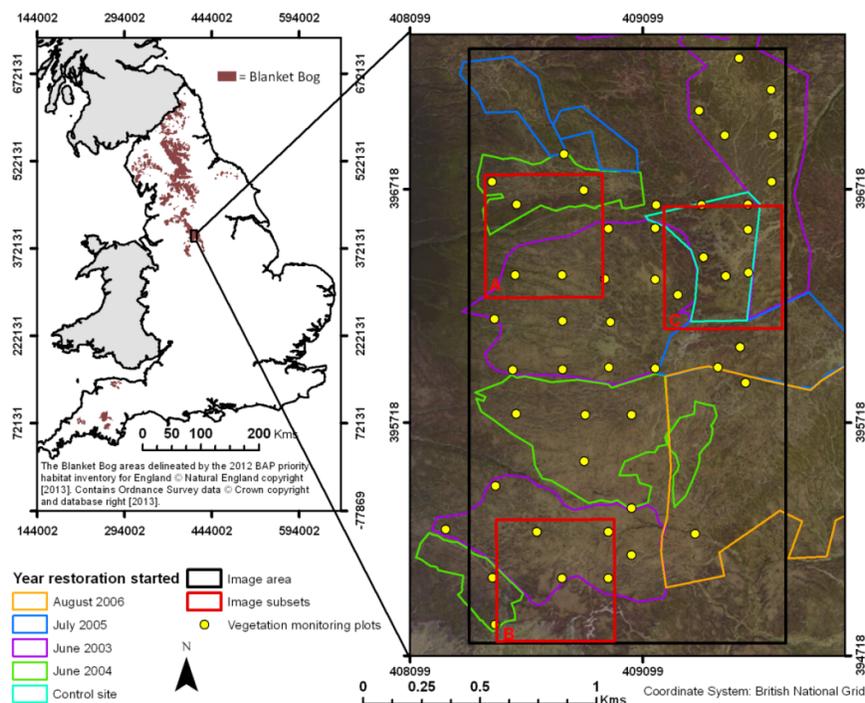
In this paper, we aim to establish the empirical relationship between hyperspectral remote sensing data and vegetation abundance, creating a model of the predictive coverage of PFTs and cover types across a landscape scale restoration site on upland blanket bog.

## 2. Method

### 2.1. Study Area

The study area is focused on the restoration work by Moors for the Future (MFF) on Bleaklow Plateau, an upland blanket peatland in the Peak District National Park (PDNP) in the southern Pennines in Northern England. The peatlands of the PDNP, are a fragile environment, subject to severe degradation, situated at the south-eastern extent of the upland peatlands in the UK, and also at the southern climatic border of blanket bog distribution [24] (Figure 1). The restoration work involves application and maintenance of a nurse crop of grasses, the application of lime to raise the pH, and fertiliser to ensure their survival. Geojute, a geotextile mesh, is used to stabilise the substrate, as well as cut heather brush which creates a microclimate and prevents loss of seeds and further surface erosion. Finally, structurally important species such as bilberry, cowberry and cotton grass are plug planted directly into the peat.

**Figure 1.** Location map showing area covered by the image, the different restoration blocks, and image subsets referred to in the paper.



The site includes peatland in various states of disturbance, from intact to degraded blanket bog and areas undergoing restoration. The age of the restoration areas covered by the airborne images spans four years, providing a temporal trajectory across a spatial area (Figure 1). However, the heterogeneous vegetation communities vary on a much smaller scale due to the complex interaction of influencing factors, such as patches of intact vegetation, local topography, and erosional processes. They are representative of the mire (bog) communities in the UK National Vegetation Classifications (NVC), with the addition of fine nurse grasses. Present are M19-*Calluna vulgaris*—*Eriophorum vaginatum* blanket mire, the slightly more impoverished M20-*Eriophorum vaginatum* blanket and raised mire—often associated with disturbances in the vegetation community—and the M3-*Eriophorum angustifolium* bog pool community, which is a typical of recent or disturbed bog pools, or often forms an early seral stage in the transition from exposed peat back to mire vegetation [25].

## 2.2. Data Acquisition

The airborne hyperspectral data was acquired using the SPECIM (Oulu, Finland) AISA Eagle and Hawk sensors. With the spatial binning settings and acquisition conditions, this resulted in an image with 482 bands with a spectral range of 400–2450 nm, after the data from both sensors was stacked. The spatial resolution was dictated by the integration time of the sensors, due to the topography of the terrain and flight restrictions. Following geocorrection, this resulted in a spatial resolution of square 0.7 m pixels, for both sensors. The rigorous pre-processing procedure included atmospheric correction using the Empirical Line Correction (ELC) method on a primary flight line, followed by radiometric normalisation using the orthogonal flight line approach (ORTHORADMOS) [26] to correct between the 12 overlapping flight lines. The ELC was carried out using five Pseudo Invariant Features, three  $6 \times 6$  m vicarious fabric targets in black, grey and white, and two natural vegetation targets, chosen as spectrally bright and dark natural targets. Radiometric correction resulted in an average RMSE of all the regressions for all the flight lines of 0.87% for the Eagle sensor and 1.00% for the Hawk sensor, tested using a validation set of pixels in the overlapping areas between flight lines. Geometric correction using a parametric approach with the aircraft navigational information along with a LiDAR DEM was verified with post processed DGPS to be accurate to 0.77 m, within 1.1 pixels.

Vegetation plots were surveyed in the 10 day window immediately after the flight, while vegetation composition was as similar as possible to conditions during the image acquisition. The 58 plots, each of size 2 by 2 m, were selected by systematic sampling of existing MFF monitoring plots, which had been previously laid out on a 100 m grid, to a 200 m spacing. This stratified approach allows characterisation of the different areas of the site and ages of restoration (Figure 1). The spectral data was extracted and averaged, from the four inner-most central pixels covering each vegetation plot. This was done to ensure the variation in vegetation composition across the plot was captured, which allowed direct comparison to the vegetation survey, and also so that any additional pixels overlapping the edge were not included in the analysis. Visual inspection of the selected pixels confirmed that edge effects were minimised. The image spectra from the plots were split into calibration and validation datasets in a 60:40 ratio, to allow validation of the regression models. They were selected by stratified random sampling, where the strata were vegetation community groups along a floristic trajectory of

restoration, as identified in a previous analysis. This ensured that calibration and validation sets contained plots from the different stages of restoration and therefore a mix of vegetation communities.

### 2.3. Sensitivity Analysis to Class Generalisation

The predictive performance of empirical regression models can depend upon the degree of class generalisation [27]. There are three main PFTs in peatlands; bryophytes, graminoids and shrubs. These represent groups of different structure and biochemistry, such as leaf structure, water and pigment content and architecture, resulting in different spectral characteristics at the leaf and canopy scale [28]. Senescent vegetation, has been shown to have a distinct effect on the spectral signal [29], so was considered here as a separate cover type, as was bare peat. Senescent vegetation was identified as above ground biomass which was not photosynthetically active during the vegetation survey. As species level data does not display such distinctive differences in biophysical characteristics and spectral response as PFTs [30], the percentage cover of species data was combined into generalised floristic variables representing these three PFTs (Table 1).

**Table 1.** List of key species, in Plant Functional Types (PFTs).

Plant Functional Type	Key Species
Graminoids	Nurse (fine) grasses
	<i>Eriophorum angustifolium</i>
	<i>Eriophorum vaginatum</i>
Shrubs	<i>Calluna vulgaris</i>
	<i>Vaccinium myrtillus</i>
Bryophytes	<i>Hypnum</i>
	<i>Campylopus</i>

Whilst it is convenient to aggregate species into PFTs, the results of the vegetation survey in this study showed that the peatland vegetation communities are not made up of spatially contiguous clusters of functional types. Vegetation, in fact, grows in more mixed patterns, with individual species favouring slightly different conditions. The spatial pattern of the key species has also been influenced by the active restoration management. Therefore, sensitivity to class aggregation of the vegetation data into PFT was tested by examining relationships of spectral indices to single key species, as well as the PFT. This was done using scatterplots, histograms of the distribution and correlations between the species abundance data and the spectral indices. For both PFT and the key species, the lack of linear relationships was explained by the non-normal distribution of the vegetation variables. They are all skewed, with many zero values. This has been reported by other studies modelling vegetation coverage [31], but may be exaggerated by the degradation and subsequent restoration leading to an unusual distribution of species, including areas of vegetation with species missing that might otherwise be expected in an undisturbed ecosystem.

Species within a PTF have similar biophysical characteristics leading to similar spectral responses [28]. The cover values of single species are more variable than averaged values of entire stands [31], therefore, by aggregating species into PFTs, the influence of any anomalies in the individual species response are reduced, giving a stronger relationship. Another reason for generalising to the PFT level

is the heterogeneous nature of the peatland environment. Despite acquisition at a high spatial resolution, the Eagle and Hawk sensor still detected a mixture of species within a single 0.7 m pixel, making relationships with single species hard to define. Vegetation mapping of single species using PLSR had previously been found to be less successful than for functional groups measured with Ellenburg indicator values, and was ascribed to the zero values and low number of occurrences, causing weak models [31]. PFTs in a peatland have been successfully mapped using PLSR, but the relationship to composition of single species has proved unsuccessful [32].

#### 2.4. Regression Modelling

Hyperspectral measurements using narrow bands contain useful information for quantifying biophysical characteristics of vegetation canopies, not available from multispectral scanners [33,34]. However, the data sets contain a large amount of redundant information [35]. The use of statistical models for the retrieval of vegetation characteristics requires that the strong collinearity in the spectral data is accounted for. Two approaches can be taken to this: either full spectrum methods using all the spectral bands simultaneously by employing data compression in the regression algorithm; or alternatively, the number of independent variables can be reduced before modelling. This is commonly done either by using spectral indicators, such as vegetation indices which exhibit a strong relationship to the main vegetation biophysical variables, or by selecting a few diagnostic spectral bands as regression factors that permit modelling of the target variable [36].

