Vegetatio **116**: 133–146, 1995.

Identification of site-types important for rare ferns in an area of deciduous woodland in northwest Spain

Javier Amigo Vázquez & Guy Norman
Laboratorio de Botánica, Facultade de Farmacia, Universidade de Santiago, E-15706 Santiago de Compostela, Spain

Accepted 17 October 1994

Key words: Discriminant analysis, Ferns, Indicator species, Plant conservation, Spain (NW), TWINSPAN, Woodland

Abstract

The Eume woods in northwest Spain are an important refuge for a number of endangered ferns, including *Culcita macrocarpa, Cystopteris diaphana, Dryopteris aemula, Dryopteris guanchica, Hymenophyllum tunbrigense* and *Woodwardia radicans*. Because some of these species are inconspicuous or difficult to identify in the field, we attempted to develop simple floristic criteria to identify sites important or potentially important for these species ('high-value sites'). Having drawn up field-layer vascular plant inventories for 44 plots in the study area, we considered two major strategies. Under strategy 1, we looked for those conspicuous species which best discriminated between plots containing the endangered species ('current-high-value plots') and all other plots. Under strategy 2, we first carried out a global plots-by-species classification using indicator species analysis (Hill et al. 1975) and then looked for those conspicuous species which best discriminated between groups containing current-high-value plots and all other groups. For our data set, the second strategy proved more effective. The best criterion identified was 'If two or more of *Luzula sylvatica, Viola riviniana* and *Rubus* spp. are present, then that plot is high-value'. We include a listing of a BASIC program for the discriminant analysis algorithm, which we consider to have a number of potential applications in vegetation science and biogeography.

Nomenclature: Plant nomenclature follows Castroviejo et al. (1986–1990) or, for species not included in this series, Tutin et al. (1964–1980).

Introduction

The seminatural deciduous woodlands of the Eume Valley, in the northwest Spanish region of Galicia, cover an area of approximately 2300 ha. The lower part of the valley has an unusual warm wet microclimate with very low frost risk, and is an important refuge for a number of endangered Tertiary-relict fern species, namely *Culcita macrocarpa, Dryopteris aemula, Dryopteris guanchica, Hymenophyllum tunbrigense* and *Woodwardia radicans*, all considered to be rare or vulnerable at Iberian or global level (Salvo 1990). *Cystopteris diaphana* (considered by Salvo to be rare at Iberian level) has also been reported from the area (J. Cremades, personal communication), but we have not been able to confirm this. In total, at least 28 pteridophytes are present.

The rare species occur mainly as small scattered subpopulations; all except *Dryopteris aemula* are found most frequently beside streams but also in other wet shaded sites. Effective conservation of these subpopulations clearly requires identification of the sites at which they occur and, ideally, of sites to which they might spread or be reintroduced. Identification of these 'high-value' sites is, however, made difficult by the extent of the area and its complex topography, by the small size of the individual subpopulations and (particularly in the case of *Dryopteris guanchica* and *Cystopteris diaphana*) by the difficulty of identification in the field. In view of these difficulties, the aim of this study was to develop simple floristic criteria (based on the presence or absence of other conspicuous species) for use in large-scale field survey for the identification of high-value sites. Our starting assumptions were a) that the endangered fern species are likely to be fairly closely associated with other more conspicuous and more easily identified species, and b) that at least some of these conspicuous species are likely to be less patchily distributed than the endangered ferns, and thus possible indicators of sites of potential importance. We decided at the outset to look for criteria of the form 'If two or more of conspicuous species A, B and C are present, then that site is high- value'. To the best of our knowledge this is the first study of this type, although multivariate methods have been used to find indicators of rare habitats (see for example Hill et al. 1975) and to determine which sites in a given location are most important for maintenance of habitat diversity (Nilsson 1986).

It should be stressed that we have deliberately focused on one particular aspect of conservation value. However, the Eume woods constitute one of the largest remaining continuous expanses of seminatural woodland in Galicia, and the existing area is being steadily reduced as a result of felling of mature trees, planting of *Eucalyptus globulus*, track laying and extension of existing hydroelectric installations: obviously there is an urgent need for protection of the whole area on these grounds alone. In addition, the Eume woods constitute an important refuge for endangered species of other major taxa, particularly bryophytes.

Study area

General descriptions of the vegetation and climate of northern Galicia can be found in Izco et al. (1990) and Retuerto & Carballeira (1991).

The Eume woods extend more or less continuously along both sides of the lower Eume river valley (Fig. 1), covering a total area of about 2290 ha. Altitude is between 10 and 600 m a.s.l. The valley is steep-sided and V-shaped; this, together with the basically westerly direction of flow of the river, leads to differences in floristic composition between the north- and south-facing sides of the valley (Losa Quintana 1974). The geological substrate is largely slate but in some cases slates alternating with quartzite and granite. Soils are thus generally acid and nutrient-poor, although (especially in lower-altitude areas with good humification) considerable soil depths may be attained.

Climatic data for the region are shown in Table 1. Precipitation is high; temperatures are mild, with relatively high coldest-month minimum temperatures at low altitudes. Interseasonal temperature variation is extremely low.

Historically, the woods have been heavily managed as a source of timber, firewood and chestnuts (with consequent selection in favour of *Castanea sativa*) and for grazing of cattle and goats: the undergrowth is often burnt to favour the latter use. In the 1960s, the construction of the Eume reservoir (Fig. 1) led to the loss of a major expanse of woodland between 200 and 300 m altitude, and fragmented the remaining woodland along this 12 km stretch of the river.

Methods

Plot identification strategies

Floristic criteria were derived on the basis of field-layer vascular plant inventories for forty-four 200 m^2 plots in the study area. We considered two alternative strategies. Strategy 1 was to decide which plots were high-value (as defined below under 'Assigning values to plots') and then to look for the conspicuous species which best discriminated between these plots and all other plots. Strategy 2 was to first carry out a global plots-by-species classification of the data, then to define high-value groups-of-plots as those with the highest proportion of high-value plots, and finally to look for the conspicuous species which best discriminated between these groups and all other groups. The rationale behind this strategy 2 was that it might allow identification of site-types of potential (as opposed to current) importance for the endangered species of interest.

