



Figure 4. Top, a section (12km x 8 km) of summer-winter composite image covering the North York Moors; below, the same image segmented with vector-outlines to the image-segments overlaid.



Figure 5. An example of the colour charts used to sort training data at the Variant level into spectral subclasses (D - deciduous wood; L - littoral (mud); Ar - arable (rape); Ab - arable (barley); Gi - grass (improved); Ga - grass (acid); H - heather; Us -urban (suburban)). For each training area (e.g. D11), the display shows four colour blocks signifying: top, the mean spectral response in the summer image; second, the spectral response of the pixel most different from the mean in the summer image; third, the mean spectral response in the winter image; fourth, the spectral response of the pixel most different from the mean in the winter image; it also records the number of pixels used in calculating the mean spectral response. The training areas were selected according spectral homogeneity and sample size, and grouped according to spectral similarity. This approach also allowed the identification of problems within the training data. In the red box there is confusion between arable barley and improved grassland which may be due to a mis-identification during field reconnaissance. In the purple box is an area probably of acid grassland which has been labelled as heather, maybe a labelling error when transferring field annotations to the GIS.

8.3 Knowledge-based correction

Knowledge-based correction (KBC) procedures (Groom & Fuller 1996) were used to identify and re-label segments which were classified with low confidence and / or with classes out of their natural context. Some KBCs were applied universally. Others were general rules, used where needed but ignored where spectral classification had been successful. Some KBCs were one-off corrections: specific to a particular land cover type, on a particular image, and perhaps in just a small part of the landscape. Interactive KBC was sometimes used where generalised contextual rules could not be applied. Because of this variety of approaches, the following descriptions give examples rather than a complete picture. All KBC alterations are coded into the attribute data, attached to the relevant segments in the GIS database. The codes are cross-referenced to a tabular record giving the exact details of a correction. This information is available to users.

A rule applied consistently across the entire classification process concerned class-probabilities. Where any maximum likelihood class was allocated with a probability <50%, the other spectral subclass probabilities were summed to see if another BH was more appropriate (i.e. its spectral subclasses cumulatively took a dominant percentage of the overall probability). If so, the lead Subclass/Variant of the BH (i.e. usually the second choice class) took precedence. As an example, if a *Built up* subclass was allocated with 35% probability but *Arable* 'wheat', 'barley', 'oats' and 'linseed' had respectively 30%, 20%, 10% and 5% probabilities, the *Arable* BH was deemed more likely to apply, with 'wheat' as the most likely Variant.

Internal context was also used: arable land, surrounded by dominant urban cover, was corrected to the more likely **Suburban** Subclass; this rule was applied widely but had no relevance in semi-natural landscapes. Bare ground in the context of coniferous plantations, evidently felled, was recorded as such; this rule applied where felled plantations were commonplace. External context masks were also used to correct classes similarly 'out of context'. Coastal masking was used to exclude inland 'littoral' habitats and shoreline 'terrestrial' types (Groom & Fuller 1996); this rule was applied to any scenes with a coastline. The DTM helped to identify erroneous areas of lowland *Fen, marsh and swamp* above the floodplain or 'urban' areas, really bare ground, in high upland; such rules were only applied as necessary and then tailored to specific circumstances. An altitude of 600 m (see Ratcliffe & Thompson, 1988) was set for the *Montane habitats* class and applied to vegetation above this threshold; this rule was used consistently. The CORINE data were used to identify land classified as 'scrub' and 'grassland' which was in reality most likely to be orchard) i.e. the *Arable and horticulture* BH (with 'orchard' recorded as the class Variant); this rule only had real relevance in areas with a significant incidence of orchards. General manual recodes were also made where known inaccuracies could not be dealt with using external or internal context. The process was effective because the operator could define a region in which to apply the correction, use the attribute data to select the relevant polygons, then apply the correction. The procedure was a 'one-off', tailored to local circumstances: however, correction of arable fields mis-classified as urban / suburban land was a commonplace form of manual recoding.

KBC rules operated sequentially. As a consequence, a polygon might have changed class more than once: *Littoral sediment* in a terrestrial context might first have changed to *Built up*; then, because it was recorded at 400 m altitude, it was deemed more likely to be *Inland rock*. All contextual corrections were recorded in the GIS database, including the input and output classes of any changed polygons, and the rule(s) used in their alteration. Multiple corrections were recorded in sequence.

Table 3. GIS attributes available for classification and post-classification analyses

Unique identifier for every segment

Modal aspect

Mean elevation

Modal slope

Codes for ancillary information e.g. terrestrial or maritime

A priori land cover type (subclass level) used for training

Information on roll over and cropping of segment

Identifying code for segments used for training

Number of pixels in the segment

Number of pixels in the shrunken segment at classification

Percentage shrinkage achieved during classification

Top class returned by MLC

Probability value of top class returned by maximum likelihood classification

Top 5 classes and probabilities returned by maximum likelihood classification

Per-pixel maximum likelihood classification, top 5 classes and class-fractions in segment

Class returned by stage-1 KBC

Record of stage-1 KBC operations

Final land cover type prior to stage-2 (final) KBC

Record of stage-2 (final) KBC operations

Widespread Broad Habitat code (e.g. 2.3.4)

Hierarchical class code (WBHclass.subclass.variant e.g. Gss)

Generalised description of operations used to classify the segment

Top 5 classes and fractions of LCMGB 1990 within segment

Majority class from Land Cover Map 1990 within segment

Majority class from Corine Land Cover 1990 in LCM2000 segment

Table 4. Broad Habitats and their relation to LCM2000 display classes and approximate colours

BH	LCM2000 display classes
22. Inshore sublittoral	Sea / Estuary
13. Standing water/canals	Water (inland)
20. Littoral rock	Littoral rock and bare sediment
21. Littoral sediment	
	Saltmarsh
18. Supra-littoral rock	Supra-littoral rock and sediment
19. Supra-littoral sediment	
12. Bogs	Bogs
	Dense dwarf shrub heath
10. Dwarf shrub heath	Open dwarf shrub heath
15. Montane habitats	Montane habitats
1. Broad-leaved woodland	Broad-leaved / mixed woodland
2. Coniferous woodland	Coniferous woodland
4. Arable & horticultural	Arable and horticultural land
5. Improved grassland	Improved grassland
	Neutral grass (grazed / rough)
6. Neutral	
7. Calcareous	Calcareous grass (grazed / rough)
8. Acid	Acid grass (grazed / rough)
9. Bracken	Bracken
11. Fen, marsh and swamp	Fen, marsh and swamp
17. Built up areas, gardens	Suburban / rural developed
	Urban and industrial areas
16. Inland rock	Inland rock and bare ground
20 relevant BHs	21 display classes

03/01/2002