Narrow band spectral indices, namely Plant Senescing Reflection Index (PSRI), Photochemical Reflectance Index (PRI), Cellulose Absorption Index (CAI), Vogelmann Red Edge Index (VOG) and Red Edge Position (REP), were used as spectral indicators in order to reduce the number of independent variables in the calibration equation. A stepwise multiple linear regression approach was rejected; data exploration of the relationships between the vegetation abundance data and spectral indices revealed that statistical assumptions required for multiple regression modelling were not met, in particular, weak relationship between the variables, the non-normal distributions, and collinearity both between independent (spectral) variables and between floristic variables. The Spearman rank correlation coefficient,  $r$ , between spectral indices and cover type, demonstrate the lack of relationship (Table 2).

**Table 2.** The Spearman rank correlation coefficient,  $r$ , between spectral indices and cover type.

	REP	VOG	CAI	PRI	PSRI
Graminoids	−0.216	−0.493 *	0.113	−0.112	−0.315
Shrubs	−0.240	−0.501 *	0.044	−0.459 *	−0.816 *
Bryophytes	−0.346	0.418 *	0.050	−0.042	0.089
Senescent vegetation	−0.293	−0.220	0.183	−0.064	−0.056
Bare peat	0.338	0.702 *	−0.170	0.397 *	0.614 *

\* Significant at 0.01 level.

Instead, a PLSR, using SPSS software, was employed to gain maximum information from the spectral information from the hyperspectral data.

Published results using Partial Least Squares Regression (PLSR) for vegetation studies in comparison to other methods of statistical modelling show it to be a powerful technique [36,37]. Direct comparison to results obtained from indices also prove PLSR to be more accurate [36,38]. PLSR applies a data compression within the regression and, as such, is a promising technique for quantitatively assessing vegetation characteristics, efficiently using the full spectral information of hyperspectral data [34,39]. It has become an established technique in vegetation remote sensing, used extensively for biochemical and biophysical modelling [36,37,40]. Other applications include biomass estimation [38], and modelling soil properties [41,42]. Application of PLSR in modelling vegetation presence and composition has previously been applied by relating ordination axes to reflectance for mapping floristic gradients [30–32,43].

PLSR is a multivariate technique designed to analyse data with strongly collinear, noisy and numerous independent, or predictor, variables [44,45]. As such, it copes with the high dimensionality and collinearity of hyperspectral data. PLSR is well suited for calibration on a small number of samples; the method can be used if the number of wavelengths is greater than the number of observations [36]. Similar to Principal Component Regression (PCR), the information content of the intercorrelated bands is reduced to a few independent latent factors, or vectors. Most important, these latent vectors are not only generated to maximise information content in the spectral information as a PCA, but are also optimised for the response variable [39,43]. This may be expected to lead to better calibration in comparison to other linear empirical-statistical methods. All the vegetation variables are entered together, in this case per cent coverage of each species; the model takes into account relationships between the dependent (floristic) variables in the formation of the latent factors. This makes PLSR a suitable method for quantifying the abundance of each of the PFTs because of the strong relationships between the proportions of each type of vegetation.

An alternative data compression modelling technique, Principal Component Regression (PCR), has been criticised when applied to hyperspectral data. The criticisms are that the reflectance data is compressed based purely on the statistical properties, so the first few principal components (PCs), used as regression factors, do not retain the factors of variability required in the predictive equation. The benefits of the full spectral data can be lost, as the higher order PCs do not preserve the spectral information in the original feature space, oversimplifying the variance [36,46,47]. PLSR avoids this by taking into account variation in the response variable, which is why it was selected for analysis in this study. Other machine learning methods may be appropriate for different datasets.

In an ideal PLSR model, a few new latent factors will explain most of the variation in both the predictors and responses. Factors that explain response variation (variation in the vegetation data in this case) provide good predictive models for new responses. Factors that explain predictor variation (variation in the spectral reflectance) are well represented by the observed values of the predictor. The use of too many latent factors can cause over fitting [48], so the optimum number of latent factors for the model was found by visual inspection of three types of plot: cumulative proportion of variance accounted for; score-plots of the latent factors; and residual variance plots. This was assessed when applied to the validation dataset. A good PLSR model will have good correlations between the X-scores and the Y-scores for the first few latent factors, with the relationship decreasing in strength with each successive factor [49].

Variable selection can enhance PLSR results [31,48,50,51]. To remove unnecessary spectral information, a PLSR model with the less important predictors removed was tested. Prior to any analysis on the spectral data, the bands from 1,356–1,456 nm and 1,792–1,962 nm around the main water absorptions in the SWIR were removed, but subsequent selection was based upon the weighted regression coefficients along with a plot of the variables' importance. Interactive variable selection reveals much about the relationship between the spectral data and the vegetation data, and the importance of different parts of the spectrum in predicting vegetation abundance. The regression coefficients are a function of the range of the original reflectance values, and so are highlighted areas with a large range in reflectance values, for example, the blue spectrum and the large peak at 1,450 nm which will be more volatile due to atmospheric scattering [34,52]. For this reason, the bands that were considered for removal from the PLSR were those that have relatively small coefficients and a small VIP value, below 0.8 [53]. Bands were only removed if they did not fall at wavelengths with known diagnostic absorbance features. For the PLSR with five latent factors, the bands removed were in the Hawk sensor between 1475 nm and 1791 nm, and for the analysis with 10 latent factors, Hawk bands from 1,166 nm–1,298 nm and 1,513 nm–1,784 nm.

The method for selecting the variables to remove may be an important factor for increased accuracy. The interactive method employed here, of interpreting the VIP and weighted regression coefficients together, has the benefit of ensuring that all the important parts of the spectrum are included; for example, the red edge known to represent chlorophyll concentration [54], and the absorption features in the SWIR related to plant cellulose [55]. Purely statistical methods of selection such as jack-knifing [51] do not always improve results, as keeping the selected bands well-spaced is important for maximising information content [31]. Only variables that form continuous and evidently unimportant regions should be eliminated, and this elimination should not be driven by the degree of fit and the prediction errors [50].

The PLSR models were validated against the test dataset of 22 independent samples, allowing an objective assessment of which model had the optimum parameters, number of latent factors and selection of variables, and therefore which model performed best. The best PLSR models were applied to the airborne images on a per pixel basis, resulting in a set of predictive images, one for each of the three PFTs or two cover types.

### **3. Results**

#### *3.1. Class Sensitivity*

Testing how appropriate aggregating the data into cover types was revealed that the single species distributions were more strongly skewed. This is because the aggregation reduces the occurrences of zero and low percentages of cover in the vegetation data, having a strong effect on the distribution. The correlation coefficient ( $r$ ) values from the Spearman rank correlation were generally lower for the key species than for the aggregated functional groups; there were fewer significant relationships between the key species variables and the spectral indices than for PFTs and cover types. Relationships between the spectral data are marginally stronger for PFTs than for species level.

### 3.2. Latent Factors of the PLSR

The optimum number of latent factors for the model to avoid over fitting, was found using the plots for cumulative proportion of variance (Figure 2); scores of the latent factors (Figure 3); and residual variance (Figure 4).

**Figure 2.** Cumulative proportion of variance accounted for by increasing latent factors in Partial Least Squares Regression (PLSR) model.