Data collection

Continuously wooded areas were first mapped at 1:50,000 (Fig. 1) on the basis of a previously published land-use map of the area (Vales 1992) and of aerial photography. Working from this map, a total of 44 sites were preselected, with the aim of obtaining even coverage of the area and of sampling from the main tree-species community types. Between April and July 1992, a 200 m^2 (14 × 14 m) plot was marked out at each site. All vascular plants occurring within the plot were recorded, with their abundance estimated on the 7-point Braun-Blanquet scale (Braun-Blanquet

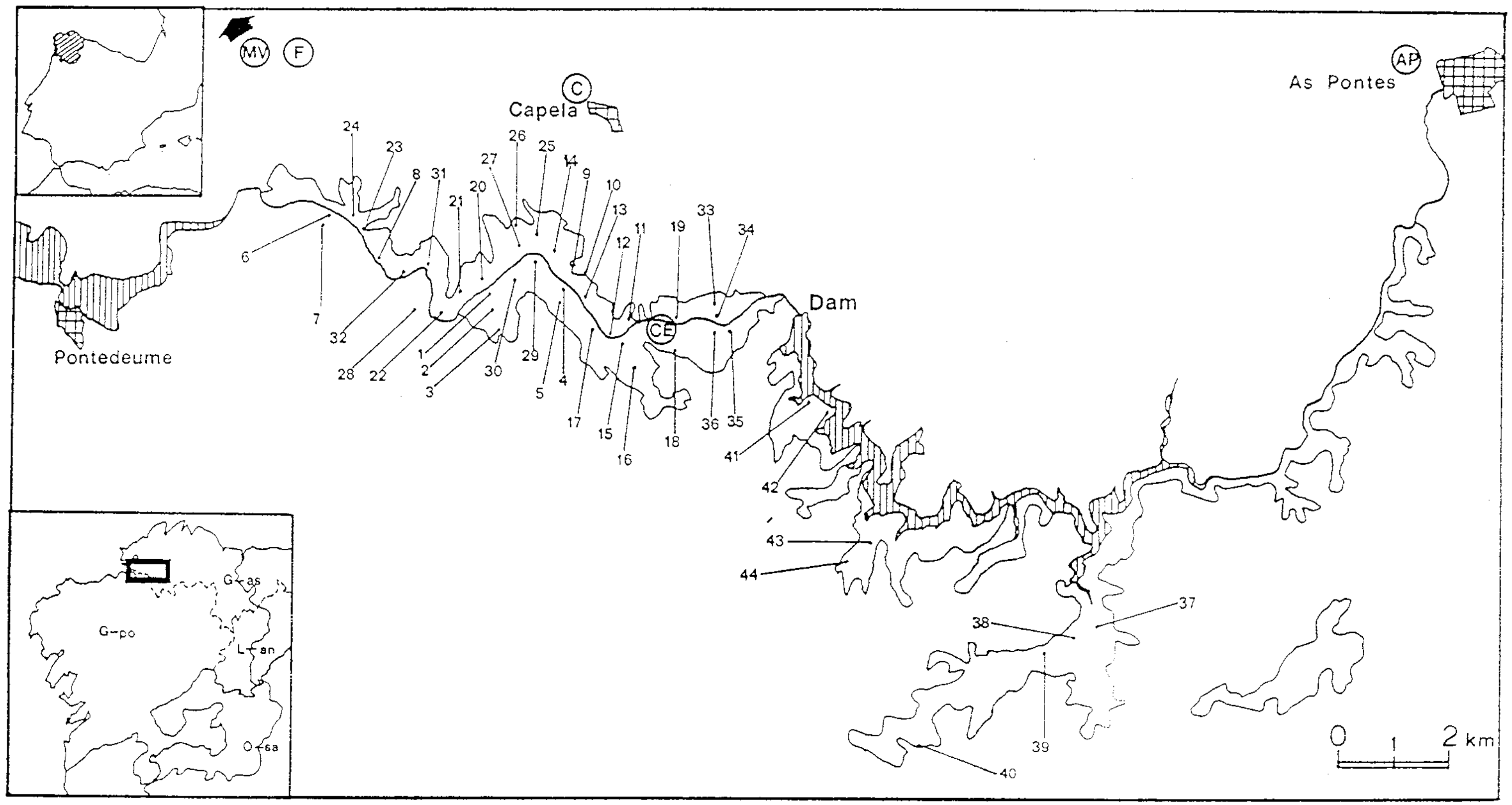

Fig. 1. Map of the study area, showing distribution of deciduous woodland location of the 44 plots and location of weather stations (see Table 1). Smaller maps showing the position of Galicia in the Iberian peninsula, and the biogeographical units referred to in the text.

1979). The number of individuals of all species in the tree/shrub layer (height greater than 1.8 m), altitude, aspect and estimated mean slope were also recorded, and a general description of the site (wetness, stoniness, signs of burning, signs of grazing etc.) was noted down. Soil pH was determined potentiometrically at 24 sites, and was between 4 and 5 in 90% of cases (results not shown). We did not specifically aim to sample from plots containing the endangered species, and none of the plots contained *Culcita macrocarpa*.

Assigning values to species

Consideration of methods for assigning 'rarity values' to species is beyond the scope of this study (see for example Klopatek et al. 1981). We based our assessment of the current status of Galician pteridophytes, and of the importance of the Galician populations of these species with respect to Iberian and global populations, on Salvo (1990), Jalas & Suominen (1972) and our own knowledge of the pteridophyte flora of Galicia. Table 2 lists all pteridophytes recorded from Galicia (from Rodríguez-García et al. 1990; Horjales et al. 1988; Soñora et al. 1992), and the rarity value assigned to each.

