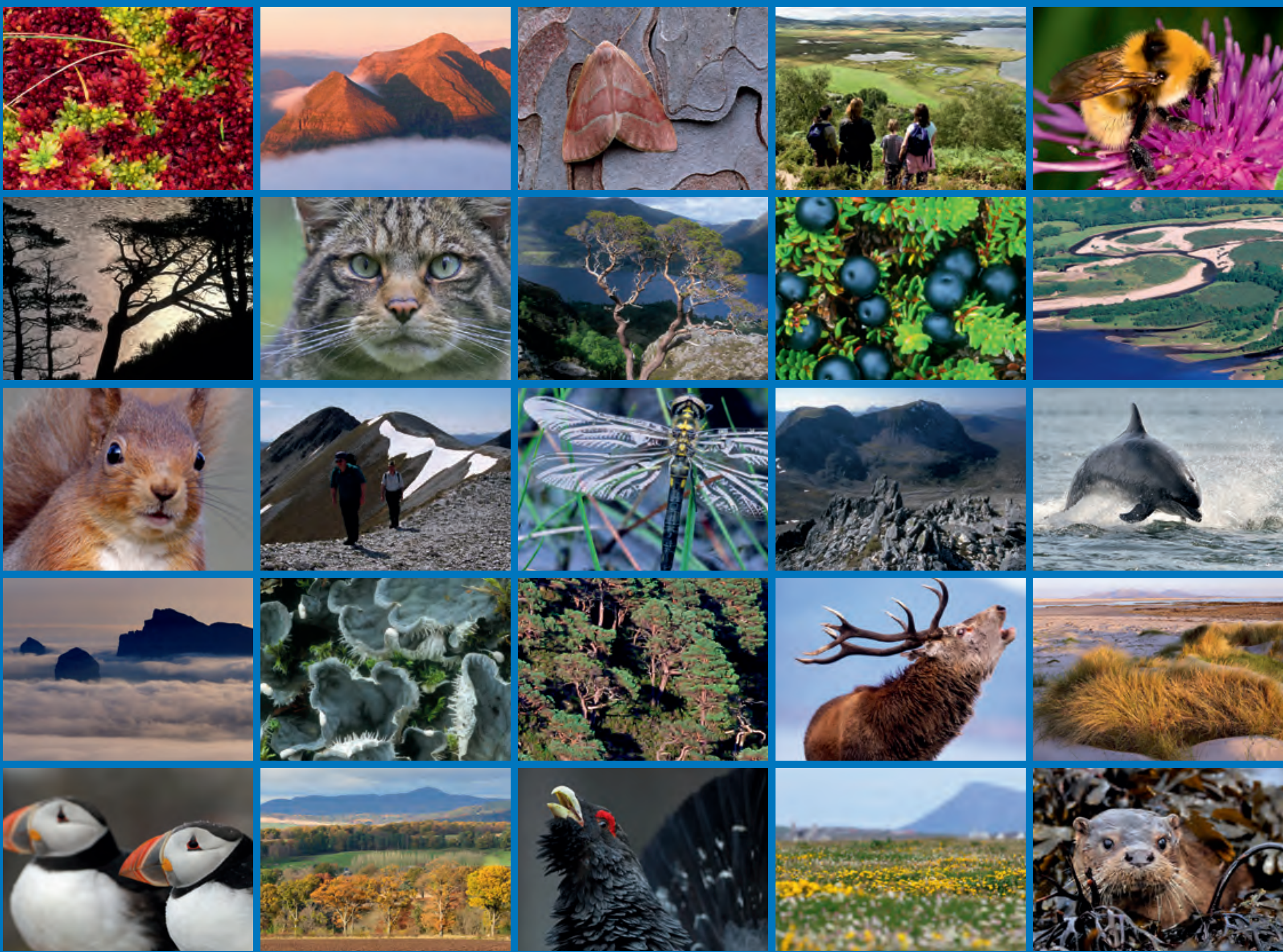
Scottish Natural Heritage
Commissioned Report No. 907

Developing a methodology for locating Annex I lowland terrestrial habitat



COMMISSIONED REPORT

Commissioned Report No. 907

Developing a methodology for locating Annex I lowland terrestrial habitat

For further information on this report please contact:

Jeanette Hall
Scottish Natural Heritage
Great Glen House
Leachkin Road
INVERNESS
IV3 8NW
Telephone: 01463 725204
E-mail: jeanette.hall@snh.gov.uk

This report should be quoted as:

Dadds, N.J. 2016. Developing a methodology for locating Annex I lowland terrestrial habitat. *Scottish Natural Heritage Commissioned Report No. 907.*

COMMISSIONED REPORT

Summary

Developing a methodology for locating Annex I lowland terrestrial habitat

Commissioned Report No. 907
Project No: 13603
Contractor: URS Infrastructure & Environment UK Ltd.
Year of publication: 2016

Keywords

Grassland; wetland; Annex 1; BAP; GIS; indicator species.

Background

This project involved the use of GIS to attempt to locate patches of Annex I lowland grassland and wetland habitats (as defined in the EC Habitats Directive) in Scotland, in two trial areas: East Dunbartonshire and Scottish Borders. Relevant UK BAP Priority habitats were also included. It was commissioned by SNH primarily as a means to expand their knowledge of Annex I habitats which they have an obligation to monitor and report under Article 17 of the Habitats Directive. The project utilised a combination of habitat data of various types and indicator species data, and assessed the usefulness of the different data sources.

Main findings

- Assessments were made of habitat data derived from aerial/satellite imagery analysis (Tweed Catchment Phase 1 and Potential Species-rich Grassland dataset) by comparison with localised detailed surveys (mainly NVC) and aerial photography. The Potential Species-rich Grassland dataset appeared to exaggerate the amount of wet habitat at the expense of dry habitat, did not contain all potentially species-rich grassland, and included some improved grassland. The Tweed Catchment Phase 1 performed better with regard to grasslands and categorisation of wet habitats, but was also prone to misclassify some habitats (most frequently grasslands) and to miss smaller patches of priority habitat. The minimum mappable unit of habitat data derived from imagery analysis means that some small priority habitat patches (e.g. grassland patches, flushes, springs) will always be missed.
- A large proportion of East Dunbartonshire lacked any Phase 1 habitat data, so BAP Broad Habitats from the LCM 2007 dataset were used to fill the gaps. This was useful but it is not recommended that the LCM 2007 is as the only habitat dataset used for these purposes, because it is a national dataset and therefore relatively coarse (the minimum mappable unit is 0.5ha), and for many grassland areas there was no distinction between acid, neutral and calcareous, and where this distinction was made it was largely derived from geology.
- A comparison of NVC data with slope steepness for habitat coded as Phase 1 dry grassland found that a slope steepness of 20o or greater corresponded best to NVC

priority grassland without including too much improved grassland or too little priority habitat. This steep grassland factor typically identified core areas of priority habitat and not the entirety of them; note also that many areas of priority dry grassland were on shallower slopes, and that some supposed steep grassland areas may correspond to unmapped non-grassland habitat. Despite these issues, the steep grass factor was considered useful in highlighting where there is a higher likelihood of Annex I H6210/H6230, Lowland Calcareous Grassland, Lowland Dry Acid Grassland or Lowland Meadow.

- Indicator species data were used to derive species scores for each priority habitat. 230 indicator species were selected and first scored according their efficacy as indicators. The data were then processed to produce a raster for each priority habitat where grid squares with indicator species held a score for the habitat dependent on the number and efficacy of indicators.
- Only indicator species records with spatial resolutions of 100m or better were used, and all such records were aggregated during raster processing to produce consistent 100m grid squares aligned to OS grid. Use of 1km-accuracy data caused an excessive reduction in clarity of potential priority habitat areas, and tetrad/hectad records were excluded from the start because this effect would be worse for such accuracies.
- There were sometimes similar indicator species scores for different priority habitats, often resulting from the fact the many species-rich indicators occur in more than one priority habitat. To help combat this issue, a short list of key species was devised, comprising species which were particularly likely to indicate a priority habitat.
- Following merging of the habitat and species data in polygon format, scores for each priority habitat were calculated based on the various contributing habitat factors, indicator species score and key species presence. The factors were weighted similarly but differently for each habitat depending on the considered importance of each factor, and to ensure that categorisation of potential as ‘Low’, ‘High’ etc. was useful.
- The process of scoring polygons for likelihood of priority habitats based on various contributing datasets means that a given polygon can and often does have potential for more than one priority habitat. In order to make an estimate of the actual areas of the priority habitats in reality, it will be necessary to undertake a field sampling survey to assess the effectiveness of the scoring procedure and to then make regional extrapolations.
- The final habitat potentials are reliant on the contributing data, and it is important to remember that areas of high potential will sometimes not contain priority habitat when field checked, and similarly some areas of high potential will not have been picked up. This is an inevitable result of the vagaries of the incorporated data, especially the known imperfections in the habitat data (in particular those derived from imagery analysis), the dependency of species data on recording effort and accuracy, and the fact that grid squares for species data will normally include irrelevant habitat or only part of a priority habitat.

For further information on this project contact:
Jeanette Hall, Scottish Natural Heritage, Great Glen House, Inverness, IV3 8NW.
Tel: 01463 725204 or jeanette.hall@snh.gov.uk
For further information on the SNH Research & Technical Support Programme contact:
Knowledge & Information Unit, Scottish Natural Heritage, Great Glen House, Inverness, IV3 8NW.
Tel: 01463 725000 or research@snh.gov.uk

Table of Contents **Page**

1. INTRODUCTION

This project involved the use of GIS to attempt to locate patches of Annex I lowland grassland and wetland habitats (as defined in the EC Habitats Directive[1]) in Scotland, in two trial areas: East Dunbartonshire and Scottish Borders (hereon referred to as Borders). Relevant UK BAP[2] (Biodiversity Action Plan) Priority habitats were also included. It was commissioned by SNH to expand their knowledge of these Annex I habitats for which they have an obligation under Article 17 of the Habitats Directive to maintain favourable conservation status and to report authoritatively on their status at six-yearly intervals, a task next due in 2018. The project will work towards an Annex I Habitat Map of Scotland.

Species-rich unimproved grasslands and wetlands in the lowlands typically survive as small, isolated patches within enclosed farmland. These tend to escape notice, unlike large upland sites. They may be known and monitored locally but there is no Scotland-wide system for locating, recording and monitoring non-designated patches of species-rich habitat. This lack of knowledge of small, valuable sites has led to significant loss of biodiversity in the past and there is a possibility of further loss of species-rich open habitats to meet woodland expansion targets. Remote sensing is not yet able to distinguish unimproved species-rich habitat reliably so a combination of data sources has been used, including habitat mapping, information derived from aerial and satellite imagery, and species records for selected indicator species.

The primary objectives of this project were to:

- develop a methodology for locating patches of Annex I lowland grasslands and wetlands within enclosed farmland landscapes
- produce a polygon dataset locating potential relevant habitats, with attributes including Annex I habitat type where possible and BAP Broad Habitat type as a minimum, and the sources of data for each polygon;
- assess the usefulness of each data source used, and provide recommendations for future work and likely costs.

2. RELEVANT ANNEX I/BAP PRIORITY HABITATS

In the UK, Annex I and BAP Priority lowland grassland habitats are identified primarily by their constituent National Vegetation Classification (NVC) communities. This is also the case for some but not all wetlands, the exception being bog habitats which can comprise identical constituent NVC types. The constituent NVC communities of Annex I and BAP Priority grassland and wetland habitats occurring in lowland Scotland are shown in Table 1 and Table 2 below. The tables also show which Phase 1 or BAP Broad Habitats we consider that the Annex I / BAP Priority habitats normally occur in, what the Phase 1/BAP Broad Habitat or NVC information is likely to indicate in terms of the Annex I/BAP Priority habitats, and comments on indicator species.

[1] EC Directive 92/43/EEC on the Conservation of Natural Habitats and of Wild Flora and Fauna
[2] The UK list of Priority Habitats is available at http://jncc.defra.gov.uk/default.aspx?page=5155

Table 1. Summary of Annex I and BAP Priority lowland grassland habitats occurring in Scotland

Annex I habitat	BAP Priority habitat	Constituent NVC communities	Appropriate Phase 1/ Broad Hab.	Indicated by Phase 1	Indicated by NVC	Comment on use of plant records
H6130 Calaminarian grasslands of the Violetalia calaminariae	*Calamin-arian grasslands (CALA)*	*OV37 and non-NVC types on metalliferous skeletal soil, river gravel and mine workings/spoil.*	*I1 (not I1.5), I2 (not I2.4), river shingle, possibly H8.1 (maritime cliff), Inland Rock*	*N/A – need NVC/ plant records*	*Annex I + BAP Priority habitat*	*Regardless of Phase 1/NVC, H6130 likely with* Thlaspi caerulescens*, increasingly likely with additional species from H6130 list.* Minuartia verna *cannot be relied upon alone since it is not rare in other habitats such as calcareous grassland.*
H6210 Semi-natural dry grasslands and scrubland facies: on calcareous substrates (Festuco-Brometalia)	*Lowland Calcareous Grassland*[3] *(LCG)*	*CG1*[4]*, CG2, CG7; also (rare) CG10 on limestone with a significant Mesobromion element*[5]	*B3, possibly H8.4 (coastal grassland) (in Scotland generally not 'rough low-productivity grassland')*	*B3 = BAP Priority habitat*	*Annex I + BAP Priority habitat*	*LCG likely if* Thymus *present, also* Alchemilla glaucescens, Botrychium lunaria, Galium sterneri, Gentianella amarella, Anacamptis pyramidalis, Dianthus deltoides, Viola hirta *and* Potentilla tabernaemontani *suggest LCG, the last four more likely in H6210 than H6230. LCG increasingly likely with other species from LCG list but many of these can also occur in e.g. Lowland Meadow and LDAG. Unlikely outside SE Scotland.*
H6230 Species-rich Nardus grassland on siliceous substrates in mountain areas[6]		*CG10 not on limestone, U4c, U5c (only these richer sub-communities of U4 and U5)*	*B3, possibly H8.4 (mostly not B1 but possible if U4c/U5c – need NVC)*	*B3 = BAP Priority habitat*	*Annex I + BAP Priority habitat*	*LCG likely if* Thymus *present. See above for other indicators more exclusive to H6210.*

[3] Annex I H6211 (important orchid sites on H6210-type habitat) is also a constituent of Lowland Calcareous Grassland but does not occur in Scotland.
[4] Only one CG1 site has been identified in Scotland, near Berwick-on-Tweed.
[5] CG3, CG4, CG5, CG6, CG8 and CG9 are also constituent NVC communities of Annex I H6210 but do not occur in Scotland.
[6] CG11 is also a constituent of Annex I H6230 but is an upland type unlikely to occur at lowland altitude except in the far north-west of Scotland.

Annex I habitat	BAP Priority habitat	Constituent NVC communities	Appropriate Phase 1/ Broad Hab.	Indicated by Phase 1	Indicated by NVC	Comment on use of plant records
(none)	*Lowland Dry Acid Grassland*[7] *(LDAG)*	*U1, U4 (except U4b which is semi-improved)*	*B1 (possibly H8.4 but require NVC), Acid Grassland, Rough low-productivity grassland*	*B1 = BAP Priority habitat (no valid Annex I)*		*Few species occur in LDAG and not also in LCG/Lowland Meadow.* Jasione montana *often occurs in LDAG, and if the habitat is not heath then* Polygala serpyllifolia, Ornithopus perpusillus *and* Hypericum humifusum *suggest LDAG.*
H6410 Molinia meadows on calcareous, peaty or clayey-silt-laden soils (Molinion caeruleae)	*Purple Moor Grass and Rush Pastures (PMRP)*	*M24, M26*[8]	*B5, E2, E3*	*N/A – need NVC/ plant records*	*Annex I + BAP Priority habitat*	*H6410 is possible if both* Molinia *and* Sanguisorba officinalis *or* Valeriana dioica *or* Cirsium dissectum *present, increasingly likely with more species from H6410 list, M26 particularly if* Trollius europaeus *also present. H6410 is rare in Scotland, and these species could coexist in a given area in different adjacent habitats.*
(none)		*M23, M25*	*B5, E2, E3*	*N/A – need NVC/ plant records*	*BAP Priority habitat (no valid Annex I)*	*PMRP normally has abundant* Molinia, Juncus acutiflorus *or* Juncus effusus, *but these occur in other habitats, and* J. effusus *rush pasture is normally poor;* Carum verticillatum *is likely to indicate PMRP, certain rare species also most likely in PMRP (*Cirsium dissectum, Epipactis palustris*). Other species that occur in PMRP can also occur in other mires/fens.*

[7] Annex I H2330 (inland dunes with open *Corynephorus* and *Agrostis* grasslands) is a constituent of Lowland Dry Acid Grassland but does not occur in Scotland. U3 and SD10b are also constituents of Lowland Dry Acid Grassland but are absent from Scotland or do not occur inland respectively.

[8] M22 is also a constituent of Purple Moor Grass and Rush Pastures but does not occur in Scotland.

H6520 Mountain hay meadows	*Upland Hay Meadows (UHM)*	*MG3*	*B2.1, Neutral Grassland*	*N/A – need NVC/ plant records*	*Annex I + BAP Priority habitat*	*If Phase 1 appropriate, H6520 moderately likely if* Geranium sylvaticum, Cirsium heterophyllum *or* Trollius europaeus *present, increasingly likely with more species from H6520 list.*
(none)	*Lowland meadows*[9] *(LM)*	*MG5, MG8*	*B2.1, Neutral grassland*	*N/A – need NVC/ plant records*	*BAP Priority habitat (no valid Annex I)*	*If Phase 1 appropriate, LM moderately likely if* Saxifraga granulata, Platanthera chlorantha, Silaum silaus *or* Sanguisorba officinalis *present (suggest MG5), increasingly likely with more species from LM list. There are no MG8 key indicators that do not also occur in other priority habitats such as PMRP, but* Caltha palustris *must be present.*

[9] Annex I H6510 (lowland hay meadows with *Alopecurus pratensis* and *Sanguisorba officinalis*) (MG4) is also a constituent of Lowland Meadows but does not occur in Scotland.

Table 2. Summary of Annex I and BAP Priority lowland wetland habitats occurring in Scotland

Annex I habitat	BAP Priority habitat	Constituent NVC communities	Appropriate Phase 1	Indicated by Phase 1	Indicated by NVC	Comment on use of plant records
H7110 Active raised bogs	Lowland raised bog (LRB)	M1, M2, M3, M17, M18, M19	E1.6.2	Annex I + BAP Priority habitat	N/A – need evidence of raised bog + active status.	Limited use: species in blanket bog generally also occur in raised bog and sometimes other habitats such as fens. Best to identify bog type and status from Phase 1, inventories (e.g. Raised Bog Inventory) and SSSI citations.
H7120 Degraded raised bogs capable of natural regeneration		M3, M15, M16, M17, M18, M19, M20, M25, woodland/scrub/ bare peat on raised bog	E1.7, E1.8, woodland/ scrub/bare peat	N/A – need evidence of raised bog + degraded status		Limited use. If raised bog known and Andromeda/ Sphagnum magellanicum occur then unlikely to be all degraded; also Campylopus introflexus suggests degradation as it favours bare peat. Best to identify bog type and status from Phase 1, inventories (e.g. Raised Bog Inventory) and SSSI citations.
H7130 Blanket bogs	Blanket bog (BB)	M1, M2, M3, M15, M17, M18, M19, M20, M25	E1.6.1, E1.7, E1.8	Annex I + BAP Priority habitat	N/A – need evidence of blanket bog	Limited use: species in blanket bog generally also occur in raised bog and sometimes other habitats such as fens. Best to identify bog type and status from Phase 1, inventories (e.g. Raised Bog Inventory) and SSSI citations. To count as bog (degraded), M15/M25 must be on deep peat.
H7150 Depressions on peat substrates of the Rhynchosporion	Most likely in LRB / BB but others possible	M1, M2, M14, M15, M16, M17, M18, M21, M29, M30	D2, D6, E1.6.1/2, Bog, Dwarf shrub heath, possibly E3 /	N/A – this Annex I is best indicated by Rhynchospora records; the BAP Priority habitat varies		Critical – not possible to know whether correct vegetation type is present from Phase 1 or NVC (it is often at edge of M1 or M2 pools and drier bog/wet heath), but H7150 strongly indicated by

Annex I habitat	BAP Priority habitat	Constituent NVC communities	Appropriate Phase 1	Indicated by Phase 1	Indicated by NVC	Comment on use of plant records
			Fen marsh & swamp			Rhynchospora *spp. presence.*
H7140 Transition mires and quaking bogs	*Lowland fen (LF)*	*M4, M5, M8, M9b*[10]*, S27*[11]*, occasionally bog NVC types particularly M18*	*E2, E3, F1, F2, Fen marsh & swamp*	*BAP Priority habitat*	*Annex I + BAP Priority habitat*	*In the Scottish lowlands,* Sphagnum warnstorfii, Calliergon giganteum *and* Calliergon cordifolium *are most likely to occur in H7140 assuming (for the last two) the habitat is not woodland. Also, there is normally always one of the key sedges present in H7140. If M9 is indicated without sub-community, lack of basicolous spp. suggests H7140 but otherwise H7230.*
H7220 Petrifying springs with tufa formation (Cratoneurion)		*M37*[12]	*E2.3, possibly within E2/E3, Fen marsh & swamp*	*BAP Priority habitat*	*Annex I + BAP Priority habitat*	Palustriella / Cratoneuron *is normally present but these mosses do occur in other habitats including on wet rock.*
H7230 Alkaline fens		*M9a*[13]*, M10*[14]	*E2, E3, F1, F2*	*BAP Priority habitat*	*Annex I + BAP Priority habitat*	*There are several species that are most likely to occur in H7230 (in the Scottish lowlands):* Carex dioica, Cinclidium stygium, Eleocharis quinqueflora, Eriophorum latifolium, Hamatocaulis vernicosus, Scorpidium scorpioides*, and* Sphagnum contortum /

[10] M9a contains basicolous species and is therefore referable to Annex I H7230.
[11] M2 and M29 can also occur in Annex I H7140, but are likely to be small components contained within or adjacent to the other listed constituent NVC communities, and (particularly M2) are common in non-H7140 vegetation. M14 is also a constituent of H7140 but does not occur in Scotland.
[12] M38 is also a constituent of Annex I 7220 but is an upland type not likely to occur at lowland altitude.
[13] M9b contains few or no basicolous species and is therefore referable to Annex I H7140.
[14] M13 is also a constituent of Annex I 7230 but is of very rare occurrence in Scotland.

Annex I habitat	BAP Priority habitat	Constituent NVC communities	Appropriate Phase 1	Indicated by Phase 1	Indicated by NVC	Comment on use of plant records
						platyphyllum / subsecundum. *If M9 is indicated without sub-community, basicolous spp. suggest H7230 but otherwise H7140.*
(none)		*In lowland Scotland: M6, M27-30, M32, swamps (except Reedbeds)*	*E2, E3, F1, F2*	*BAP Priority habitat (no valid Annex I)*		*The above key indicators for H7140 and H7230 are also appropriate for Lowland Fen generally. Another (which can occur in H7220/H7230 and non-Annex I springs/rills) is* Sedum villosum. *If* Phragmites *is present then may be Reedbed BAP Priority habitat, but as that is not Annex I and can only be confirmed if* Phragmites *is dominant, no attempt was made to segregate fen with* Phragmites *except by excluding S4/S26 from Lowland Fen.*
(none)	*Reedbeds (RDBD)*	*S4, S26, and occasionally S27 where dominated by Phragmites*[15]	*E3, F1, F2*	*N/A – need NVC or know-ledge of Reedbed*	*BAP Priority habitat likely*	*This habitat was not distinguished for this project (see above). Must know whether* Phragmites *is dominant to identify Reedbed (if not dominant then Lowland Fen likely). NVC S4/S26 are normally dominated by* Phragmites *so indicate Reedbed.*

[15] The other NVC types where *Phragmites* is either sometimes dominant (S24) or often so (S25) are not likely to occur in Scotland.

Recorded occurrences of Calaminarian grasslands in Scotland are mostly in the uplands, except for those in Shetland which are designated features of Sites of Special Scientific Interest (SSSIs). However, since it is possible that some Calaminarian grassland may have gone unrecorded in the trial areas, it has been included in this study. M24 (*Molinia caerulea-Cirsium dissectum* fen-meadow), a component of Annex I H6410, has so far only been recorded from Taynish NNR in Scotland, but is included for the same reason.

Annex I H7210 (Calcareous fens with *Cladium mariscus* and species of the Caricion davallianae) is also a constituent of Lowland Fen but is not included in this study. The reason for this is that although several swamp and mire communities, and two sand dune communities, are possible constituents of Annex I H7210, and some of these occur in Scotland, for inclusion in H7210 they must contain a significant amount of basicolous species similar to those of M10. This floristic composition is very highly localised in the UK and does not occur in Scotland[16], where stands of *Cladium* are acidic or neutral without the required basicolous elements. *C. mariscus* has been included as an indicator species because occurrences in Scotland may nevertheless correspond to Lowland Fen.

To determine whether CG10 was on limestone, for transfer to possible Annex I H6210 rather than H6230, reference was made to the freely-available 1:625,000 UK bedrock geology dataset. This dataset is rather coarse but we did not have access to better resolution geology data. The only part of the study area in which this geology dataset contained any type of limestone (the search also included metalimestone, dolomite, dolostone and any rock type described as a mixture with some form of limestone even if subordinate) was a coastal strip north of Berwick-on-Tweed, described as 'limestone with subordinate sandstone and argillaceous rock'. This strip contained no mapped CG10 (it does contain mapped CG2).

3. SUMMARY OF DATA SOURCES

The following data sources were utilised during this project.

Supplied by The Wildlife Information Centre (TWIC):

- Tweed Catchment Phase 1, which has a complete coverage of the Borders;
- Indicator vascular plant and bryophyte records for Borders (TWIC also supplied a small number of monophagous invertebrate indicator records but we were not able to utilise them – see below).

Supplied by Glasgow Museum Biological Records Centre:

- Indicator vascular plant and bryophyte records for East Dunbartonshire.