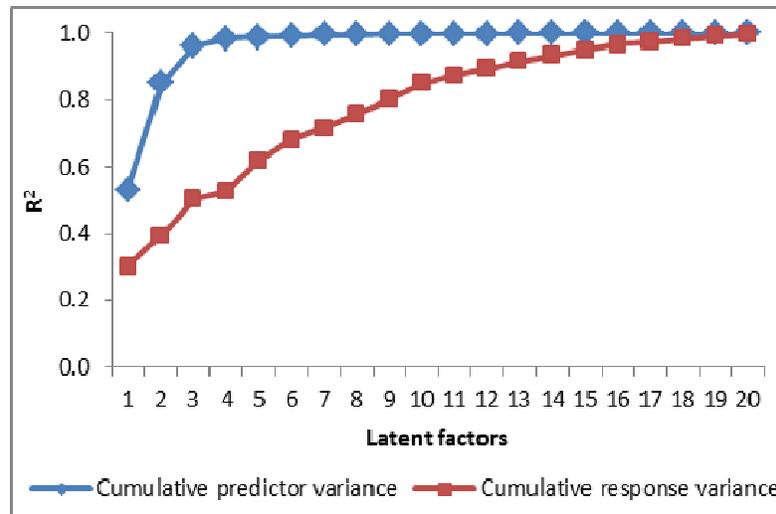
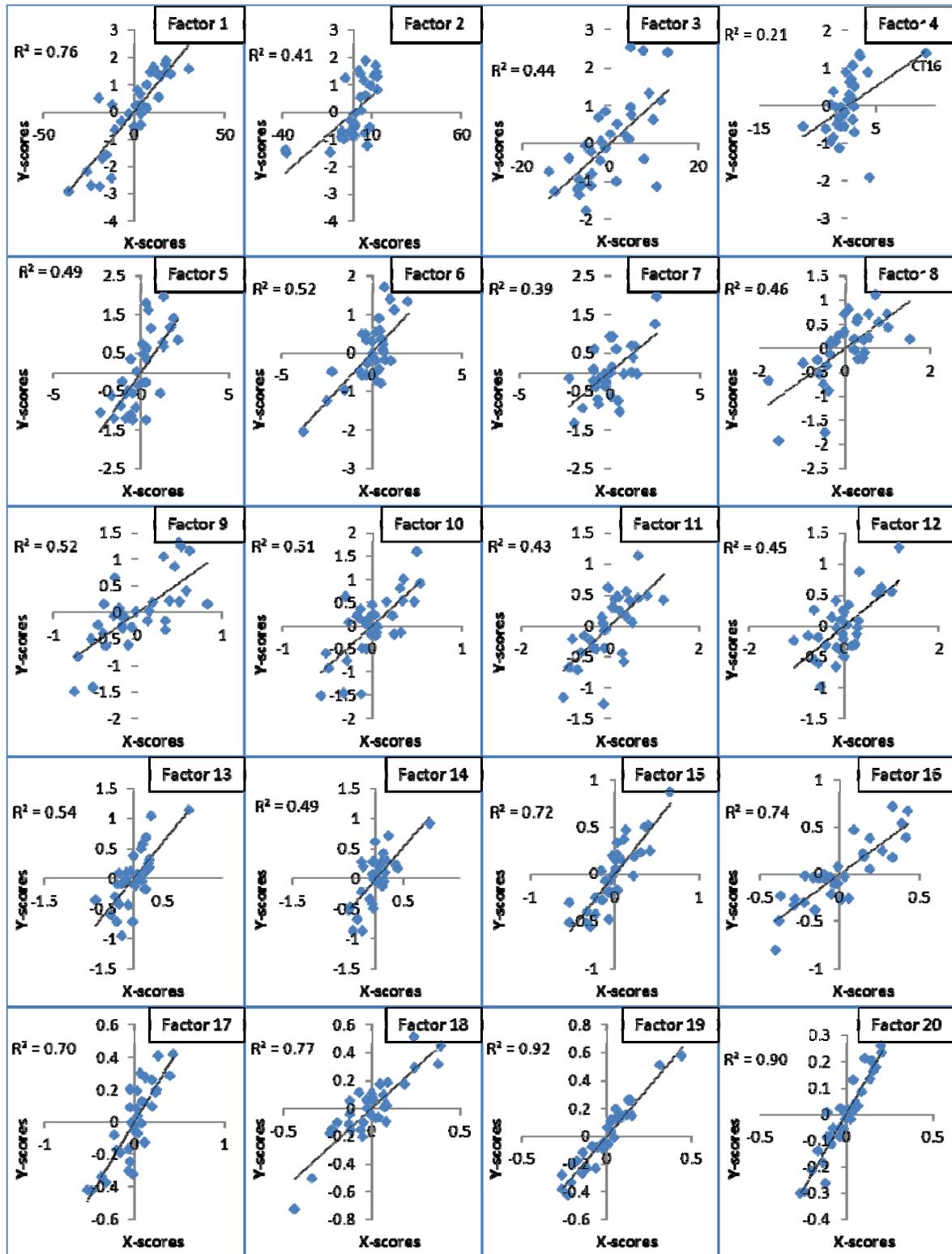


Figure 2 shows the amount of variation explained by each additional latent factor; 85% of the predictor variation is explained by the first two factors, rising to 96% with three, and plateauing out at 99% by five latent factors. However, the response variation is less well explained by the first few factors; only 39% in the first two factors, rising to 62% by factor five and 85% by factor 10, suggesting that more latent factors are required to explain both sets of variance and predict the response of the vegetation data.

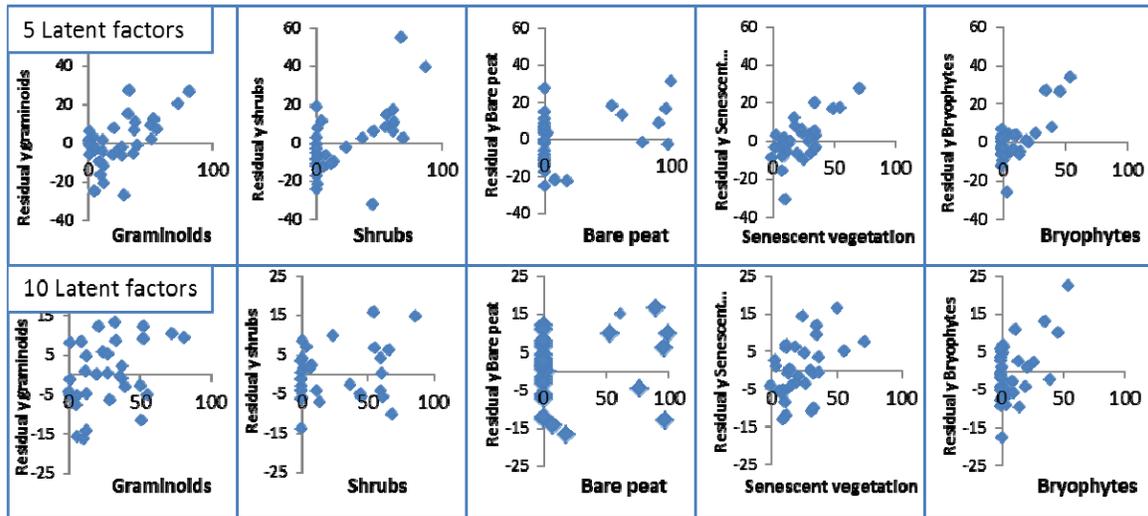
The correlations in the latent factors (Figure 3) followed the expected pattern of a strong relationship for the first few factors to an extent. The first latent factor showed one of the strongest relationships, which then drops until latent factor 4, when the  $R^2$  values remained relatively constant between 0.39 and 0.54. The stable correlations for the X and Y scores, not dropping off after the first few latent factors, again suggest that a larger number of factors should be included. Above 15 the values rose again, suggesting that above this level the model would suffer from over fitting.

Residual plots, and plots of distance to response and predictor models, did not identify any outliers in the data, so all the vegetation plots were included in the model. The 10 latent factor model performed better on these two counts as well. All of the samples in the 10 latent factor model fitted closer to the model in both dimensions, but particularly in terms of the modelled predictors, so the residuals were reduced. There was an upward trend in the distribution of the residuals in the five latent factor model, underestimating the cover in plots with low percentages and over estimating those with high ones. This is symptomatic of data without a normal distribution. The effect was eliminated with 10 latent factors (Figure 4).

Figure 3. Plots of X-scores against Y-scores for each of the latent factors in the PLSR model.  $R^2$  values shown in each one.



**Figure 4.** Plots of PLSR residual against observed values for each PFT.



3.3. PLSR Variable Selection

Selection of variables to be removed was based on the weighted regression coefficients and the plot of variables importance (Figures 5 and 6). The regression coefficients take into account differences in the range of variation of the original values and are an expression of the relative importance of the predictors—in this case spectral bands—in predicting the response. The Variable Importance for Projection (VIP) statistic summarises the contribution of each predictor in fitting the PLSR model for both predictors and response [53].

**Figure 5.** Regression parameter profile and variable importance plots for the PLSR using five latent factors. (a) Variable importance plot for all bands; (b) Variable importance plot with spectral bands removed; (c) Regression parameter profile for all bands; (d) Regression parameter profile with spectral bands removed.

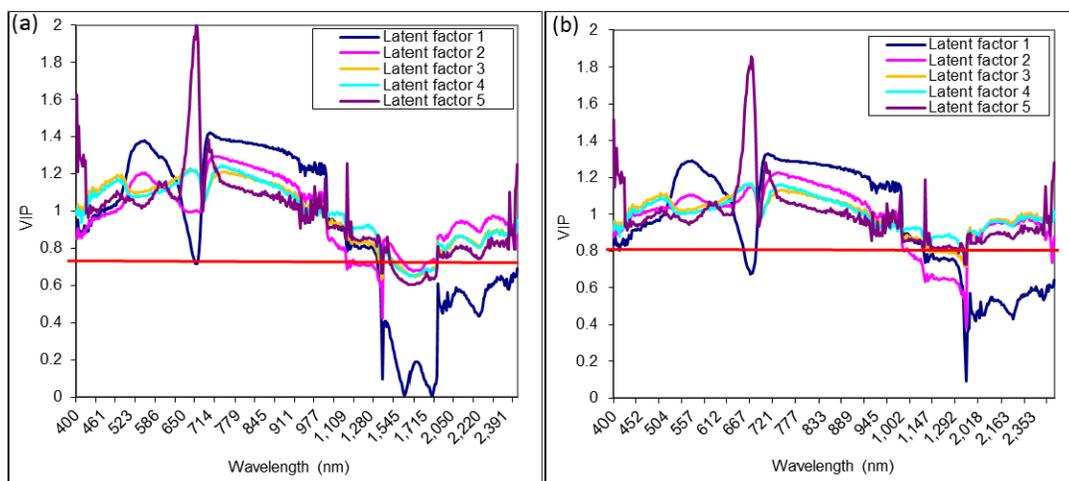


Figure 5. Cont.