Assigning values to plots

Three alternative definitions of high-value plots were compared:

1. 'A plot is high-value if it contains any fern species with rarity value 4 or higher'.

2. 'A plot is high-value if it contains any fern species with rarity value 3 or higher'.

3. 'A plot is high value if the sum of the rarity values of all fern species present is greater than 7'. We use the abbreviation FVI ('fern value index') for this index. A threshold of seven was subjectively chosen as being indicative of rich fern communities. Note that in other circumstances it might be more useful to use a logarithmic (rather than linear) scale of rarity values for calculating summed indices of this type.

Where values had to be assigned to groups of plots (under strategy 2; see above), the highest-value group or groups were defined as those containing the highest proportion of high-value plots on definition 1, 2 or 3.

Data analysis

Discriminant analysis

We wrote a simple BASIC program called CRITERIO (listed in Table 3) for discriminant analysis, since we do not know of any standard statistics or vegetation classification package which contains this algorithm. The program requires presence-absence data for a 'target' plot or group of plots and a 'non-target' plot or group of plots. It searches for that floristic criterion which best discriminates between the target and non-target plots or groups of plots. The format of the criterion is predefined as 'If two or all of species A, B and C are present, then that plot is a member of the target group'. Note that other criterion formats (e.g. 'If one or more of species A, B, C and D are present ...') might be more appropriate in other circumstances. The program finds that three-species combination which identifies the maximum proportion of target-group plots on the above criterion; if there is more than one such combination, it then finds that or those which identify the minimum proportion of non-target groups. Note that for other uses it might be more useful to first minimize the number of non-target plots identified and then maximize the number of target plots identified, or alternatively to minimize the total number of target and non-target plots misidentified.

For this part of the analysis – since our aim was to identify discriminant species for use in rapid large-scale field survey – we eliminated from the data matrix all ferns and Gramineae (since most members of these groups are relatively difficult to identify) and all non-wintergreens (so that the criteria would be usable throughout the year); in addition species with estimated abundance of 'r' on the Braun-Blanquet scale were classed as absent from that plot (because species present at low abundance are probably not useful indicators).

TWINSPAN classification

The 44 plots were classified on the basis of field-layer species composition, using indicator species analysis (Hill et al. 1975) as contained in the program TWINSPAN/DECORANA (Hill 1979a, b). Tree- - and shrub-layer data were not included in the data matrix, though 'tree' species occurring in the herb layer were included. Any species which had been recorded with the lowest abundance estimate ('r' on the Braun-Blanquet scale, generally corresponding to a single individual) was classed as absent from that plot.

Table 1. Climatic data from weather stations near the study area (see Fig. 1).

Stations	MV	F	C	CE	AP
Longitud W	8° 18′	8° 09′	8° 04′	8° 02′	7° 51′
Latitud N	43° 29′	43° 28′	43° 26′	43° 24′	43° 27′
Altitud (m a.s.l.)	216	30	387	73	360
Annual mean temperature °C	13.3	13.2	13.8	15.1	11.7
Thermal amplitude	9.3	9.5	11.5	12	10.7
Mean annual precipitation (mm)	970	1388	1915	1910	1684
Summer precipitation: Jun-Aug	132	152	228	194	173

Table 2. Checklist of Galician pteridophytes, showing rarity values assigned to each, and whether or not they occur in woodland habitats (Y – yes, N – no, ? – possibly). Species recorded from the study area are marked with an asterisk.

Taxon	Woodland habitat	Rarity value
Adiantum capillus-veneris	N	3
**Anogramma leptophylla*	?	1
**Asplenium adiantum-nigrum*	?	3
Asplenium adiantum-nigrum var. *corunnense*	N	4
**Asplenium billoti*	?	0
Asplenium marinum	N	3
**Asplenium onopteris*	?	1
Asplenium septentrionale	N	4
Asplenium trichomanes	N	3
**Asplenium trichomanes* subsp. *quadrivalens*	?	0
**Athyrium filix-femina*	Y	1
Blechnum spicant	Y	1
Blechnum spicant var. *homophyllum*	Y	4
**Ceterach officinarum*	N	3
Cheilanthes acrostica	N	4
Cheilanthes guanchica	N	4
Cheilanthes hispanica	N	3
Cheilanthes tinaei	N	4
Christella dentata	N	5
Cryptogramma crispa	N	3
**Culcita macrocarpa*	Y	4
Cystopteris dickieana	Y	4
**Cystopteris fragilis*	Y	3
**Cystopteris diaphana*	Y	4
**Davallia canariensis*	?	2
**Dryopteris aemula*	Y	3
**Dryopteris affinis*	Y	1
**Dryopteris affinis* subsp. *borreri*	Y	2
Dryopteris affinis subsp. *stilluppensis*	?	4
Dryopteris carthusiana	Y	4
**Dryopteris dilatata*	Y	2
Dryopteris expansa	N	4
**Dryopteris filix-mas*	Y	2
**Dryopteris guanchica*	Y	4
Dryopteris oreades	N	3
**Equisetum arvense*	N	1
Equisetum hyemale	N	4
Equisetum palustre	N	4
Equisetum ramosissimum	N	2
Equisetum telmateia	N	2
Gymnocarpium dryopteris	Y	4
Huperzia selago	N	4
**Hymenophyllum tunbrigense*	Y	4
Hymenophyllum wilsonii	Y	5
Isoetes histrix	N	2
Isoetes longissimum	N	3
Isoetes velatum	N	2
Isoetes velatum subsp. *asturicense*	N	3
**Lastraea limbosperma*	Y	2
Lycopodiella inundata	N	4
Lycopodium clavatum	N	4
Notholaena marantae	N	4
Ophioglossum lusitanicum	N	4
Ophioglossum vulgatum	N	3
**Osmunda regalis*	Y	1
**Phyllitis scolopendrium*	Y	2
Pilularia globulifera	N	4
Polypodium cambricum	?	3
**Polypodium interjectum*	?	0
**Polypodium vulgare*	?	0
Polystichum aculeatum	Y	3
**Polystichum setiferum*	Y	2
**Pteridium aquilinum*	?	0
Pteris cretica	N	4
Pteris vittata	N	4
Selaginella denticulata	N	4
Thelipteris palustris	Y	4
Vandenboschia speciosa	Y	4
**Woodwardia radicans*	Y	3