Supplied by East Dunbartonshire Council:

- Localised Phase 1 data for Sites of Importance for Nature Conservation (SINCs);
- Hand-drawn rough locations of other species-rich sites in East Dunbartonshire with limited indication of habitat type.

Supplied by SNH:

- Potential Species-rich Grassland dataset (Environment Systems Ltd. In press)
- SNH NVC dataset (very limited in East Dunbartonshire and localised but scattered throughout Borders, including many SSSIs and non-designated grassland sites);
- NVC data for non-designated sites from the 2010-2011 Lowland Grassland Review surveys (43 sites scattered throughout Borders) (Dadds & Averis, in press);
- SNH Phase 1 dataset (extremely limited coverage in Borders and sparse in East Dunbartonshire);

[16] http://jncc.defra.gov.uk/publications/JNCC312/habitat.asp?FeatureIntCode=H7210

- Land Cover Map (LCM) 2007 (used to give BAP Broad Habitats to the parts of East Dunbartonshire which lack better habitat information, and to replace localised Tweed Catchment Phase 1 habitat patches with only scattered habitat and no underlying main habitat with semi-natural BAP Broad Habitat);
- Raised Bog Inventory (Lindsay & Immirzi 1996)
- Ordnance Survey Digital Terrain Modelling (DTM) data (for the whole study area);
- Ordnance Survey Master Map (OSMM) data (for the whole study area);
- Aerial photography (for the whole study area, but only used selectively where required for double-checking survey data);
- 1:250,000 soils data.

Supplied by the British Bryological Society recorder for Borders:
- Indicator bryophyte records for Borders.

Supplied by Botanical Society of the British Isles recorders:
- Notes on wetland sites not included in the Tweed Wetland Strategy for the vice-counties of Roxburgh and Selkirk;
- A small number of selected indicator species records for the part of vice-county Midlothian that is within the current Scottish Borders local authority area.

Downloaded from the National Biodiversity Network (NBN) Gateway (https://data.nbn.org.uk/) on the recommendation of the British Bryological Society recorder for East Dunbartonshire:
- Indicator bryophyte records for East Dunbartonshire.

Downloaded from the British Geological Society (BGS) (http://www.bgs.ac.uk/downloads/browse.cfm?sec=6&cat=11):
- 1:625,000 bedrock geology.

Other on-line resources consulted for site information:
- SNH SiteLink (SSSI notified features and citations) (http://gateway.snh.gov.uk/sitelink/);
- Tweed Wetland Strategy (Magee *et al* 2010);
- 'A Botanical Tour of Berwickshire', a report by the BSBI recorder for Berwickshire obtained from the BSBI website (Braithwaite 2013).

Other resources were consulted not for datasets but for information on the priority habitats and potential of various species as indicators of those habitats, the principle ones being:
- JNCC (Annex I habitat definitions and distributions, UK BAP Priority Habitat definitions, and other information) (http://jncc.defra.gov.uk/default.aspx?page=2);
- Published NVC volumes (Rodwell 1991; 1992; 1995; 2000)
- NVC review (Rodwell *et al* 2000)
- An Illustrated Guide to British Upland Vegetation (Averis *et al* 2004)
- Plants and Habitats (Averis 2013)
- New Atlas of the British and Irish Flora (Preston *et al* 2002)
- Mosses and Liverworts of Britain and Ireland (Atherton *et al* 2010)
- BSBI Axiophyte data (http://bsbi.org.uk/axiophytes.html).

Plant nomenclature follows Stace (1997) for vascular plants.

4. ASSESSMENT OF HABITAT DATA

4.1 Description of habitat datasets

4.1.1 NVC data

Available NVC data came from two sources: an SNH general NVC dataset including the results of numerous NVC surveys scattered across the survey area, and NVC data from the SNH Lowland Grassland Review (Dadds & Averis, 2014) which included sites in Borders but not East Dunbartonshire. Attributes for the latter project, having been recently undertaken for SNH, are set out in a uniform standardised manner in accordance with SNH guidelines, and cover more than 40 grassland sites in Borders. NVC descriptions within the SNH NVC data are not uniform, having been undertaken at different times in the past and to differing standards, and there are also overlapping surveys which required additional work to resolve (see below). Some of the earlier NVC surveys incorporated into the SNH NVC data are rather crudely mapped compared to more recent surveys. Many polygons in both datasets are mosaics of NVC communities, and the attribute formatting of mosaics was inconsistent in the SNH NVC data (although it has since been standardised). There is very little NVC data in East Dunbartonshire, and as a whole the NVC data occupy a very small fraction of the study area, though many important sites are covered.

4.1.2 Phase 1 data

Phase 1 habitat mapping in East Dunbartonshire is rather sparse, with widely scattered patches of mapping originating from either East Dunbartonshire Council or SNH. Phase 1 information from the former existed as separate GIS data files for Sites of Interest for Nature Conservation (SINCs), and additionally as hand-drawn rough maps with limited habitat information for a number of other sites. Phase 1 data from SNH existed as a single GIS data file containing data for a small number of sites. Both the East Dunbartonshire Council and SNH Phase 1 data are presumed to have resulted from various past field surveys, rather than from aerial imagery analysis, which is advantageous since small habitat patches can be picked up during field surveys, and habitats which are difficult to separate from aerial imagery (such as semi- and un-improved grasslands) can be reliably identified. For the large parts of East Dunbartonshire for which there is no Phase 1 habitat information, the Land Cover Map (LCM) 2007 data (discussed further below) can be used to provide estimated BAP Broad Habitat type.

For Borders, Phase 1 habitat mapping is available for the entire local authority area. This Phase 1 mapping originates from the Tweed Catchment Phase 1 (for the associated report see Medcalf 2010), which was in large part the result of an analysis by Environment Systems of aerial photographs taken in 2007 and 2009. Important points about the Tweed Catchment Phase 1 habitat data are that:

i) The minimum mappable unit (MMU) was 0.25ha ($2500m^2$), so patches smaller than this would have been subsumed into the surrounding habitat(s) during the imagery analysis (though it is assumed that smaller patches from existing field-based surveys were retained);
ii) Assignation of improved grassland was carried out automatically, along with arable land and coniferous plantation;
iii) Assignation of other habitats was carried out with the aid of a small team of ecologists who manually inspected the aerial photography.
iv) The seasonal timing of the aerial photography and availability of RGB data only (i.e. not including infra-red data) caused difficulties in some locations and for some habitats (e.g. bracken).
v) The results of field-based surveys to which the contractor had access were incorporated into the dataset (which included a number of NVC site surveys).

The first two points in the above list are of particular importance for non-improved grasslands (which have the potential to contain Annex I grassland habitats) because they can occur in patches smaller in area than 0.25ha, and may, if grazed, appear very similar to improved grassland in aerial imagery particularly if appropriate seasonal imagery is not available. The first point is also relevant to flushes (having the potential to contain Annex I wetland habitats) which are often also smaller in area than 0.25ha.

In addition to this Phase 1 data originating from the Tweed Catchment Phase 1, a very limited amount of additional Phase 1 data was available from SNH, comprising detailed mapping (not without errors, however – see below) for a single large SSSI.

4.1.3 LCM 2007 data

The LCM 2007 dataset is a national one (for the associated report see Morton *et al* 2011). This was also derived from analysis of satellite imagery, and has a larger MMU of 0.5ha. It maps the entire country in accordance with likely UK BAP Broad Habitat. The LCM methodology included use of sample field visits across the country to improve habitat assignation accuracy, but since the analysis was undertaken on a national scale, and could not include such extensive manual inspection of aerial photographs and incorporation of local field surveys as was undertaken for the Tweed Catchment Phase 1, and has a larger MMU, it is to be expected that the LCM data should be less accurate at large scales than Phase 1 data. An additional issue is that distinction of grassland as acid, neutral or calcareous was based largely on geology, and in the study area grassland is frequently coded as ‘Rough low-productivity grassland’ where this distinction could not be made.

The LCM 2007 data was primarily used for East Dunbartonshire, to fill the extensive habitat gaps between patches of Phase 1 habitat data with estimated BAP Broad Habitat type. It was also used to replace polygons of scattered habitat with no underlying dominant habitat in the Tweed Catchment Phase 1, where a semi-natural BAP Broad Habitat type was present in the LCM 2007 data.

4.1.4 Potential Species-rich Grassland dataset

In addition to the Phase 1 and LCM 2007 datasets described above, there also exists a Potential Species-rich Grassland dataset recently produced for SNH (Environment Systems 2013). This was also derived from various aerial imagery sources including infrared data, and also made use of other data including soil type, slope and supporting evidence of habitat quality. This work was a pilot exercise using the areas of Borders and East Dunbartonshire. The techniques used were based on and build upon those set out in recent literature (e.g. Medcalf *et al* 2011). Areas identified as Potential Species-rich Grassland (i.e. not improved grassland, and not urban, water, heath, woodland etc.) were classed as wet or dry, and a level of likelihood given from Class 1 to Class 3, based on heterogeneity (i.e. unevenness, more being better), productivity (less being better) and whether there is corroborating evidence (such as knowledge of good habitat from SSSI citations or other sources). The wet/dry grassland areas thus delimited are not identified with certainty, but are assigned a level of probability that wet or dry species-rich grassland is present (Class 1 being most likely), since field checking would be required for confirmation of habitat type (except where already known).

4.2 Reliability assessment of habitat data derived from imagery analysis

4.2.1 Tweed Catchment Phase 1 and LCM 2007

In order to explore the accuracy of the Tweed Catchment Phase 1 and LCM 2007 datasets, a comparison was made with NVC surveys recently undertaken for the SNH Lowland Grasslands Review, in which Annex I/BAP Priority grasslands and mires dominated by grasses/rushes were identified to NVC community. This data was used for the comparison because: i) it included 43 sites scattered across Borders (though none in East Dunbartonshire), including some wetland communities as well as dry grasslands (the BAP/Annex I constituent NVC communities comprised CG2, CG7, CG10, MG5, U1, U4, M23 and M25, and one CG1 stand near Berwick-on-Tweed); ii) a high level of certainty can be assumed for these 43 sites because all but two were surveyed by well-known very skilled and experienced NVC surveyors whose expertise is acknowledged by SNH and others; and iii) this data was not available for incorporation into the Tweed Catchment Phase 1 at the time it was produced. Note that only polygons with constituent NVC communities (i.e. representing Annex I/BAP Priority grassland) were mapped and digitised, but non-constituent communities also appeared in mosaic polygons and were sometimes dominant. This does conveniently mean that every NVC polygon in that dataset contains some amount of Annex I/BAP Priority dry or wet grassland.

To compare these NVC surveys with the Tweed Catchment Phase 1 and LCM 2007 datasets, each polygon was initially assigned a corresponding Phase 1 habitat and BAP Broad Habitat based on the dominant NVC community. An intersection of all three datasets was then made, and attributes added to record whether the Phase 1/BAP Broad Habitats assigned to the NVC matched those given by the Tweed Catchment Phase 1/LCM 2007. Assignment of Phase 1/BAP Broad Habitats to NVC in this way is rather unsatisfactory because more than one Phase 1/BAP Broad Habitat may be 'correct' for a given NVC community, and NVC polygons are often mosaics, so the following manual alterations were made to ameliorate these issues: i) improved/poor semi-improved grassland were accepted for NVC MG6, since it can retain limited acid or basicolous floristics; ii) improved grassland was accepted for improved/acid grassland mosaics if the former was dominant; iii) rough low-productivity grassland was accepted as a possible correct BAP Broad Habitat for acid grassland/coarse neutral grassland NVC communities; iv) if Phase 1/BAP Broad Habitat grassland other than calcareous grassland actually contained calcareous grassland NVC communities at 1% or greater the former was not accepted (so as not to hide unmapped calcareous grassland in the former); v) calcareous grassland was accepted for NVC polygons dominated by rock if a calcareous grassland NVC community was also present, but otherwise not; vi) Phase 1/BAP Broad Habitat acid grassland was accepted for NVC mosaics with bracken unless the amount of NVC bracken (U20) was greater than 10%. The data could be analysed in a more complex and accurate manner but this was not feasible in the time available.

This comparison produced correspondences in terms of areas of habitats (it is irrelevant to consider numbers of polygons, because an intersection of three habitat layers commonly results in a large number of small polygons). The Tweed Catchment Phase 1 dataset was confirmed as more accurate in general than the LCM 2007 dataset, which was not unexpected. The overall correspondence accuracy of the Tweed Catchment Phase 1 dataset at NVC scale, for the NVC data compared here and using the above method, was 38%, and the LCM 2007 correspondence was 21%. These figures should not however be taken as indicative of accuracy at scales more appropriate to these data: it is generally unfair to expect high accuracy for Phase 1/LCM mapping at NVC scale, and contentious to apply single Phase 1/BAP Broad Habitat categories to some NVC mosaics. Additionally, the analysis involved known sites with some grassland interest, and accuracy of the Phase 1/LCM 2007 data would be expected to be higher more generally where interesting habitats are more sparse or absent.

Despite these issues, the comparison was useful for this project in showing that many patches of Annex I/BAP Priority grassland habitat (often but not always small) have been missed in the Tweed Catchment Phase 1, through misclassification or merging with surrounding poorer grasslands, especially where grazed. Further comments regarding the

Tweed Catchment Phase 1 are given in the following section alongside comments on the Potential Species-rich Grassland dataset.

Specifically for acid grassland, the LCM 2007 data achieved a somewhat higher correspondence with the NVC than the Tweed Catchment Phase 1 in the above exercise. This suggested that the LCM data could be used to extract acid grassland habitat in those situations where the LCM data indicated acid grassland and the Phase 1 indicated improved/poor semi-improved grassland. However, since the Tweed Catchment Phase 1 showed a much higher correspondence for improved grassland[17], and inspection of aerial photography showed that the LCM data sometimes indicated acid grassland where it was most likely improved, the idea of using the LCM 2007 data for this purpose was abandoned.

4.2.2 Potential Species-rich Grassland data, with Tweed Catchment Phase 1 comments

In order to gain some understanding of the effectiveness of the techniques used to develop the Potential Species-rich grassland dataset and its usefulness for the current project, we again used the NVC data gathered during the SNH Lowland Grasslands Review (see above for brief description), together with normal RGB aerial photography supplied by SNH and aerial photography available online, to make some comparisons. Since this comparison employs NVC data believed to be accurate (for the reasons given above), it represents a basic field test of the Potential Species-rich Grassland data.

Owing to time constraints, we have not undertaken a rough mathematical assessment as above for the Phase 1/LCM 2007 data[18], but make some observations based on visual comparison alone. These are illustrated in the example area shown in Figure 1 below, in combination with the notes beneath, which also incorporate further comments on the Tweed Catchment Phase 1 (the Phase 1 habitats are not shown, however, to avoid illegibility). This particular area was chosen as the example for comments because it encompasses a clear range of habitat types whilst also overlapping four SNH Lowland Grassland Review sites. Several of these observations are based on firm knowledge from the Lowland Grassland Review, but where this is not the case they are based instead on a combination of knowledge of nearby grassland surveys and extensive familiarity with using aerial photography in the field for Phase 1 and NVC surveys.

- *Note 1:* The part of the surveyed grassland site shown here is known to be dominated by extensive M23 which is a Potential Species-rich Grassland wet grassland habitat, but is not indicated as such in the Potential Species-rich grassland data. The Phase 1 is incorrect in mostly recording semi-improved acid grassland, and in the southern corner of the site records acid/neutral flush where there is in fact species-poor grazed MG10 not referable to flush vegetation.
- *Note 2:* Vegetation most likely also dominated by rushes is visible here, connected to the vegetation highlighted in Note 1, and has been successfully included in the Potential Species-rich grassland 'wet' category. The Phase 1 has unimproved acid grassland here, which would appear to be at least partly incorrect.
- *Note 3:* This surveyed grassland site is mostly dry with several rocky ridges with species-rich dry grassland containing species such as field gentian *Gentiana*

[17] 62% correspondence with NVC polygons dominated by MG6/MG7 where they do not also contain small amounts of calcareous grassland. In the latter case the Phase 1 was marked as incorrect to highlight unmapped calcareous grassland, so the percentage is higher because in those cases much of the polygon was still improved. The Tweed Catchment Phase 1 report (Medcalf & Williams, 2010) states an 85% accuracy (assumed to be an average) for automatic attribution of improved grassland, arable and coniferous woodland, and the exact figure for all improved grassland is probably closer to this given the high proportion of pure improved grassland outside the grassland sites.
[18] A similar area correspondence exercise would be useful in helping to determine more concretely the success of the Potential Species-rich Grassland dataset in locating known species-rich grassland.

campestris, but the field is erroneously included in the wet category of the Potential Species-rich Grassland data, and even the wider species-rich grassland strips (which are mostly on dry slopes steeper than 15°) are not included in the Potential Species-rich Grassland dry grassland data. The Phase 1 records improved grassland here which, whilst correct for the larger part of the site, misses even the wider species-rich dry grassland ridges and slopes (c.25m wide).

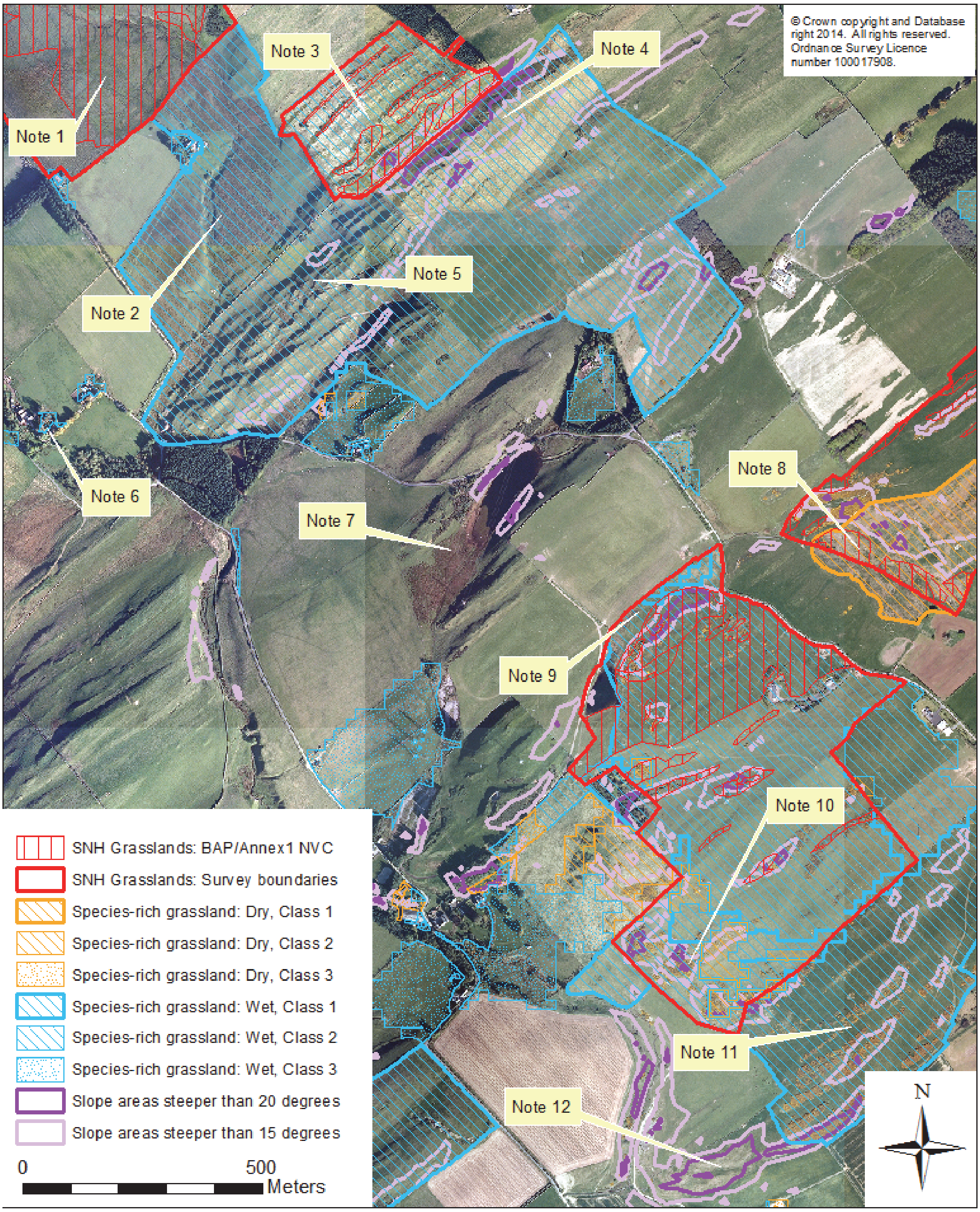


Figure 1. Notes on the reliability of the Species-rich and Tweed Catchment Phase 1 data (see the accompanying text of notes below).

- *Note 4:* The steep slope facing the grassland site to the north-west, which exceeds 20^{o}, is dry and was observed to be potentially species-rich like the slopes in the surveyed grassland site, but again is included in the Potential Species-rich Grassland wet grassland data (the shaded aspect may have made identification from aerial imagery difficult). The adjacent yellowish grassland is also probably dry and not wet, particularly on the indicated steeper slopes. The Phase 1 has this area as poor semi-improved grassland, which is possible for part of it but incorrect on the steep slope just mentioned, which is most likely unimproved U4 with possible CG10.
- *Note 5:* This area has rocky ridges which are apparently similar to but more pronounced than those in the grassland site to the north, and again are most probably dry and not wet, particularly on the indicated steeper slopes, and therefore inappropriately placed in the Potential Species-rich Grassland wet grassland data. The Phase 1 has this area as semi-improved acid grassland, which is probably broadly correct, though there could be unrecognised small patches of species-rich dry grassland as in the grassland site to the north.
- *Note 6:* Although very small, this patch in the Potential Species-rich Grassland wet grassland data appears to be a small patch of tall trees with some bare ground or hard-standing, judging by close inspection of aerial photography (not visible in the figure) and the shadows thrown on the adjacent field. The Phase 1 maps this patch as built-up, which may be partly true.
- *Note 7:* This large patch of vegetation is probably rushy and therefore potentially species-rich, judging by similarity of colour and texture to known rushy areas nearby and elsewhere, and the facts that the lower parts are adjacent to a stream draining a small loch and there is evidence on closer inspection (not visible in the figure) of old parallel drainage channels, but the area is not within the Potential Species-rich Grassland wet (or dry) grassland data. The area is mistakenly mapped as improved grassland in the Phase 1 despite the large size of the patch and contrast with nearby improved grassland.
- *Note 8:* This part of a surveyed grassland site is known to be dominated by wet M23, but is mistakenly included in the Potential Species-rich Grassland dry grassland data. The ground immediately to the north is steeply-sloping and probably correctly included (assuming there was no site knowledge) in the Potential Species-rich Grassland dry grassland data, although in this instance we know (because no NVC polygons exist here) that there is no BAP/Annex I interest, and heterogeneity in this area may be due to gorse scrub clearance which is known to have occurred locally. The Phase 1 surprisingly incorporates all this variation into improved grassland, despite the visual dissimilarity to adjacent improved grassland and localised steep slopes.
- *Notes 9:* The strip of the Potential Species-rich Grassland wet grassland data outside the adjacent NVC polygon is correctly identified, since it is known to be wet and drains to a small loch. However, the rest of the surveyed site is predominantly dry and inappropriately included in the wet grassland data. The majority of the site is MG6 (satisfactorily identified in the Phase 1 as poor semi-improved), but there are localised small patches of more Potential Species-rich Grassland dry grassland (not identified in the Phase 1). Note the correspondence here of steeper slopes to BAP/Annex I-containing NVC polygons, or (see next note) to slopes with gorse.
- *Note 10:* The steeper slopes in this part of the grassland site partly support gorse (correctly identified in the Phase 1) but like the majority of the site are erroneously included in the Potential Species-rich Grassland wet grassland data.
- *Note 11:* This field is also in the Potential Species-rich Grassland wet grassland data, but, like several other sloping fields with ridges in the area, it is most likely dry improved grassland with patches of gorse (both correctly identified in the Phase 1) and, as elsewhere nearby, there may be small species-rich dry patches (hypothetical and not in the Phase 1).

- *Note 12:* This area is mostly steep (over 20^{o}) and appears to be grassy, and therefore potentially species-rich, but is not included in the Potential Species-rich Grassland dataset. The Phase 1 maps this incorrectly as arable, which is highly unlikely even given the slope information alone. It is surmised that there may have been gorse clearance here as elsewhere nearby, which leads to heterogeneity and perhaps spectral similarity to certain states of arable land.

From the above comments, it is clear that the Potential Species-rich grassland dataset is prone to overestimate the amount of wet grassland at the expense of dry grassland. This apparent overestimation of wet grassland may in part be the result of incorporated soil data. Soil data supplied by SNH, originating from the James Hutton Institute (formerly Macauley Land Use Institute), is based on mapping at the rather crude scale of 1:250,000 derived from data gathered from the 1940s to the 1980s. This is likely to miss fine-scale soil variation which can be critically important for plant communities, and may not account for changes in drainage which may have occurred since the soil data was gathered. The Potential Species-rich grassland data does show some success in highlighting heterogenous areas and avoiding many improved/arable fields, although it sometimes fails to pick up non-improved grassland and on some other occasions includes it within supposed heterogenous areas. For these reasons we did not incorporate the wet/dry categorisation of the Potential Species-rich Grassland data into our analysis, and although we used the overall class, since it often indicates heterogeneity and likely non-improvement, we gave it a low weight in the final scoring (see below).

The comments in this section regarding the Tweed Catchment Phase 1 provide further evidence that it is sometimes unsuccessful in assigning correct habitat, though a proportion of the erroneous habitat attributions will be due to small habitat patch size below the minimum used during aerial imagery analysis (see above). It should also be noted that the area shown in Figure 1 and used for making this assessment is an upland fringe area, which is more complex than more low-lying areas and inherently more difficult for aerial imagery analysis. Despite these problems, as noted above the Tweed Catchment Phase 1 has a reasonably high success rate for attribution of improved grassland (as well as arable and forestry), and for the area shown in Figure 1 it was more successful at suggesting improved grassland than the Potential Species-rich grassland dataset.