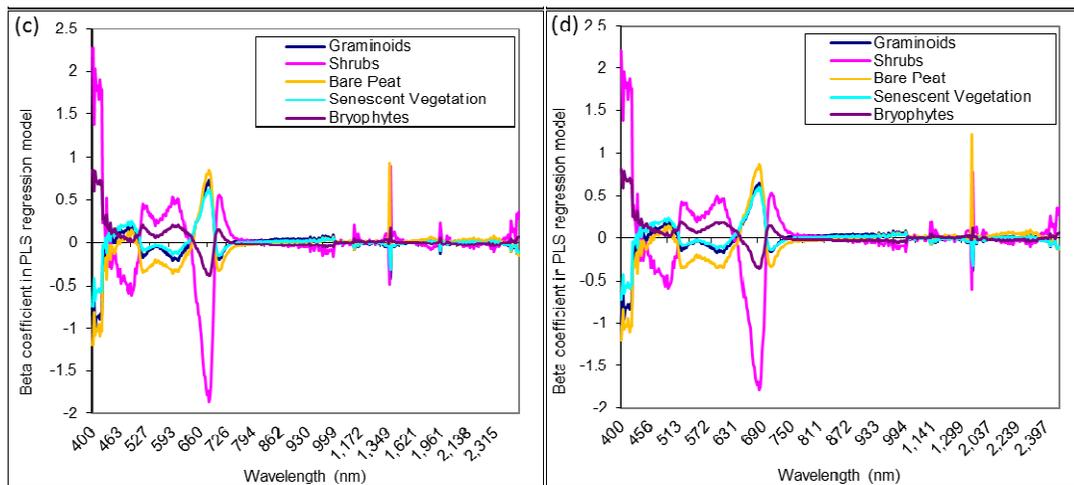
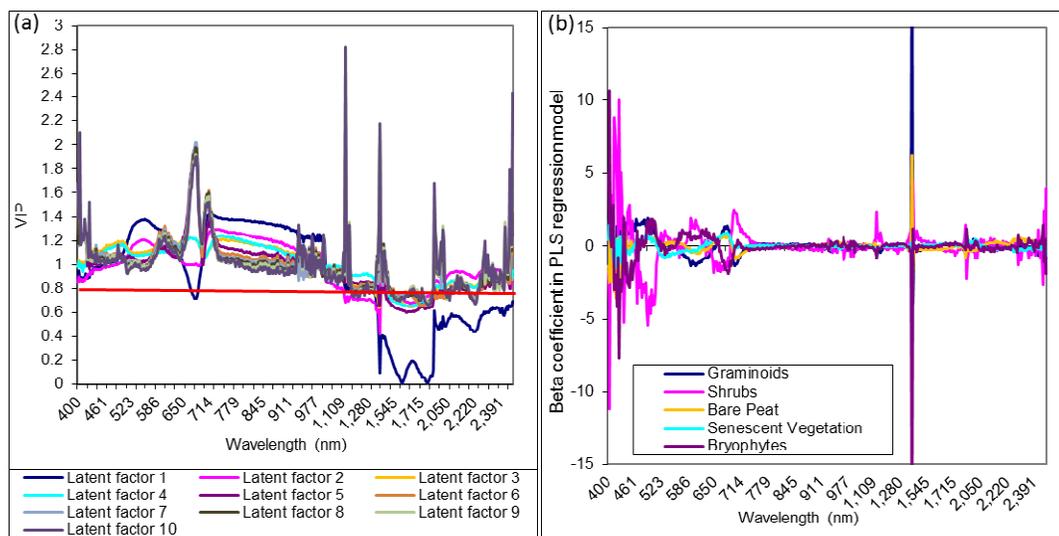


Figure 6. Regression parameter profile and variable importance plots for the PLSR using 10 latent factors. (a) Variable importance plot for all bands; (b) Regression parameter profile all bands.



### 3.4. PLSR Validation

The PLSR models were validated against a test dataset. The accuracy of the PLSR was assessed in terms of absolute prediction accuracy (RMSE) and the amount of explained variance (coefficient of determination  $R^2$ ). Table 3 shows the final results.

Reducing the number of variables reduced the RMSE on the validation dataset by 2%– 4.5%, with the largest improvement in the shrubs obtained using 10 latent factors. The validation results confirm that 20 latent factors do over fit the model. The lowest overall  $RMSE_{val}$ , for the complete model, 13.19 is the PLSR with five latent factors and a reduced number of spectral bands. This has to be considered the best overall model for predicting all the cover types in the same manner. The best model for each individual cover type or PFTs do favour the 10 latent factor models selected for

graminoids, bare peat and bryophytes, as indicated in Table 3, and Figure 7. The greater amount of variation in the vegetation classes explained by 10 factors should lead to greater improvements in the final results.

**Table 3.** Results for the PLSR models.  $R^2_{cal} = R^2$  in calibration,  $R^2_{val} = R^2$  between the predicted and the validation data,  $RMSE_{cal} = RMSE$  in calibration,  $RMSE_{val} = RMSE$  between the predicted and the validation data. LF = latent factors.

		Overall	Graminoids	Shrubs	Bare Peat	Senescent Vegetation	Bryophytes
5 LF all variables	$R^2_{cal}$	0.62	0.64	0.64	0.84	0.55	0.41
	$R^2_{val}$		0.33	0.62	0.76	0.56	0.19
	$RMSE_{cal}$	20.09	18.41	27.68	22.75	14.89	13.14
	$RMSE_{val}$	16.73	16.20	16.80 *	16.81	10.02 *	21.73
5 LF reduced variables	$R^2_{cal}$	0.62	0.63	0.63	0.85	0.55	0.41
	$R^2_{val}$		0.34	0.61	0.76	0.56	0.18
	$RMSE_{cal}$	16.72	12.88	17.01	13.20	10.66	11.23
	$RMSE_{val}$	13.19	16.08	17.04	16.54	10.11	21.76
10 LF all variables	$R^2_{cal}$	0.85	0.85	0.94	0.93	0.79	0.73
	$R^2_{val}$		0.44	0.55	0.88	0.32	0.54
	$RMSE_{cal}$	7.74	8.15	6.70	8.74	7.29	7.64
	$RMSE_{val}$	17.58	14.97 *	24.24	13.62	13.13	19.36
10 LF reduced variables	$R^2_{cal}$	0.84	0.84	0.92	0.92	0.76	0.74
	$R^2_{val}$		0.41	0.64	0.89	0.32	0.56
	$RMSE_{cal}$	8.36	8.54	8.06	9.73	7.76	7.52
	$RMSE_{val}$	15.46	15.93	19.78	10.96 *	13.03	16.11 *
20 LF all variables	$R^2_{cal}$	1.00	1.00	1.00	1.00	1.00	0.99
	$R^2_{val}$		0.39	0.58	0.90	0.53	0.44
	$RMSE_{cal}$	1.51	1.04	1.95	2.00	1.10	1.11
	$RMSE_{val}$	23.66	30.66	24.00	21.55	13.54	25.19

\* most successful model for each cover type.

**Figure 7.** Scatterplots of the PLSR predicted vs. observed values for the vegetation abundance in the validation dataset.

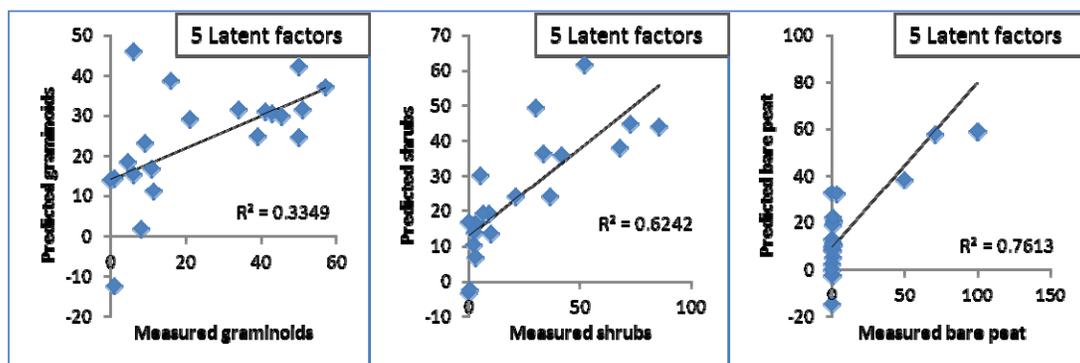
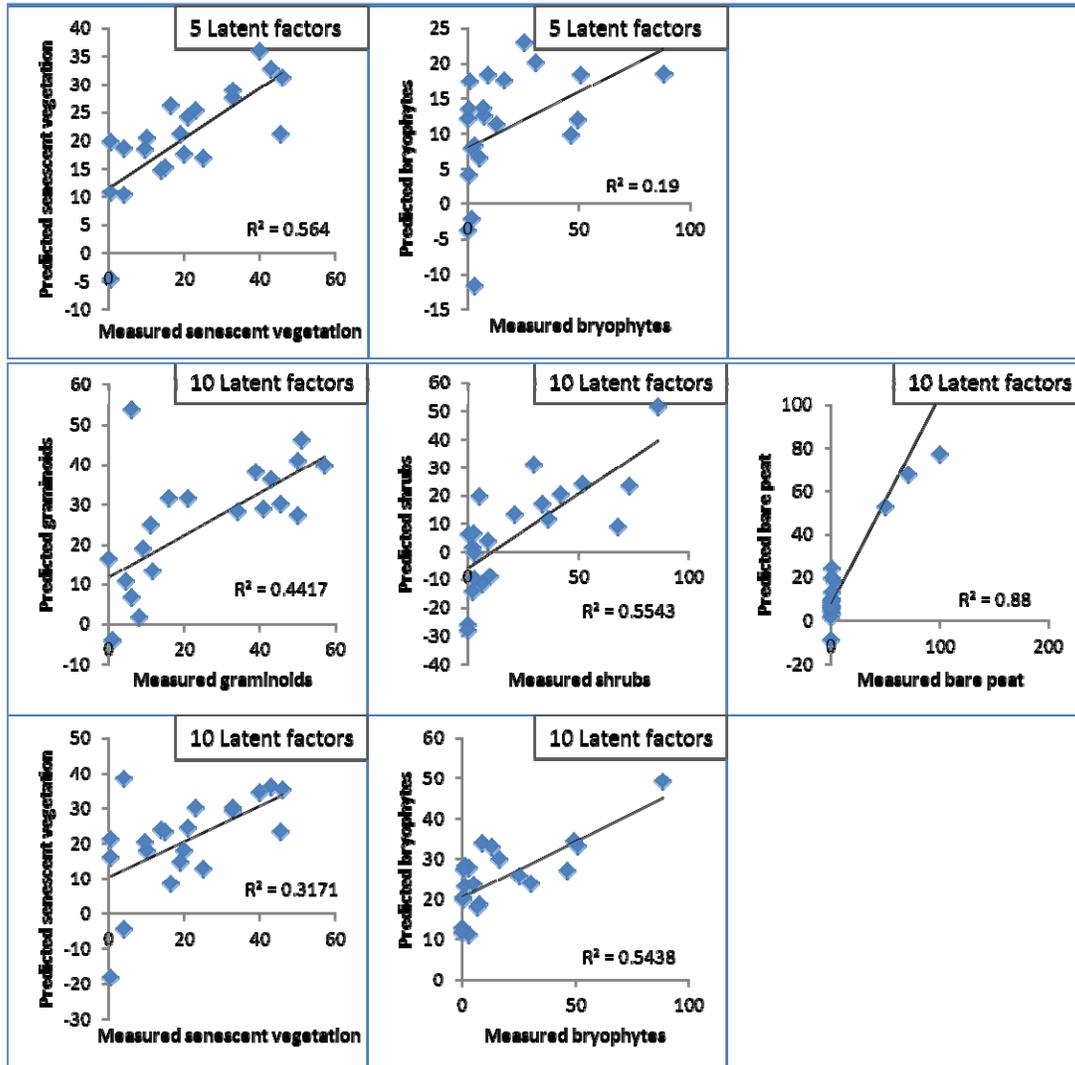


Figure 7. Cont.