Table 3. Listing of the GW-BASIC program CRITERIO. The DATA lines of the program should be modified to incorporate your data; as an example, the data shown are for presence/absence of 16 species in five plots (lines 0830–0880), and the program is set to search for those three species which best discriminate between the target group (plots 2 and 3) and all other plots (lines 0890–0910). The names of the species in your data set should also be inserted, from line 930 on. Note the questions at lines 0020–0040 which you will be asked when you run the program.

```
0000 PRINT "Program CRITERIO"
0010 REM This GW-BASIC program requires data in DMATRIX, TARGET and SPECIES$
0020 INPUT "How many species in your data matrix?"; NS
0030 INPUT "How many plots/sites in your data matrix?"; NP
0040 INPUT "How many plots/sites in your target group?"; NP1
0050 NP2=NP-NP1
0060 PRINT "Program running"
0070 DIM DMATRIX(NS,NP): DIM TARGET(NP)
0080 DIM BEST(5,3): DIM SPECIES$(NS)
0090 FOR I=1 TO NP
0100  FOR J=1 TO NS
0110   READ DMATRIX(I,J)
0120  NEXT
0130 NEXT
0140 FOR I=1 TO NP
0150  READ TARGET(I)
0160 NEXT
0170 FOR I=1 TO NS
0180  READ SPECIES$(I)
0190 NEXT
0200 S1MAX=0: S2MIN=1: NB=0
0210 FOR I=1 TO 5
0220  FOR J=1 TO 3
0230   BEST(I,J)=0
0240  NEXT
0250 NEXT
0260 REM Start calculation
0270 FOR I=1 TO NS
0280  FOR J=1 TO NS
0290   FOR K=1 TO NS
0300    IF I=J THEN GOTO 0510
0310    IF J=K THEN GOTO 0510
0320    IF K=I THEN GOTO 0510
0330    SCORE1=0: SCORE2=0
0340    FOR P=1 TO NP
0350     IF TARGET(P)=1 THEN GOSUB 0780
0360     IF TARGET(P)=0 THEN GOSUB 0800
0370    NEXT
0380    S1=SCORE1/NP1: S2=SCORE2/NP2
0390    IF S1<S1MAX THEN GOTO 0510
0400    IF S1>S1MAX THEN GOTO 0510
0410    IF S1=S1MAX AND S2<S2MIN THEN GOTO 0430
0420    IF S1=S1MAX AND S2=S2MIN THEN GOTO 0490
0430    FOR B=1 TO 5
0440     BEST(B,1)=0:BEST(B,2)=0:BEST(B,3)=0
0450    NEXT
0460    BEST(1,1)=I: BEST(1,2)=J: BEST(1,3)=K
0470    S1MAX=S1:S2MIN=S2:NB=1
0480    GOTO 0510
0490    NB=NB+1: IF NB=6 THEN GOTO 0510
0500    BEST(NB,1)=I: BEST(NB,2)=J: BEST(NB,3)=K
0510   NEXT
0520  NEXT
0530 NEXT
0540 PRINT "Result:"
0550 IF NB=0 THEN PRINT "There are no three-species criteria which successfully discriminate
between the target and non-target groups."
0560 IF NB=0 THEN GOTO 0760
0570 IF NB>1 THEN GOTO 0640
0580 PRINT "The three species which best discriminate between the target and non-target groups
are:"
0590 PRINT SPECIES$(BEST(1,1))
0600 PRINT SPECIES$(BEST(1,2))
0610 PRINT SPECIES$(BEST(1,3))
0620 PRINT "The presence of two or three of these species correctly identifies ";S1MAX*100;"%
of your target-group plots, and incorrectly identifies ";S2MIN*100;"% of your non-target-group
plots."
0630 GOTO 0760
0640 PRINT "There are"; NB;" three-species combinations which discriminate equally well
between your target and non-target groups."
0650 IF NB>5 THEN PRINT "Only the first five of these combinations found by the program are
listed here."
0660 PRINT "Combination 1 is:"
0670 PRINT SPECIES$(BEST(1,1))
0680 PRINT SPECIES$(BEST(1,2))
0690 PRINT SPECIES$(BEST(1,3))
0700 PRINT "The presence of two or three of these species correctly identifies ";S1MAX*100;"%
of your target-group plots, and incorrectly identifies ";S2MIN*100;"% of your non-target-group
plots."
0710 IF NB=2 THEN PRINT "The following combination is equally effective (or ineffective!):"
0720 IF NB>2 THEN PRINT "The following combinations ae equally effective (or ineffective!):"
0730 FOR B=2 TO NB
0740  PRINT SPECIES$(BEST(B,1));" - "; SPECIES$(BEST(B,2));" - "; SPECIES$(BEST(B,3))
```

```
0750 NEXT
0760 PRINT "PROGRAM TERMINATED"
0770 END
0780 IF DMATRIX(P,I)+DMATRIX(P,J)+DMATRIX(P,K)>=2 THEN SCORE1=SCORE1+1
0790 RETURN
0800 IF DMATRIX(P,I)+DMATRIX(P,J)+DMATRIX(P,K)>=2 THEN SCORE2=SCORE2+1
0810 RETURN
0820 REM DMATRIX
0830 REM 16 species in 5 plots
0840 DATA 0,1,1,1,0,0,1,1,1,0,1,0,1,1,0,0
0850 DATA 0,1,1,0,0,0,1,1,1,1,1,0,0,0,0,1
0860 DATA 0,1,1,1,1,0,1,0,0,1,1,1,0,0,1,1
0870 DATA 0,1,1,0,1,1,0,0,1,1,0,1,1,0,0,0
0880 DATA 1,0,1,0,1,1,0,0,0,0,1,0,1,0,0,1
0890 REM TARGET
0900 REM The target group comprises plots 2 and 3
0910 DATA 0,1,1,0,0
0920 REM SPECIES$
0930 REM 16 species
0940 DATA "Luzula","Hedera","Geranium","Umbilicus"
0950 DATA "Anemone","Blechnum","Euphorbia","Omphalodes"
0960 DATA "Oxalis","Ruscus","Viola","Cardamine"
0970 DATA "Crepis","Saxifraga","Stellaria","Tamus"
```