4.3 SSSIs

All SSSIs in the study area with notified SSSI grassland features have NVC data, so will have all been accounted for. SSSIs with wetland features were covered through using SNH SiteLink and (for Borders) the Tweed Wetland Strategy, at the same time double-checking Phase 1 habitat attribution to make occasional alterations where deemed necessary (for example, a small area of Phase 1 dry heath where the SSSI citation described a small active raised bog).

4.4 Slope steepness as an additional factor for dry grasslands

Steeper ground makes agricultural improvement more difficult and eventually impossible, and can also lower grazing intensity by discouraging grazing access. Thus grasslands on steeper slopes are often richer than on surrounding flatter ground, and particularly in areas dominated by improved pastures they often represent important islands of more species-rich habitat. Since the probability of the occurrence of certain types of species-rich grassland is higher on steeper ground, and the Tweed Catchment Phase 1 data was found to have classed some species-rich grassland as improved by either misclassification or merging of small patches with surrounding habitat, a basic analysis of slope was undertaken to highlight those areas of grassland where agricultural improvement was less likely to have occurred and species-rich habitat was therefore possible, regardless of Phase 1 grassland classification.

This was done by generating angle of slope from Ordnance Survey Digital Terrain Modelling (DTM) data at 5m resolution, using ESRI Spatial Analyst, and comparing this with the 43 Borders grassland sites surveyed during the SNH lowland grasslands review. A cursory inspection by eye suggested that slopes steeper than around 17° corresponded best with NVC polygons containing BAP/Annex I grassland (called constituent NVC polygons), so analyses at different slope angles were undertaken starting at this point. This was done by calculating the percentage (by area and number) of slopes steeper than the given gradient that intersected with the constituent NVC community polygons, and also the percentage of whole grassland sites surveyed in the Borders that intersected these slopes. The results of this analysis are shown in the graph below.

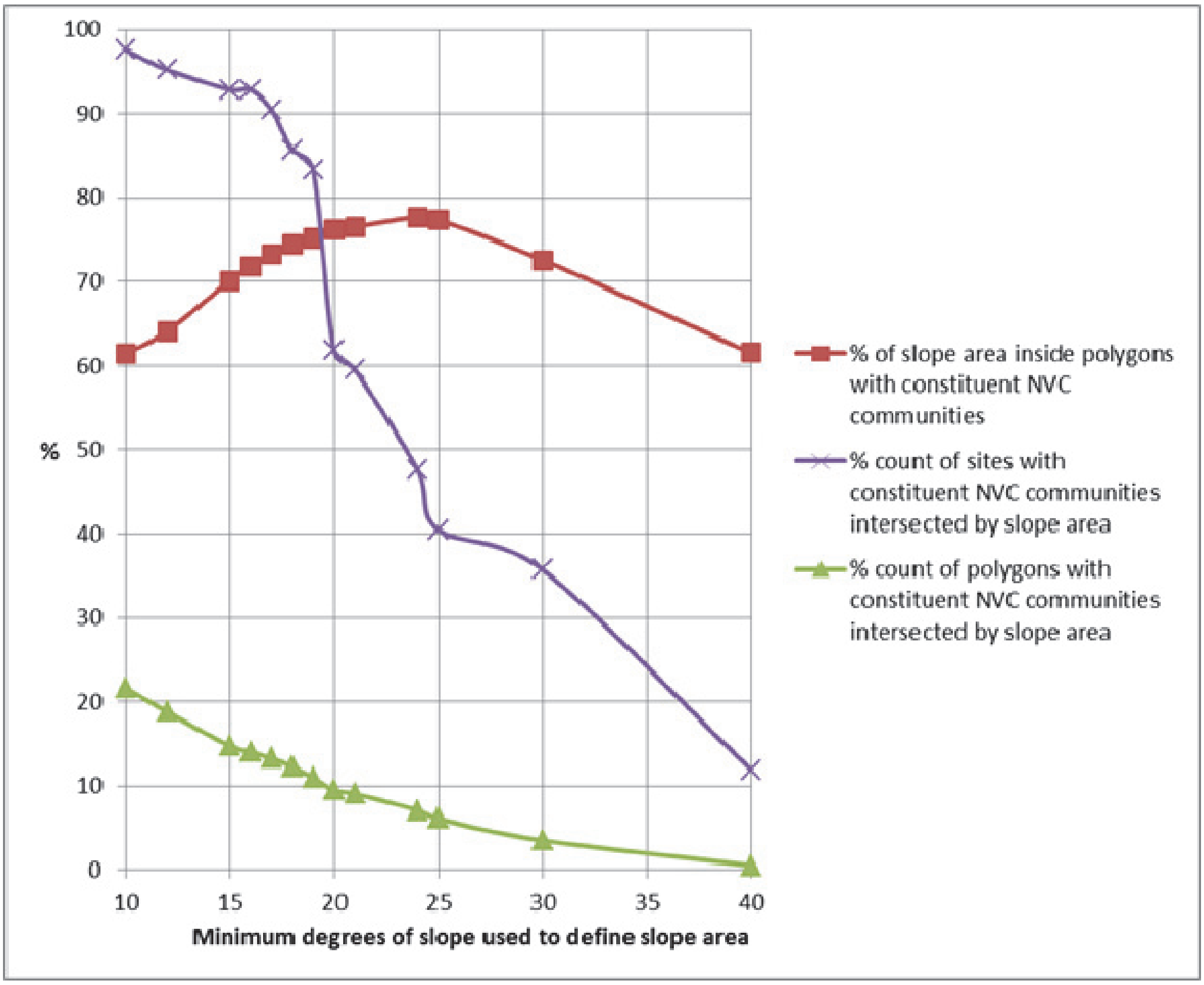


Figure 2. Slope steepness compared to grassland NVC data.

The percentage intersection by number of slope polygons against constituent NVC polygons is not very revealing, showing a steady and predictable drop from 10° to 40° minimum slopes (which naturally have increasingly smaller areas of defined slopes). The percentage intersection by area is more interesting, showing that the peak correspondence of defined slopes to constituent NVC polygons (with this data) is at approximately 24° minimum slope. On the other hand, the percentage count of intersecting whole sites drops drastically from 19° to 24° minimum slope, meaning that had a 24° minimum slope been used to attempt to find these sites amongst known grassland habitat, then over 40% less sites would have been identified compared to using a 19° minimum slope. However, the unevenness of the curve for percentage count of intersecting whole sites also suggests that a proportion of the sudden drop from 19^o to 20^o minimum slope may be an artefact of the data. A good compromise would appear to be 20^o minimum slope, for which there is only a slightly lower correspondence with constituent NVC polygons by area than at 24^o minimum slope, but for which the number of identified sites is much greater.

We therefore used slope areas of minimum 20^o steepness within any kind of Phase 1 dry grassland as an additional factor contributing to the scores for calcareous grasslands, Lowland Meadow and Lowland Acid Grassland. Although 20^o may not appear particularly steep, it is easier to visualise by considering that it is steeper than 1 in 3, a gradient few if any public roads approach. It is important to remember that these steep slope areas do not pick up all sites with notable grassland, and even when they are successful in this they tend to indicate 'core' areas, around which notable grassland can extend further, and often (for the Lowland Grassland Review sites inspected) priority grassland can also occur separately elsewhere in the vicinity on nearby gentler slopes. It is also possible that some steep grass areas may contain small unmapped non-grassland habitat patches such as scrub, trees or bracken. Thus field checking of areas with indicated steep grass should include a wider field search than just those areas marked as steep grass.

5. TREATMENT OF HABITAT DATA

5.1 Borders-specific data

5.1.1 Tweed Catchment Phase 1 data

As described above, there were a large number of Phase 1 polygons with scattered habitat (scattered trees, scrub or bracken) but no indication of underlying dominant habitat, for which we took the LCM 2007 BAP Broad Habitat if it was semi-natural (commonly acid grassland, rough low-productivity grassland or heath). The intersection of the Phase 1 with the LCM 2007 for these potential scattered habitat replacements resulted in many small and/or thin polygons because field boundaries etc. generally did not coincide, and such small/thin polygons were likely to be irrelevant. Having first eliminated the extremely small replacement polygons, we then calculated degree of thinness by using the following formula in the field calculator: *4 x 3.14 x shape area / (shape length x shape length)*; this formula yields 1 if the polygon area is contained within the smallest possible perimeter length (i.e. a circle), and increasingly small fractions of 1 for increasingly thin polygons. Since it was found that some of the remaining thin polygons were large in area and not insignificant, we multiplied the thinness factor by area to give a single guide to combined area/thinness. By manual inspection of the polygons it was apparent that a good cut-off for this area/thinness factor (with this data and for this purpose) was 100, and potential replacement polygons with values less than this were eliminated. 841 polygons of LCM 2007 BAP Broad Habitat remained and were used to replace Tweed Catchment Phase 1 scattered habitats. Though most of these scattered habitat polygons were small, some were very large, for example some areas of Phase 1 scattered bracken.

Further localised changes were made to the Tweed Catchment Phase 1 data to correct obviously erroneous habitat attributions. There were a small number of coastal but still obviously terrestrial polygons, and one large polygon slightly further inland, which were erroneously coded as 'intertidal'. A separate large extent of golf course was also incorrectly coded as 'hedge'. Errors of this magnitude would have arisen during data input or later processing. Most of the coastal polygons were maritime cliffs/slopes and were corrected accordingly; the nearby inland area was reclassed as neutral grassland with a small strip of coastal grassland, basing this on the adjacent Phase 1 categories. The open golf course 'hedge' area was clearly meant to be amenity grassland and was changed to this. There may have been other such errors but it was not feasible to verify the entire dataset.

5.1.2 SNH Phase 1 data for Borders

The SNH Phase 1 dataset contained only one significant area of Phase 1 information in Borders, concerning a large SSSI. The level of detail for this site was much greater than the Tweed Catchment Phase 1 and was therefore used to replace the Tweed Catchment Phase 1 data entirely for that site (although it should be noted that the Tweed Catchment Phase 1, though much less detailed and missing small flushes etc., was broadly correct in identifying the raised bog and extensive heath on an adjacent hill). Two corrections were necessary before this was undertaken. The first concerned a very large area of heavily muirburnt heather (very clear on aerial imagery and described as such in the SSSI citation, and correctly coded as heath in the Tweed Catchment Phase 1) that was incorrectly coded as neutral/marshy grassland. This error would also have arisen during data input/processing, since obvious muirburnt heather moorland is not confusable with neutral grassland. Polygons of neutral/marshy grassland on this site overlaying clear muirburnt heather were reclassified as dry dwarf shrub heath. The other more minor correction was the alteration of small patches of scattered bracken with no underlying habitat to the likely dominant habitat, which we assumed with the aid of aerial imagery to be acid grassland.

5.2 East Dunbartonshire-specific data

5.2.1 East Dunbartonshire digital Phase 1 data

The Phase 1 data from East Dunbartonshire Council (EDC) originated as separate files for each site, which were merged following a data validation exercise involving standardisation of habitat coding, merging of fields of similar data type but used differently for different sites, correction of a few polygon overlaps, and deletion of polygons with zero area (which also contained no data). Some additional Phase 1 information for East Dunbartonshire was provided by SNH. Where there were overlaps between the SNH and EDC Phase 1 data, the apparently more detailed source was accepted and the other rejected. Overlaps were few except on the lower edge of the Campsie Fells, where the SNH data appeared to be more detailed and were preferred; for overlaps elsewhere the EDC data was usually more detailed and therefore preferred.

5.2.2 LCM 2007 data

Since the combined SNH and EDC Phase 1 datasets still left the majority of East Dunbartonshire without indication of habitat, the obvious recourse was to utilise the LCM 2007 data to fill the gaps with BAP Broad Habitats, this being the next best alternative. However, the LCM 2007 data, designed as a national dataset, is necessarily simplified such that fields of identical adjacent BAP Broad Habitat are merged into single polygons, resulting (at least for East Dunbartonshire) in many very large polygons and a lack of interior OS compartments/features. These characteristics are desirable for the LCM 2007 itself for which a degree of simplification was necessary, but not for the purposes of this project where the rural small compartments and features in the OS data are potentially helpful in narrowing down potential habitat patches. The necessary simplification and reduction of vertices in the LCM 2007 data means that polygon boundaries often do not exactly coincide with OS lines, such that a simple intersection of the LCM 2007 and OS Master Map datasets to combine the habitats of the first with the detail of the second would result in an enormous number of thin artefact polygons. For these reasons we carried out a procedure to add the most extensive overlapping BAP Broad Habitat to OS Master Map polygons in rural areas (excluding roads/buildings etc.). Since OS Master Map polygons were infrequently properly split by BAP Broad Habitat data (at least in lowland East Dunbartonshire), we judged that the gain in small OS Master Map feature details and avoidance of excessive artefact slivers outweighed the loss of occasional BAP Broad Habitat divisions.

There is not a straightforward way to incorporate information into one polygon layer from the largest overlaps of another polygon layer: the spatial join tool selects a random overlapping polygon from a join layer when more than one overlaps a target layer polygon, with no option to select the largest overlap. A more cumbersome method was therefore used, but prior to this we eliminated from the OS Master Map data all buildings, all roads/paths, and the majority of building curtilages representing small urban/sub-urban gardens etc. The purpose of this was to remove likely irrelevant areas and to make the following procedure much less onerous in terms of computer processing (in this case reducing the number of OS Master Map polygons from about 150,000 to about 18,000). We also eliminated areas covered by existing East Dunbartonshire Phase 1 data. The procedure was to: i) intersect the OS Master Map and LCM 2007 data; ii) use the summary statistics tool to get the area of the largest resulting polygon within each OS Master Map polygon, by calculating the 'maximum' statistic on the intersection shape area for each OS Master Map object identifier (thus providing the area of the largest BAP Broad Habitat overlap for each OS Master Map polygon, ignoring small overlaps); iii) carry out an attribute join to join the area figure from the previous step to the polygons in the OS Master Map data, by using the OS Master Map object identifier as the linkage between the summary statistics table and the OS Master Map data; iv) carry out a second attribute join to join the dominant BAP Broad Habitat to the polygons in the OS Master Map data, by using the area figure joined in the previous step as the linkage between the correct polygon in the intersection shapefile (with the largest overlap, and having that same area) and the OS Master Map data. We found it was necessary to convert the shape areas to strings in another field in both the intersection shapefile and the results table of the summary statistics, in order to make the subsequent attribute joins work. The final result was a dataset of OS Master Map data with dominant LCM 2007 BAP Broad Habitat for all rural areas outside the existing East Dunbartonshire Phase 1 data.

5.2.3 East Dunbartonshire non-digital Phase 1 data

Having followed the procedure in the previous paragraph, we made further alterations to the BAP Broad Habitat layer by manually adding the hand-drawn SINC information from East Dunbartonshire Council, where possible changing to Phase 1 habitat instead of BAP Broad Habitat if the provided information was good enough. We also ensured that all detailed OS Master Map water features were returned to open water habitat, since the simplified nature of the LCM 2007 data tends to partially or completely eliminate many such features.

5.3 Final treatment of combined Phase 1 data

5.3.1 Merging and indication of potential priority habitats

When the above procedures were complete, the Phase 1/BAP Broad Habitat datasets for Borders and East Dunbartonshire were combined so that subsequent processing could be undertaken more efficiently on one file.

Attribute fields were then added for each Annex I/BAP Priority habitat. These were populated with values of '1' where polygons contained habitat deemed appropriate to the Annex I/BAP Priority habitat in question, and '0' if not appropriate (using the field calculator). The appropriate Phase 1/BAP Broad Habitats (those in which the priority habitats are likely to occur) are given in Table 1 above. This process was straightforward for all priority habitats except bog habitats, treatment of which is more complex and is described in the next section.

5.3.2 Classification of bog habitats

In many cases the habitats relevant to this project can be defined by NVC data, but this is not the case in particular for bogs for which topography and location are more important than the exact vegetation type. Since the key bog NVC types (M17, M18, M19 and M20) can occur in either raised or blanket bog, they cannot be used alone to separate them. Neither does their presence certainly imply bog habitat, since they can occur in fens (e.g. basin mires), particularly the wetter M18. Phase 1 habitat categories are much more useful (if correct) because intact bog vegetation is split into raised bog (E1.6.2) and blanket bog (E1.6.1). However, Phase 1 wet/dry modified bog (E1.7/E1.8/E4) also does not distinguish raised/blanket bog. Whilst the BAP Priority raised/blanket bog habitats and Annex I blanket bog (H7130) take no account of status, raised bogs are split for Annex I purposes into active and degraded (H7110 and H7120). Thus further information is required even beyond accurate Phase 1 habitat mapping where degraded bog is shown.

To help confirm the presence and status of raised bogs, we made use of the Raised Bog Inventory (Lindsay & Immirzi 1996). This was obtained as a scanned pdf, requiring slow manual entry of the data for the relevant regions into a spreadsheet prior to import into ArcGIS. To count as ‘degraded raised bog’, the site must be capable of natural regeneration such that management could be expected to lead to recovery of peat-forming vegetation within 30 years (European Commission 2007). Thus habitats such as woodland (including plantation), scrub (often birch), bare peat and possibly grassland can be included where this expectation is reasonable.

Accordingly, all degraded raised bog sites in the Raised Bog Inventory which now comprise (using the available habitat information) modified bog, woodland, scrub, bare peat or (rarely) marshy/unimproved grassland were labelled as H7120 and Lowland Raised Bog. For raised bog sites known to be active through either the Raised Bog Inventory and/or SSSI citations, areas of bog habitat were labelled as H7110 and Lowland Raised Bog, but peripheral scrub/woodland on such sites has not been included because it is not possible to tell without detailed site information how much of this is natural lagg woodland (included in H7110), how much is encroachment through drying (H7120) or how much is woodland/scrub not directly connected to the raised bog (neither). Occasionally, the Raised Bog Inventory indicated active status but the Phase 1 suggested otherwise; in these instances we assumed degraded status because the Phase 1 is much more recent. A small number of alterations were made to the Phase 1 based on SSSI citation/NVC information (for example, where there was clear evidence of dominant fen vegetation rather than bog). For one site (Adderstonlee Bog SSSI), the NVC data lacked bog communities despite the acknowledgment of raised bog in the citation; the Phase 1 was relied upon in this instance, although the area it defines as raised bog appears too large given that the citation states that the raised bog is a small part of the SSSI. In one other instance, a small area of Phase 1 raised bog did not correspond to an inventory location or SSSI; again, the Phase 1 was relied upon since we had no information to clearly rule out raised bog at that location.

Areas of Phase 1 intact blanket bog (E1.6.1), which expectedly occupy little area in the lowlands, were also manually checked to verify that blanket bog was not improbable at the given locations. Often such areas of ‘lowland’ bog were actually at the upland edge of the study area. For those areas of Phase 1 degraded bog (E1.7 and E1.8) that had not been identified in the previous steps as degraded raised bog, it was assumed that they were degraded blanket bog and labelled as H7130 / Blanket Bog.

Annex I H7150 is also a bog habitat, though it can also occur in wet heath. This habitat cannot be defined by NVC or Phase 1 habitat data, and requires knowledge of the presence of otherwise of *Rhynchospora* spp. This vegetation type typically occurs at the junction of bog pools and drier bog vegetation, and may occur in multiple NVC mire types and Phase 1 habitats.

5.3.3 Non-relevant Phase 1/BAP Broad Habitats

The finalised Phase 1/BAP Broad Habitat data for Borders and East Dunbartonshire naturally still contained many patches of habitats not relevant to the priority habitats in question, such as woodland, dense scrub/bracken, dry heath, arable and improved grassland. We purposefully did not remove these at this stage, however, because the reliability issues described above meant that small relevant habitat patches would exist within some irrelevant habitats, and some larger irrelevant habitat areas were likely to have been misclassified. Elimination of irrelevant Phase 1 habitats was undertaken after merging with all other data sources and only where no other scoring factors existed, so that polygons with irrelevant Phase 1 habitat but other data suggestive of priority habitat(s) (e.g. key species, steep grass, Potential Species-rich Grassland data) would retain the Phase 1 data.

5.4 Treatment of the NVC data

Initially, we considered treating the NVC data as the most reliable habitat data source, using it to completely replace Phase 1 data where NVC data existed. However, this was inappropriate for two reasons. The first, which would apply to any similar project, is that although many of the priority habitats can be defined by NVC data, this is not the case for bog habitats, whose dominant NVC communities can occur in raised bog, blanket bog and fen. Good Phase 1 information is much better at distinguishing these because it specifically categorises them as such, with the exception of modified (degraded) bog, for which further information from elsewhere is required to determine if the bog is raised or blanket. The second reason is that, in this instance, some of the SNH NVC data (not the recent 2010/2011 Lowland Grassland Review data) showed rather simple mapping with clear locational inaccuracy, to the degree that in at least one case (involving calcareous grassland and adjacent lowland fen habitat) the Phase 1 data was more helpful for the purposes of this project. In such cases boundaries derived from aerial imagery are likely to be more accurate, but it would be a great task to check locational accuracy of all NVC data by comparison against Phase 1 data and aerial imagery, and would in many cases require familiarity with the sites. For these reasons, areas of Phase 1 data overlapping NVC data were kept.

The SNH NVC dataset had other issues which required some time to alleviate. There were blank but significantly-sized polygons, which on comparison with the Phase 1 and aerial imagery were found to largely correspond (though placement was often quite inaccurate) to open/running water or built-up land, and very occasionally acid grassland/bracken. Since this habitat information was available and (in these instances) more accurate in the Phase 1 data, these blank polygons were eliminated.

Overlapping surveys also existed in the SNH NVC data. On inspection of the overlapping surveys, we found that one particular survey ID was involved in nearly all the overlaps and was also much less detailed and apparently less accurate than the surveys it overlapped. Therefore all the SNH NVC polygons from that survey ID were manually checked against the other NVC data: if they overlapped an apparently better survey they were marked as superseded. Any SNH NVC polygons from any survey ID were also marked as superseded if there was NVC data from the recent Lowland Grassland Review data (which is of likely higher quality and accuracy for the reasons described above), except for the rare occasions on two sites where an SNH NVC polygon contained swamp/mire not mapped by the Lowland Grassland Review (such habitats were not mapped by the Lowland Grassland Review unless part of a mosaic with NVC types relevant to that project). On those rare occasions the retained SNH NVC polygons were clipped to the Lowland Grassland Review polygons, again because the accuracy of the latter is probably higher. As a result of this inspection, 397 polygons were marked as superseded and eliminated from the SNH NVC data, which in the majority of cases corresponded to overlapping surveys from the one survey ID already observed to be less detailed and probably less accurate. Following these procedures, the Lowland Grassland Review data were merged with the SNH NVC data.

Similarly to the combined Phase 1 data, attribute fields were added to the combined NVC data for each priority habitat. These were used to indicate presence of constituent NVC communities of priority habitats, using values or '1' or '0'. In this way values of '1' would indicate presence of a priority habitat (assuming the NVC data was correct), with the exception of bog habitats which cannot be defined by NVC type as explained above. Within the NVC data, the relatively few mapped locations of the key bog NVC communities (M17, M18, M19 and M20) were compared against the Raised Bog Inventory, SSSI data and Phase 1, to verify likely status as raised bog, blanket bog or fen in a similar manner to the checks made on the Phase 1 data itself (see above).

6. TREATMENT OF SPECIES DATA

Plant species records were obtained from TWIC, Glasgow Museum and BBS. BSBI records were included in the data from TWIC and Glasgow Museum. Few bryophyte records were present in the Glasgow Museum data, and on the recommendation of the BBS we also obtained BBS bryophyte records from the NBN. Although the TWIC data included numerous bryophyte records, the BBS gave us a separate dataset on the basis that it was more up-to-date. In both cases some TWIC/Glasgow Museum bryophyte records were duplicates of the BBS data, but conversely some TWIC/Glasgow Museum records were not in the BBS data. Duplicate records were not a problem because the eventual treatment of the species data involved production of raster maps for each species with non-zero squares indicating presence of the species and not how many times it was recorded (see below). With very few exceptions, we did not use any records older than 1970, which corresponds to one of the key BSBI date classes, and has been used by both the BSBI and BBS for atlas production.

The plant records from Glasgow Museum were received as point data, with one of the attributes indicating resolution (10m, 100m, 1km or tetrad). Those from TWIC were received as polygon data where each polygon represented a square of the appropriate resolution (10m, 100m, 1km or tetrad). We converted the latter to points located at the centroids of the polygon squares so that they could be merged with the Glasgow Museum data for subsequent raster processing (see below), rather than devising two separate treatments. The plant records included any species from a list of 230 potential indicators. This list of indicator species, each appropriate for one or more of the relevant Annex I habitats or UK BAP priority habitats, was devised through consulting various sources (see below). In producing this list of indicators, we tried to avoid those species which are largely upland, those which do not occur in Scotland, and those which we judged were too common or more common in non-relevant grasslands or heaths.

Point data of species records could be used in a very simple way by producing sets of point data containing the indicator species relevant to a given habitat, and then using them to add an attribute to habitat polygons intersecting them, indicating that a number of appropriate indicators are present. However, use of points in this way will often select underlying habitat polygons that are very much larger than the actual habitat patch, particularly where the habitats are frequently small (as for certain grasslands and flushes), and will only pick up the polygons that happen to lay under the exact location of the points, excluding others within the relevant grid square which could easily include the true location. It is preferable therefore to produce distribution rasters of the species records at an appropriate grid resolution, so that each grid square is aligned to the Ordnance Survey grid and encompasses the actual possible square area of location of the records. This is the method we used (see further below for details).

