7. Results from Cors Caron

7.1 Aims of the Cors Caron study

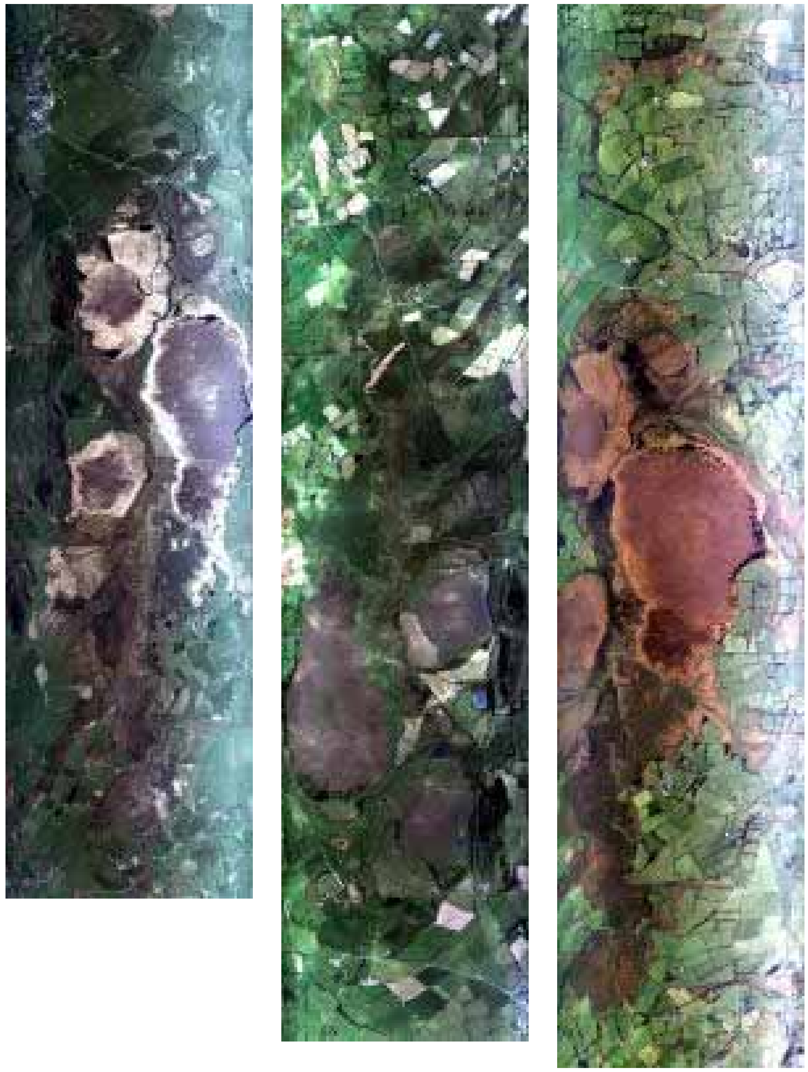
The results from Wedholme Flow show that satisfactory class definition for the purpose of the LRBI/EUHD is achievable using a remote sensing methodology based upon a hybrid approach:

- visual interpretation of lidar and Ikonos data to identify the primary/secondary habitats, and
- maximum likelihood classification of high spatial resolution data collected in visible and near infra-red wavelengths in late-summer to identify the plant communities present.

The aim of the Cors Caron work was threefold. First, to test the methodology on a raised bog with vegetation communities, terrain type and management history that differed significantly from those at Wedholme Flow; second, to determine how suitable data collected at other times of the year would be to perform the same task; and third, to investigate whether spectral information in the longer wavelength infra-red and thermal regions improved the classification.

7.2 Influence of time of the year

ATM data from Cors Caron were available for three periods during the year, roughly corresponding to spring, mid-summer and late-summer. The data were flown by NERC specifically for a PhD project investigating the vegetation history of Cors Caron and these dates were chosen in relation to our knowledge of the phenological stages of the communities present. Logistical constraints due to cloud cover and aircraft availability meant that the precise dates of data acquisition were determined by practical reasons. Visible band colour composites of the data acquired are shown in Figure 23.