### 3.5. Application to the Image; Spatial Prediction

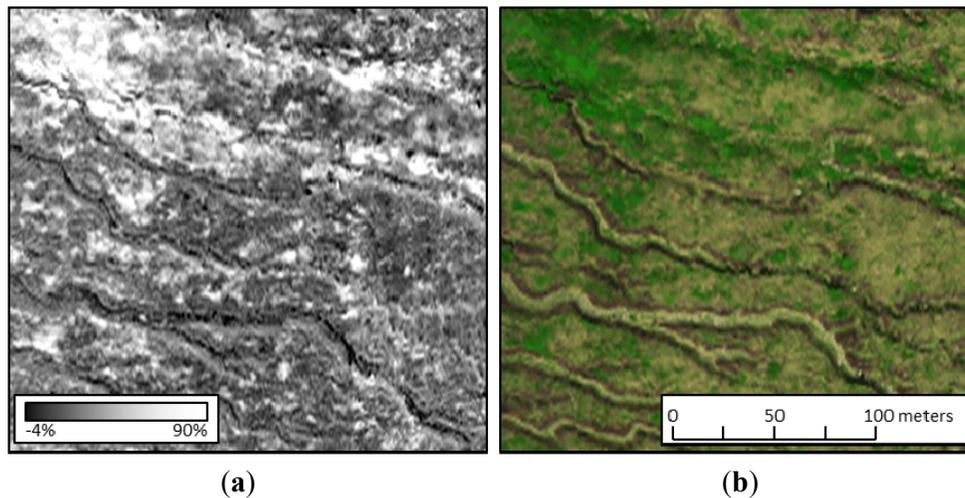
The predictive images produced by applying the PLSR show patterns of abundance that are in accordance with field observations and knowledge of the site. The success lies in the level of detail provided as a result in the high resolution data over a large area. Particular examples are picked out for illustrative purposes. From the predicted shrub image (Figure 8a) *Vaccinium myrtillus* can clearly be picked out in white. It is seen as a distinctive bright green colour in June on the tops of hags and along the tops of the gullies on the Eagle true colour composite (Figure 8b).

The success of the predicted bare peat image is demonstrated by the ability of the PLSR model to correctly identify the bare gully sides between the grass lined bottoms and the vegetated tops (Figure 9). These are narrow linear features, approximately 1–2 m wide.

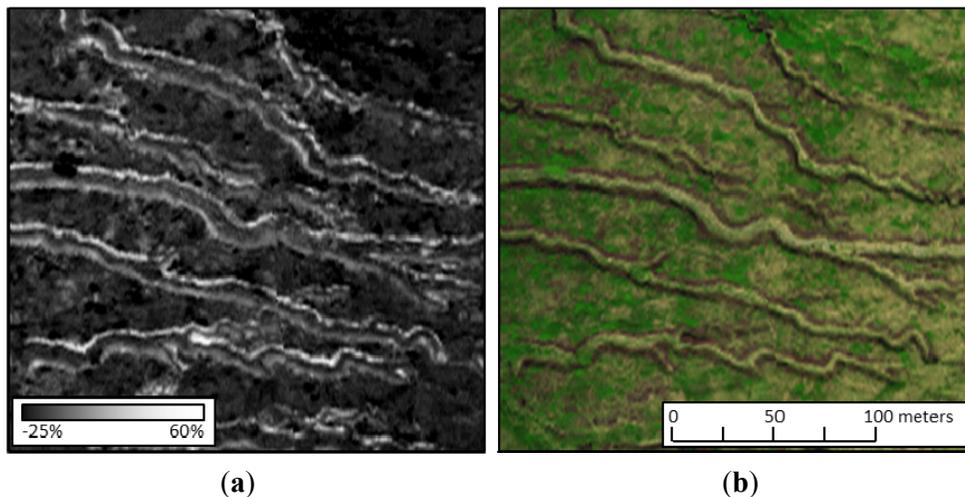
The images have predicted values with a wider range than the predicted plots calculated in the model validation because the model was built on a sample, whereas the whole image has some pixels with a wider range of values. However, histograms of PLSR predicted cover show there is a normal

distribution, and the majority of the pixels predict a percentage coverage in the range expected from the RMSE of the model validation, as confirmed by the mean values (Table 4). There are only few very extreme pixels in this case at the tails of the histograms, and so the sample pixels in the calibration dataset are considered representative.

**Figure 8.** (a) Subset of the abundance of shrub image predicted from PLSR; (b) Eagle sensor true colour image (bands 640 nm, 550 nm and 460 nm), of the same area.



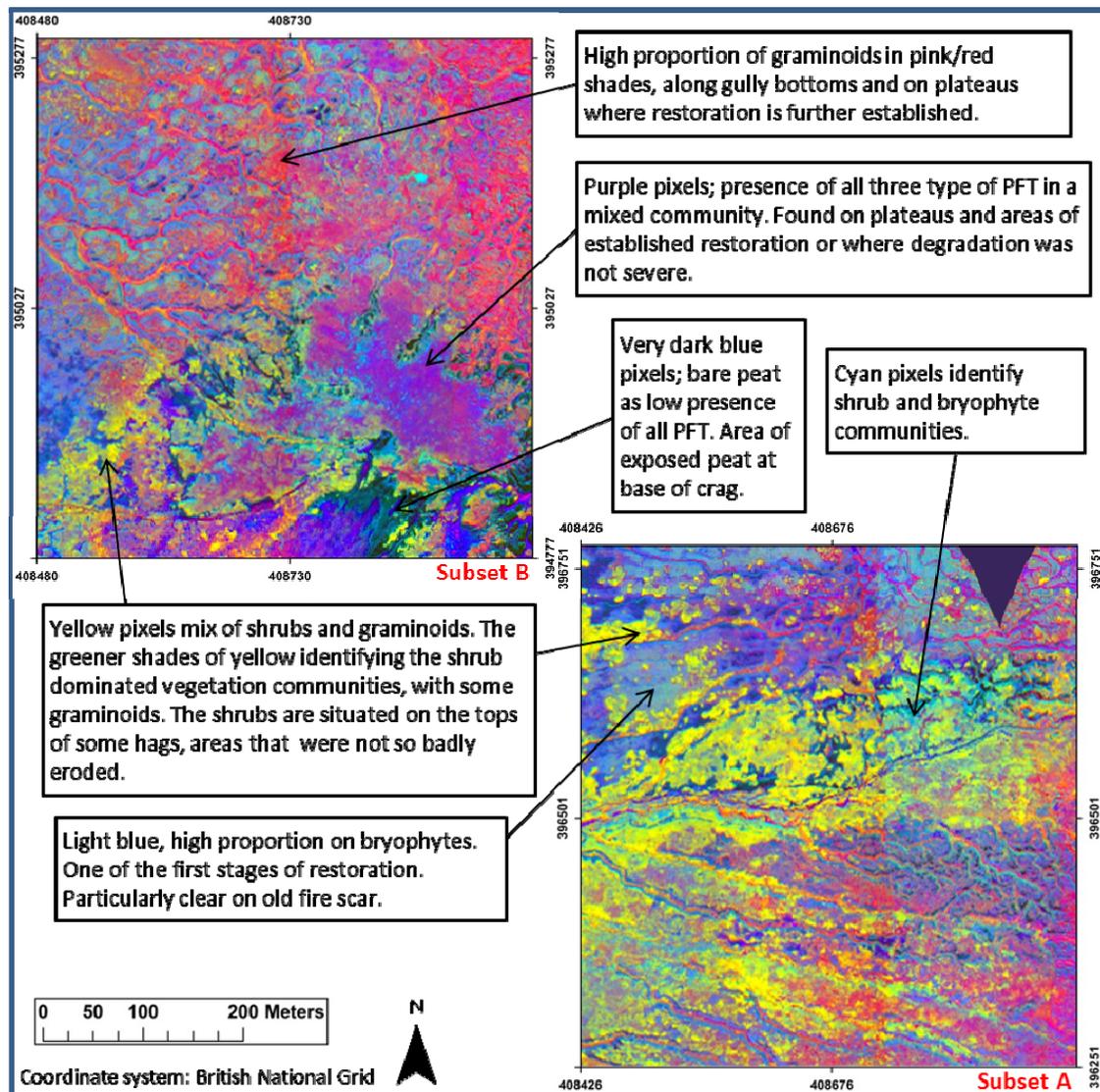
**Figure 9.** (a) Subset of the abundance of bare peat image predicted from PLSR; (b) Eagle sensor true colour image (bands 640 nm, 550 nm and 460 nm) of the same area.