Table 4. Characteristics of plots assigned to each of the eight TWINSPAN groups. S.R. = species richness; F.V.I. = fern value index (see text).

Group		Sample n°	Altitude (m)	Slope (°)	Aspect	S.R.	F.V.I.	Group		Sample n°	Altitude (m)	Slope (°)	Aspect	S.R.	F.V.I.
A1	**1**	13	50	15	SW	30	6	**B1**	**5**	2	100	20	SW	11	0
		25c	160	20	S	32	7			9	260	35	S	14	0
		27c	90	30	SE	44	7			11	100	20	S	20	0
		31abc	70	35	SE	48	14			17	50	20	NE	33	0
		35	110	30	N	40	5			19	150	20	S	26	1
										20	60	25	S	26	0
	2	1	30	30	N	41	4			22	50	25	SW	19	0
		8	30	20	SW	20	3			32	80	30	S	22	0
		12	50	15	SW	29	1								
		30	130	25	NW	24	6		**6**	10	270	30	S	8	1
		33	200	25	SW	32	3			21	140	20	E	8	0
		34	90	20	S	46	5			26	270	20	SW	23	1
										28	150	30	N	13	0
A2	**3**	3abc	240	35	N	25	19								
		15c	90	10	NE	33	8	**B2**	**7**	29	120	30	N	21	1
		16c	160	15	NE	32	7			36	140	30	NE	33	3
		24abc	40	15	NE	34	10			43b	340	35	NE	20	6
		38bc	315	40	NE	30	10			44	480	25	E	23	1
	4	4	35	15	N	31	4		**8**	5	200	15	NW	14	4
		6	40	15	N	18	6			7	180	20	NW	16	3
		14c	150	10	S	26	7			37	335	10	E	13	0
		18	250	30	N	22	2			39bc	350	25	W	27	9
		23	20	15	SW	14	4			40bc	530	20	N	24	10
										41	330	30	NW	16	2
										42	340	25	N	16	5

a = high value on definition 1; b = high value on definition 2; c = high value on definition 3.

Table 5. Most frequent herbs in each of the eight groups obtained by TWINSPAN. Trees are those which were most frequently dominant or co-dominant in plots of that group. Constant species are those present in 80% or more of plots of that group. Frequent species are those present in 50% or more of plots of that group. Underlined species are those characteristically present with high relative abundance (median estimated abundance > 1, on a scale of 1–6 corresponding to the Braun-Blanquet scale of + to 5). Of the total of 157 herb species found in the 44 plots, 107 (68%) were neither constant nor frequent in any group. Trees, and herbs with estimated abundance 'r' on the Braun-Blanquet scale, were not included in the classification.

GROUP 1

Trees: *Corylus, Castanea, Fraxinus, Ulmus*

Constant species

Athyrium filix-foemina
Dryopteris affinis
Euphorbia dulcis
Geranium robertianum
Hedera helix
Holcus mollis
Omphalodes nitida
Oxalis acetosella
Polystichum setiferum
Rubus sp.
Ruscus aculeatus
Umbilicus rupestris
Viola riviniana

Frequent species

Anthoxanthum odoratum
Asplenium onopteris
Brachypodium sylvaticum
Cardamine flexuosa
Carex caryophyllea
Carex remota
Crepis capillaris
Dryopteris dilatata
Lonicera periclymenum
Saxifraga spathularis
Stellaria holostea
Tamus communis
Teucrium scorodonium

GROUP 2

Trees: *Quercus, Corylus, Laurus, Castanea*

Constant species

Hedera helix
Laurus nobilis (herb)
Physospermum cornubiense
Polypodium vulgare
Polystichum setiferum
Pteridium aquilinum
Ruscus aculeatus
Tamus communis
Teucrium scorodonium
Viola riviniana

Frequent species

Asphodelus albus
Asplenium onopteris
Brachypodium pinnatum
Carex caryophyllea
Dryopteris affinis
Digitalis purpurea
Geranium robertianum
Holcus mollis
Hypericum pulchrum
Lonicera periclymenum
Polygonatum odoratum
Rubia peregrina
Stellaria holostea
Quercus robur (herb)
Rubus sp.

GROUP 3

Trees: *Corylus, Castanea, Laurus*

Constant species

Anemone nemorosa
Athyrium filix-foemina
Blechnum spicant
Dryopteris affinis
Dryopteris dilatata
Hedera helix
Lonicera periclymenum
Luzula sylvatica
Oxalis acetosella
Rubus sp.
Polypodium vulgare
Saxifraga spathularis

Frequent species

Carex caryophyllea
Euphorbia dulcis
Holcus mollis
Laurus nobilis (herb)
Polystichum setiferum
Ruscus aculeatus
Viola riviniana

GROUP 4

Trees: *Quercus, Corylus, Laurus, Pyrus*

Constant species

Dryopteris affinis
Hedera helix
Laurus nobilis (herb)
Lonicera periclymenum
Polypodium vulgare
Rubus sp.
Ruscus aculeatus
Teucrium scorodonium

Frequent species

Anemone nemorosa
Asphodelus albus
Blechnum spicant
Carex caryophyllea
Holcus mollis
Ilex aquilinum (herb)
Omphalodes nitida
Polystichum setiferum
Pteridium aquilinum
Pyrus pyraster

Table 5. Continued.