8 May 2001 27 June 2001 12 October 2001

Figure 23. Visible band colour composites of the Airborne Thematic Mapper data acquired over Cors Caron

It is clear from the images above that their suitability for vegetation mapping varies considerably throughout the year. The spring image is dominated by the high reflectance from the *Molinia* and shows very little information for the mire surface. The mid-summer image is more informative, but has cloud shadows present which would make its use in a digital classification very problematic. The late-summer image shows the most discrimination between the vegetation classes of interest, confirming our choice of this time of the year for the Wedholme Flow data. Having inspected imagery from the three dates it was decided that only the data from the late-summer flight from NERC would be purchased.

The classification procedure adopted for Cors Caron was very similar to that used for Wedholme Flow, except that lidar data were not available for this site. This meant that the Primary / Secondary interpretation was achieved in a slightly different fashion. First, the ATM data had to undergo a series of pre-processing steps.

7.2.1 Pre-processing the ATM data

Unlike the Ikonos data, the ATM data had to have several pre-processing techniques applied before they were suitable for digital classification. First, a correction had to be applied for the cross-track shading because the flight line lies at an angle to the solar principal plane. This is seen in Figure 23 as a progressive brightening of the image from right to left. Second, the ATM also had to be screened for data quality to exclude any bands with sensor errors or any that were very noisy due to scattering in the atmosphere. The signal-to-noise ratio of ATM band 1 was very poor and so this was excluded from further analysis.

Airborne scanner data are notoriously prone to geometric distortions. These make it difficult to compare image data with maps and ground data. There are two types of geometric distortions: those associated with the attitude or orientation of the platform (roll, pitch and yaw), and those related to the optical characteristics of the sensor, and both of these have the potential to introduce distortions into the output image. In essence, platform motion and attitude instabilities will cause the sampled data to be projected in a shifted orientation in the reconstructed image. Therefore, before any of the classification techniques could be applied, it was necessary to register the image data to a geographical co-ordinate system and to correct for the motion of the aircraft during data acquisition.

Airborne data collected by the NERC ARSF were geometrically corrected to British National Grid co-ordinate systems post-flight using data collected from an on-board flight recorder called the "Integrated Data System" (IDS). The IDS has been developed to integrate the imagery provided by the airborne scanner with the navigation and attitude data from the aircraft. The AZGCORR post-processing software package provided by the NERC ARSF makes it possible to correct for these various errors and provide fully geo-referenced digital data without the use of ground control points. Precise geometric correction of the ATM data was not possible, however, as we did not have access to a digital elevation model (DEM) of the area. To overcome this problem we took the data as corrected for aircraft motion using AZGCORR and then used a number of ground control points to perform a conventional geometric correction to the OS National Grid.

7.2.2 Generating the mire mask

Masking is the process whereby areas of an image are eliminated from subsequent processing stages. A mask is a binary image consisting of values of 0 and 1 only. When a mask is used in a processing function, the areas with values of 1 are processed, and the areas with values of 0 are not. However, before a spatial mask can be applied to the data, it has to be constructed.

Using the Environment for Visualising Images (ENVI) software package, Regions of Interest (ROI's) were defined around the boundary of Cors Caron. Ms. Schulz was consulted during construction of this mask, which was defined with the aid of a map provided by the Countryside Council for Wales detailing the delineation of the NNR boundary. The ROI was then converted into a binary mask for use in subsequent processing stages (Figure 24).

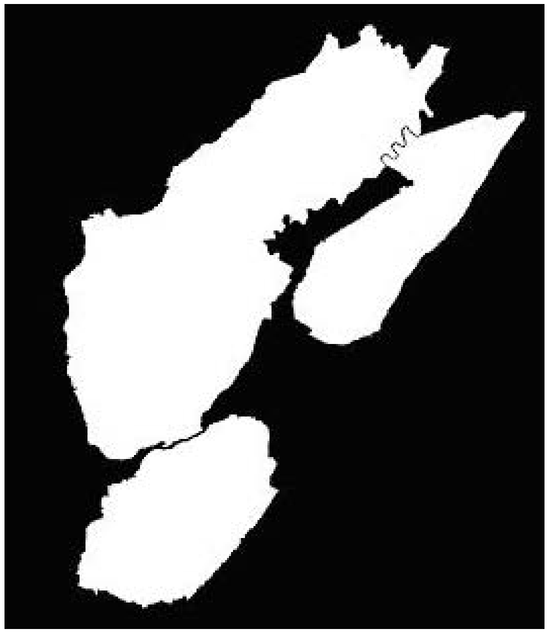


Figure 24. Binary mask of Cors Caron Region of Interest (ROI)