6.1 Scoring of indicator species

The potential for a given habitat from indicator species records could be achieved in a simple way just by counting the number of indicators at a given location, but it is obvious that indicators vary in their capacity in that role, since some are more constrained in habitat preference than others. We therefore concluded that it would be better to first give each indicator a rough score representing its 'indicator efficacy', for which we used the following procedure undertaken in Microsoft Excel[19]:

i) we noted how many of the relevant Annex I habitats each indicator is normally found in (for lowland Scotland);
ii) we noted how many of the relevant UK BAP priority habitats each indicator is normally found in (for lowland Scotland);
iii) we noted how many of the following BAP Broad Habitats and (to emphasise efficacy for indicating Annex I habitat) non-Annex I habitats each species can reasonably be expected to occur in (for lowland Scotland): calcareous grassland, neutral grassland, acid grassland, fen/marsh/swamp, bog, inland rock, sand dunes[20], non-Annex I swamp, non-Annex I mire/heath, non-Annex I neutral grassland and non-Annex I rock/waste ground[21];
iv) we calculated the reciprocals of the preceding three steps, so that values closer to one indicated restricted occurrence and therefore better indicator efficacy;
v) we added together these three reciprocals for each species.

The result of this process is that a species with maximum indicator efficacy (which means it occurs in one relevant BAP Broad Habitat, one relevant priority habitat and one relevant Annex I habitat) has a score of 3, and species that occur more widely and are likely to be increasingly less efficient at indicating an Annex I or BAP priority habitat have an increasingly lower score to start with. This is effectively a weighting system for the indicator species; weighting can be applied later during raster analysis but it would be time-consuming to manually adjust weights for a large number of species at that stage. The indicator species scores were joined to the species records in ArcGIS, using as the linkage an attribute added to the Excel and GIS tables comprising the first five characters of the genus and first six of the species epithet, with any remaining non-alphabetical characters (spaces, hyphens) subsequently removed (e.g. *Diantdeltoi*, *Airapraeco*). No account was taken of sub-specific taxa except for *Weissa brachycarpa* var. *obliqua* and *Weissa controversa* var. *densifolia* where the variety is critical for indication of calcareous and calaminarian grasslands respectively, so records of *W. brachycarpa* and *W. controversa* with no variety or other varieties were eliminated beforehand.

6.2 Treatment of different record resolutions

The problem then arises of how to deal with species records at different spatial resolutions in a practical way that is also helpful in narrowing down potential species-rich habitat. Since 10km square (hectad) records are very vague and offer little help in this respect, we eliminated all such records (which were in any case not supplied by TWIC or Glasgow Museum, and only by BBS). Tetrad records (2000m resolution) are also coarse for these purposes, and in many cases (though by no means all) records of the same species existed at better resolution, so tetrad records were also excluded from further analysis. In theory it would be feasible to use the site detail data (if good enough) from hectad/tetrad records to

[19] We used the following as the main sources for likely habitats of each indicator species: Rodwell 1991, 1992, 1995 and 2000; Averis 2013; Preston *et al* 2002; Atherton *et al* 2010.
[20] This is included to help account for the frequent occurrence of some of the bryophyte indicators in sand dunes (not relevant to this project).
[21] This is included to help account for the fact that rock exposure/waste ground can support many species, and mostly they do not constitute Annex 1 6130 (Calaminarian grassland).

give better grid references for those locations that are sufficiently well-defined and localised, particularly where other records exist at better resolution from the same sites. This process has been undertaken by institutions seeking to map herbaria specimens, which naturally lack grid references unless recent, though it requires considerable effort and preferably local knowledge, and was not feasible for this project.

It was originally envisaged that only records of high accuracy (100m or better resolution) would be used for this project, but on finding that many records exist at 1km square resolution we initially considered using these as well. We therefore retained 1km records but treated them separately from more accurate records because of the better potential of the latter to isolate species-rich habitat patches. The majority of the more accurate records were at 100m resolution. There were a smaller number of 10m and (rarely) 1m resolution records, mainly in Borders; whilst these are in theory good for isolating habitat patches, they have the distinct disadvantage that summing the scores of indicators at 10m/1m resolution will commonly give low scores simply by virtue of the infrequency of records at these accuracies[22]. Primarily for this reason the 10m and 1m resolution records were merged with the 100m records (which were far more numerous) for subsequent raster manipulation as described below. This also resulted in fewer sets of rasters to manipulate which is a very significant factor in terms of processing time.

As just noted, we initially retained 1km records, and spent a considerable period of time processing this data separately from the 1-to-100m records. However, it became apparent at the point of calculating the final scores for habitat polygons (which incorporate both habitat and species data) that the species scores derived from 1km indicator records were clouding the resulting mapping as a result of the large size of a 1km square, reducing the clarity with which high potential locations for priority habitats were identified. We then considered whether instead the 1km scores could be added to the 1-100m scores, but this is also problematic: it risks adding 1km records to 1-to-100m record locations that come from elsewhere in the 1km square, producing falsely high scores. Such an eventuality can of course also occur within 1-100m record locations, but not as frequently because of their much smaller size. Ultimately, we therefore only used species data with spatial resolutions of 100m or better.

6.3 Processing of indicator species data

Processing of the species records first involved conversion to point format where necessary (for Borders, by calculating centroid XY coordinates of the supplied square polygon records), and then the following steps were followed:

i) the point data were split into groups each comprising one Annex I or BAP priority habitat at the correct resolution (100m);
ii) records for each species in each habitat were converted to single rasters showing distribution squares at 100m resolution[23], the value of each square being the indicator efficacy score (see above) for the species in question;
iii) all the 100m species rasters for each habitat were added together (using the weighted sum tool with no weighting, effectively adding up the scores of the indicators where present);

[22] Other issues with using 10m/1m resolution records at those resolutions are that: i) the grid reference will be often incorrect by 10-20m if obtained from a standard GPS unit, so the accuracy is often not as good as suggested; and ii) they cannot be shown on maps at true size except at high magnification.

[23] This would be extremely time-consuming if carried out manually; a model was constructed to automate it, which is shown in the appendix as a possible aid to future such work, and which can be used regardless of ArcGIS licence level (with ArcGIS version 10 or above).

iv) the resulting rasters for each habitat were divided by the total possible score for each habitat (which would occur if all indicators were present in a 100m square).

The final product of this procedure is one raster map showing indicator species scores for each Annex I/BAP priority habitat, where each area of occurrence has a score reflecting both the number and efficacy of indicator species present. The higher the score (maximum '1', an unlikely occurrence since it means all possible indicators for the habitat are present in a 100m square), the more likely that the Annex I/BAP priority habitat is present. There are obvious but important caveats to these indicator species scores:

- a low or zero indicator score cannot mean on its own that the priority habitat in question is certainly unlikely or absent, because species data are dependent on recording effort and biased towards accessible and well-known sites, so indicator species will often not have been recorded in that particular location, or have only been partially recorded, or have only been recorded in part of the habitat extent.
- conversely, although a high indicator score suggests that the priority habitat in question is present, it does not guarantee it because some species-rich indicators occur in more than one habitat and in a given 100m grid square might be from more than one habitat.

The indicator rasters were subsequently converted to polygon shapefiles for later integration with the habitat shapefiles. Conversion of the former to shapefiles rather than conversion of the latter to rasters was chosen in order to maintain as much information as possible in the final data regarding the information used to make each final polygon, which was a condition of the project. Conversion of rasters to polygons is not straightforward when the raster grid values are float data (i.e. with decimal places) instead of integers, as was the case here, because the raster-to-polygon tool only works with integer rasters. Conversion to integer raster can be achieved by multiplying it by another constant raster of a large enough multiple of ten, followed by the integer tool to remove the decimal figures. We used an alternative method by converting each raster grid to both a polygon shapefile (with values of '1' in each resulting polygon) and a point shapefile (which can deal with float data, holding the true raster scores for each polygon square in each point), and then used a spatial join to add the scores in the points to the containing square polygons. Whichever method is used, the initial conversion of raster to polygon causes adjacent grid squares to be joined into single polygons; splitting these into component squares matching the raster can be achieved by intersecting the layer with a polygon fishnet of the correct grid size (in this case, 100m).

6.4 Key species

Having carried out the previous steps, it became apparent that the distribution of potential for several habitats was often similar (e.g. H6210 and H6230, and Lowland Calcareous Grassland / Lowland Meadow), which is a reflection of the closeness with which these vegetation types approach each other florisitically, and the appearance of some species-rich indicators in multiple species-rich habitats. We therefore used certain 'key species' to help identify habitat potential with more certainty. The 'key species' are selected indicators falling into one of the following two groups:

a) species that are most likely to occur (in the Scottish lowlands) in one of the Annex I or BAP Priority habitats, and so confer higher likelihood that one of these is actually present regardless of overall species score (which could be low just through under-recording) or habitat score (which could suffer from incorrect habitat attribution);
b) species which occur more widely but are always or normally present in an Annex I / BAP Priority habitat, so increase likelihood of correct habitat if they are present (but to a lesser extent than the first group of key species).

The first category of key species can be weighted heavily in the final scoring procedure because such species strongly suggest the habitat in question. The second category of key species cannot be weighted heavily because the species often occur in other habitats, but being normally necessary components of the habitat in question they can be used to add a small amount of extra score since likelihood of the habitat is slightly higher. The following key species were utilised for the different Annex I /BAP Priority habitats:

- H6130/Calaminarian grassland: *Thalspi caerulescens* is most likely to occur in this Annex I/BAP Priority habitat;
- H6210: *Anacamptis pyramidalis, Dianthus deltoides, Viola hirta* and *Potentilla tabernaemontani* are most likely to occur in this Annex I habitat, and *Thymus polytrichus* is normally present (*D. deltoides* also occurs but less often in U1 acid grassland);
- H6230: there are no species that are significantly more associated with this Annex I habitat than H6210 (or in some cases that do not also occur in forms of LDAG and Lowland Meadow), but *Thymus polytrichus* is normally present;
- Lowland Calcareous Grassland: the key species for H6210 plus *Alchemilla glaucescens*, *Botrychium lunaria, Galium sterneri, Gentianella amarella* and *Thymus polytrichus* have a high likelihood of occurring in this BAP Priority habitat in the lowlands;
- Lowland Dry Acid Grassland: if the habitat is not heath, then *Jasione montana, Polygala serpyllifolia* and *Hypericum humifusum* are most likely to occur in this BAP Priority habitat;
- H6410: there are no species that are significantly more associated with this Annex I habitat than PMRP generally or other habitats, but *Molinia* is required plus at least one of *Sanguisorba officinalis, Valeriana dioica, Cirsium dissectum* or *Trollius europaeus*;
- Purple Moor-grass & Rush Pasture: there are no species that are significantly more associated with this BAP Priority habitat than others, but one of *Molinia, Juncus acutiflorus* or *Juncus effusus* is required (*J. effusus* was not an indicator species for which data was requested, because it is commonly associated with species-poor MG10 and forms of M23 which are also commonly poor; it is probably best not to use it as an indicator of PMRP to avoid highlighting poor MG10, which is very common and does not constitute PMRP); *Carum verticillatum* is most likely to occur in this BAP Priority habitat;
- H6520/Upland Hay Meadow: there are no species that are significantly more associated with this Annex I/BAP Priority habitat than others, but one of *Geranium sylvaticum, Cirsium heterophyllum* or *Trollius europaea* is required; normally *G. sylvaticum* is present;
- Lowland Meadow: there are few species that are significantly more associated with this habitat (MG5/MG8) than others, but *Saxifraga granulata, Platanthera chlorantha, Silaum silaus* and *Sanguisorba officinalis* are most likely to occur in MG5 in the lowlands, although scarce (*S. officinalis* also occurs in H6410 but this is much rarer in Scotland than MG5; *S. granulata* and *P. chlorantha* can also occur in woodland but this is accounted for in the final scores by the lack of habitat score if the habitat is woodland); MG8 requires Caltha palustris;
- H7110/H7120/Lowland Raised Bog: there are no species that are significantly more associated with these Annex I/BAP Priority habitats than blanket bog or other habitats; however, presence of *Andromeda polifera* or *Sphagnum austinii/fuscum/magellanicum* suggests that at least part of the raised bog is indeed active and not degraded (we did not include *Sphagnum papillosum* in this list because it is more common and less demanding, occurring in a wider ranger of bog conditions which may include partially degraded bog and fen); on the other hand, presence of *Campylopus introflexus* (which favours bare peat) helps to confirm some

degree of degradation, although it also occur in heaths and other places; these habitats are best identified from Phase 1/bog inventories/SSSI citations etc.;

- H7130/Blanket Bog: there are no species that are significantly more associated with this Annex I/BAP Priority habitat than raised bog or other habitats, particularly since both active and degraded blanket bog are included, but the same key species as for raised bog can be used to indicate better quality; this habitat is best identified from Phase 1/bog inventories/SSSI citations etc.;
- H7150: presence of *Rhynchospora* spp. is critical to identifying this habitat;
- H7140: *Sphagnum warnstorfii, Calliergon giganteum* and *Calliergon cordifolium* are most likely to occur in this Annex I habitat in the lowlands if the habitat is not woodland (which is accounted for in the final scores by lack of habitat score for woodland); certain sedges are also normally present but in most cases also occur in non-H7140 swamp;
- H7220: there are no species that are significantly more associated with this Annex I habitat than others, but *Palustriella / Cratoneuron* spp. are normally present;
- H7230: *Carex dioica, Cinclidium stygium, Eleocharis quinqueflora, Eriophorum latifolium, Hamatocaulis vernicosus, Scorpidium scorpioides, Sphagnum contortum, Sphagnum platyphyllum* and *Sphagnum subsecundum* are most likely to occur in this Annex I habitat, because they are generally restricted to alkaline fen conditions which is the exact definition of this habitat;
- Lowland Fen: the species mentioned above for H7140 and H7230, and also *Sedum villosum* (which may occur in H7220/H7230 but also in non-Annex I springs/rills), are most likely to occur in this BAP Priority habitat in the lowlands.

The following points should be noted about the key species:

- key species data are also subject to the issues of botanical recording effort and bias towards accessible or well-known sites, so whilst key species presence can be used to indicate higher likelihood of priority habitat presence, key species absence cannot alone mean that a priority habitat is certainly unlikely or absent;
- patches of existing priority habitat for which there are possible key species will occur which receive a high indicator species score but no key species, if the key species have not been recorded or only recorded at worse than 100m resolution (1km, tetrad and hectad records were not used – see above for reasoning). This is most likely to occur when a key species is unlikely to be notable to a botanist; for example, *Juncus acutiflorus* is a key species for Purple Moor-grass and Rush Pasture, but not being rare or scarce it could often go unrecorded.

We also considered the use of 'negative' key species for H6210 where presence of *Anthoxanthum odoratum, Agrostis capillaris* and *Potentilla erecta* in calcareous grassland indicates H6230 (CG10, U4c or U5c) rather than H6210 (CG1, CG2, CG7 or rarely CG10 on limestone). However, we abandoned this idea because we know from existing NVC surveys that H6210 can occur immediately adjacent to U4, and patches of calcareous and acid grasslands can be very small, so ruling it out on this basis could eliminate actual H6210 that happens to be in close proximity to acid grassland.

6.5 Monophagous invertebrates

In theory, invertebrate species that are monophagous (consume only one food plant) and utilise a relevant plant indicator species could be used as proxy indicators for those plant species and therefore also for priority habitat. A short list of potential monophagous invertebrates with appropriate larval foodplants was supplied by SNH, and was expanded using the following on-line sources of information: Natural History Museum HOSTS database (http://www.nhm.ac.uk/research-curation/research/projects/hostplants/), UK Butterflies (http://www.ukbutterflies.co.uk/index.php), UK Moths (http://ukmoths.org.uk/), and Butterfly

Conservation (http://butterfly-conservation.org/). The TWIC data included records of some of theses species for Borders, but we did not have time to utilise them. However, the list of potential monophagous invertebrates was short at 18 species, and there were other problems with this source of information, which was not thought likely to be of great use: i) half of the species were uncommon, scarce or rare; ii) on further research most were found not to be monophagous; iii) the foodplants would in most cases not indicate a particular Annex I or BAP Priority habitat, but only a degree of species-richness. An additional general issue is that mature stages of invertebrates that are sufficiently well-recorded and whose habits are sufficiently well-known (principally Lepidoptera) to be possibly useful in this way are highly mobile and may therefore be recorded at some distance from the priority habitat they potentially indicate.

7. MERGING OF DATA AND SCORING PROCEDURE

7.1 Merging of habitat and species datasets

Prior to merging the Phase 1, NVC and Potential Species-rich Grassland datasets, the latter was processed to smooth out finely serrated polygon edges. These result from raster imagery analysis, and if left intact cause the production of excessively numerous (and spurious) very small polygons during union operations wherever they are closely coincident with straighter Phase 1/NVC polygon edges. Smoothing of fine serrations in the Potential Species-rich Grassland dataset was achieved using topology rules (allowing vertices to move up to 1m whilst prohibiting overlaps)[24], by eliminating areas corresponding to roads/buildings/water in OS Master Map or Phase 1 built-up habitat[25], and by removing remaining polygons of very small size or thinness. The Phase 1, NVC and Potential Species-rich Grassland datasets were then combined using the union tool.

The only attribute field retained from the Potential Species-rich Grassland dataset was the 'Candidate' field which codes degree of likely interest from 1 to 3 based on heterogeneity and other evidence; the wet/dry classification was not used because of the apparent level of error (as noted above). A final union was used to combine the finalised species score data (existing as 100m square polygons with species scores for the different priority habitats) with the habitat data forming one shapefile containing all scoring attributes.

Following the scoring procedure (see below), almost one quarter of the polygons had no score for any priority habitat. These corresponded to polygons for which there was Phase 1 habitat information but the habitat was not relevant to any of the priority habitats (e.g. woodland, dense scrub/bracken, ruderal vegetation, dry heath, arable, improved grass, intertidal), and for which there was also no NVC data, no steep grass, no Potential Species-rich Grassland category, no species score and no key species. Having no score of any kind and no Phase 1 habitat relevant to any of the priority habitats, these polygons were also removed.

The number of polygons naturally increases rapidly when multiple datasets are combined in this way, and some of these are excessively small, particularly when there are closely but not exactly coincident edges (mainly resulting here from the smoothed Potential Species-rich Grassland dataset). In order to minimise these and keep the dataset manageable, the eliminate tool was used to merge all very small polygons (less than $1m^2$) into adjacent

[24] The vertex movement topology rule was set to a large distance only for this purpose and not elsewhere. Note that the generalise tool can achieve a similar result but with this data was found to occasionally cause very narrow sliver gaps to appear.

[25] The Potential Species-rich Grassland dataset overlaps some built-up and road habitat because blocky edges from raster analysis tend to repeatedly cross the edges of these habitats, and perhaps because there was overzealous identification of 'species-rich' habitat in gardens and urban areas.

polygons with longest coincident edges. The combination of this process and the smoothing process for Potential Species-rich Grassland (described above) reduced the number of polygons from approximately 750,000 to 176,000.

7.2 Final score components

All the score component attributes were initially set to have an equal maximum value of ‘1’ where the attribute was applicable. The components were subsequently weighted as deemed appropriate for each priority habitat (see next section). We did not utilise the wet/dry categorisation from the Potential Species-rich Grassland dataset because of the previously noted issues (see above). The following attributes were used as components of the final habitat scores for each polygon:

i) whether the Phase 1/BAP Broad Habitat is appropriate for each Annex I/BAP Priority habitat (integer; 0 or 1);
ii) whether the NVC is a constituent community of each Annex I/BAP Priority habitat, other than Annex I/BAP Priority bog habitats which cannot be identified by NVC alone (integer; 0 or 1);
iii) whether the area is a type of dry Phase 1 grassland and also steeper than 20 degrees, and therefore more likely to be interesting regardless of Phase 1 grassland type (integer; 0 or 1);
iv) whether the area is in the Potential Species-rich Grassland dataset (float[26]; 0 if not in the dataset; 0.33, 0.67 or 1 if in the dataset, corresponding to the candidate classes 3, 2 and 1 respectively, the maximum of ‘1’ suggesting the most heterogeneity and least productivity);
v) a species score derived from the number and indicator efficacy of indicator species (float; 0 to 1, where 0 means no indicator records at 100m accuracy or better, and 1 theoretically means, but not achieved in practice, that all utilised indicators are present);
vi) whether any 'key species' are present (integer; 0 or 1, but sometimes comprising more than one element for a habitat as detailed above, requiring different usage per habitat type as described below).

NVC data (or, for bog habitats, Phase 1 and other information, as described above) were used to generate a separate attribute for each priority habitat indicating where each one was accepted to be present (with the assumption that the NVC and other data used for this purpose was correct). This element was not incorporated numerically into the final scores so that the results of the scoring process for accepted locations of priority habitat would be available for assessment.

A final attribute was added for each priority habitat which gave the potential of that habitat as (in most cases) ‘Very Low’, ‘Low’, ‘Moderate’, ‘High’ or ‘Very High’. These categories were assigned following appropriate weighting of the final score components and designation of numerical boundaries between these categories, which differed depending on the habitat in question (see below).

7.3 Score banding and weighting of score components

It is important to appropriately weight the components of a scoring system of this type otherwise (in this case) too much habitat will be shown as potentially poor or too much as potentially good. Through a combination of trial and error, consideration of what the final scores would be with one, two or multiple score components, and bearing in mind

[26] A number data type with decimal place, having smaller range than the ‘double’ data type but taking less file space.

subsequent categorisation of habitat potential as ‘Low’, ‘High’, etc., we arrived at the following weightings (details for each habitat are given in Appendix 1):

- Appropriate Phase 1 habitat – generally given a weight of 5, but lower where the Phase 1 was less reliable for indicating potential priority habitat, and higher for bog habitats which are more certainly identified by the Phase 1 category (especially given that we had double-checked bog Phase 1 habitats against the Raised Bog Inventory, SSSI citations and – where applicable – Tweed Wetland Strategy).
- Steep grass factor – only applied to calcareous grasslands, Lowland Meadow and Lowland Dry Acid Grassland, the steep grass factor was given a weight of 3 (greater than the Potential Species-rich Grassland score but lower than the Phase 1 since it cannot identify grassland type, only that there is a higher likelihood of some type of unimproved grass).
- Potential species-rich dataset – the score derived from the ‘candidate’ class was not given any further weight (remaining as 0, 0.33, 0.67 or 1), because it was regarded as less important and less reliable than the other score components, particularly when occurring on its own without any other score component, particularly as it cannot indicate on its own any particular habitat; however, this component can move a score into a higher category, and ensures that all areas identified in the species-rich dataset are retained (unless corresponding to accepted built-up areas/roads – see above).
- Indicator species score – this component was variably weighted such that the maximum achievable score for any habitat was 3 or 4.
- Key species – a high weight was given to presence of the first category of key species (that have a very high likelihood of occurring in the priority habitat in question), often 9; a low weight was given to presence of the second category of key species (that are normally or always present in a given habitat but do occur commonly elsewhere), as low as 1.

Following calculation of the final priority habitat scores using the above components and weightings, the categories of ‘Very Low’, ‘Low’, ‘Moderate’, ‘High’ and ‘Very High’ were assigned for each priority habitat. To ensure that the banding for each habitat was useful, the break values for these scoring bands differed amongst the priority habitats depending on the number of score components and their weights (particularly whether there were one or two key species components and whether one was highly weighted). Typical break values were 0, 1, 4, 7 and 9 or similar. Details of the banding for each priority habitat can be found in Appendix 1.

8. DISCUSSION OF THE FINAL DATASET

8.1 Area calculation issues

The process of scoring polygons for likelihood of priority habitats based on various contributing datasets means that a given polygon can and often does have potential for more than one priority habitat. Although area information is integral to the GIS data and can be extracted for any polygon or group of polygons, it is unwise to assume that the priority habitat with the highest score at a given location is the most likely and therefore the one to use for area calculations, because: i) indicator species scores can be similar for priority habitats with several indicator species in common (e.g. H6210/H6230; Lowland Calcareous Grassland/Lowland Meadow/Lowland Dry Acid Grassland); ii) the habitat present in reality could be the lower-scoring of overlapping potential habitats (particularly through accidents of species recording, and also because some of the contributing habitat datasets are not completely reliable – see above); iii) the resolution of the indicator species scores and key species presence is 100m, within which more than one priority habitat could be present; iv) the resolution of some of the contributing habitat datasets is such that small patches of priority habitat are likely to be missed (see above); and v) a degree of misclassification is known to have occurred in some of the habitat data.

An additional complication for area analysis where older NVC data exists from the SNH NVC data (not the recent Lowland Grassland Review data, which adheres to SNH GIS protocol) is that the mosaic coding is highly variable and not uniform. It would take a significant amount of time to standardise and correct the mosaic coding in order to be able to automatically split the mosaic components and percentages into separate attribute fields for subsequent NVC area calculations.[27]

These issues mean that the final dataset is unlikely to be useful for meaningful and reliable area calculations of priority habitats as it stands. In order to do this, it will at least be necessary to gain some idea of the success of the different scoring bands ('Low', 'High', etc.) in locating the priority habitats, through a fieldwork-based sampling survey.

8.2 Dry grasslands

Many locations in Borders, but few in East Dunbartonshire (for which there were surprisingly almost no *Thymus polytrichus* records, probably reflecting under-recording) scored highly for Lowland Calcareous Grassland in general. It is difficult to separate H6210 and H6230 with habitat data (except NVC) and/or species data because of the number of common indicator species, unless any of the key species are present that are more likely in H6210 (such as *Dianthus deltoides, Viola hirta* or *Anacamptis pyramidalis*), which are unfortunately all scarce and not always present in H6210. Polygons labelled as 'High' or 'Very High' for Lowland Calcareous Grassland have a good chance of actually containing H6210 or H6230, but it must be remembered that owing to floristic convergence and similarity of indicator species, some high scoring areas for Lowland Calcareous Grassland could turn out to be forms of U4 or MG5 (Lowland Dry Acid Grassland or Lowland Meadow respectively), although this still means they are likely to be at least a BAP Priority habitat. Where key species for H6210 are recorded there is a high likelihood of it presence, but it will also be present in other locations where H6230 may have higher potential but the key species have gone unrecorded or are absent.