**Table 4.** Mean predicted abundance value from the whole PLSR images for each cover type.

Cover Type	Mean Predicted Abundance (%) from PLSR Image
Graminoids	20.69
Shrub	25.52
Bare Peat	12.95
Senescent vegetation	15.96
Bryophytes	10.87

**Figure 10.** False colour composite images, predicted abundance from PLSR for Graminoids (Red) Shrubs (Green) and Bryophytes (Blue). Subset areas marked on Figure 1.



False colour composites of combinations of the five modelled images are useful representations of the vegetation composition at a landscape scale. Figures 10 and 11 illustrate this, using some example subsets representing different areas and conditions on the site. Figure 10 is a composite of PLSR predicted abundance images for graminoid, shrub and bryophytes, the three main vegetation types on the site. The clear differences in colour demonstrate the heterogeneous nature of the vegetation communities and the small spatial scale on which they vary. The relative dominance of the three PFTs in each pixel dictates the colour representation, following the colour additive principle. Different topographical features, stages of restoration and resulting vegetation communities along the floristic trajectory of restoration can be identified (Figure 10). Figure 11 shows bare peat, bryophytes, and graminoids PLSR images, displayed as RGB respectively, and demonstrates the flexibility of using the various combinations of PLSR images to highlight different vegetation communities, elements of the ecosystem, and restoration processes.

**Figure 11.** False colour composite image, predicted abundance from PLSR for bare peat (Red) Bryophytes (Green) and Graminoids (Blue). Subset areas marked on Figure 1. Colours indicate relative predicted abundance of the three cover types.

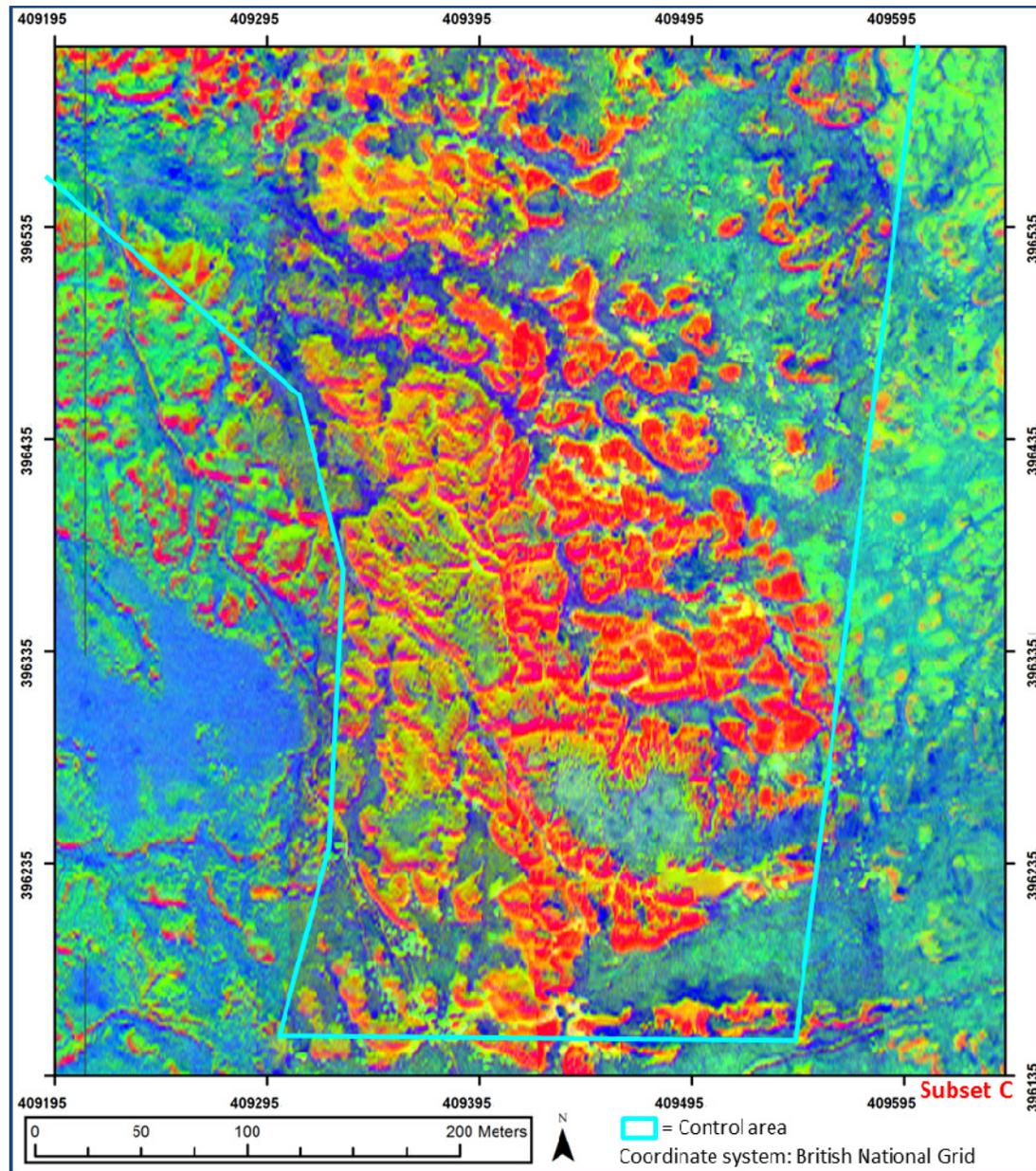
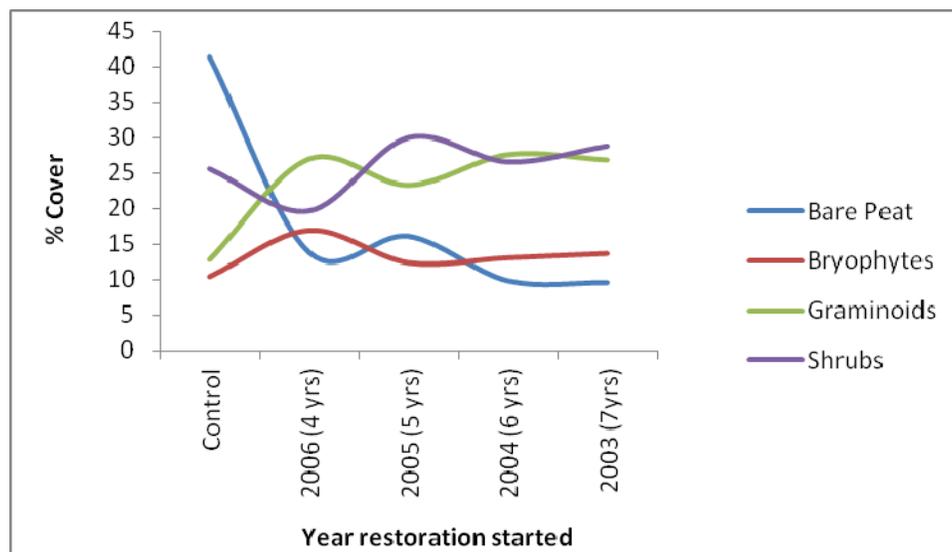


Figure 11 is centred on the control area, where there has been no restoration, but it also includes some of the restored areas (Figure 1). The control area itself is dominated by exposed peat hags, which are easily identified as red tones in the image. The edges of the peat hags are being recolonized by two different PFT in two separate processes, which can also be identified on the FCC image (Figure 11). The pink tones are predominantly bare peat but with a proportion of graminoids. This re-colonisation of the hags with grasses is happening in the main control area, where grass seeds have washed into the gully bottoms from the surrounding restoration blocks. This is especially the case on the more sheltered south facing gully sides, but different illumination may be contributing here. The other

process is bryophyte establishment, predominately *Campylopus introflexus*, a pioneer species after disturbance and one of the first re-colonisers in the restoration process. In Figure 11, this process is identified by the graduation from orange to yellow tones. The orange tones are pixels with a high proportion of bare peat and a presence of bryophytes and, as the proportion of bryophytes increases, the colour shifts towards yellow. From the colour additive principle, yellow indicates areas with approximately equal predicted abundance of bare peat and bryophytes, with very low values for graminoids pixels, this is present on a few of the hags. This second process is occurring more predominantly on the western side of the control area which borders one of the more established and oldest restoration blocks (Figure 1). The yellow tones graduate to greenish-blue colours as the proportion of bare peat is reduced and bryophytes and graminoids increase. The other main features in Figure 11 are the dark blue gullies with a high percentage of graminoids where the grass seed has been washed in and germinated. Away from the hag and gully topography, the paler blues represent areas of established graminoids with *Eriophorum* spp, such as the plateau area seen at the western side of Figure 11.

The application in the PLSR model to the whole image has provided a quantitative measure for the abundance of cover types across the whole restoration site. This landscape scale model is an invaluable tool for operational restoration monitoring. By calculating the mean of each cover type in the blocks of restoration seeded in each year between 2003 and 2006 (Figure 1) we can see some general trends in succession (Figure 12). There is a sharp decline in bare peat between the area that has had no restoration and those seeded most recently in 2006. In this four-year time period, the graminoids and bryophytes also increase. It takes a further year for the shrubs to start increasing coverage. This happens at the expense of bryophytes, which start to drop off as they become out-competed and the other species take hold.

**Figure 12.** Mean predicted coverage of the PFT in all the restoration blocks seeded between 2003 and 2006, from the PLSR images. The age of the restoration at time of imaging is given in brackets after the year.