7.2.3 Training site location

Once the geocorrection had been implemented, it was necessary to locate training sites within the bog boundary upon which the classification would later be based. This was conducted through direct liaison with Jenny Schulz and Richard Lindsay. Training sites were positioned in areas of marsh corresponding to Ms. Schulz's classification map of the surface condition classes (Figure 25). Training sites were also positioned in other areas where there were clearly spectral differences in the plant communities present, but where these small-scale patterns had not been highlighted by the classification map. The training sites used for classification purposes are listed in Table 12.

Table 12. Training sites used for the Cors Caron classification

Training Site Name	Number of Pixels
Molinia cut	2693
Molinia uncut	1706
Uncut *Sphagnum* surface	4878
Standing water	193
Carr 1	522
Cut *Sphagnum*/Cottongrass	688
Disturbed bog surface	191
Dense *Calluna*	2098
Carr 2	1171
Wetter *Sphagnum*	699

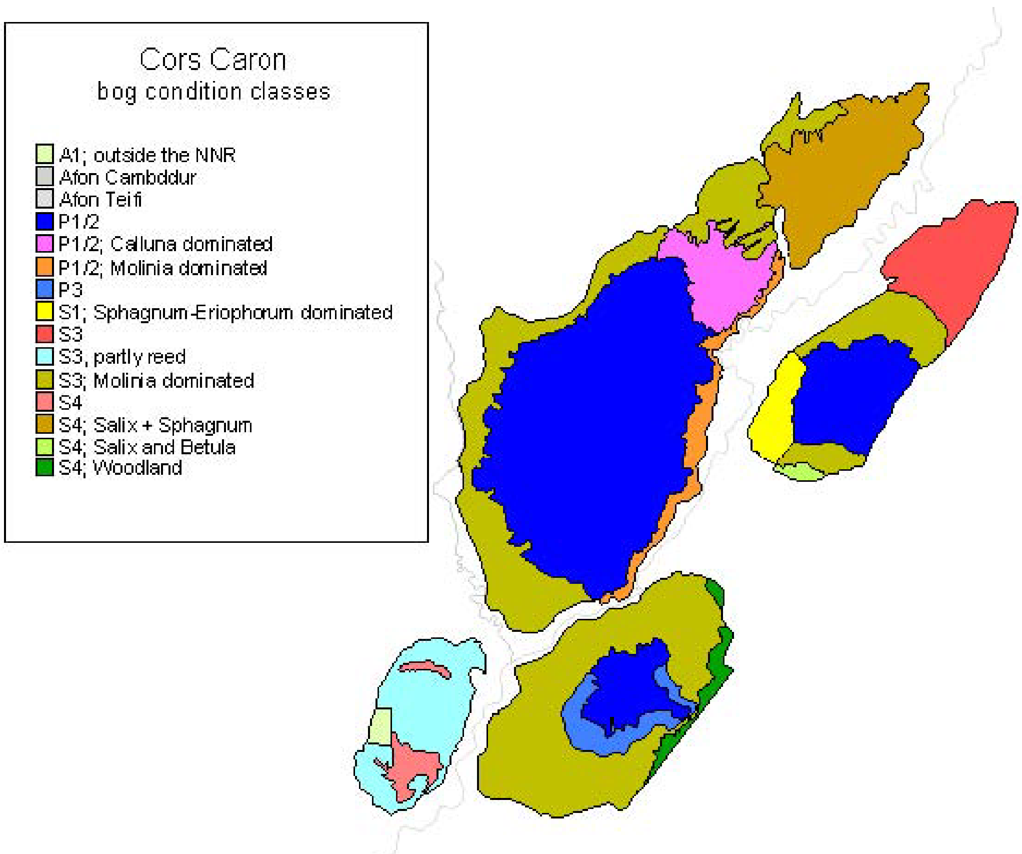


Figure 25. Bog condition map of Cors Caron derived from field survey (larger version in Appendix, after Table 24)

7.3 Results from the classification

7.3.1 Visible and Near Infra-red wavelengths

In order to provide a comparison with the Wedholme Flow results, the initial classifications were based upon data from those ATM bands which most closely matched the Ikonos bands: ATM bands 2, 3, 5 and 7. Many classifications were produced, and the best is shown in Figure 26.