For H6520/Upland Hay Meadow, only one site actually satisfied the criterion of having both unimproved neutral grassland and key species present (in this case *Geranium sylvaticum*, by Easter Branxolme Loch near Hawick) and was therefore labelled as 'Very High' potential. However, an existing NVC survey identified not MG3 (the constituent NVC community) but MG5 and CG10 in the small relevant area, and it may be that the *G. sylvaticum* was from a small adjacent woodland patch; alternatively the NVC survey may have missed a small patch of MG3. MG3 was not recorded in any of the available NVC data, but it would be surprising, despite its rarity, if this habitat was absent completely from Borders and East Dunbartonshire. It would therefore be worth field checking other slightly less high-scoring areas, particularly those that do not appear to be woodland; this includes a few patches along river banks (e.g. River Tweed and tributaries), a likely situation for surviving fragments of MG3. Some other higher-scoring patches also correspond to known MG5, which reflects the similarity of indicator species and identical Phase 1 habitat type.

A good number of sites in the study area scored 'High' for Lowland Meadow which reflects the fact that the MG5 key species (*Saxifraga granulata, Platanthera chlorantha, Silaum silaus* and *Sanguisorba officinalis*), although scarce individually, together provide scattered points throughout the study area. Only one site near Torrance achieved 'Very High' because

---

[27] We have provided, via manual inspection of the SNH NVC data, an attribute for number of mosaic components, which would allow crude area analysis by dividing NVC polygon areas by this number.

it was the only location with mapped unimproved neutral grassland and key species present. Owing to floristic convergence with other communities, particularly U4 and CG10, it is possible that 'High' scoring areas for Lowland Meadow may instead contain U4 or CG10; it was found during the Lowland Grassland Review (Dadds and Averis, in press) that many previously-recorded MG5 patches did actually comprise U4 and/or CG10. However, this still means that high-scoring areas for Lowland Meadow have a good chance of containing at least a BAP Priority habitat and possibly Annex I calcareous grassland. There is a low to moderate chance that high-scoring areas for Lowland Meadow are woodland, since *S. granulata* and *P. chlorantha* sometimes occur in woodland. We did not eliminate woodland from this (and other) priority habitats because it is not possible to be sure (particularly given the previously discussed deficiencies with the Tweed Catchment Phase 1) that woodlands do not contain unwooded banks or clearings, which could be small but still support Lowland Meadow or other priority open habitat.

LDAG in the form of U4 (less often and much less extensively U1) is widespread in the study area particularly on hilly ground at the fringe of the uplands. Polygons marked as 'Moderate' or higher have a high probability of actually being LDAG, though a proportion of this area, possibly large, will be U4b which is species-poor. Areas marked as 'Low' and even 'Very Low' (often only scoring from the species-rich dataset indicating a degree of heterogeneity) may sometimes also be LDAG because it is relatively common. Since the Tweed Catchment Phase 1 is known to have missed some areas of acid grassland, some will not be picked up at all. Some 'Moderate' LDAG areas in East Dunbartonshire, and occasionally in Borders where replacing Phase 1 scattered habitats, derive from the BAP Broad Habitats of acid grassland and rough low-productivity grassland, and in the latter case in particular a proportion, probably small, may be neutral grassland. Areas of 'High' or 'Very High' LDAG potential are likely to be most interesting especially where corresponding to steep slopes. As noted for Lowland Calcareous Grassland and Lowland Meadow, there is floristic convergence between these habitats and LDAG, and a proportion of the extensive LDAG area, most likely very small since LDAG is much more extensive, may actually correspond to these habitats. Where the potential for Lowland Calcareous Grassland or Lowland Meadow is higher, there is no guarantee that these habitats are actually more likely, especially where the scores are low, because similarity of final scores could reflect floristic convergence, or, for example, variability in recording of indicator species.

8.3 Purple moor-grass & rush pastures

H6410 is very rare in Scotland, and the available NVC data did not contain any records of the constituent NVC communities (M24 and M26) in the study area. Neither were there any records of *Cirsium dissectum* (a strong indicator of M24), and although the key M26 indicators of *Valeriana dioica, Sanguisorba officinalis* and *Trollius europaeus* were recorded, these species occur more often in other habitats (e.g. M23, MG5). We have maintained five classes of 'likelihood' for this habitat (from 'Very Low' to 'Very High') in order to highlight where H6410 has the highest probability of occurring based on the data available to us. However, even the highest potential H6410 patches are still in reality unlikely to actually contain H6410 for the aforementioned reasons. Such patches are certainly still worth investigating because they are likely to include at least a BAP Priority type (particularly Purple Moor-grass & Rush Pasture) and possibly a different Annex I type.

The PMRP final data will be more reliable than the H6410 data, because unlike H6410 PMRP as a whole is not rare. Where the potential for H6410 is higher than for PMRP, this will often be because one of the key species is present, as described in the previous paragraph. As for other habitats, localised indicator species records will raise the score in the relevant 100m square, and it should be noted that lower-scoring areas immediately adjacent to a high-scoring 100m square may also be higher-scoring in reality but lack existing species records. In other cases, the high-scoring habitat could in reality be only a small part of a higher-scoring 100m square. We considered removing the key indicator scores where the habitat was known to be bog; however, we did not do this because it risks losing PMRP occurring within bog systems, such as on the lagg of raised bogs.

8.4 Bogs

Blanket bog and active/degraded raised bog, corresponding to H7130 and H7110/H7120 respectively, were identified with high certainty using the Raised Bog Inventory, SSSI citations and Tweed Wetland Strategy in addition to Phase 1 data. All polygons labelled as 'High' or 'Very High' for active raised bog and degraded raised bog *do* contain active or degraded raised bog to the best of our knowledge. Most other areas for active / degraded raised bog are indicated as 'Very Low' and are *not* raised bog to the best of our knowledge; small areas indicated as 'Low' or 'Moderate' have some indicators that are common with raised bog but as far as we know these are also *not* raised bog.

Similarly, areas coded as 'High' or 'Very High' for H7130/Blanket Bog are thought to contain blanket bog, with the assumption that the Phase 1 data is correct in showing bog habitat (which may occasionally not be the case). Such areas are all towards or at the edge of the study area, as expected, and mostly on hilly ground. Areas which are not rated 'High' or 'Very High' for blanket bog are not, as far as we know, blanket bog; occasional 'Moderate' patches contain indicators common to raised and blanket bog but are not the latter to the best of our knowledge.

Note that the terms raised and blanket bog, both in general and for the purposes of Annex I / BAP Priority habitats, describe hydrologically distinct and often large areas of peat within which there may be other habitat types, often subordinate in active bogs but potentially including other priority habitats such as transition mires, flushes and swamps, for example in pools or on the lagg of raised bogs. Thus some areas within a stated raised / blanket bog may not be bog habitat, and bogs can validly contain patches with a high likelihood of fen or other habitat. Note also that some degraded raised and blanket bogs are dominated by Phase 1 modified bog habitat, and in the case of degraded raised bog the dominant habitat may itself be non-bog habitat such as plantation and occasionally grassland in the most degraded cases.

H7150, which is essentially defined by the presence of *Rhynchospora* spp., and occurs at the interface of wet pools and drier bog, achieved a rating of 'High' and 'Very High' in only one location. This was at Low Moss in East Dunbartonshire, which contains the only record of *Rhynchospora* in the species data. It is surprising that there is even one record in the study area, since this species is most characteristic of undisturbed western bog/wet heath. We know through having visited this site that the bog was previously quite dry with plentiful birch encroachment, but has been managed by tree removal and drain blocking to rewet it. Assuming that the record is accurate, the location of the *Rhynchospora* is near a retail park at the edge of which deep metal sheet piling appeared to have been effectively retaining water, with large pools against the sheeting. It would be worth checking the site to see if *Rhynchospora* is still present, which would help to suggest that management of the bog has been successful. Areas rated as 'Low' for H7150 contain appropriate habitat (mostly bog) but are not likely to contain H7150 unless *Rhynchospora* is present but has not been recorded, an unlikely eventuality given the preference of *Rhynchospora* for undisturbed western bog/wet heath.

8.5 Fens

Polygons labelled as 'Very High' for H7140 (transition mires) are highly likely to contain H7140 because the Phase 1 habitat is appropriate and key species are present. There are small patches marked as 'Very High' sparsely scattered across Borders and very rarely in

East Dunbartonshire, with noticeable aggregations near Hawick and Selkirk. This probably reflects the true scarcity of this habitat, but may also reflect under-recording of appropriate Phase 1 habitats or indicator species. By comparison with known H7140 determined from NVC surveys, a high proportion match with a 'High' / 'Very High' polygon either on it or close by, but there are some known H7140 areas that are within 'Moderate' polygons, and a smaller number within even lower-scoring polygons, indicating that the Phase 1 habitat data is occasionally wrong and that species records are lacking. There is a possibility, particularly for 'Moderate' or lower-scoring areas, of non-H7140 swamp occurring instead of H7140. It is difficult to rule out non-H7140 swamps or complex habitat mosaics lacking H7140 because of the similarity of species between H7140 and certain swamps/fens. Whilst H7140 often occurs with non-H7140 swamp in succession from water, the latter is more common and often occurs without H7140. However, at least all areas marked 'High' or 'Very High' for H7140 should be checked because they are likely to contain some type of notable vegetation, even if not H7140.

H7220 is problematic because the most strictly appropriate Phase 1 habitat is 'spring' which is normally target-noted and not mappable, and many springs are missed during Phase 1 habitat mapping. For this reason we expanded the appropriate habitats to include any fen type, since springs are commonly associated with other fens, especially down-hill flushes. The other problem is that the key species which are normally present (*Palustriella / Cratoneuron* spp.) can occur in other fens or on wet rock. H7220 is therefore one of the most difficult habitats to identify without NVC information. Polygons with 'Very High' H7220 potential (there are none with 'High' potential, and only a few with 'Moderate') are highly likely to contain the key species, and especially if this is on an open hillside they are worth investigating, since there is a good chance of H7220 being present, or alternatively another fen type. Only two locations with H7220 exist in the available NVC data, both in Borders, one at the very edge of the 300m limit and another on a coastal slope; neither of them are picked up without the NVC data, because the key species were not recorded and, in these cases, the Phase 1 habitat was not fen. Locating H7220 is very reliant on recording of *Palustriella / Cratoneuron*, and as such most locations will not be picked up.

There are a number of indicator species for H7230, as previously mentioned, that are more or less restricted to the alkaline fens defining this habitat. This makes it easier to locate by species, and the results more trustworthy when such species are present. All polygons marked as 'Very High' have these key species present and are highly likely to contain H7230. Occurrence of 'Very High' potential H7230 are scattered throughout Borders, especially near Hawick and Selkirk, but are extremely rare in East Dunbartonshire. By comparison with existing NVC data, there is a good correlation between known H7230 and 'Very High' potential H7230 polygons, if not directly on the relevant NVC area then in close proximity. This bodes well for 'Very High' polygons outside of the sparse NVC survey data. There are still some known H7230 areas, however, that fall in lower-scoring polygons, mostly 'Moderate' (there are very few 'High' potential polygons) but occasionally 'Low' or 'Very Low', again indicating lack of species records. All 'High' or 'Very High' polygons for H7230 should be inspected because of the high probability of H7230 being present, and as many 'Moderate' polygons as possible.

The overall Lowland Fen BAP Priority habitat has the advantage of greater breadth of floristic composition, meaning more indicator species can be used to help predict its occurrence. The distribution of 'Very High' and 'High' potential Lowland Fen polygons is similar to a combination of H7140 and H7230, as would be expected. There is again a reasonable correspondence between known Lowland Fen from NVC data and 'Very High' or 'High' potential Lowland Fen, but as for the fen Annex I habitats, some Lowland Fen known from NVC data falls in 'Moderate' or lower-scoring polygons, again indicated shortfalls in the Phase 1 and species data. It is possible that a small proportion of the Lowland Fen area suggested by the data is actually the Reedbed BAP Priority habitat, but to be classified as such *Phragmites* must be known to be dominant and not just present. We made no attempt to distinguish Lowland Fen from Reedbed except to exclude mapped areas of NVC types S4 and S26 from Lowland Fen (S27 can also sometimes constitute Reedbed but this cannot be told without knowing the abundance of *Phragmites*)[28]; however, we would expect Reedbed outside the known and excluded S4 and S26 areas to occupy only a small fraction of the overall fen area, particularly since a number of important wetlands in the study area have NVC data.

9. RECOMMENDATIONS

9.1 Using the data from this project

The attribute fields in the dataset are described in Appendix 4. The critical attributes for symbolising priority habitat potential are those with labels (titles) in the format 'cat*xxxx*', where '*xxxx*' contains abbreviations of the priority habitats (see Appendix 4 for a list of abbreviations, which are also used in Table 1 above). Figure 3 below shows an example area of the dataset, for Lowland Calcareous Grassland (LCG), and using a symbology scheme of light yellow to dark brown for the categories of habitat potential. In order to maximise visual effectiveness of the dataset, we recommend that locations of accepted priority habitat (from consitutent NVC communities or, for bogs, other sources) and steep grass are separately symbolised on top of the habitat potential, as demonstrated in the figure (the relevant attributes for symbolising accepted habitat are labelled 'def*xxxx*', and presence/absence of steep grass can be symbolised using the 'Steep Grass' attribute).

As already noted, accepted priority habitat derived from mapped constituent NVC communities (or, for bog habitats, from other information as explained above) was not incorporated into the final scores to so that calculated habitat potentials would be retained for accepted priority habitat. However, it is clearly useful to be able to see accepted priority habitat (if only to avoid arranging NVC surveys of already well-surveyed sites). In Figure 3 these areas are symbolised using transparent polygons with thick red boundaries, so that the habitat potential is still visible within them.

For dry grasslands, we also suggest separate symbolisation of the steep grass factor, because although it suggests that unimproved grassland could be present (it does not guarantee it because of the resolution and reliability of the Phase 1 data – see above), it cannot indicate what sort of unimproved grassland that might be. Thus it could not be weighted too heavily in the final scores, and if other information is lacking the final scores will still be fairly low for 'steep grass' areas even though there might be unimproved grassland of some type present. Figure 3 demonstrates use of an overlaid hatching symbol for steep grass, which makes it clear to persons inspecting the data (including field surveyors) that low-scoring areas nevertheless have potential to be some type of unimproved grassland if the steep grass symbol is present.

Interpretation of the data in Figure 3, but similarly applicable to all the priority habitats in the dataset, and also applicable to other such projects using the same methods, is discussed below:

- Areas with relevant species records are typically obvious by the incorporation of 100m grid squares.[29] The squares are 100m because that was the coarsest resolution of species records used, and the most common resolution of those

[28] S24 and S25 are not likely to occur in Scotland. S4/S26 are distinguishable from S27 dominated by *Phragmites* by the much poorer diversity of the associated flora in S4/S26.
[29] There are a few 100m squares without species scores; this is not an error but results from removal of a small number of older (mainly 19th century) species records that escaped earlier exclusion.

records. The 100m squares for species data necessarily correspond to Ordnance Survey grid squares for those species data, and not to habitat patches, but subdivisions within the 100m squares originate from habitat data. If the indicator species score was high or especially if there were key species that are much more likely to occur in the priority habitat than elsewhere, then this causes those 100m squares to have higher habitat potential. If species records existed, but these did not include such key species, and the general indicator species score for the priority habitat was low, then this causes those 100m squares to have low habitat potential, unless other factors raise it. These effects are obvious in several places in Figure 3.

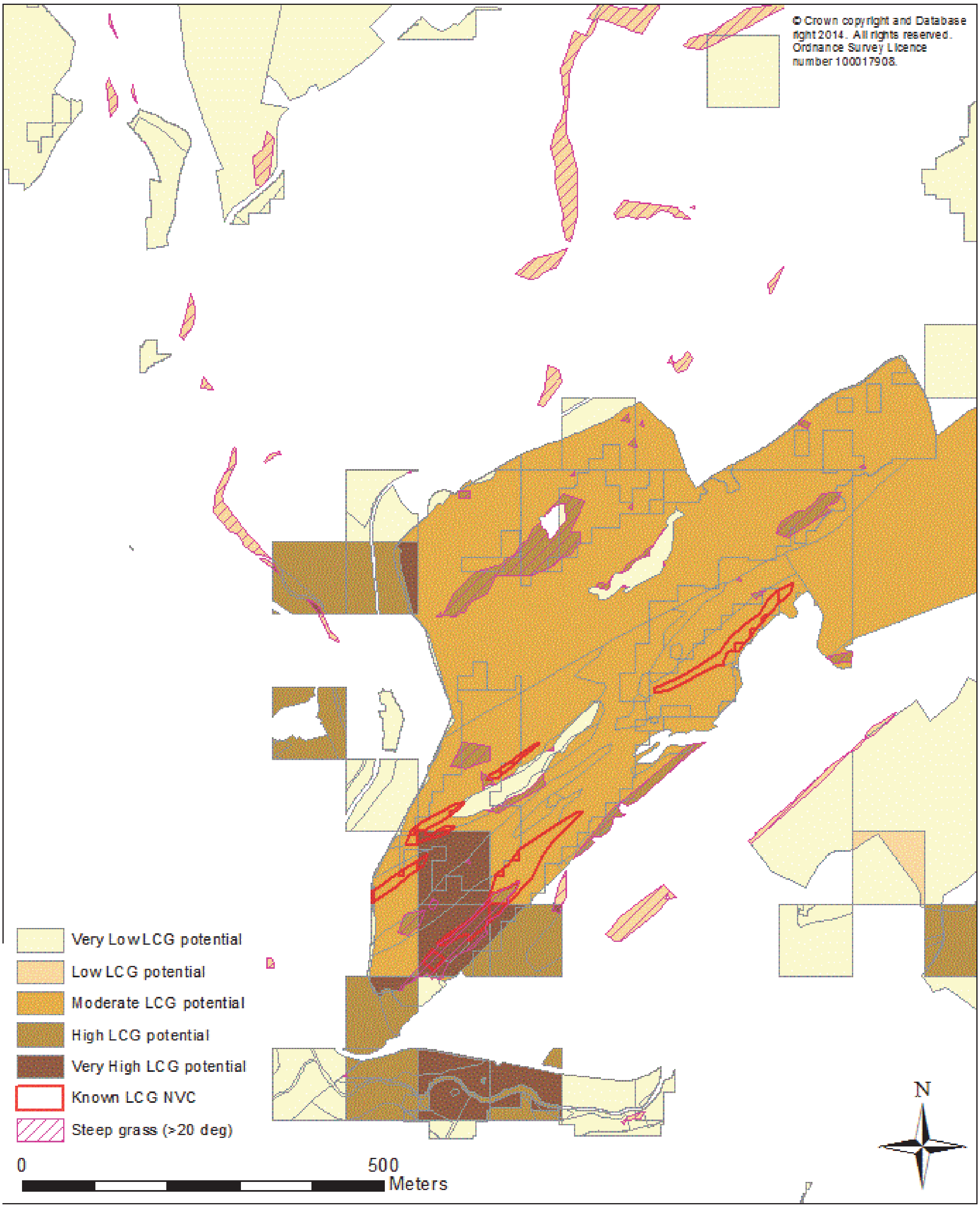


Figure 3. An example of the data, using Lowland Calcareous Grassland (LCG).

- Square jagged or serrated outlines often stand out, and these originate from the Potential Species-rich Grassland dataset (the blocky appearance results from raster analysis). Since we did not regard this data as highly reliable (as explained above), and therefore gave this component of the scoring process a low weight, it often has little impact on final habitat potential, but it sometimes raises it to a higher class particularly where the Potential Species-rich Grassland dataset indicated highest likelihood of species-rich habitat. This effect can be seen in the bottom of Figure 3, where a 100m square originating from species data has lower- and higher-scoring segments resulting from differing scores from the Potential Species-rich Grassland dataset (the missing parts of the 100m squares in this vicinity correspond to roads/built-up areas).
- Areas comprising a Phase 1 habitat regarded as appropriate to the priority habitat in question (this would obviously be calcareous grassland for Figure 3, since the subject is Lowland Calcareous Grassland) have higher potential. This is the underlying cause for the large expanse of ‘Moderate’ potential in Figure 3. Areas that have ‘Low’ or ‘Very Low’ potential do not contain an appropriate Phase 1 habitat for the priority habitat in question (according to the data used), and are visible because of other factor(s) which on their own indicate lower potential. Where one of the other contributing factors is present within an area of appropriate Phase 1 habitat, the habitat potential may be raised. This can occur where, for example, there are relevant species data (e.g. just left of the centre of Figure 3 where a small part of a 100m square for species data falls on appropriate Phase 1 habitat), or where there is thought to be steep grass (see next point). Note that if Phase 1 calcareous grassland had not been mapped in this area, then there would be little or no ‘Moderate’ habitat potential and a lot more ‘Low’ and ‘Very Low’, but there would still be localised ‘High’ or ‘Very High’ patches resulting from species data, and the steep grass strips (see next point) would still stand out.
- Presence of the steep grass factor (applied only to calcareous grasslands, Lowland Meadow and Lowland Dry Acid Grassland) is made very apparent by separately symbolising it as suggested above. The steep grass factor causes areas of any dry grassland type to score more highly; this includes ostensibly non-relevant dry grasslands such as improved grass because, as explained previously, the Tweed Catchment Phase 1 in particular has a tendency to miss patches of notable grassland within non-notable grassland. There are numerous strips and small patches in Figure 3 which score ‘Low’ rather than ‘Very Low’ because although the Phase 1 habitat was not appropriate (e.g. improved grass), the habitat was a dry grassland with slopes of at least 20^o and therefore potentially some type of unimproved grassland. The steep grass factor similarly causes appropriate Phase 1 habitat to be raised to ‘High’ rather than ‘Moderate’ potential, as shown in the centre of Figure 3.
- Areas shown as ‘Very Low’ potential will in many cases not contain the priority habitat in question. They represent areas which have no known appropriate Phase 1 habitat, no steep grass, and no records of the key species that are most likely to occur in that priority habitat than elsewhere. They do have a low general indicator species score or are present in the Potential Species-rich Grassland data, which alone cannot indicate a particular habitat. It is possible to remove ‘Very Low’ potential areas from visibility by simply removing the symbolisation for that category; it would of course also be possible to make copies of the dataset for each priority habitat and delete the ‘Very Low’ potential polygons in each one (we did not do this at SNH’s request).
- The final habitat potentials are reliant on the contributing data, and it is important to remember that areas of high potential will sometimes not contain priority habitat when field checked, and similarly some areas of high potential will not have been picked up. This is an inevitable result of the vagaries of the incorporated data,

especially the known imperfections in the habitat data (in particular those derived from imagery analysis, see above), the dependency of species data on recording effort and accuracy, and the fact that 100m squares for species data will normally include irrelevant habitat as well as priority habitat or only part of a priority habitat (if it is large). These issues will apply to all such projects, but some could be alleviated by field checks (see below).

9.2 Recommendations

9.2.1 Estimated time for similar projects

A significant amount of time was spent during this project on assessing the reliability of the habitat datasets, on establishing a suitable minimum slope that would increase likelihood of species-rich dry grassland, and on devising a list of 230 indicator species with individual habitat preferences and scores reflecting their efficacy as indicators. It is envisaged that these tasks will not be undertaken again for similar projects unless refinements to the indicator species are sought or it is considered necessary to make other comparisons of habitat data. The following are therefore assumed to be the main steps required for future similar work:

Combine existing Phase 1 datasets into one file with standardised attributes.
This step could be very quick if there are few datasets involved and the required attributes for each are identical in format, and there are no obvious errors to correct. If the attribute formats are different between datasets but particularly if there are many datasets to combine (both of which applied to this project) then considerably more time could be required. Where multiple Phase 1 datasets overlap, a decision must be made as to which one should be kept at the expense of the other, and geoprocessing carried out accordingly. If there is only one Phase 1 dataset then the only requirement will be to ensure that there are no obvious errors in attribution and that attribute formats are standard. If some data will be hand-drawn and requiring digitising (as was the case for some East Dunbartonshire data in this project) then the time required could dramatically increase since digitising is slow.
Estimated time: 0.5 to 3 days, potentially considerably more if hand-drawn maps must be digitised.

Add attributes to the Phase 1 data and populate to indicate presence of appropriate Phase 1 habitat for each priority habitat.
For non-bog habitats, this step only requires that the added attribute fields are selectively calculated according to presence or otherwise of Phase 1 habitat appropriate to each priority habitat (appropriate habitats are given in Table 1). Bog habitats could be treated similarly but this will fail to distinguish raised/blanket bog in areas of Phase 1 modified bog, assumes that Phase 1 attribution of intact raised/blanket bog is correct (which may not be the case because distinguishing these in the field can sometimes be difficult), and may not identify badly degraded raised bog that is now a non-bog Phase 1 habitat (e.g. woodland). If the more thorough method used in this project is followed, it will be necessary to consult the Raised Bog Inventory, SSSI citations and any other documentation which might help in distinguishing raised/blanket bog and active/degraded status for raised bogs (such as wetland strategies or biodiversity action plans). For H7150, potential presence can only be attributed after species records of *Rhynchospora* spp. have been obtained.
Estimated time: 0.5 day plus 0.5 to 1 day for thorough investigation of bog habitats depending on amount of bog habitat. Comparison against the Raised Bog Inventory is easiest if the relevant data are digitised which may require another 0.5 days.

Combine NVC datasets into one file and check for overlaps etc.
In some cases the only NVC data available will be from SNH, which will be in one file. If there are overlapping NVC surveys it is advisable to consider which ones should be

regarded as superior (usually the more recent, if that data is available, or the more detailed, as determined by manual inspection), and to eliminate the inferior of the overlapping surveys. It is advisable to also check for polygons with no NVC data and to delete them, ideally checking that they correspond to non-NVC areas (e.g. built-up, roads, open water, rock), and to check for obvious data input errors in the NVC codes
Estimated time: 0.5 to 1 day depending on number of overlaps

Add attributes to the NVC data and populate to indicate presence of constituent NVC communities.
This step is similar to step 2 above. It only requires that appropriate data selections are made for each priority habitat, and that the relevant attribute fields are populated accordingly to indicate presence or otherwise of constituent NVC communities for each priority habitat. This does not apply to bog habitats which cannot be defined by NVC. If possible (it may not be if there is a great quantity of bog), we recommend that occurrences of M17, M18, M19 and M20 are checked against the Raised Bog Inventory, SSSI citations and other documentation, particularly as some areas of bog may have been misclassified in the Phase 1 data, and some occurrences of these communities may not be bog (some may be better classed as fen).
Estimated time: 0.5 day plus 0.5 to 1 day for thorough investigation of bog NVC depending on amount of bog NVC.