GROUP 5

Trees: *Quercus, Laurus, Castanea, Pyrus, Erica*

Constant species	Frequent species
Asphodelus albus	*Brachypodium pinnatum*
Cytisus scoparius	*Holcus mollis*
Hedera helix	*Lonicera periclymenum*
Laurus nobilis (herb)	*Physospermum cornubiense*
Pteridium aquilinum	*Polygonatum odoratum*
Pyrus pyraster (herb)	*Rubia peregrina*
Quercus robur (herb)	
Rubus sp.	
Teucrium scorodonium	
Ruscus aculeatus	

GROUP 6

Trees: *Eucalyptus, Ulex*

Constant species	Frequent species
Pteridium aquilinum	*Anthoxanthum odoratum*
Pyrus pyraster (herb)	*Carex caryophyllea*
Rubus sp.	*Erica arborea* (herb)
Ulex europaeus (herb)	*Frangula alnus* (herb)
	Pseudarrhenaterum longifolium
	Teucrium scorodonium
	Quercus robur (herb)
	Ulex galli

GROUP 7

Trees: *Castanea, Quercus, Ilex*

Constant species	Frequent species
Anemone nemorosa	*Agrostis capillaris*
Anthoxanthum odoratum	*Asphodelus albus*
Brachypodium pinnatum	*Blechnum spicant*
Deschampsia flexuosa	*Carex caryophyllea*
Erica arborea (herb)	*Cytisus scoparius*
Frangula alnus	*Daboecia cantabrica*
Hedera helix	*Dryopteris filix-mas*
Holcus mollis	*Euphorbia dulcis*
Physospermum cornubiense	*Ilex aquilinum* (herb)
Pteridium aquilinum	*Lonicera periclymenum*
Rubus sp.	*Melampyrum pratense*
Teucrium scorodonium	*Quercus robur* (herb)

GROUP 8

Trees: *Quercus, Castanea, Erica, Betula, Eucalyptus*

Constant species	Frequent species
Deschampsia flexuosa	*Blechnum spicant*
Hedera helix	*Carex caryophyllea*
Holcus mollis	*Dryopteris dilatata*
Pteridium aquilinum	*Dryopteris filix-mas*
Quercus robur (herb)	*Erica arborea* (herb)
Vaccinium myrtillus	*Lonicera periclymenum*
	Melampyrum pratense
	Rubus sp.
	Polypodium vulgare
	Teucrium scorodonium

Results

Characteristics of the 44 plots

A total of 157 field-layer species were found. Ten species were present in 50% or more of plots: *Hedera helix* (86% of plots), *Rubus* spp. (84%), *Pteridium aquilinum* (73%), *Holcus mollis* (70%), *Teucrium scorodonia* (68%), *Polystichum setiferum* (67%), *Lonicera periclymenum* (64%), *Ruscus aculea-*

Table 6. TWINSPAN indicator species for each division (in all cases indicator values are 1 or −1 and indicator thresholds 0).

1) **First-level division:**

Splitting the total group of 44 plots into groups A (21 plots; final groups 1 - 4) and B (23 plots; final groups 5 - 8): Polystichum setiferum, Dryopteris dilatata, Dryopteris affinis and Hedera helix (group A), and Deschampsia flexuosa, Pteridium aquilinum and Asphodelus albus (group B).

2) **Second-level divisions:**

a) Splitting group A into groups A1 (11 plots; final groups 1 and 2) and A2 (10 plots; final groups 3 and 4): Polystichum setiferum, Omphalodes nitida, Ruscus aculeatus, Rubia peregrina and Tamus communis (group A1) and Lonicera periclymenum, Dryopteris affinis, Luzula sylvatica, Blechnum spicant and Rubus sp. (group A2).

b) Splitting group B into groups B1 (12 plots; final groups 5 and 6) and B2 (11 plots; final groups 7 and 8): Asphodelus albus (group B1) and Deschampsia flexuosa and Holcus mollis (group B2).

3) **Third-level divisions:**

a) Splitting group A1 into final groups 1 (5 plots) and 2 (6 plots): Dryopteris affinis (group 1)

b) Splitting group A2 into final groups 3 (5 plots) and 4 (5 plots): Athyrium filix-femina, Dryopteris dilatata and Luzula sylvatica (group 3) and Teucrium scorodonia (group 4).

c) Splitting group BI into final groups 5 (8 plots) and 6 (4 plots): Asphodelus albus, Ruscus aculeatus and Hedera helix (group 5) and Quercus robur sapling (group 6)

d) Splitting group B2 into final groups 7 (4 plots) and 8 (7 plots): Physospermum cornubiense, Brachypodium pinnatum, Anthoxanthum odoratum, Asphodelus albus (group 7).

tus (59%), *Castanea sativa* sapling (57%) and *Polypodium vulgare* (52%). Of the remaining 147 species, three were present in 41–50% of plots, 12 in 31–40% of plots, 19 in 11–20% of plots and 113 in 10% or less of plots. Mean species richness was 25 (sd 10; range 8–47); mean FVI was 4.2 (sd 4.2; range 0–19). Species richness and FVI were significantly correlated ($r = 0.47$, $p < 0.01$). Fern species with rarity value 4 were present in three plots: plots 24 and 31 (*Dryopteris guanchica*) and plot 3 (*Hymenophyllum tunbrigense*). Fern species with rarity value 3 were present in plot 31 (*Cystopteris fragilis*), plots 3, 24, 36, 37, 38 and 39 (*Dryopteris aemula*) and plot 3 (*Woodwardia radicans*): a total of seven plots thus contained one or more species with a rarity value of 3 or 4. Twelve plots had FVI values of 7 or more (plots 3, 5, 14, 15, 16, 24, 25, 27, 31, 38, 39 and 40).