Unlike Wedholme Flow, where we had the lidar image to identify the primary and secondary areas of the bog, for Cors Caron we had to include this distinction within the selection of training sites. Overall, this strategy was successful. However, it was found necessary to choose training sites from each of the three bogs as spatial extrapolation of the spectral signatures was not successful in general. The reason for this can be seen in Figure 27. This figure shows some residual haziness affecting the area along the southern edge of the swathe, which the cross-track shading correction had been unable to account for.

7.3.2 Incorporation of other spectral bands

Many earlier studies on terrestrial vegetation communities have utilised broad-band spectral data from instruments such as the Landsat Thematic Mapper (TM). Remote sensing at higher spectral and spatial resolutions, such as those offered by airborne scanner systems, offers the prospect of high quality land cover classification and mapping. However, one of the problems facing the analyst presented with data from these more advanced sensors is the wide choice of bands available. Decisions must be made about which of the bands might be suitable, and how many of these should be used. Simply using all the available bands for a classification is not only inefficient, in terms of the time taken to process the data, it also risks introducing error and uncertainty into the results. For every classification task there is an optimum number and combination of spectral bands; the question is how to determine this.

The pattern recognition technique known as 'feature selection' tackles this problem. This can be performed on the basis of probabilistic distance measures such as the *Transformed Divergence* which is calculated on the basis of the class mean vectors and the covariance matrices measured from a training sample (i.e. the spectral qualities of the LRBI classes). The transformed divergence measure therefore reflects the separability between the classes of interest. Feature selection differs from techniques such as principal components analysis in that it does not transform the bands into a new feature space, it simply analyses the classes to be separated (in this case, the bog condition classes) and finds the best original bandset combination to separate them.

Different feature selection methods give different results for the same data. Therefore, in order to test which method gives the best result, a combined approach was adopted. In this study, feature selection was followed by a maximum likelihood classification, whereby the classification error for each method gave a measure of the accuracy of the feature selected bands used. The best feature subset is therefore the one with the highest classification accuracy.

The feature selection algorithm was implemented on the bog class data from the October flightline, in order to determine the best band subset for spectrally separating the classes of interest. The training data input to the algorithm were derived through consultation with Jenny Schulz and Richard Lindsay, the consultant ecologists for this project. The map produced by Ms Schulz from field survey is shown in Figure 25.