Process DTM data to obtain slope, and process with Phase 1 data to obtain steep grass areas.
This step is required to identify dry grassland (according to available Phase 1 information) that is at least 20^{o} steep and therefore more likely to be unimproved regardless of Phase 1 classification. Processing DTM data can take a significant period of time for a large region or with a lower specification computer. Once the slope data is produced, it is relatively simple to process it together with the Phase 1 data to obtain areas of dry Phase 1 grassland of at least 20^{o} steepness.
Estimated time: 0.5 day

Use of documentation to check for notable habitat
For this project this mainly involved use of the Raised Bog Inventory (which is covered in previous steps) but also use of the Tweed Wetland Strategy and other material from BSBI recorders to check (as far as was feasible in the time available) that notable sites were in the Phase 1 data and coded correctly. Time required for this would be variable for other projects depending on what information was available.
Estimated time: 1 to 3 days

Process indicator species data to make rasters with species scores for each priority habitat, followed by conversion to polygon files
This process requires use of a GIS model such as the one given in Appendix 2, or scripting to achieve the same, to avoid potentially very much greater time requirements. If the species scoring procedure used for this project is followed, the indicator efficacy scores for individual indicator species devised for this project could be quickly applied to other species data in GIS using an attribute join, having first created a suitable common linkage such as a genus/species abbreviation. If, as in this project, only 100m accuracy or better species data are used, and it is acceptable to aggregate them into one resolution or 100m as for this project, then the process of making rasters for each priority habitat need only be carried out once for each one. The time required for the GIS model in Appendix 2 to run for each priority habitat depends on the number of species records and power of the computer, but it would not be unexpected for processing for each habitat to take large fractions of an hour. Other processing is required after the model has finished to make the scores a fraction of the possible maximum, and to convert the rasters to feature classes with separated 100m square polygons where indicator species are present.

Estimated time: 2 to 3 days

Combine the habitat and species datasets, check the combined data and remove polygons with no score components
This requires geoprocessing to union the Phase 1, NVC and species data together. If data is available similar to the Potential Species-rich Grassland dataset used in this project, this can also be combined at this point, but will probably require processing to smooth fine-scale serrations (otherwise excessive very small polygons will be produced when combined with the other data). The geoprocessing may not be quick if the datasets are large. Polygons with no Phase 1 habitat appropriate to a priority habitat, no NVC data, no species data and (if used) no inclusion in a Potential Species-rich Grassland dataset can be removed at this point. This process will generate numerous very small or very thin polygons (much smaller than the minimum mappable unit of the component datasets) which should be merged with adjacent polygons using the eliminate tool or similar (this will also prevent the number of polygons becoming unmanageable).
Estimated time: 0.5 to 1 day

Calculating final scores and categories of habitat potential for each priority habitat.
This first requires that three attributes are added for every priority habitat: one for the final score, one for presence of accepted priority habitat (from NVC or, for bog habitat, other sources) and one for categorisation of priority habitat potential as 'Low', 'High', etc. The weightings for the final scores used for this project may or may not be applicable depending on whether the same types of dataset were used.
Estimated time: 1 day if the formulae and bandings used for this project are suitable and can be copied; 2 to 3 days if revising the formulae and bandings for each priority habitat to suit the available data.

The above estimations suggest that similar projects (assuming they do not undertake the data comparisons undertaken for this project, and largely copy the methodology) will require around 10 to 15 days depending on the complexity and number of datasets used. This does not include reporting time, nor any contingency time to cover complications that may arise. It also does not cover use of BAP Broad Habitat data from the LCM 2007 dataset. If this is used, as we used it, to fill gaps with no Phase 1 data and for replacing Phase 1 scattered habitats with no underlying habitat, then at a few days more time would be required.

The above time estimation also excludes time for obtaining data, and this may be appreciable. To minimise this aspect with most probably little impact on the project, data requests would be best restricted to SNH, local record centres and the NBN. Data in the BSBI's databases are also on the NBN and often with local records centres, so it should not normally be necessary to contact the BSBI for species records. If there is no relevant local record centre or they do not hold BSBI data, the NBN can be supplied with a custom list of species and will extract data for the required area, though they may require some time to do this. Data from the British Bryological Society may or may not be on the NBN or held in local record centres, and it would be prudent to check with them. For greatest efficiency we would strongly recommend that all required data is obtained *before* a project is begun, and that a generous amount of time is allowed for obtaining it beforehand.

9.2.2 General recommendations for similar projects

With regard to the LCM 2007 data, we would recommend that it is used as a supplement to Phase 1 habitat data and not as the only habitat layer. This is because the minimum mappable unit is large at 0.5ha meaning that more patches of vegetation will be merged into one, and moreover there is likely to have been greater investment of effort, checking and incorporation of other information for Phase 1 habitat data derived from imagery analysis for a specific region (such as the Tweed Catchment Phase 1). For example, the LCM 2007 data did not include any calcareous grassland in Borders or East Dunbartonshire, reflecting the difficulty in identifying it from imagery analysis and the small patch size of some calcareous grassland, whereas there were calcareous grassland areas in the Tweed Catchment Phase 1 presumably reflecting incorporation of knowledge additional to that resulting from imagery analysis. Calcareous grassland patches in the LCM 2007 data would mostly likely have been labelled as 'Rough low-productivity grassland' or have been subsumed into dominant surrounding habitats such as 'Acid grassland'.

A possible alteration to the methodology used here would be to incorporate accepted priority habitat (from NVC or, for bog, other sources) into the scoring process. To do this, a large number could be added to the final score where this is accepted priority habitat, such that final scores could not reach a certain magnitude unless accepted priority habitat was present. In this case, for non-bog habitats where the NVC is deemed reliable, the NVC data could replace the Phase 1 data completely because it can identify the relevant Annex I and BAP Priority habitats with high certainty. For priority bog habitats, the Phase 1 data would always need to be retained to aid distinction of raised/blanket and active/degraded bog, since NVC type cannot do this.

Where Phase 1 habitat data exists that is derived from field survey, we would normally recommend that it replaces Phase 1 habitat data derived from imagery analysis, unless the latter has made use of the former to refine habitat attribution. One of the positive arguments for imagery analysis is that habitat boundaries tend to be more accurate than those from field surveys. Whilst this is generally true[30] if such boundaries are perceivable to the imagery analysis, a major advantage for a field survey is that even if it does result in somewhat inaccurate habitat boundaries, it is more likely to pick up small Annex I and BAP Priority habitats and to correctly identify them. Generally, a small error in area or location of an Annex I or BAP Priority habitat is preferable to being unaware of its presence at all.

With regard to the Potential Species-rich Grassland dataset and similar data that might be available elsewhere, such data has the disadvantage compared to a Phase 1 dataset that it does not distinguish habitats beyond degree of wetness and heterogeneity/productivity. This makes it less useful (for the purpose here) because a given area of heterogeneity/wetness could be suitable for many Annex I / BAP Priority habitats. Thus it was only used to add further general potential to final habitat scores, and its contribution for this project was limited because of the reliability issues discussed above. However, the Potential Species-rich Grassland dataset was still able to distinguish areas of slightly higher priority habitat potential, and sometimes patches within otherwise zero potential habitats, including patches around or within polygons categorised as improved grassland or arable in the Tweed Catchment Phase 1. It is not possible to tell which dataset was more correct (this could be checked during field sampling), but techniques used to develop the Potential Species-rich Grassland dataset might be used to improve imagery analysis for Phase 1 habitats, although the errors noted above in attribution of 'wet' habitat suggest that methods used for the Potential Species-rich Grassland dataset also need improvement. Given the choice between commissioning future imagery analysis for potential species-rich data or Phase 1 habitat data, it would probably be a better use of resources (but more expensive) to commission a Phase 1 habitat analysis, and for this to incorporate techniques and lessons from the Potential Species-rich Grassland project where these might improve Phase 1 attribution.

---

[30] The difference should be slight if the field survey was undertaken recently using aerial photography (preferably in the field) superimposed with Ordnance Survey mapping. This is the method we commonly use for site surveys. It is normally more accurate than using typical GPS units to locate features/boundaries as long as those features/boundaries are detectable in the aerial photograph.

It is probably only practicable to use 100m or better resolution species data. Using 1km resolution species records tends to reduce the clarity with which higher potential priority habitat patches are identified, since much wider areas will be given higher indicator species scores, causing potential priority habitat patches to be enlarged well beyond (in most cases) the real-world priority habitat patch size. It is certainly not recommended that tetrad or hectad resolution records are used because the same reasons apply to a much worse extent. If the proportion of 10m/1m resolution records is significant they could be treated separately from the 100m records, for example to make 20m resolution rasters (using a 20m grid will help to account for GPS error that will be frequent at 10m/1m resolution); however, this will have two important consequences: i) it will double the amount raster processing required; and ii) 20m species scores could be lower than the surrounding 100m squares where there are more 100m species records than 10m/1m records.

A possible way to accelerate the process would be to not use the indicator species score method at all, but only to use the key species. There are far fewer key species and their application is simpler by the method we used, which only requires indication of presence or absence of any key species, and that this component is weighted appropriately in producing the final score.

There remains the issue that real-world habitat patches from which indicator species records originate will sometimes extend outside the grid square of the species records, and at other times will exist only as small patches within that square. In theory this problem would be solved if habitat data were so reliable that there was high certainty of small habitat patches being mapped and habitat identification being correct, since this would allow all irrelevant habitat areas within species record grid squares to be eliminated, but in practice this will never occur except for localised NVC surveys.

As noted above, some priority habitats have similar distributions of habitat potential, primarily because they converge floristically and have indicator species in common. This suggests that there may be little point in attempting to separate certain priority habitats by the means used in this project, except where there are key species present. Situations where this applies include:

- Calcareous grasslands: it would probably be best to calculate potential for Lowland Calcareous Grassland but not the component H6210 and H6230 Annex I habitats, which show very similar habitat potential, except where key species are present that are much more likely in H6120 than H6230 (see above). The proportions of H6210/H6230 could be established by field sampling.
- Neutral/acid grasslands: separation of potential for priority neutral grassland types (corresponding to MG3 and MG5) is probably not worthwhile unless there are key species present to distinguish potential more clearly. Given the close floristic convergence of some MG5 with acid grassland (mainly U4) the potentials for Lowland Meadow and Lowland Dry Acid Grassland can also be similar.
- Wet grasslands: potential for H6410 (equating to M24/M26, which are very rare in Scotland) is often similar to Purple Moor-grass & Rush Pasture and probably not worth separating unless key indicators are present (although even if key species are recorded there is greater likelihood of species-rich M23/MG5 since these can rarely contain these species).

Other issues for future similar work are:

- There is potential for H7150 wherever suitable peatland occurs but high potential (which in this case equates to virtual certainty) requires knowledge of presence of *Rhynchospora* spp., and the quantity of H7150 in a given area will be difficult to estimate because it exists at the periphery of bog pools which are often small and generally not mapped.

- H7220 is very difficult to deal with because the springs which comprise it are usually too small to have been mapped and the key species can occur in other fen or on wet rock.
- Distinction of raised/blanket bog and degraded/active status requires both Phase 1 information *and* reference to other sources such as the Raised Bog Inventory, wetland strategies, etc. (Phase 1 modified bog does not distinguish whether raised/blanket, and NVC type cannot reliably make these separations).

9.2.3 Addition of EUNIS categories

EUNIS habitat categories could be added to polygons in the dataset. However, since the dataset only gives the *potential* for the priority habitats, it can also only give the potential for EUNIS habitat categories, and cannot state that a given EUNIS habitat is certainly present or absent. The exceptions to this are those areas accepted as being priority habitat because they contain corresponding constituent NVC communities or, for bogs, because there is supporting information from other sources. A problem for polygons with NVC data is that mosaics are often recorded which will often mean that more than one EUNIS category applies (e.g. U4/CG10/M23 would represent three EUNIS categories – types of acid and calcareous grassland, and rush-pasture).

9.2.4 Field sampling recommendations

Derivation of Phase 1 habitats from aerial/satellite imagery analysis should ideally be supplemented by field surveys to: i) check certain habitats that have been located and classified as potentially notable by imagery analysis but cannot be identified reliably by that means alone (in particular non-improved grasslands); and ii) locate and identify habitat patches whose compactness is beyond the lower resolution of imagery analysis (notably flushes, springs and small non-improved grassland patches). It may, for smaller study areas, be feasible to check the majority of habitat belonging to the first group, but it is unlikely to ever be feasible to search all appropriate habitat in sufficient detail to find every small calcareous grassland patch or flush. Whilst some species records can indicate high likelihood of certain priority habitats, or at least increase their potential, this depends on recording effort which is biased to more easily-accessible areas and known 'hotspots', and such records are of little use for calculating priority habitat area which could be much larger or much smaller than the records' spatial resolutions. This means that notable small habitats will be (potentially greatly) underestimated in terms of both area and number, and some habitats (particularly non-improved grasslands) will be frequently misclassified, if Phase 1 habitat data is derived largely from imagery analysis and not supplemented by field sampling.

A possible means of estimating how many and what area of lowland priority habitats there are in a given region, including small patches, would be to thoroughly search a statistically valid number of randomised grid squares (e.g. 1km or 100m). If a sampling strategy is devised which incorporates the habitat potentials produced for this project (from very low to very high) for the various priority habitats, this could determine the reliability of those habitat potentials; sampling of areas without priority habitat should be included to establish how much priority habitat occurs outside of the currently mapped habitat potentials. The sampling strategy might be made more accurate by also stratifying according to dominant habitat (e.g. improved grassland/arable-dominated and unimproved grassland/moorland-dominated areas), and perhaps also according to altitude (e.g. 0-100m, 100-200m, 200-300m), since the types and quantity of priority habitat will differ according to these factors (many priority grassland/wetlands would be more common in the 200-300m zone and in areas not dominated by improved grassland or arable Phase 1 habitats). This would need to be undertaken with sufficient effort that all or most occurrences of small priority habitat patches within the sample squares were accurately mapped and identified, to minimise error extrapolating to regional scale.

The dataset could be processed to show the potentials for the relevant priority habitats in grid form (converting to raster), for which a grid resolution of 100m would probably be most apt (corresponding to the indicator species grid resolution). Merging polygons into a regular grid for this purpose would entail loss of the precise spatial distribution of habitat details (such as recorded Phase 1 habitat and NVC) and the precise area of habitat potentials. However, simplifying the data in this way could be helpful for sampling purposes, and would make it visually more obvious where small higher habitat potentials exist (since small higher habitat potentials within a grid square would be transferred to the whole grid square).

9.2.5 Other general recommendations

It would be very helpful for other habitat work to undertake the following procedures on the SNH NVC data: i) standardise the mosaics, ensuring they follow SNH GIS protocol (i.e. 'x/y/z (nx/ny/nz)' where x/y/z are the mosaic components and nx/ny/nz the respective percentage covers, in exactly that format); and ii) provide an attribute indicating where polygons are from NVC site surveys that overlap other NVC site surveys, and another attribute giving the survey years of such overlaps, so that, if desired, the older overlapping site surveys can be quickly ignored or removed.

It would also be very useful to produce a digital version of the Scottish data in the Raised Bog Inventory, by creating a point shapefile with the relevant data. This would avoid other projects having to do this with raised bog data applying to their region of interest. If the Inventory only exists as a scanned document and not an editable format then this will require manual data entry. Preferably also check the data as they are now nearly twenty years old: active/degraded status may now differ from the descriptions in the Inventory, and there are some omissions (a few small or degraded sites in the Tweed Wetland Strategy were not in the Raised Bog Inventory). Definition of active/degraded status also appears to be different in the Inventory to current Annex I guidance31 published three years after it (for example, the Inventory places burnt, dry and drained bogs in the 'active' category but Annex I guidance requires that active bogs support 'a significant area of vegetation that is normally peat-forming' which may not be present on burnt, dry or drained bogs, and which may be better classed as 'degraded' – Phase 1 data, SSSI citations and other sources can help in this respect).

Regarding use of geology and soil data in habitat projects, it would be helpful to obtain access to better resolution data, if it exists. 1:625,000 geology data is too crude to show small outcrops. Better resolution geology data would be particularly useful for locating small outcrops of basic rock, e.g. for locating potential small patches of calcareous grassland.
1:250,000 soil data is rather coarse and may lead to oversimplification or misclassification of habitats. In particular there may be some unreliability using 1:250,000 soil data to assign 'wet' or 'dry' categories of grassland (see above).

[31] *Interpretation Manual of European Habitats* (European Commission DG Environment, 1999)

10. REFERENCES

Atherton, I., Bosanquet, S. & Lawley, M. 2010. *Mosses and Liverworts of Britain and Ireland: a field guide.* British Bryological Society.

Averis, A.M., Averis, A.B.G., Birks, H.J.B., Horsfield, D., Thompson, D.B.A. & Yeo, M.J.M. 2004. *An Illustrated Guide to British Upland Vegetation*. JNCC, Peterborough.

Averis, B. 2013. *Plants and Habitats*. Ben Averis.

Braithwaite, M.E. 2013. *A Botanical Tour of Berwickshire*. M.E. Braithwaite, Hawick.

Dadds, N.J. & Averis, A.B.G. 2014. The extent and condition of species-rich lowland grasslands in Scotland*. Scottish Natural Heritage Commissioned Report No. 571.*

Environment Systems Ltd. in press. Locating potential species-rich grassland areas using remote sensing techniques. *Scottish Natural Heritage Commissioned Report No. 906.*

European Commission 2007. *Interpretation Manual of European Union Habitats.* European Commission DG Environment.

Joint Nature Conservation Committee, 2010. *Handbook for Phase 1 habitat survey.* JNCC, Peterborough.

Lindsay, R. & Immirzi, P. 1996. An inventory of lowland raised bogs in Great Britain. *Scottish Natural Heritage Research, Survey and Monitoring Report No. 078.*

Magee, M., Badenoch, C. & Bisset, N. ed. 2010. *Tweed Wetland Strategy.* Tweed Forum, Melrose.

Medcalf, K.A. & Williams, J. 2010. *Tweed Aerial Survey Phase 2 Aerial Photography Interpretation Land Cover Classification & Habitat Mapping*. Scottish Borders Council & Tweed Forum Consortium.

Medcalf, K.A., Parker, J.A., Turton, N. & Finch, C. 2011. *Making Earth Observation Work for UK Biodiversity Conservation – Phase 1.* Report to the JNCC and Defra.

Morton, D., Rowland, C., Wood, C., Meek, L., Marston, C., Smith, G., Wadsworth, R. & Simpson, I.C. 2011. *Final Report for LCM2007 – the new UK Land Cover Map*. Centre for Ecology & Hydrology (Natural Environment Research Council).

Preston, C.D., Pearman, D.A. & Dines, T.D. 2002. *New Atlas of the British and Irish Flora.* Oxford University Press.

Rodwell, J.S. ed. 1991. *British Plant Communities Volume 2 Mires and heaths*. Cambridge University Press.

Rodwell, J.S. ed. 1992. *British Plant Communities Volume 3 Grassland and montane communities.* Cambridge University Press.

Rodwell, J.S. ed. 1995. *British Plant Communities Volume 4 Aquatic communities, swamps and tall-herb fens*. Cambridge University Press.

Rodwell, J.S. ed. 2000. *British Plant Communities Volume 5 Maritime communities and vegetation of open habitats.* Cambridge University Press.

Rodwell, J.S., Dring, J.C., Averis, A.B.G., Proctor, M.C.F., Malloch, A.J.C., Schaminée, J.N.J., & Dargie T.C.D. 2000. *Review of the coverage of the National Vegetation Classification.* JNCC Report No. 302.

Stace, C.A. 1997. *New Flora of the British Isles*, 2nd edition. Cambridge University Press.

Streeter, D., Hart-Davies, C., Hardcastle, A., Cole, F. & Harper, L. 2009. *Collins Flower Guide.* HarperCollins, London.

APPENDIX 1: FINAL SCORE DERIVATION FOR THE PRIORITY HABITATS

H6210

The weights for the scores of the different contributing factors were reached by considering what the final scores would be with factors alone or in combination, giving sufficient emphasis to those that we deemed more critical. The Phase 1 habitat, if appropriate (i.e. calcareous grassland), is very important and was treated accordingly, but the steep grass factor is less so because although it suggests unimproved grassland it does not indicate what type. We had less confidence in the Potential Species-rich Grassland dataset but allowed it to have a maximum score of '1' with the potential to shift scores to a higher band. The species score derived from the indicator species records was set to achieve a maximum of '4' (multiplying by 8 since the maximum achieved was 0.5), which is of similar magnitude to but lower than the Phase 1 and is intended to reflect the fact that species scores for different calcareous grasslands are similar and can also be similar to species scores for richer acid and neutral grasslands. However, the key species for this habitat (although scarce) more strongly suggest that it is present, so this factor was set so that if it was present the final score would be high even without habitat scores. The formula we arrived at to calculate final scores for H6210 was:

*(appropriate Phase1 * 5) + (steep Phase 1 grass * 3) + (Potential Species-rich Grassland score [0, 0.33, 0.67 or 1]) + (indicator species score * 8) + (key species presence * 9)*

The following table shows what the scores from this formula are for one factor alone or two factors.

	Phase 1	Steep grass	Spp-rich dataset	Species score	Key species
Phase 1	5				
Steep grass	8	3			
Spp-rich dataset	Min 5.3 Max 6	Min3.3 Max 4	Min 0.3 Max 1		
Species score	Min 5.001 Max 9	Min 3.001 Max 7	Min 0.301 Max 5	Min 0.001 Max 4	
Key species	14	12	Min 9.3 Max 10	Min 9.001 Max 14	9

This weighting means that scores higher than 7.000 from only two factors must have key species, or be Phase 1 calcareous grassland with either steep slope or very high species score, so polygons achieving this were labelled 'High'. Scores higher than 4.000 but smaller than 7.000 from one or two factors would include several slightly lower potential conditions including polygons with Phase 1 calcareous grassland but no other factor, and combinations of steep grass with indicator species, and these were labelled 'Moderate'. Scores higher than 1.000 but lower than 4.000 include areas with a moderate species score but no key species, and areas with a higher score from the species-rich dataset but no other factors, and were labelled 'Low'. Polygons scoring up to 1.000, which would include areas with only scores from the species-rich dataset, were labelled 'Very Low', except for zero scores which received no label. We also defined a 'Very High' label for polygons scoring higher than 11.000, which would mean that (from two factors) as a minimum there would have to be key species present together with Phase 1 calcareous grassland, steep grass or a high species score.

H6230

This habitat is a type of calcareous grassland like H6210, so the formula used to calculate the final scores was the same with the exception that key species were not included, because none could be defined for H6210 (see above – no species is significantly more associated with this habitat than H6210 or in some cases forms of acid and neutral grassland). To keep the same maximum of 4 for the species score the weighting factor needed to be 10 since the maximum achieved was 0.4. The formula was therefore:

*(appropriate Phase1 * 5) + (steep Phase 1 grass * 3) + (Potential Species-rich Grassland score [0, 0.33, 0.67 or 1])+ (indicator species score * 10)*

The absence of key species for this habitat ruled out the 'Very High' category, but the other categories were the same as for H6210.