#### 4. Discussion

Initial investigation revealed a lack of clear linear relationships between the spectral indices and the PFT abundance data (Table 2), suggesting that PLSR using the full spectral information of the hyperspectral data was the more appropriate modelling approach. Investigation into the data revealed the primary problem in attempting to model abundance of species from spectral data as the non-normal distribution of the vegetation data, as seen in the histograms and scatterplots of relationships with spectral indices. The skewed distributions are a result of vegetation plot data matrices with zero or low values. Aggregation into PFT rather than single species reduced this effect. This same situation, of low numbers of occurrences causing weak models has been reported with the modelling of species abundance [31,32]. A sensitivity test on the relationships between the spectral data and the vegetation abundance data at PFT and single species level confirmed that aggregating the data to PFTs was the correct scale for analysis, and is supported by studies in the literature [31,32,56]. This test suggests that aggregating vegetation survey results into PFT would be sensible for operational monitoring. It would make the ground vegetation survey results more robust, as it is very unlikely that any misidentification of individual species would be severe enough to change result between PFT. Using PFT also conforms to the way that the restoration practitioners think and communicate about the plant communities.

The interactive method for selecting the variables for analysing the VIP statistic along with the weighted regression coefficients did make some improvements in the results. Schmidtlein and Sassin [31] found this method to be more effective by further reducing the number than in this study and ensuring the spacing between bands. Ensuring the full spectral data is represented and that known diagnostic features are represented in the PLSR avoids the pitfalls identified in the PCR method. The VIP statistic did select bands in parts of the spectrum known to be diagnostic of biophysical properties of the vegetation. Spectral bands from the Eagle sensor (VIS-NIR, 400–1,000 nm) were dominant, particularly the blue proportion of the visible spectrum (400–500 nm), showing the value of this region where all the vegetation pigments absorb. The chlorophyll absorption features were identified with the VIP statistics, with a second peak at the chlorophyll absorption maxima, 660 nm, (Figures 5 and 6). The rapid rise in vegetation reflectance between the low red chlorophyll reflectance to the high near infrared reflectance was identified in the PLSR as an important feature, marked by a peak in the significance of all of the bands around the red edge, between 670 and 780 nm. The subsequent peaks in the SWIR at 1,160 nm, 1,450 nm, and 1,960 nm, were all at wavelengths with water absorptions [52]. The peak in importance of bands at the longer wavelengths of 2,290 and 2,448 are a result of cellulose absorption features, with the shrubs showing the highest values in these peaks. The weighted regression coefficients of the individual spectral bands showed which bands are most important in the model. The bands in the SWIR, Hawk sensor, showed a considerably lower contribution (Figures 5 and 6). These results suggest that the SWIR has less influence in the landscape-scale monitoring of the vegetation on peatland from the hyperspectral image than initially predicted [57].

One possible reason for the high RMS error of the validation data is the effect of complex field conditions on the PLSR technique, which was originally designed for lab analysis where confounding variables can be controlled. Plant canopy reflectance is affected by either the observing conditions; variable viewing and illumination geometry; or variations in the target such as variable LAI or internal shade from the canopy structure. Feilhauer *et al.* [39] showed that a brightness-normalised modification

to the PLSR significantly improved the prediction of leaf chemistry from normal PLSR, but made little improvement on the effects of variable LAI and variable viewing angle. A further refinement of the model in the future could be to test this adapted algorithm.

The models are weakest at predicting coverage in plots with zero or very small percentages of the particular type of vegetation, sometimes giving negative cover values in these cases (Figure 7). The negative cover values are within the RMSE margin of the model and so were considered to be unavoidable output from the statistical model.

Despite the PLSR validation errors being larger than reported in other studies [31,32,43], the model does manage to predict the abundance of the different cover types to a level that is useful in an operational context. The lowest RMSE values for each cover type range from 10%–16%, with senescent vegetation, bare peat, graminoids, bryophytes and then shrubs being predicted the most accurately. The ability to predict how much bare peat there is in an area of restoration with a 10% accuracy is extremely useful for restoration managers, as would being able to predict the proportion of shrubs to graminoids to within 16%. Traditional vegetation survey employs a visual assessment of the percentage cover of each species, which involves subjective decisions and therefore inherent error [58]. One of the main assumptions under which multivariate modelling is applied to remotely sensed data states that the modelled variables are measured accurately. Curran [52] suggests that violation of this assumption would cause the most difficulty, so it is the uncertainty of the vegetation abundance data that has the biggest impact upon the regression model. When viewed in this context, the  $RMSE_{val}$  of the PLSR is reasonable. Furthermore, the percent abundance of the dominant species in each of the vegetation communities, along the trajectory of restoration, as defined in previous analyses of vegetation composition data, varies by a magnitude that is greater than the error in the PLSR prediction models, therefore making this an operationally useful tool.

The predictive images show patterns of abundance that are in accordance with field observations and knowledge of the site. Their success is demonstrated by the ability of the model to allow features at a single pixel level as a result of the high spatial resolution. The widespread pattern of a decrease in bare peat and bryophytes and an increase in graminoids and shrubs is evident in the mean coverage of each cover type in the restoration blocks (Figure 12), despite the more complex small scale processes and the extremely heterogeneous nature of existing vegetation and topological features within each block. This is why high spatial resolution is also a requirement in order to see these more complex variations.

## 5. Conclusions

The output from the PLSR model produced images of predicted coverage for each PFT or cover type. This is a validated quantitative measure of abundance of vegetation at a landscape scale. Monitoring of the restoration site has not been achieved on this scale with this accuracy before. It provides an invaluable operational tool for restoration managers. Complete coverage at a landscape scale is impractical by traditional surveying methods. The PLSR images allow operational decisions by restoration managers to be informed by the spatial patterns. Applications could include land management decisions on further treatment, such as patch reseeded, fertilizer application or plug planting, without the need for extensive ground-based survey. The PLSR predictive images will help

inform judgment about the effectiveness of the restoration, for example by identifying areas where reversion has occurred. The high spatial resolution data makes the images particularly effective, as the detail in the vegetation response has been captured and not averaged over areas larger than the actual scale of change. The PLSR used the majority of the spectral bands in the model; so it was equally the high spectral resolution of the data which contributed to the success of the vegetation cover predictions. These PLSR predicted images provide a monitoring tool that would not be achievable using lower spatial and spectral resolution remote sensing.

### Acknowledgments

The project was funded by a NERC CASE studentship with Natural England (NE/F013787/1), held at The University of Manchester. The authors would like to acknowledge the support of Moors for the Future Partnership for facilitating the work and in their role as CASE partner. The hyperspectral images were collected by the NERC Airborne Research and Survey Facility (ARSF), and the ASD field spectrometer was on loan from the NERC Field Spectroscopy Facility (FSF).

### Author Contributions

The research for this paper was conducted by all the authors. The writing of the paper was conducted by Dr. B Cole with revisions and edits by both J. Mcmorrow and Prof. M. Evans.

### Conflicts of Interest

The authors declare no conflict of interest.

### References

1. Davies, A. Review of the Historical Environmental Changes in the UK Uplands Relevant to Management and Policy. In *Foundations for the Future: Learning from the Past*; University of Sterling: Stirling, UK, 2008; p. 62.
2. Bain, C.G.; Bonn, A.; Stoneman, R.; Chapman, S.; Coupar, A.; Evans, M.; Gearey, B.; Howat, M.; Joosten, H.; Keenleyside, C.; et al. *IUCN UK Commission of Inquiry on Peatlands*; IUCN UK Peatland Programme: Edinburgh, UK, 2011.
3. Moore, P.D. The future of cool temperate bogs. *Environ. Conserv.* **2002**, *29*, 3–20.
4. Pearce-Higgins, J.; Grant, M.; Beale, C.M.; Buchanan, G.; Sim, I. International Importance and Drivers of Change of Upland Bird Populations. In *Drivers of Environmental Change in Uplands*; Bonn, A., Allott, T., Hubacek, K., Stewart, J., Eds.; Routledge: London, UK, 2009; pp. 209–227.
5. Thompson, D.B.A.; MacDonald, A.J.; Marsden, J.H.; Galbraith, C.A. Upland heather moorland in Great Britain: A review of international importance, vegetation change and some objectives for nature conservation. *Biol. Conserv.* **1995**, *71*, 163–178.
6. Ramchunder, S.J.; Brown, L.E.; Holden, J. Environmental effects of drainage, drain-blocking and prescribed vegetation burning in UK upland peatlands. *Prog. Phys. Geogr.* **2009**, *33*, 49–79.
7. Tallis, J.H. Growth and degradation of British and Irish blanket mires. *Environ. Rev.* **1998**, *6*, 81–122.