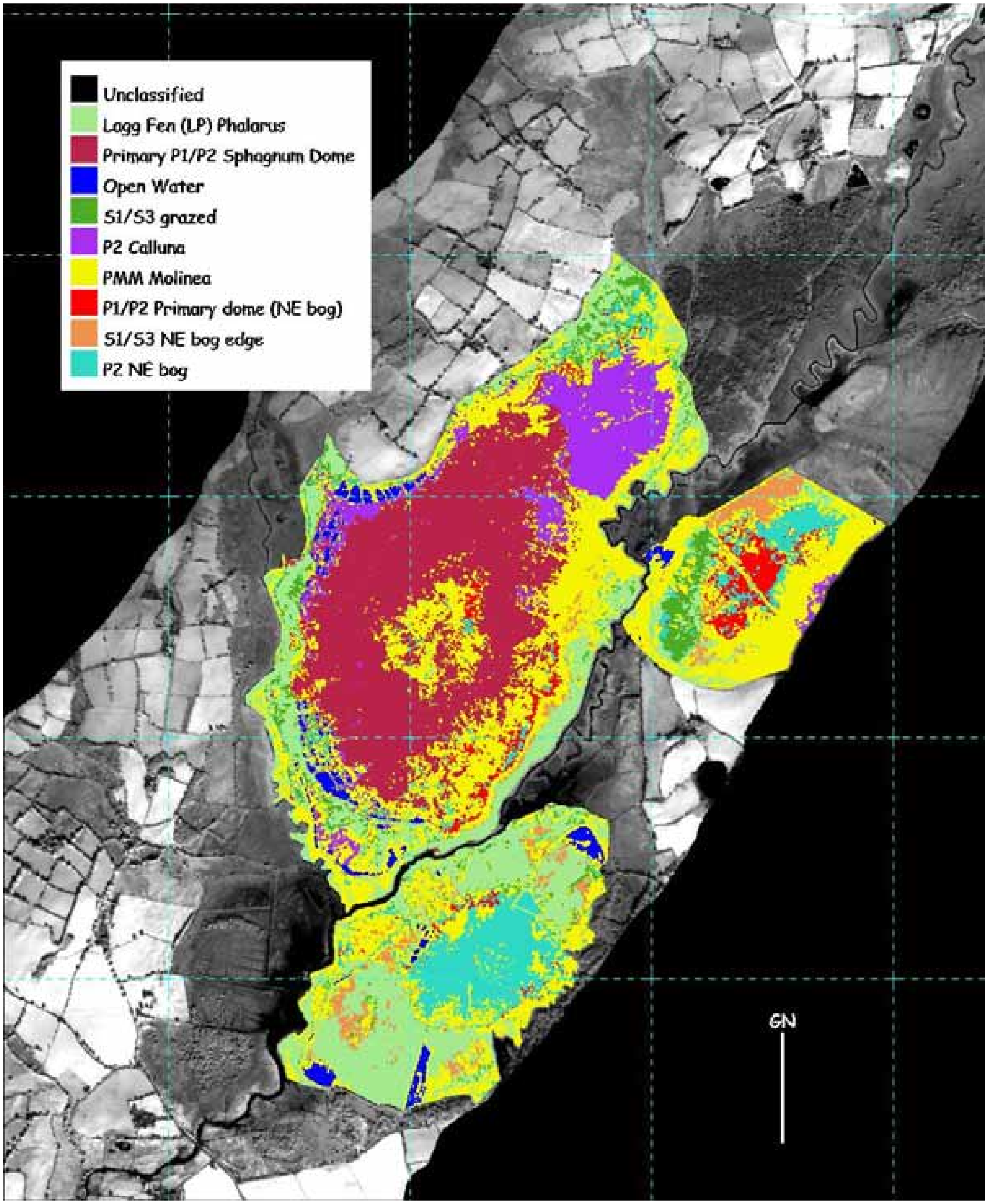


Figure 26. Maximum likelihood classification of Cors Caron based upon ATM bands 2,3,5 and 7 (Ikonos bands) [ATM data@NERC, 2001]