Lowland Calcareous Grassland

This habitat comprises H6210 and H6230, and because it is phytosociologically wider the key species incorporate those for H6210 plus others that have a high likelihood of occurring in lowland calcareous grassland in general (see above). The key species are weighted highly so that where present the score will be at least 'High', and 'Very High' if combined with Phase 1/steep grass/high species score. The species score was set to a maximum of 4 as for H6210/H6230 by multiplying by an appropriate factor (10). The formula is thus similar to H6210:

*(appropriate Phase1 * 5) + (steep Phase 1 grass * 3) + (Potential Species-rich Grassland score [0, 0.33, 0.67 or 1]) + (indicator species score * 10) + (key species presence * 9)*

	Phase 1	**Steep grass**	**Spp-rich dataset**	**Species score**	**Key species**
Phase 1	5				
Steep grass	8	3			
Spp-rich dataset	Min 5.3 Max 6	Min 3.3 Max 4	Min 0.3 Max 1		
Species score	Min 5.001 Max 9	Min 3.001 Max 8	Min 0.301 Max 5	Min 0.001 Max 4	
Key species	14	12	Min 9.3 Max 10	Min 5.001 Max 13	9

VL 0 – 1 (species-rich only or low species score only)
L >1 – 4 (includes steep grass only)
M >4 – 7 (includes Phase 1 alone or steep grass/high spp.-rich plus species score)
H >7 – 11 (key species alone, steep grass plus good species score)
VH >11 (key species plus Phase 1/steep grass/high species score or combination)

H6410

To ensure *Molinia* alone remained in the 'very low' category, along with the species-rich dataset alone, *Molinia* was set to score only 1 if present and the 'very low' band was expanded up to 2. This is because although *Molinia* is necessary in H6410, it is more likely in the lowlands to occur in non-Annex I PMRP. Appropriate Phase 1 habitat was reduced to 3 because the possible habitats encompass a wide range of other habitats and do not indicate H6410 in same way as, for example, calcareous grassland obviously indicates Lowland Calcareous Grassland. The species score was reduced to the same maximum level as the Phase 1. The key species were also reduced in value because although they must generally be present in H6410 they also occur more often elsewhere, but were regarded as slightly more important than the Phase 1 or general species score. The final score formula was therefore:

*(appropriate Phase1 * 3) + (Potential Species-rich Grassland score [0, 0.33, 0.67 or 1]) + (indicator species score * 6) + (key species presence * 4) + (Molinia presence * 1)*

	Phase 1	Spp-rich dataset	Species score	Key species	Molinia
Phase 1	3				
Spp-rich dataset	Min 5.3 Max 4	Min 0.3 Max 1			
Species score	Min 5.001 Max 6	Min 0.301 Max 4	Min 0.001 Max 3		
Key species	7	Min 4.3 Max 5	Min 4.001 Max 7	4	
Molinia	4	Min 1.3 Max 2	Min1.001 Max 4	5	1

VL 0 – 2 (ensures species-rich data/Molinia alone or together still score very low)
L >2 – 4 (Phase 1 / species score alone on with Molinia)
M >4 – 6 (Phase 1 / species score plus key species)
H >6 – 9 (Species score / species-rich plus key species)
VH >9 (requires more than two factors)

Purple Moor-grass & Rush Pasture

This formula and scoring for this habitat was similar to H6410. As for H6410, the appropriate Phase 1 habitats could encompass non-PMRP habitats such as MG10 and fens, so it was kept at the same level as for H6410. The species score was regarded as slightly more important than the Phase 1 and was scored accordingly. The key species include *Molinia* as well as *Juncus acutiflorus* and *Carum verticillatum*, and although *Molinia* also occurs in wet heaths and bogs, it is more likely in the lowlands to occur in some form of PMRP because this habitat is much more common, and this is also balanced by the fact that bog/heath receive no score from the Phase 1 for this habitat; it was scored slightly higher than the general species score. The final score formula was therefore:

*(appropriate Phase1 * 3) + (Potential Species-rich Grassland score [0, 0.33, 0.67 or 1]) + (indicator species score * 8) + (key species presence * 4)*

	Phase 1	Spp-rich dataset	Species score	Key species
Phase 1	3			
Spp-rich dataset	Min 5.3 Max 4	Min 0.3 Max 1		
Species score	Min 5.001 Max 7	Min 0.301 Max 5	Min 0.001 Max 4	
Key species	8	Min 5.3 Max 6	Min 5.001 Max 9	5

VL 0 – 1 (species-rich only or low species score only)
L >1 – 4 (Phase 1 / species score alone)
M >4 – 6 (species-rich plus key species, key species alone)
H >6 – 8 (Phase 1 plus species score)
VH >8 (Phase 1/species score plus key species or more than two factors)

H6520 / Upland Hay Meadow

This habitat is scarce, occurring only in unimproved neutral grassland and normally where *Geranium sylvaticum* is present, or *Cirsium heterophyllum* or *Trollius europaeus* instead. The formula for this habitat was therefore weighted so that 'Very High' potential would only occur if both the Phase 1 were appropriate (unimproved neutral grassland) and key species were recorded. Since unimproved grassland will commonly be something other than MG3, the banding was set so that Phase 1 alone would be 'Low'. The key species were set to a maximum of 5; setting them to higher possible values resulted in too much woodland being given high scores (*G. sylvaticum* is common in some woodlands). The species score was set to be a maximum of 3 by multiplying by 5. The formula was:

*(appropriate Phase1 * 5) + (Potential Species-rich Grassland score [0, 0.33, 0.67 or 1]) + (indicator species score * 5) + (key species presence * 5)*

	Phase 1	Spp-rich dataset	Species score	Key species
Phase 1	5			
Spp-rich dataset	Min 5.3 Max 6	Min 0.3 Max 1		
Species score	Min 5.001 Max 6	Min 0.301 Max 4	Min 0.001 Max 3	
Key species	10	Min 5.3 Max 6	Min 5.001 Max 8	5

VL 0 – 1 (species-rich only or low species score only)
L >1 – 5 (includes Phase 1 only)
M >5 – 6 (includes Phase 1 plus species score/species-rich)
H >6 – 9 (key species plus good species score)
VH >9 (minimum Phase 1 plus key species)

Lowland Meadow
This habitat is another type of unimproved neutral grassland and was treated in a similar way to Upland Hay Meadow. As for UHM, not all unimproved neutral grassland is relevant so the banding was similarly set so that Phase 1 alone would be 'Low'. There are two types of key species for this habitat, corresponding to the two constituent NVC communities: those that are most likely (in lowland Scotland) to occur in MG5, and one species (*Caltha palustris*) that is required in MG8 but is more common in other habitats. Since the former (although rare) we regarded as highly indicative of MG5, we weighted this factor highly but ensuring that the highest scores were not likely without also having unimproved neutral grassland. The *Caltha* factor could only be given a low weight because it is more common outside Lowland Meadow. The formula was therefore similar to that for Upland Hay Meadow but with two key species components, appropriately weighted, and the general indicator species score weight adjusted to achieve the same maximum of 3. We also added in the factor of steep grass which will help to identify better quality habitat:

*(appropriate Phase1 * 5) + (steep Phase 1 grass *3) + (Potential Species-rich Grassland score [0, 0.33, 0.67 or 1]) + (indicator species score * 6) + (key species [not Caltha] presence * 7) + (key species [Caltha] presence * 1)*

	Phase 1	Steep grass	Spp-rich dataset	Species score	Key spp. (MG5)	Key sp. (Caltha)
Phase 1	5					
Steep grass	8	3				
Spp-rich dataset	Min 5.3 Max 6	Min 3.3 Max 4	Min 0.3 Max 1			
Species score	Min 5.001 Max 6	Min 3.001 Max 7	Min 0.301 Max 4	Min 0.001 Max 3		
Key spp. (MG5)	12	10	Min 9.3 Max 8	Min 9.001 Max 10	7	
Key sp. (Caltha)	6	4	Min 1.3 Max 2	Min 1.001 Max 4	8	1

VL 0 – 1 (species-rich only or low species score only)
L >1 – 5 (includes Phase 1 only)

M	>5 – 7 (includes Phase 1 plus species score/species-rich/Caltha)
H	>7 – 12 (MG5 key species)
VH	>12 (minimum MG5 key species, necessitating species score, plus Phase 1)

Lowland Dry Acid Grassland
The formula for LDAG was similar to Lowland Calcareous Grassland and Lowland Meadow, but with reduced key species weighting because the key species for LDAG are not quite as effective as for LCG / Lowland Meadow.

*(appropriate Phase1 * 5) + (steep Phase 1 grass * 3) + (Potential Species-rich Grassland score [0, 0.33, 0.67 or 1]) + (indicator species score * 5) + (key species presence * 5)*

	Phase 1	Steep grass	Spp-rich dataset	Species score	Key species
Phase 1	5				
Steep grass	8	3			
Spp-rich dataset	Min 5.3 Max 6	Min 3.3 Max 4	Min 0.3 Max 1		
Species score	Min 5.001 Max 6	Min 3.001 Max 7	Min 0.301 Max 4	Min 0.001 Max 3	
Key species	10	8	Min 5.3 Max 6	Min 5.001 Max 8	5

VL	0 – 1 (species-rich only or low species score only)
L	>1 – 4 (includes Phase 1 / steep grass only)
M	>4 – 7 (includes Phase 1 / steep grass plus species score)
H	>7 – 9 (key species / steep grass plus good species score)
VH	>9 (includes Phase 1 plus key species)

H7110
For this habitat (active raised bog) we were very confident in its identification because we had checked occurrences of all bog types in the Phase 1 data against the Raised Bog Inventory, SSSI citations and Tweed Wetland Strategy, and checked that these other sources were accounted for in the Phase 1. We therefore heavily weighted the first part of the formula. Although we could have used only this aspect to identify H7710, we calculated a formula in a similar fashion to the previous habitats in order to highlight any particularly good areas within active raised bogs resulting from indicator species.

*(confirmed H7110 * 8) + (Potential Species-rich Grassland score [0, 0.33, 0.67 or 1]) + (indicator species score * 6) + (key species indicating non-degradation * 5)*

	Confirmed 7110	Spp-rich dataset	Species score	Key species
Confirmed 7110	8			
Spp-rich dataset	Min 8.3 Max 9	Min 0.3 Max 1		
Species score	Min 8.001 Max 12	Min 0.301 Max 5	Min 0.001 Max 4	
Key species	14	Min 5.3 Max 6	Min 5.001 Max 9	5

VL	0 – 1 (species-rich only or low species score only)
L	>1 – 4 (includes species score only)
M	>4 – 7 (includes species-rich plus species score, key species only)
H	>7 – 9 (confirmed 7110, species score plus key species)
VH	>9 (confirmed 7110 plus species score/key species)

H7120

For this habitat (degraded raised bog) as for H6210 we had checked possible occurrences in the Phase 1 data against the Raised Bog Inventory, SSSI citations and Tweed Wetland Strategy, and checked that these other sources were accounted for in the Phase 1. However, since the Raised Bog Inventory is now 18 years old, we considered that where supposed active raised bog in the Raised Bog Inventory was mapped as modified bog in the Phase 1, and there was no other source of information, then it probably was now degraded. We checked aerial photography to help decide this. In one case the Tweed Wetland Strategy indicated a degraded bog which was not mentioned elsewhere, and the exact location was not fully certain. However, confidence in assignment of degraded bog was generally still high and we again heavily weighted the first part of the formula accordingly. As for active raised bog, we calculated a formula in a similar fashion to the previous habitats in order to highlight areas within degraded raised bogs where indicator species were recorded. The key species component was given a low weighting because the species involved (*Campylopus introflexus*), although favouring bare peat, can also occur in heaths and other habitats.

*(confirmed H7120 * 8) + (Potential Species-rich Grassland score [0, 0.33, 0.67 or 1]) + (indicator species score * 6) + (key species indicating degradation * 1)*

	Confirmed 7120	**Spp-rich dataset**	**Species score**	**Key species**
Confirmed 7120	8			
Spp-rich dataset	Min 8.3 Max 9	Min 0.3 Max 1		
Species score	Min 8.001 Max 12	Min 0.301 Max 5	Min 0.001 Max 4	
Key species	9	Min 1.3 Max 2	Min 1.001 Max 5	1

VL 0 – 1 (species-rich only or low species score only)
L >1 – 4 (includes species score only)
M >4 – 7 (includes species-rich plus species score)
H >7 – 9 (confirmed 7120)
VH >9 (confirmed 7120 plus species score)

H7130 / Blanket Bog

As we had thoroughly checked possible occurrences of raised bog in the Phase 1 data against the Raised Bog Inventory, SSSI citations and Tweed Wetland Strategy, and checked that these other sources were accounted for in the Phase 1, we assumed that all other bog in the Phase 1 was blanket bog. It includes modified blanket bog as well as intact blanket bog. Confidence in assignment of degraded bog was high and we again heavily weighted the first part of the formula accordingly. The formula was similar to that for active raised bog but the first component was raised slightly and the key species score lowered slightly to avoid giving high scores to active raised bog with key species.

*(confirmed H7130 * 8) + (Potential Species-rich Grassland score [0, 0.33, 0.67 or 1]) + (indicator species score * 6) + (key species indicating non-degradation * 5)*

	Confirmed 7130	Spp-rich dataset	Species score	Key species
Confirmed 7130	10			
Spp-rich dataset	Min 10.3 Max 11	Min 0.3 Max 1		
Species score	Min 10.001 Max 14	Min 0.301 Max 5	Min 0.001 Max 4	
Key species	15	Min 5.3 Max 5	Min 5.001 Max 8	4

VL 0 – 1 (species-rich only or low species score only)
L >1 – 4 (includes species score only)
M >4 – 9 (includes species-rich plus species score)
H >9 – 14 (confirmed 7130)
VH >14 (confirmed 7130 plus species score/key species)

H7150

This habitat, which is most frequent in wet bogs, was treated differently to the other bog habitats because it is best identified by the presence of *Rhynchospora* spp., which were used as the only indicator species for this habitat. The Phase 1 was weighted less because in most cases within the study area (and frequently elsewhere) appropriate Phase 1 habitats lack *Rhynchospora* and are therefore not H7150. The species score was weighted highly because it is effectively also the key species score. The 'Moderate' banding was not used for this habitat; 'High' or 'Very High' indicates H7150 and 'Low' indicates potentially correct habitat but no known *Rhynchospora*. The formula was:

*(appropriate Phase 1 * 5) + (Potential Species-rich Grassland score [0, 0.33, 0.67 or 1]) + (indicator species score * 20)*

	Phase 1	Spp-rich dataset	Species score
Phase 1	5		
Spp-rich dataset	Min 5.3 Max 6	Min 0.3 Max 1	
Species score	14.8	Min 10.1 Max 10.8	9.8 if Rhynchospora present

VL 0 – 4 (species-rich only)
L >4 – 7 (Phase 1 with/without species-rich)
H >7 – 11 (*Rhynchospora* with/without species-rich)
VH >14 (*Rhynchospora* plus Phase 1)

Lowland Raised Bog

This habitat combines both H7110 active raised bog and H7120 degraded raised bog. The formula used was identical to that used for H7110 except that the first component referred to confirmed Lowland Raised Bog in general rather than only active raised bog. The species score used was the H7110 species score.

*(confirmed LRB * 8) + (Potential Species-rich Grassland score [0, 0.33, 0.67 or 1]) + (indicator species score * 6) + (key species indicating non-degradation * 5)*

	Confirmed LRB	Spp-rich dataset	Species score	Key species
Confirmed LRB	8			
Spp-rich dataset	Min 8.3 Max 9	Min 0.3 Max 1		
Species score	Min 8.001 Max 12	Min 0.301 Max 5	Min 0.001 Max 4	
Key species	14	Min 5.3 Max 6	Min 5.001 Max 9	5

VL 0 – 1 (species-rich only or low species score only)
L >1 – 4 (includes species score only)
M >4 – 7 (includes species-rich plus species score, key species only)
H >7 – 9 (confirmed 7110, species score plus key species)
VH >9 (confirmed 7110 plus species score/key species)

H7140

Several fen Phase 1 habitats are relevant to H7140 transition mires, but none exclusively indicate it unless key indicators are present, so the Phase 1 is given moderate weighting and the banding set so that Phase 1 alone cannot score higher than 'Low'. There are two sets of key species: those that are not *Carex* species which are more closely associated with H7140 (if not woodland, but woodland will receive no score contribution from habitat), and which are given moderate weighting, and those that are *Carex* spp. which are expected in transition mires but which can occur in non-H7140 swamp, which are given low weighting such that the potential will be 'Very Low' with these *Carex* indicators alone.

*(appropriate Phase1 * 5) + (Potential Species-rich Grassland score [0, 0.33, 0.67 or 1]) + (indicator species score * 5) + (key species [non-Carex] presence * 5) + (key species [Carex] presence * 1)*

	Phase 1	Spp-rich dataset	Species score	Key spp. (not Cx)	Key spp. (Carex)
Phase 1	5				
Spp-rich dataset	Min 5.3 Max 6	Min 0.3 Max 1			
Species score	Min 5.001 Max 8	Min 0.301 Max 4	Min 0.001 Max 3		
Key species (not Carex)	10	Min 5.3 Max 6	Min 5.001 Max 8	5	
Key species (Carex)	7	Min 2.3 Max 3	Min 2.001 Max 5	7	2

VL 0 – 2 (species-rich with/without low species score)
L >2 – 5 (includes Phase 1 only)
M >5 – 7 (includes Phase 1 plus species score/species-rich)
H >7 – 9 (species score plus key species)
VH >9 (Phase 1 plus key species)

H7220

For this habitat, the appropriate Phase 1 habitat (fen) is much less important than the key species, because fen also includes many other vegetation types. It was considered unwise to limit the appropriate Phase 1 to specifically 'spring' because this is normally target-noted and not mapped, and many are likely to go unnoticed during Phase 1 habitat mapping. The formula was set so that 'High' or 'Very High' are only likely if the key species are present. However, note that the key species are not exclusive to this habitat.

*(appropriate Phase1 * 3) + (Potential Species-rich Grassland score [0, 0.33, 0.67 or 1]) + (indicator species score * 6) + (key species presence * 9)*

	Phase 1	**Spp-rich dataset**	**Species score**	**Key species**
Phase 1	3			
Spp-rich dataset	Min 3.3 Max 4	Min 0.3 Max 1		
Species score	Min 3.001 Max 6	Min 0.301 Max 4	Min 0.001 Max 4	
Key species	12	Min 9.3 Max 10	Min 6.001 Max 13	9

VL 0 – 2 (species-rich with/without low species score)
L >2 – 5 (includes Phase 1 only)
M >5 – 7 (Phase 1 plus species score)
H >7 – 9 (key species alone)
VH >9 (key species plus other factor)

H7230
Unlike the previous fen habitats, H7230 has the distinct advantage that there are several species almost exclusively found in this habitat which are not all scarce (though some are), and these key species have been heavily weighted accordingly. Since the appropriate Phase 1 habitats of fen and swamp may or may not contain H7230, the banding was set so that Phase 1 alone could not score 'Moderate'.

*(appropriate Phase1 * 5) + (Potential Species-rich Grassland score [0, 0.33, 0.67 or 1]) + (indicator species score * 6) + (key species presence * 10)*

	Phase 1	**Spp-rich dataset**	**Species score**	**Key species**
Phase 1	5			
Spp-rich dataset	Min 5.3 Max 6	Min 0.3 Max 1		
Species score	Min 5.001 Max 9	Min 0.301 Max 5	Min 0.001 Max 4	
Key species	15	Min 10.3 Max 11	Min 10.001 Max 14	10

VL 0 – 2 (species-rich with/without low species score)
L >2 – 5 (includes Phase 1 only)
M >5 – 7 (includes Phase 1 plus species score/species-rich)
H >7 – 9 (species score plus Phase 1)
VH >9 (key species necessary)

Lowland Fen
The indicator species for this habitat are numerous and combine those of H7140, H7220 and H7230, plus some additional indicators that occur more widely in Lowland Fen but not in these Annex I habitats. The key species are also a combination of those for the Annex I habitats plus *Sedum villosum*. The formula used was the same as for H7230, but the key species were lowered slightly in weight to account for the fact that the H7140 / H7220 key species are not quite as reliable those for H7230. The species score weight was adjusted to achieve the same maximum of 4. The banding was set so that appropriate Phase 1 habitat would be 'Moderate' or higher, because (unlike for H7140, H7220 and H7230), any type of fen or swamp in the lowlands indicates Lowland Fen.

*(appropriate Phase1 * 5) + (Potential Species-rich Grassland score [0, 0.33, 0.67 or 1]) + (indicator species score * 8) + (key species presence * 8)*

	Phase 1	**Spp-rich dataset**	**Species score**	**Key species**
Phase 1	5			
Spp-rich dataset	Min 5.3 Max 6	Min 0.3 Max 1		
Species score	Min 5.001 Max 9	Min 0.301 Max 5	Min 0.001 Max 4	
Key species	13	Min 9.3 Max 9	Min 10.001 Max 12	8

VL 0 – 2 (species-rich with/without low species score)
L >2 – 4 (includes Phase 1 only)
M >4 – 7 (includes Phase 1 with/without species score/species-rich)
H >7 – 9 (species score plus Phase 1)
VH >9 (key species plus Phase 1 / high species score necessary)

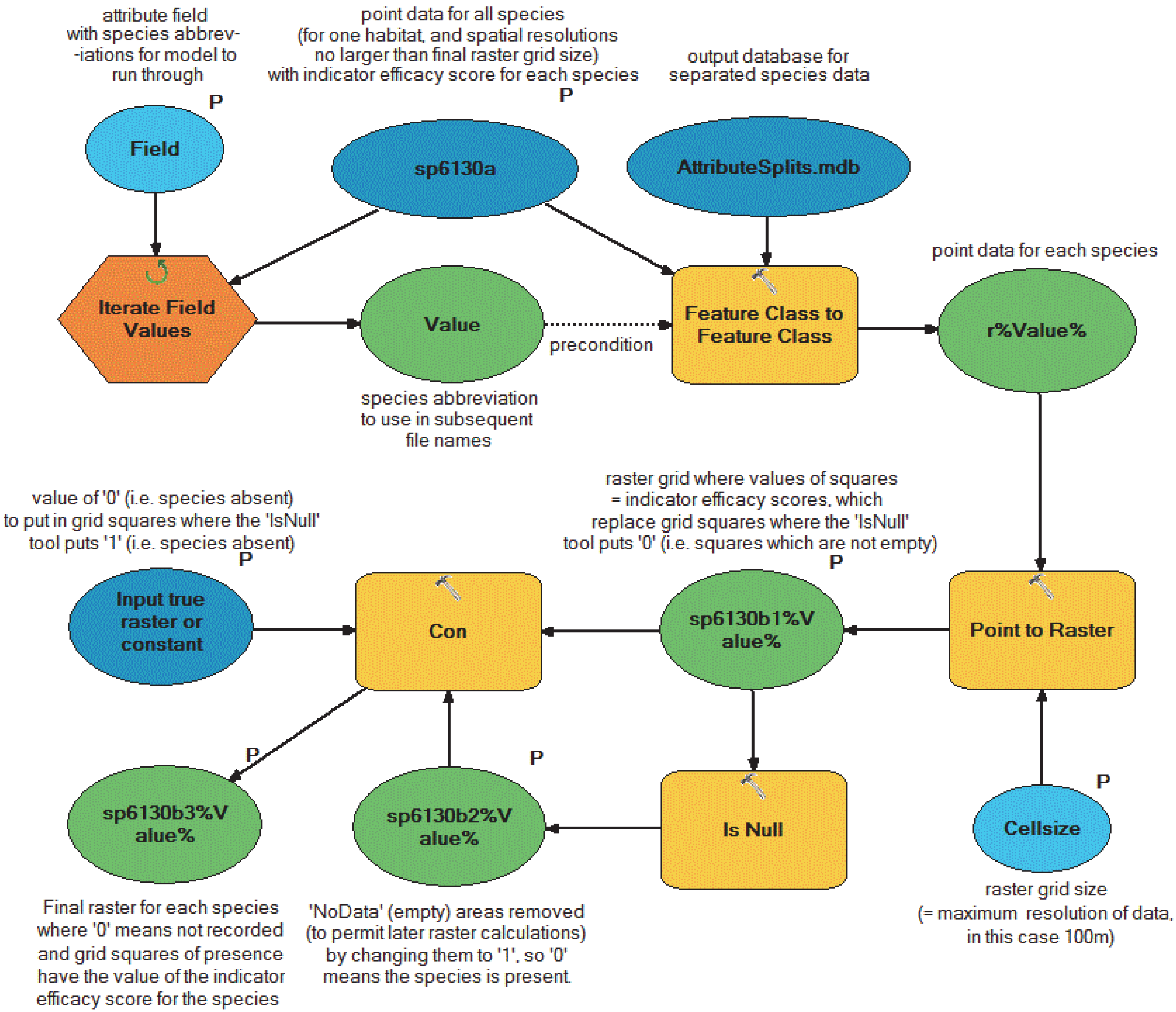
APPENDIX 2: ARCGIS MODEL FOR AUTOMATIC ATTRIBUTE SPLITTING AND RASTER PRODUCTION

The model is shown on the next page. It can be set up with any ArcGIS license level, but version 10 or above is required to use the ‘iterator’. Oval shapes indicate some kind of input and rectangular shapes are tools; the iterator is hexagonal. Bold ‘P’ symbols show that the input is set as a parameter, i.e. it will be visible and editable by the user when the model is opened prior to running it.

This model produces distribution rasters of (in this case) different species at an appropriate grid resolution, where (in this case) each raster grid square for species presence has the value of the indicator efficacy score of the species, or a value of zero where it has not been recorded from a grid square. It is actually only the top half of the model and the ‘Point to Raster’ tool that do this; the rest of the lower half of the model is only to set grid squares of absence to ‘zero’ instead of the default ‘NoData’. This is critical because the latter mathematical manipulation of rasters (adding rasters of appropriate indicator species together for each priority habitat) will not work wherever ‘NoData’ exists.

The resulting raster grid squares need to be correctly aligned with the Ordnance Survey grid, and one way of achieving this is to set a processing extent covering the study area in Geoprocessing/Environments and set the coordinates of at least the lower left corner of this rectangle to a 10km grid line intersection.