TWINSPAN classification of the 44 plots

The program was set to carry out four divisions (i.e. to produce 16 groups of plots). We selected the classification obtained at the third division (i.e. 8 groups) as being most useful, for two reasons. First, six of the groups produced at the fourth division were single-plot 'splinter groups'. Second, the third division was that which clearly (though not entirely) separated *Eucalyptus*-dominated sites from all other sites. The characteristics of each of the eight final groups are summarized in Tables 4 and 5. TWINSPAN indicator species (not to be confused with our 'three-species criteria'; see Discussion) are listed in Table 6.

The environmental correlates of the classification are considered only briefly here, since they are not directly relevant to the aims of this study and since we lack reliable data on several factors likely to be very important in determining floristic composition at our sampling scale (particularly recent history as regards burning, grazing and tree-felling, microclimatic factors such as insolation regime, and edaphic factors). The following characteristics of groups were considered:

1. mean altitude,
2. mean estimated slope,
3. proportion of south-facing plots,
4. proportion of plots beside streams ('streamside' plots),
5. proportion of plots with *Corylus* dominant or co-dominant in the tree layer, and
6. proportion of plots with *Quercus* dominant or co-dominant in the tree-layer.

Groups were compared pairwise using Student's t test for continuous variables (altitude and estimated slope), the normal approximation to the binomial distribution (Zar 1984, p. 396) for proportions when n was greater than 30 and Fisher's exact test (Zar 1984, p. 393) for proportions when n was less than 30. Only those characteristics which differ significantly ($p < 0.05$) between the pairs of groups considered are mentioned. Neither mean estimated slope nor proportion of *Eucalyptus*-dominated plots differed significantly between any of the pairs of groups considered, in the latter case clearly because of small sample size.

In comparison to group B, group A contains significantly more plots close to streams ('streamside' plots; 43% of A plots versus 0% of B plots, $p< 0.01$), significantly more plots in which *Corylus* is dominant (52% versus 9%, $p<0.01$), and significantly fewer plots in which *Quercus* is dominant (14% versus 61%; $p<0.01$). The mean altitude of group-A plots is significantly lower than that of group-B plots (96 m versus 218 m; $p<0.01$), though altitude range in both groups is high (see Table 4).

Group A was divided into one group (A1) of 11 plots and another (A2) of 10 plots. In comparison to group A2, group A1 has significantly more south-facing plots (73% of A1 plots versus 20% of A2 plots; $p<0.05$).

Group B was divided into one group (B1) of 12 plots and another (B2) of 11 plots. Again the proportion of south-facing plots differs between the two groups (83% of B1 plots and 27% of B2 plots; $p<0.05$), and mean altitude is significantly lower in B1 than in B2 (140 m versus 305 m; $p<0.01$). B1 contains all *Eucalyptus*-dominated sites, though some sites at which *Eucalyptus* is present at low density are included in group B2.

Group A1 was divided into one group (final group 1) of five plots and another (final group 2) of six plots. Group 1 has significantly more streamside plots than group 2 (80% of group-1 plots versus 0% of group-2 plots; $p<0.05$).

Group A2 was divided into two groups of five plots (final groups 3 and 4). Again, group 3 has significantly more streamside plots than group 4 (100% of group-3 plots versus 0% of group-4 plots; $p<0.01$).

Group B1 was divided into one group of eight plots (final group 5) and another of four plots (final group 6). None of the group characteristics considered differ significantly between these plots, though group 6 contains all three *Eucalyptus*-dominated plots.

Group B2 was divided into one group of four plots (final group 7) and another of seven plots (final group 8). Again, none of the group characteristics considered differ significantly between these groups.

Did TWINSPAN group high-value plots together?

The three plots which were high-value on definition 1 (because they contained species with rarity value 4; see Methods) were assigned to group 1 (plot 31) and group 3 (plots 3 and 24).

The seven plots which were high-value on definition 2 (because they contained species with rarity value 3 or 4) were assigned to group 1 (plot 31), group 3 (plots 3, 24 and 38), group 7 (plot 43) and group 8 (plots 39 and 40). The additional plots (with respect to those which were high-value on definition 1) are all rated as high-value because they contain *Dryopteris aemula*: as noted above, this species (unlike the other high-value species) is not limited to very wet sites or to low-altitude sites.

The 12 plots which were high-value on definition 3 (because they had FVI values of 7 or more) were assigned to group 1 (plots 25, 27 and 31), group 3 (plots 3, 15, 16, 24 and 38), group 4 (plot 14) and group 8 (plots 39 and 40). FVI varies significantly among groups (one-way ANOVA: the factor group is a significant source of variance in FVI; $F_{(7, 36)}=9.2$, $p<0.001$). This is as expected given the co-occurrence of the endangered ferns (except *Dryopteris aemula*). Comparison of multiple means (95% confidence level) separates four 'supergroups' (i.e. groups of groups) which have homogeneous FVI; FVI is highest in the supergroup containing groups 1 and 3, and FVI in the group with the highest mean value (group 3) is significantly higher than in all other groups except group 1.

Thus, if *Dryopteris aemula* is excluded, TWINSPAN grouped the high-value plots fairly consistently (in groups 1 and 3, together containing ten of the 44 plots, including all nine streamside plots). With the exception of plot 38 (the only high-altitude streamside plot, assigned by TWINSPAN to group 3), all plots containing *Dryopteris aemula* were assigned to groups 7 and 8 (i.e. the high-altitude groups): we have relatively few samples from higher-altitude areas, and our data are thus probably insufficient to allow accurate identification of site-types important for this species.