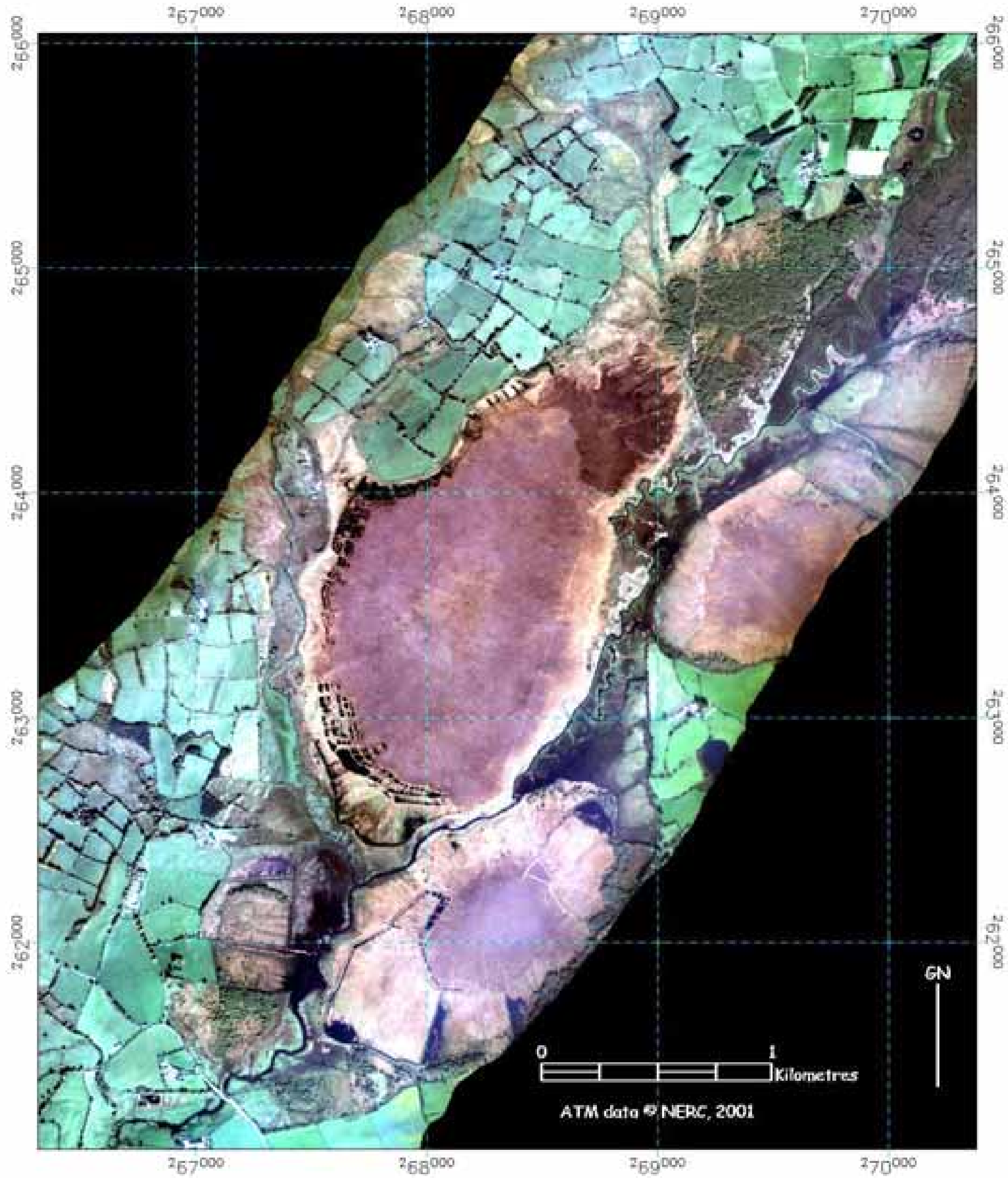


Figure 27. Simulated true colour composite of ATM bands 5, 3, and 2. Note the residual haziness affecting the two smaller bogs, despite the data having been corrected for cross-track shading. [ATM data © NERC, 2001]

Table 13 provides the results of the feature selection algorithm, when applied to all 11 bands of the October ATM image. The table lists the bands chosen from the original bandset in order of their importance for separating the classes defined during training site selection. It is apparent that this algorithm has highlighted a different bandset to the IKONOS bandset of 2, 3, 5, 7 used previously and suggests that a band combination of 3, 6, 9, 7 would better separate the classes, producing a higher classification accuracy. This ties in more closely with previous work by McMorrow *et. al.* (2002) on upland peat surfaces and Milton *et. al.* (2002) on valley mire communities in the New Forest. In these studies, spectral features in the 1500-1600nm wavelength region were shown to be indicative of depth to water table, and hence reflectance in this region is thought to act as a surrogate measure for the presence of species tolerant of waterlogged conditions.

Table 13. Results of the feature selection procedure applied to the October 2001 Cors Caron ATM data

Order of importance	ATM band	Centre Wavelength (nm)	Bandwidth (nm)
1	3	560	80
2	6	722.5	55
3	9	1650	200
4	7	830	140
5	8	980	140
6	5	660	60
7	2	485	70
8	11	10750	4500
9	4	615	20
10	10	2215	270
11	1	435	30

Figure 28 shows a colour composite created from the ‘best’ three feature selected bands. The first thing to note is that the haziness over the smaller bogs noted on Figure 27 is absent in this image, and this fact alone is likely to result in a more reliable classification. Furthermore, the distinction between a number of the classes of interest, such as the S1/S3 grazed area on the north-east bog and the boundaries between *Molinia* and lagg Fen on the main bog are much clearer in these spectral bands.