Processing time could be substantial (several minutes to a large fraction of an hour or even longer) depending on the number of species and records involved, and the power of the computer.



attribute field with species abbrev- -iations for model to run through
P
Field
point data for all species (for one habitat, and spatial resolutions no larger than final raster grid size) with indicator efficacy score for each species
P
sp6130a
output database for separated species data
AttributeSplits.mdb
Iterate Field Values
Value
precondition
Feature Class to Feature Class
point data for each species
r%Value%
species abbreviation to use in subsequent file names
value of '0' (i.e. species absent) to put in grid squares where the 'IsNull' tool puts '1' (i.e. species absent)
P
Input true raster or constant
Con
raster grid where values of squares = indicator efficacy scores, which replace grid squares where the 'IsNull' tool puts '0' (i.e. squares which are not empty)
P
sp6130b1%Value%
Point to Raster
P
sp6130b3%Value%
P
sp6130b2%Value%
Is Null
P
Cellsize
Final raster for each species where '0' means not recorded and grid squares of presence have the value of the indicator efficacy score for the species
'NoData' (empty) areas removed (to permit later raster calculations) by changing them to '1', so '0' means the species is present.
raster grid size (= maximum resolution of data, in this case 100m)

APPENDIX 3: DATASET ATTRIBUTES

To avoid excessive repetition in the table below, where an attribute is required for each priority habitat the stated prefix is followed by a variable suffix (‘*xxxx*’) corresponding to the Annex 1 code(s) or abbreviation of BAP Priority habitat(s) as follows:

H*nnnn* = Annex I habitat where *nnnn* is the Annex I code number.
CALA = Calaminarian grassland BAP Priority habitat
LCG = Lowland Calcareous Grassland BAP Priority habitat
LDAG = Lowland Dry Acid Grassland BAP Priority habitat
UHM = Upland Hay Meadow BAP Priority habitat
LM = Lowland Meadow BAP Priority habitat
PMRP = Purple Moor-grass & Rush Pasture BAP Priority habitat
LRB = Lowland Raised Bog BAP Priority habitat
BB = Blanket Bog BAP Priority habitat
LF = Lowland Fen BAP Priority habitat

Attribute title & data type	Description
LA *Text* *2 characters*	Two-letter abbreviation of the Local Authority. The possibilities are ‘ED’ or ‘SB’ for East Dunbartonshire and Scottish Borders respectively.
P1BHcode *Text* *100 characters*	Phase 1 habitat alphanumeric code, except for locations without Phase 1 data in East Dunbartonshire where the full written BAP Broad Habitat is given instead (semi-natural BAP Broad Habitats are also given locally in Borders where they replace Phase 1 scattered habitat with no underlying dominant habitat).
P1BHdesc *Text* *100 characters*	Full name of the Phase 1 habitat alphanumeric code in the previous attribute field, or (where applicable – see description of previous attribute) the full BAP Broad Habitat.
P1BHsource *Text* *50 characters*	Source of the Phase 1 habitat or BAP Broad Habitat polygon. The possible sources are: Tweed Catchment Phase 1; EDC Phase 1 (East Dunbartonshire Phase 1); SNH Phase 1 (with survey ID number); OSMM + LCM2007 (only in East Dunbartonshire for areas with no Phase 1 data); OSMM Water theme (only in East Dunbartonshire because of limited Phase 1 data and only retained where species or NVC data present, because open water not relevant to the study); LCM2007 (replacing areas of scattered Phase 1 habitat with no underlying dominant habitat). Where the entry says ‘+ KBE’, this means the habitat type was changed based on information in the attribute fields ‘P1BHkbe’ and ‘NVCkbe’.
P1BHkbe *Text* *100 characters*	Phase 1 Knowledge-Based Enhancement. Description of why changes to the Phase 1 / BAP Broad Habitat type were made or (where relevant) why potential for one or more of the bog priority habitats was assigned
NVCdesc *Text* *100 characters*	Description of the NVC components (where locally applicable). For data whose source is the Lowland Grassland Review 2010-2011, the format follows standard SNH protocol for mosaics, i.e. *x/y/z (xn/yn/zn)* where *x/y/z* are the NVC mosaic components and the numbers in brackets are their respective estimated percentage covers. For data whose source is the SNH NVC Scotland dataset, the format of mosaics is variable as is inclusion of component proportions.
NVCsource *Text* *50 characters*	Source of the NVC data (where locally applicable). The possible sources are ‘SNH NVC Scotland’ (the same data that is freely available on the Natural Spaces section of the SNH website), or ‘Lowland Grassland Review 2010-2011 (Borders)’ which is the data from the Borders surveys undertaken for that project. Where the source is the SNH NVC Scotland data, the survey ID number is also given in brackets.
NVCkbe	NVC Knowledge-Based Enhancement. Description of why (where locally

Text *100 characters*	applicable) potential for one or more of the bog priority habitats was assigned, or whether fen was regarded as more appropriate than bog.
NVCmosaics *Short integer*	Count of the number of NVC mosaic components (‘1’ implying that the NVC is not a mosaic, and ‘0’ that there is no NVC information).
SteepGrass *Short integer*	Contains either ‘1’ or ‘0’ where ‘1’ means that the polygon contains a Phase 1 / BAP Broad Habitat dry grassland and the slope is 20^{o} or steeper (implying an increased likelihood of priority dry grassland as explained in the report; 20^{o} was chosen following comparison of slope with known NVC grassland in Borders).
PSR *Float*	Values of 0.33, 0.67 and 1 corresponding to heterogeneity values in the original Potential Species-rich Grassland dataset of 3, 2, and 1 respectively (where 1 represented the highest potential and 3 the lowest). The values were converted to make logical sense during final score calculations.
spSum *Float*	Sum of the indicator species scores, which if greater than zero shows at a glance that indicator species for certain habitat(s) exist.
kspSum *Short integer*	Sum of the key species values, which if greater than zero shows at a glance that key species for certain habitat(s) exist.
P1*xxxx* *Short integer*	Contains either ‘1’ or ‘0’ where ‘1’ means that the Phase 1 habitat is appropriate for the Annex I or UK BAP habitat abbreviated in the suffix (see above).
P1BHsum *Short integer*	Sum of the P1*xxxx* values, which shows how many priority habitats the Phase 1 habitat is considered appropriate for.
NVC*xxxx* *Short integer*	Contains either ‘1’ or ‘0’ where ‘1’ means that the NVC is appropriate for the Annex I or UK BAP habitat abbreviated in the suffix (see above).
NVCsum *Short integer*	Sum of the NVC*xxxx* values, which shows how many priority habitats the NVC is considered appropriate for.
sp*xxxx* *Float*	Contains the indicator species score for the Annex I or UK BAP habitat abbreviated in the suffix (see above)..
ksp*xxxx* *Short integer*	Contains either ‘1’ or ‘0’ where ‘1’ means that key species are present for the Annex I or UK BAP habitat(s) abbreviated in the suffix (see above). Additional characters are added to the suffix where there is more than one type of key species (scored differently) for the habitat in question; these show NVC type or species.
def*xxxx* *Short integer*	Contains either ‘1’ or ‘0’ where ‘1’ means that the priority habitat abbreviated in the suffix (see above) is considered very likely to be actually present (where NVC exists and corresponds directly to the priority habitat, or where bog type has been confirmed by reference to other information).
fs*xxxx* *Float*	Contains the final score for the priority habitat abbreviated in the suffix (see above), derived by a formula using relevant score component attributes with weightings considered appropriate (see report).
cat*xxxx* *Text* *2 characters*	Contains the category of potential for the priority habitat abbreviated in the suffix (see above), either ‘VL’ (Very Low), ’L’ (Low), ‘M‘ (Medium), ‘H’ (High) or ‘VH’ (Very High), derived from the final score using banding levels considered appropriate (see report).

APPENDIX 4: TABLE OF INDICATOR SPECIES

The table shows the indicator species developed for this project to derive indicator species scores, one of the contributing factors of the final scores. It also shows what habitats the species were considered to occur in (in lowland Scotland), and what the indicator efficacy score of each one was calculated to be using this information. In producing this list of indicators, we tried to avoid those species which are largely upland, those which do not occur in Scotland, and those which we judged were too common or more common in non-relevant grasslands or heaths.

The abbreviated habitat column headings are explained below:

H*nnnn* = Annex I habitat

CALA = Calaminarian grassland BAP Priority habitat
LCG = Lowland Calcareous Grassland BAP Priority habitat
LDAG = Lowland Dry Acid Grassland BAP Priority habitat
UHM = Upland Hay Meadow BAP Priority habitat
LM = Lowland Meadow BAP Priority habitat
PMRP = Purple Moor-grass & Rush Pasture BAP Priority habitat
LRB = Lowland Raised Bog BAP Priority habitat
BB = Blanket Bog BAP Priority habitat
LF = Lowland Fen BAP Priority habitat

BH CG = Calcareous Grassland BAP Broad Habitat
BH NG = Neutral Grassland BAP Broad Habitat
BH AG = Acid Grassland BAP Broad Habitat
BH DSH = Dwarf Shrub Heath BAP Broad Habitat
BH FMS = Fen, Marsh & Swamp BAP Broad Habitat
BH B = Bog BAP Broad Habitat
BH IR = Inland Rock BAP Broad Habitat

npD = non-priority Dunes
npS = non-Annex I swamp
npMH = non-priority moor/heath
npRW = non-priority rock/waste
npNG = non-priority neutral grassland

The last five categories were added to appropriately reduce the scores of indicators that do occur in relevant Annex I/BAP Priority habitats but that were scoring too highly because they can also occur in related non-priority habitat (e.g. several of the bryophyte indicators also occur in non-relevant sand dunes and on non-relevant rock; many indicators that occur in calaminarian grassland also occur on non-calaminarian grassland rock exposures; many acid grassland/bog indicators also occur in non-relevant heath; some indicators of Annex I swamp also occur in Annex I swamp; some indicators of lowland meadow also occur in non-priority neutral grassland).

Species	H6130	H6210	H6230	H6410	H6520	H7110	H7120	H7130	H7140	H7150	H7220	H7230	CALA	LCG	LDAG	UHM	LM	PRMP	LRB	BB	LF	BH CG	BH NG	BH AG	BH DSH	BH FMS	BH B	BH IR	npD	np S	npMH	npRW	npNG	Score
Achillea ptarmica				Y								Y						Y			Y					Y					Y			1.5
Agrostis capillaris																																		0
Aira caryophyllea															Y									Y				Y				Y		1.333
Aira praecox															Y									Y				Y				Y		1.333
Alchemilla filicaulis			Y		Y									Y		Y						Y	Y										Y	1.333
Alchemilla glabra			Y		Y									Y		Y	Y					Y	Y										Y	1.167
Alchemilla glaucescens			Y											Y								Y												3
Alchemilla glomerulans					Y											Y							Y										Y	2.5
Alchemilla wichurae			Y		Y									Y		Y						Y	Y										Y	1.333
Alchemilla xanthochlora			Y		Y									Y		Y	Y					Y	Y										Y	1.167
Anacamptis pyramidalis		Y												Y								Y												3
Anagallis tenella				Y								Y						Y			Y					Y					Y			1.5
Andromeda polifolia						Y		Y											Y	Y							Y							2
Angelica sylvestris				Y					Y			Y						Y			Y					Y					Y		Y	1.167
Antennaria dioica			Y											Y	Y							Y		Y	Y			Y				Y		1.75
Anthoxanthum odoratum																																		0
Anthyllis vulneraria		Y												Y	Y							Y		Y				Y				Y		1.75
Arabis hirsuta		Y	Y											Y	Y							Y		Y				Y				Y		1.25
Astragalus danicus		Y												Y	Y							Y		Y										2
Blindia acuta											Y	Y									Y					Y								2.5
Blysmus compressus												Y									Y					Y					Y			2.5

Botrychium lunaria		Y	Y											Y								Y												2.5
Briza media		Y	Y	Y	Y							Y		Y	Y	Y	Y	Y			Y	Y	Y	Y		Y					Y			0.567
Calamagrostis canescens									Y												Y					Y				Y				2
Calliergon cordifolium				Y					Y									Y			Y					Y								2
Calliergon giganteum				Y					Y									Y			Y					Y								2
Caltha palustris				Y					Y			Y						Y			Y					Y					Y			1.333
Campanula rotundifolia		Y	Y		Y									Y	Y	Y	Y					Y	Y	Y									Y	0.833
Campyliadelphus chrysophyllus			Y											Y								Y						Y	Y			Y		2.25
Campylium stellatum				Y							Y	Y						Y			Y					Y								1.833
Campylopus introflexus						Y	Y	Y											Y	Y					Y		Y							1.833
Carex appropinquata				Y								Y						Y			Y					Y				Y				1.5
Carex caryophyllea		Y	Y		Y									Y	Y	Y	Y					Y	Y	Y									Y	0.833
Carex curta									Y									Y			Y					Y					Y			2
Carex diandra									Y			Y									Y					Y				Y				2
Carex dioica											Y	Y									Y				Y	Y								2.5
Carex disticha				Y														Y			Y					Y					Y			2
Carex echinata				Y					Y			Y						Y			Y					Y					Y			1.333
Carex hostiana				Y								Y						Y			Y					Y								2
Carex lasiocarpa									Y			Y									Y					Y				Y				2
Carex limosa									Y			Y									Y					Y				Y				2
Carex paniculata									Y												Y					Y				Y				2.5
Carex pulicaris				Y								Y						Y			Y					Y					Y			1.5
Carex rostrata									Y			Y									Y					Y				Y	Y			1.833
Carex vesicaria									Y			Y									Y					Y				Y	Y			1.833
Carex viridula											Y	Y						Y			Y				Y	Y					Y			1.5
Carum																		Y								Y					Y			1.5

verticillatum																																		
Cicuta virosa									Y												Y					Y				Y				2.5
Cinclidium stygium												Y									Y					Y								3
Cirsium dissectum				Y														Y								Y								3
Cirsium heterophyllum				Y	Y											Y		Y					Y			Y								1.5
Cladium mariscus				Y					Y									Y			Y					Y					Y			1.5
Clinopodium vulgare		Y												Y								Y	Y										Y	2.333
Coeloglossum viride		Y	Y		Y									Y		Y	Y					Y	Y										Y	1
Conopodium majus					Y										Y	Y	Y						Y	Y									Y	1.667
Cratoneuron filicinum											Y	Y									Y					Y								2.5
Crepis paludosa				Y	Y				Y		Y					Y	Y	Y			Y					Y								1.5
Ctenidium molluscum		Y	Y	Y							Y	Y		Y				Y			Y	Y				Y		Y			Y	Y		0.733
Cynoglossum officinale		Y												Y								Y	Y											2.5
Dactylorhiza fuchsii		Y	Y	Y	Y							Y		Y		Y	Y	Y			Y	Y	Y			Y					Y		Y	0.6
Dactylorhiza incarnata			Y	Y	Y							Y		Y		Y		Y			Y	Y	Y			Y					Y		Y	0.7
Dactylorhiza maculata				Y								Y			Y			Y			Y			Y		Y					Y			1.167
Dactylorhiza majalis				Y								Y						Y			Y					Y					Y			1.5
Dactylorhiza purpurella				Y	Y							Y				Y		Y			Y		Y			Y					Y		Y	0.917
Dactylorhiza traunsteineri				Y								Y						Y			Y					Y					Y			1.5
Dianthus deltoides		Y												Y	Y							Y		Y										2
Ditrichum flexicaule		Y	Y	Y										Y								Y						Y	Y			Y		1.583
Ditrichum gracile		Y	Y	Y										Y								Y						Y	Y			Y		1.583
Drosera anglica						Y		Y				Y							Y	Y	Y				Y	Y	Y				Y			1

Drosera intermedia						Y		Y	Y										Y	Y	Y				Y	Y	Y				Y			1
Drosera rotundifolia						Y		Y	Y			Y							Y	Y	Y				Y	Y	Y				Y			0.917
Dryopteris carthusiana				Y					Y									Y			Y					Y					Y			1.5
Eleocharis quinqueflora												Y									Y					Y								3
Empetrum nigrum						Y	Y	Y											Y	Y					Y		Y				Y			1.333
Epipactis palustris				Y														Y								Y								3
Equisetum hyemale												Y									Y					Y					Y			2.5
Equisetum palustre				Y					Y			Y						Y			Y					Y					Y			1.333
Equisetum variegatum												Y									Y					Y								3
Erica tetralix						Y	Y	Y										Y	Y	Y					Y	Y	Y				Y			1
Eriophorum latifolium												Y									Y					Y								3
Eriophorum vaginatum						Y	Y	Y											Y	Y							Y							1.833
Eupatorium cannabinum				Y					Y									Y			Y					Y					Y			1.5
Euphrasia arctica subsp. borealis			Y		Y									Y	Y	Y	Y					Y	Y	Y									Y	1
Euphrasia confusa			Y											Y	Y		Y					Y	Y	Y										1.667
Euphrasia nemorosa		Y	Y											Y	Y		Y					Y	Y	Y									Y	1.083
Euphrasia officinalis agg.		Y	Y	Y	Y							Y		Y	Y	Y	Y	Y			Y	Y	Y	Y	Y	Y							Y	0.567
Euphrasia rostkoviana subsp. montana			Y		Y									Y	Y	Y	Y					Y	Y	Y									Y	1
Euphrasia rostkoviana subsp. rostkoviana			Y		Y									Y	Y	Y	Y					Y	Y	Y									Y	1
Euphrasia scottica												Y						Y			Y					Y					Y			2
Filipendula ulmaria				Y					Y			Y						Y			Y					Y					Y		Y	1.167

Fissidens dubius		Y												Y								Y						Y	Y			Y		2.25
Galium boreale			Y		Y									Y		Y						Y	Y											1.5
Galium palustre				Y					Y			Y						Y			Y					Y					Y			1.333
Galium sterneri		Y	Y											Y								Y												2.5
Galium uliginosum				Y					Y									Y			Y					Y					Y			1.5
Galium verum		Y	Y		Y									Y	Y	Y	Y					Y	Y	Y									Y	0.833
Genista tinctoria															Y		Y						Y	Y									Y	0.833
Gentianella amarella		Y	Y											Y								Y												2.5
Gentianella campestris			Y											Y	Y		Y					Y	Y	Y									Y	1.583
Geranium pratense																							Y										Y	0.5
Geranium sylvaticum					Y											Y		Y					Y			Y					Y			1.833
Geum rivale				Y					Y									Y			Y					Y					Y			1.5
Gymnadenia conopsea			Y	Y	Y							Y		Y		Y		Y			Y	Y	Y		Y	Y								0.833
Hamatocaulis vernicosus												Y									Y					Y								3
Hammarbya paludosa								Y												Y							Y							3
Helianthemum nummularium		Y	Y											Y	Y							Y		Y	Y								Y	1.333
Helictotrichon pratense		Y	Y											Y	Y							Y		Y										1.5
Helictotrichon pubescens		Y			Y									Y		Y	Y					Y	Y										Y	1.167
Homalothecium lutescens		Y												Y								Y						Y	Y			Y		2.25
Hydrocotyle vulgaris				Y					Y									Y			Y					Y					Y			1.5
Hymenostylium recurvirostrum											Y	Y									Y					Y		Y				Y		1.833
Hypericum humifusum															Y									Y	Y									2
Hypnum lacunosum			Y											Y								Y												3
Jasione montana															Y									Y				Y				Y		1.333

Juncus acutiflorus				Y														Y			Y					Y					Y		Y	1.833
Juncus compressus				Y														Y			Y					Y								2.5
Juncus subnodulosus				Y								Y						Y			Y					Y					Y			1.5
Knautia arvensis		Y												Y			Y					Y	Y										Y	1.833
Koeleria macrantha		Y												Y	Y							Y		Y										2
Lathyrus linifolius															Y		Y						Y	Y	Y								Y	0.833
Leontodon hispidus		Y												Y	Y							Y		Y									Y	1.833
Leontodon saxatilis		Y												Y	Y							Y		Y										2
Linum catharticum		Y	Y	Y								Y		Y				Y			Y	Y				Y								1.083
Listera cordata						Y		Y											Y	Y					Y		Y				Y			1.5
Lotus corniculatus		Y	Y		Y									Y	Y	Y	Y					Y	Y	Y	Y								Y	0.833
Lotus pedunculatus				Y					Y									Y			Y					Y					Y			1.5
Lychnis flos-cuculi				Y					Y			Y						Y			Y					Y					Y			1.333
Lycopus europaeus				Y					Y									Y			Y					Y					Y			1.5
Lysimachia vulgaris				Y														Y			Y					Y					Y			2
Lythrum salicaria				Y					Y									Y			Y					Y					Y			1.5
Mentha aquatica				Y					Y			Y						Y			Y					Y					Y			1.333
Menyanthes trifoliata									Y												Y					Y				Y	Y			2.333
Meum athamanticum															Y		Y						Y	Y									Y	0.833
Minuartia verna	Y		Y										Y	Y								Y						Y				Y		1.333
Molinia caerulea																																		0
Myosotis ramosissima		Y												Y	Y							Y		Y				Y				Y		1.75
Myosotis stolonifera				Y								Y						Y			Y					Y								2
Narthecium						Y		Y											Y	Y					Y		Y							2

ossifragum																																		
Oenanthe fistulosa									Y												Y					Y				Y				2.5
Ophioglossum vulgatum			Y											Y	Y		Y					Y	Y	Y									Y	1.583
Orchis mascula			Y		Y									Y		Y						Y	Y											1.5
Origanum vulgare		Y												Y								Y	Y										Y	2.333
Ornithopus perpusillus															Y									Y										2
Palustriella commutata											Y	Y									Y					Y		Y				Y		1.833
Parnassia palustris			Y						Y			Y		Y							Y	Y				Y								1.333
Pedicularis palustris				Y					Y			Y						Y			Y				Y	Y					Y			1.333
Pedicularis sylvatica															Y									Y	Y						Y			1.5
Pellia endiviifolia											Y	Y									Y					Y		Y	Y		Y	Y		1.7
Philonotis calcarea											Y	Y									Y					Y		Y	Y			Y		1.75
Pilosella officinarum		Y	Y											Y	Y							Y		Y				Y				Y		1.25
Pimpinella saxifraga		Y	Y											Y			Y					Y	Y										Y	1.333
Pinguicula vulgaris						Y		Y	Y		Y	Y							Y	Y	Y				Y	Y	Y				Y			0.867
Platanthera bifolia			Y											Y	Y		Y					Y	Y	Y									Y	1.583
Platanthera chlorantha			Y											Y			Y					Y	Y										Y	1.833
Polygala serpyllifolia															Y									Y	Y									2
Polygala vulgaris		Y	Y											Y			Y					Y	Y										Y	1.333
Polytrichum strictum						Y		Y											Y	Y							Y							2
Potentilla erecta																																		0
Potentilla palustris				Y					Y									Y			Y					Y					Y			1.5
Potentilla tabernaemontani		Y												Y								Y												3

Primula veris		Y			Y									Y		Y	Y					Y	Y										Y	1.167
Pseudorchis albida			Y		Y									Y	Y	Y	Y					Y	Y	Y									Y	1
Ranunculus bulbosus		Y			Y									Y		Y	Y					Y	Y										Y	1.167
Ranunculus lingua									Y												Y					Y				Y				2.5
Rhinanthus minor		Y	Y	Y	Y							Y		Y		Y	Y	Y			Y	Y	Y			Y							Y	0.65
Rhynchospora alba						Y		Y		Y									Y	Y					Y		Y							1.833
Rhynchospora fusca						Y		Y		Y									Y	Y					Y		Y							1.833
Sagina nodosa											Y	Y									Y					Y								2.5
Sanguisorba minor subsp. minor		Y												Y	Y		Y					Y	Y	Y									Y	1.583
Sanguisorba officinalis				Y	Y											Y	Y	Y					Y			Y					Y		Y	1.083
Saxifraga granulata					Y											Y	Y						Y											2.5
Scabiosa columbaria		Y												Y								Y											Y	2.5
Scapania aspera												Y									Y					Y		Y				Y		2.333
Scorpidium revolvens									Y		Y	Y									Y				Y	Y								2.333
Scorpidium scorpioides												Y									Y					Y								3
Scutellaria galericulata				Y					Y									Y			Y					Y					Y			1.5
Sedum acre		Y												Y	Y							Y		Y										2
Sedum anglicum															Y									Y				Y				Y		1.333
Sedum villosum												Y									Y					Y								3
Selaginella selaginoides			Y								Y	Y		Y							Y				Y	Y					Y			1.333
Senecio aquaticus				Y					Y			Y						Y			Y					Y					Y			1.333
Sherardia arvensis		Y												Y	Y							Y		Y				Y				Y		1.75
Silaum silaus																	Y						Y										Y	1.5

Solidago virgaurea			Y											Y	Y							Y		Y										2
Sphagnum austinii						Y		Y											Y	Y							Y							2
Sphagnum capillifolium						Y		Y											Y	Y					Y		Y				Y			1.5
Sphagnum contortum												Y									Y					Y								3
Sphagnum cuspidatum						Y		Y											Y	Y							Y							2
Sphagnum denticulatum				Y		Y		Y			Y	Y						Y	Y	Y						Y	Y				Y			0.867
Sphagnum fallax						Y		Y	Y									Y	Y	Y	Y				Y	Y	Y				Y			0.917
Sphagnum fuscum						Y		Y											Y	Y							Y							2
Sphagnum magellanicum						Y		Y											Y	Y							Y							2
Sphagnum papillosum						Y		Y										Y	Y	Y					Y	Y	Y				Y			1.167
Sphagnum platyphyllum												Y									Y					Y								3
Sphagnum squarrosum									Y		Y										Y					Y					Y			2
Sphagnum subsecundum												Y									Y					Y								3
Sphagnum tenellum						Y		Y											Y	Y					Y		Y				Y			1.5
Sphagnum teres									Y			Y									Y					Y								2.5
Sphagnum warnstorfii									Y												Y					Y								3
Spiranthes romanzoffiana																		Y								Y					Y			1.5
Stachys officinalis		Y												Y	Y		Y					Y	Y	Y									Y	1.583
Stellaria palustris				Y								Y						Y			Y					Y								2
Succisa pratensis			Y	Y	Y				Y			Y		Y	Y	Y	Y	Y			Y	Y	Y	Y		Y					Y		Y	0.533
Thalictrum minus			Y											Y			Y					Y	Y										Y	1.833
Thlaspi caerulescens	Y												Y															Y						3
Thuidium assimile		Y	Y											Y								Y						Y	Y			Y		1.75

Thymus polytrichus	Y	Y	Y									Y	Y	Y							Y	Y			Y									1.583
Tofieldia pusilla			Y									Y		Y							Y	Y				Y								1.5
Tortella tortuosa		Y	Y											Y								Y						Y				Y		1.833
Trifolium arvense															Y							Y		Y				Y				Y		1.25
Trifolium striatum		Y												Y	Y							Y		Y				Y				Y		1.75
Triglochin palustre				Y							Y	Y						Y			Y				Y	Y					Y			1.333
Trisetum flavescens		Y	Y		Y									Y		Y	Y					Y	Y										Y	1
Trollius europaeus				Y	Y											Y	Y	Y					Y			Y					Y		Y	1.083
Vaccinium microcarpum						Y		Y											Y	Y							Y							2
Vaccinium oxycoccos						Y		Y											Y	Y							Y							2
Valeriana dioica				Y					Y			Y						Y			Y					Y								1.833
Valeriana officinalis				Y					Y			Y						Y			Y		Y			Y					Y			1.167
Valerianella locusta		Y												Y	Y							Y		Y				Y				Y		1.75
Veronica scutellata									Y			Y						Y			Y					Y								2
Viola hirta		Y												Y								Y												3
Viola lutea			Y		Y									Y	Y	Y	Y					Y	Y	Y									Y	1
Viola palustris				Y					Y			Y						Y			Y					Y					Y			1.333
Viola riviniana		Y	Y		Y									Y	Y	Y	Y					Y	Y	Y	Y						Y		Y	0.783
Viola tricolor															Y									Y					Y					1.5
Weissia brachycarpa var. obliqua		Y	Y											Y														Y				Y		1.833
Weissia controversa var. densifolia	Y												Y															Y						3

www.snh.gov.uk

ISBN: 978-1-78391-380-0

Policy and Advice Directorate, Great Glen House,
Leachkin Road, Inverness IV3 8NW
T: 01463 725000

You can download a copy of this publication from the SNH website.

recycle

Scottish Natural Heritage
Dualchas Nàdair na h-Alba
All of nature for all of Scotland
Nàdar air fad airson Alba air fad