8. Bonn, A.; Rebane, M.; Reid, C. Ecosystem Services. A New Rationale for Conservation of Upland Environments. In *Drivers of Change in Upland Environments*; Bonn, A., Allott, T., Hubacek, K., Stewart, J., Eds.; Routledge: London, UK, 2009; pp. 448–474.
9. Holden, J.; Shotbolt, L.; Bonn, A.; Burt, T.P.; Chapman, P.J.; Dougill, A.J.; Fraser, E.D.G.; Hubacek, K.; Irvine, B.; Kirkby, M.J.; *et al.* Environmental change in moorland landscapes. *Earth-Sci. Rev.* **2007**, *82*, 75–100.
10. Van der Wal, R.; Bonn, A.; Monteith, D.T.; Reed, M.; Blackstock, K.; Hanley, N.; Thompson, D.; Evans, M.; Alonso, I. *Mountains, Moorlands and Heaths*; United Nations Environment Program-World Conservation Monitoring Centre: Cambridge, UK, 2011.
11. Worrall, F.; Burt, T.; Adamson, J.; Reed, M.; Warburton, J.; Armstrong, A.; Evans, M. Predicting the future carbon budget of an upland peat catchment. *Clim. Chang.* **2007**, *85*, 139–158.
12. Bellamy, P.H.; Loveland, P.J.; Bradley, R.I.; Lark, R.M.; Kirk, G.J.D. Carbon losses from all soils across England and Wales 1978–2003. *Nature* **2005**, *437*, 245–248.
13. Sieben, E.J.J. Plant functional composition and ecosystem properties: The case of peatlands in South Africa. *Plant Ecol.* **2012**, *213*, 809–820.
14. Lavorel, S.; McIntyre, S.; Landsberg, J.; Forbes, T.D.A. Plant functional classifications: From general groups to specific groups based on response to disturbance. *Trend. Ecol. Evol.* **1997**, *12*, 474–478.
15. Schumann, M.; Joosten, H. *Global Peatland Restoration Manual*; Institute of Botany and Landscape Ecology, Greifswald University: Greifswald, Germany, 2008; p. 64.
16. Aplin, P. Remote sensing: Ecology. *Prog. Phys. Geogr.* **2005**, *29*, 104–113.
17. Gillespie, T.W.; Foody, G.M.; Rocchini, D.; Giorgi, A.P.; Saatchi, S. Measuring and modelling biodiversity from space. *Prog. Phys. Geogr.* **2008**, *32*, 203–221.
18. Turner, W.; Spector, S.; Gardiner, N.; Fladeland, M.; Sterling, E.; Steininger, M. Remote sensing for biodiversity science and conservation. *Trend. Ecol. Evol.* **2003**, *18*, 306–314.
19. Medcalf, K.A.; Parker, J.A.; Turton, N.; Finch, C. *Making Earth Observation Work for UK Biodiversity Conservation—Phase 1*; Project Code: CR 0477; Joint Nature Conservation Committee: Peterborough, UK, 2011.
20. Petchey, S.; Brown, K.; Hambidge, C.; Porter, K.; Rees, S. *Operational Use of Remote Sensing for Environmental Monitoring*; Joint Report JP001; Natural England: Sheffield, England, 2011.
21. Mehner, H.; Cutler, M.; Fairbairn, D.; Thompson, G. Remote sensing of upland vegetation: The potential of high spatial resolution satellite sensors. *Glob. Ecol. Biogeogr.* **2004**, *13*, 359–369.
22. Aplin, P. On scales and dynamics in observing the environment. *Int. J. Remote Sens.* **2006**, *27*, 2123–2140.
23. Mehner, H.; Fairbairn, D.; Cutler, M. Defining the Optimal Spatial Resolution for Mapping Upland Vegetation. In proceedings of RSPSoc Annual Conference 2004: Remote Sensing and Photogrammetry Society: Mapping and Resources Management, Aberdeen, UK, 6–10 September 2004.
24. Tallis, J.H. Peat erosion in the Pennines: The badlands of Britain. *Biologist* **1997**, *44*, 277–279.
25. Elkington, T.; Dayton, N.; Jackson, D.L.; Strachan, I.M. *National Vegetation Classification Field Guide to Mires and Heaths*; Joint Nature Conservation Committee: Peterborough, UK, 2002.

26. Asmat, A.; Milton, E.J.; Atkinson, P.M. Empirical correction of multiple flightline hyperspectral aerial image mosaics. *Remote Sens. Environ.* **2011**, *115*, 2664–2673.
27. McMorrow, J. Linear regression modelling for the estimation of oil palm age from Landsat TM. *Int. J. Remote Sens.* **2001**, *22*, 2243–2264.
28. Schaepman-Strub, G.; Limpens, J.; Menken, M.; Bartholomeus, H.M.; Schaepman, M.E. Towards spatial assessment of carbon sequestration in peatlands: Spectroscopy based estimation of fractional cover of three plant functional types. *Biogeosciences* **2008**, *6*, 275–284.
29. Cole, B.; McMorrow, J.; Evans, M.G. Spectral monitoring of moorland plant phenology to identify a temporal window for hyperspectral remote sensing of peatland. *ISPRS J. Photogramm. Remote Sens.* **2013**, in press.
30. Schmidtlein, S. Imaging spectroscopy as a tool for mapping Ellenberg indicator values. *J. Appl. Ecol.* **2005**, *42*, 966–974.
31. Schmidtlein, S.; Sassin, J. Mapping of continuous floristic gradients in grasslands using hyperspectral imagery. *Remote Sens. Environ.* **2004**, *92*, 126–138.
32. Schmidtlein, S.; Feilhauer, H.; Bruelheide, H. Mapping plant strategy types using remote sensing. *J. Veg. Sci.* **2012**, *23*, 395–405.
33. Curran, P.J. Imaging spectrometry. *Prog. Phys. Geogr.* **1994**, *18*, 247–266.
34. Kumar, L.; Schmidt, K.; Dury, S.J.; Skidmore, A. Imaging Spectrometry and Vegetation Science. In *Imaging Spectrometry: Basic Principles and Prospective Applications*; van der Meer, F.D., de Jong, S.M., Eds.; Kluwer Academic Publishers: Dordrecht, The Netherlands, 2001; pp. 111–155.
35. Schlerf, M.; Atzberger, C.; Hill, J. Remote sensing of forest biophysical variables using HyMap imaging spectrometer data. *Remote Sens. Environ.* **2005**, *95*, 177–194.
36. Atzberger, C.; Guerif, M.; Baret, F.; Werner, W. Comparative analysis of three chemometric techniques for the spectroradiometric assessment of canopy chlorophyll content in winter wheat. *Comput. Electron. Agric.* **2010**, *73*, 165–173.
37. Huang, Z.; Turner, B.J.; Dury, S.J.; Wallis, I.R.; Foley, W.J. Estimating foliage nitrogen concentration from HYMAP data using continuum removal analysis. *Remote Sens. Environ.* **2004**, *93*, 18–29.
38. Cho, M.A.; Skidmore, A.; Corsi, F.; van Wieren, S.E.; Sobhan, I. Estimation of green grass/herb biomass from airborne hyperspectral imagery using spectral indices and partial least squares regression. *Int. J. Appl. Earth Obs. Geoinf.* **2007**, *9*, 414–424.
39. Feilhauer, H.; Asner, G.P.; Martin, R.E.; Schmidtlein, S. Brightness-normalized partial least squares regression for hyperspectral data. *J. Quant. Spectrosc. Radiat. Transf.* **2010**, *111*, 1947–1957.
40. Hansen, P.M.; Schjoerring, J.K. Reflectance measurement of canopy biomass and nitrogen status in wheat crops using normalized difference vegetation indices and partial least squares regression. *Remote Sens. Environ.* **2003**, *86*, 542–553.
41. Gomez, C.; Rossel, R.A.V.; McBratney, A.B. Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study. *Geoderma* **2008**, *146*, 403–411.
42. Viscarra Rossel, R.A.; Walvoort, D.J.J.; McBratney, A.B.; Janik, L.J.; Skjemstad, J.O. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma* **2006**, *131*, 59–75.

43. Feilhauer, H.; Faude, U.; Schmidlein, S. Combining Isomap ordination and imaging spectroscopy to map continuous floristic gradients in a heterogeneous landscape. *Remote Sens. Environ.* **2011**, *115*, 2513–2524.
44. Rosipal, R.; Krämer, N. Overview and Recent Advances in Partial Least Squares. In *Subspace, Latent Structure and Feature Selection Techniques*; Saunders, C., Grobelnik, M., Gunn, S., Shawe-Taylor, J., Eds.; Springer: Berlin Heidelberg, Germany 2006; pp. 35–51.
45. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130.
46. Cheriyyadat, A.; Bruce, L.M. Why Principal Component Analysis is not an Appropriate Feature Extraction Method for Hyperspectral Data. In Proceedings of the 2003 IEEE International Geoscience and Remote Sensing Symposium, IGARSS '03, Toulouse, France, 21–25 July 2003; pp. 3420–3422.
47. Onojeghuo, A.O.; Blackburn, G.A. Optimising the use of hyperspectral and LiDAR data for mapping reedbed habitats. *Remote Sens. Environ.* **2011**, *115*, 2025–2034.
48. Höskuldsson, A. Variable and subset selection in PLS regression. *Chemom. Intell. Lab. Syst.* **2001**, *55*, 23–38.
49. SAS. The PLS Procedure. In *SAS/STAT (R) 9.2 User's Guide*, 2nd ed.; SAS Institute Inc.: Cary, NC, USA, 2012.
50. Kubinyi, H. Evolutionary variable selection in regression and PLS analyses. *J. Chemom.* **1996**, *10*, 119–133.
51. Martens, H.; Martens, M. Modified Jack-knife estimation of parameter uncertainty in bilinear modelling by partial least squares regression (PLSR). *Food Qual. Prefer.* **2000**, *11*, 5–16.
52. Curran, P.J. Remote-sensing of foliar chemistry. *Remote Sens. Environ.* **1989**, *30*, 271–278.
53. Wold, S. PLS for multivariate linear modelling. *Chemom. Methods Mol. Des.* **1995**, *2*, 195–218.
54. Curran, P.J.; Dungan, J.L.; Macler, B.A.; Plummer, S.E. The effect of a red leaf pigment on the relationship between red edge and chlorophyll concentration. *Remote Sens. Environ.* **1991**, *35*, 69–76.
55. Nagler, P.L.; Daughtry, C.S.T.; Goward, S.N. Plant litter and soil reflectance. *Remote Sens. Environ.* **2000**, *71*, 207–215.
56. Ustin, S.L.; Gamon, J.A. Remote sensing of plant functional types. *New Phytol.* **2010**, *186*, 795–816.
57. Cutler, M.; McMorro, J.; Evans, M. Remote sensing of upland peat erosion in the Southern Pennines. *North West Geogr.* **2002**, *2*, 21–30.
58. Cherrill, A.; McClean, C. Between-observer variation in the application of a standard method of habitat mapping by environmental consultants in the UK. *J. Appl. Ecol.* **1999**, *36*, 989–1008.