The best maximum likelihood classification created from these three ATM bands is shown in Figure 29. All of the major classes required for the LRBI/EU Habitats Directive are identified in this classification.

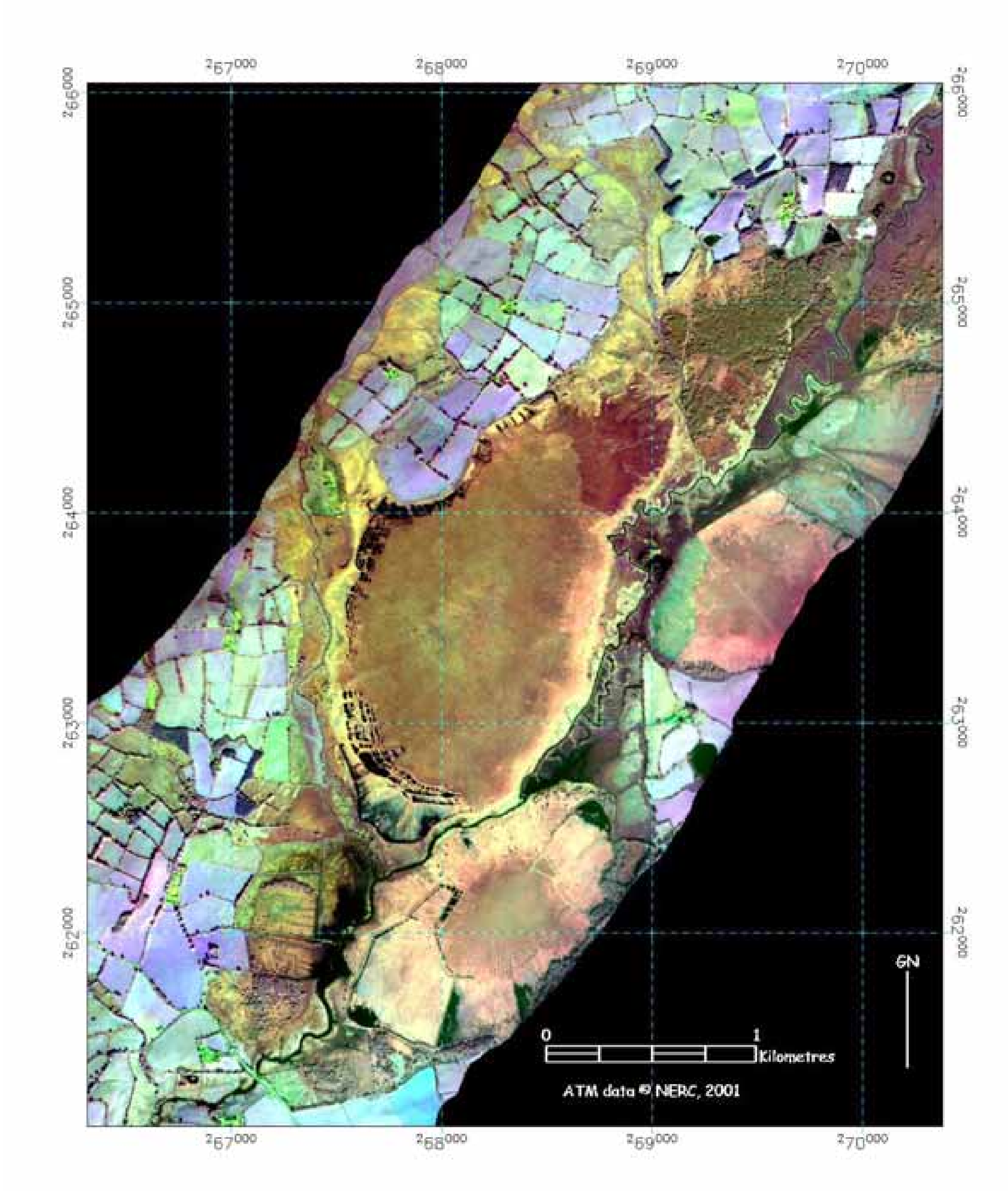


Figure 28. Colour composite of ATM bands 9, 3, and 6, the best three bands identified by the feature selection algorithm to separate the different raised bog habitat types. [ATM data © NERC, 2001]

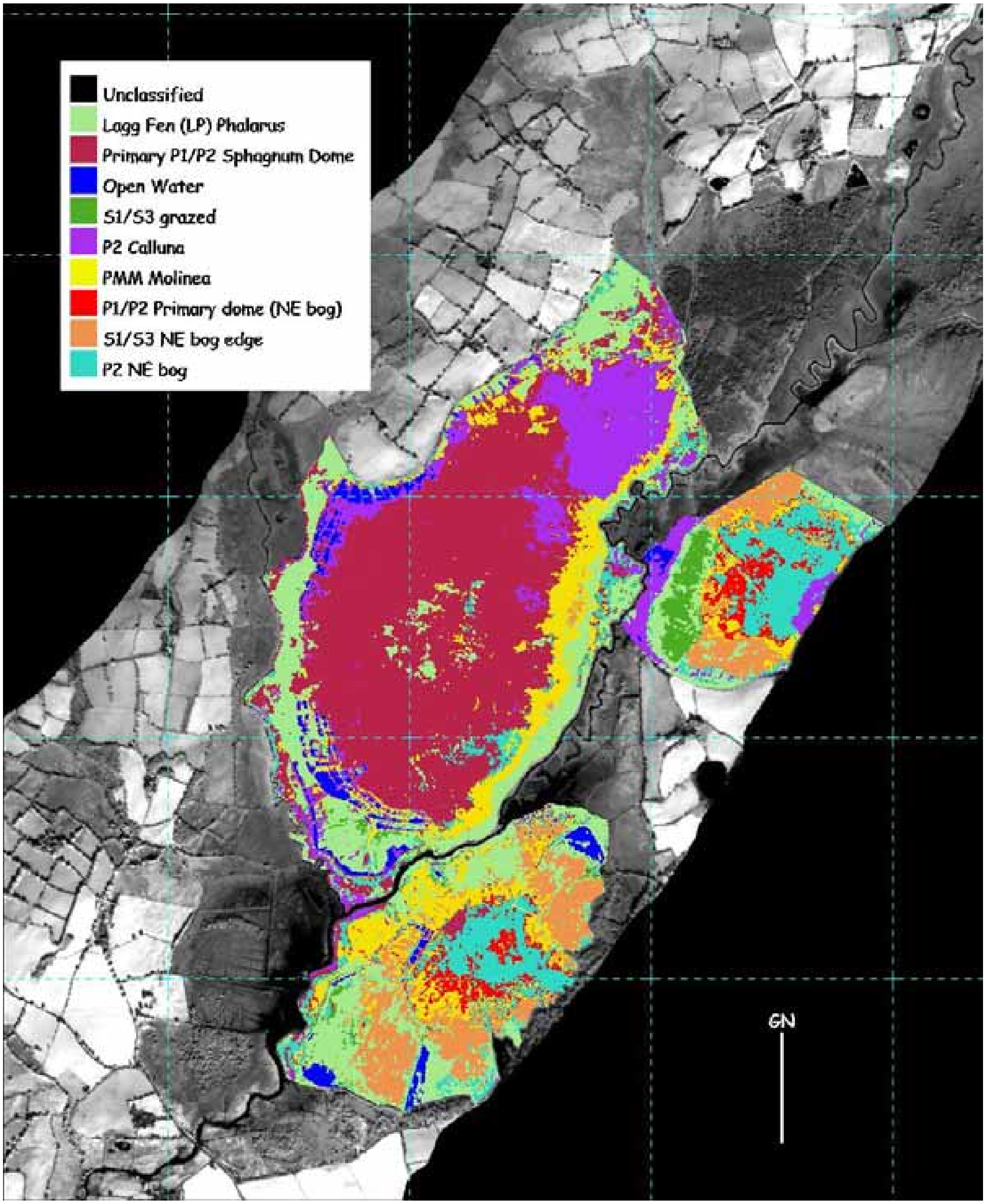


Figure 29. Maximum likelihood classification of the best three feature selected ATM bands. [ATM data © NERC, 2001]