Discriminant analysis between high-value plots ('strategy 1')

As noted in the preceding section, definition 2 (of high-value plots) was of dubious value for our data, because it identified two basically different groups, i.e. plots without *Dryopteris aemula* and plots with *Dryopteris aemula*. Between-plots discriminant analysis (program CRITERIO; see Methods) was therefore applied using definition 1, and the additional definition 'Presence of *Dryopteris aemula*?', to determine high-value plots. The best three-species combinations in each case were as follows.

1. Target group: plots which were high-value on definition 1 (plots 3, 24 and 31). Best combinations: (*Ceratocapnos claviculata* + *Veronica montana* + *Cardamine pratense*) or (*Ceratocapnos claviculata* + *Cardamine flexuosa* + *Laurus nobilis* sapling). Both combinations identify 100% of target plots and 2% of non-target plots.
2. Target group: plots containing *Dryopteris aemula* (plots 3, 38, 39, 40, 43). Best combinations: (*Luzula sylvatica* + *Erica arborea* + *Melampyrum pratense*) or (*Luzula sylvatica* + *Fraxinus excelsior* + *Vaccinium myrtillus*) or (*Luzula sylvatica* + *Melampyrum pratense* + *Vaccinium myrtillus*). All three combinations identify 100% of target plots and 7% of non-target plots.

As discussed below, the discriminant analysis is ineffective when the target group contains a single plot; for example, when the target plot is that containing *Hymenophyllum tunbrigense* (plot 3), over 20 'best combinations' are identified, most of which are useless. For example, eight of the combinations include *Aquilegia vulgaris*, a woodland-fringe species which, in our sampling programme, occurred only in plot 3, but which is clearly not consistently associated with this habitat type.

Discriminant analysis between TWINSPAN groups ('strategy 2')

Between-group discriminant analysis was applied firstly with TWINSPAN groups 1 and 3 as the target group (since these were the highest-value on both definitions 1 and 3). The best three-species combinations was (*Viola riviniana* + *Luzula sylvatica* + *Rubus*). This combination identifies 100% of target plots and 17% of non-target plots.

Which criteria are the most useful?

As discussed below, we consider that the strategy 1 methods were probably ineffective for our data set. We therefore carried out a preliminary trial of the criteria derived by strategy 2, in which we deliberately imitated a rapid field survey by walking up a stream gully known to contain subpopulations of all the highest-value species at various points along its length. At each point at which one of these species was observed, we noted down whether the TWINSPAN-derived criterion (two or more of *Viola riviniana* + *Luzula sylvatica* + *Rubus*) was met. The criterion species were taken to be present if they were 'visible nearby', again in imitation of rapid field survey. One or more of the rare ferns were observed in 41 of the 56 10-metre stretches along the estimated 560 m length of the gully: the TWINSPAN criterion 'worked' (largely because of the presence of *Luzula* and *Rubus*) in all but 12 of these 41 stretches, and in no case was a fern subpopulation more than 30 m from the nearest *Luzula*. The criterion 'failed' (i.e. two or more of the criterion species, but no rare ferns, were present) in six of the 15 stretches which did not contain any rare ferns. About halfway up this main stream, another stream of similar size joins it: this gully has been affected by severe fire within the last three or four years. A small group of *Woodwardia radicans* about 20 m upstream of the main stream constitutes the only subpopulation of endangered ferns present along its length. *Luzula* was present within 10 m of this subpopulation but not at any other point along the length of this stream, again providing support for the TWINSPAN-derived criterion.

Discussion

We have compared a number of related methods for determining which species are the most useful indicators of site-types important for rare ferns in an area of woodland in northwest Spain. Our problem is to some extent artificial: the majority of high-value sites are streamside sites, and the most straightforward conservation strategy would be to give special protection to all streams. However, not all streams in the area support subpopulations of endangered ferns, and some streams are low-value along much of their length (often as a result of fire). In addition, some stretches along the banks of the River Eume (not included in any of our samples) support important subpopulations of *Hymenophyllum tunbrigense* and/or *Woodwardia radicans*. Finally, important subpopulations of both *Dryopteris aemula* and *Culcita macrocarpa* are located in non-streamside sites (in the latter case usually on steep rocky slopes with subsurface water flow; personal observations). Floristic criteria may thus be more accurate than gross habitat-type criteria.

One of the aims of this study was to determine whether it is more effective to look for those conspicuous species which best discriminate between high-value plots and all other plots (strategy 1), or to first carry out a global plots-by-species classification and then look for those conspicuous species which best discriminate between high-value groups and all other groups (strategy 2).

The main problem encountered with the first strategy – attributable to the fact that we had only a few plot inventories containing the high-value species – was that some of the discriminant species selected were probably present in those plots 'by chance' (i.e. are species which do not consistently occur in the same site-types as the endangered species of interest): a probable example of this is the selection of *Ceratocapnos claviculata* as one of the species discriminating between plots 3-24-31 and the rest. Where only one plot was considered (as for example plot 3 for *Hymenophyllum tunbrigense*) this problem was greatly magnified, and the discriminant analysis produced a large number of equally ineffective combinations, many reflecting obviously atypical occurrences such as *Aquilegia vulgaris*, *Vaccinium myrtillus* or *Erica arborea*. Clearly, the ideal solution would be to collect inventories from a large number of sites containing the endangered species of interest: however, since endangered species are often present in only a very few known sites in a given area of known conservation value, this may not be possible. In addition, as was the case in the present study with *Dryopteris aemula*, the endangered species may occur in site-types with quite different floristic compositions. For these reasons we consider that the second strategy may be more useful.

The main problem encountered with the second strategy was that of deciding which classification method to use. A single-linkage cluster analysis of the data set (results not shown) grouped plots very differently to TWINSPAN, and it is difficult to decide which classification is the most useful for the purposes of this study. The nature of the TWINSPAN algorithm (see Hill et al. 1975; Kent & Coker, 1992) makes it most highly suited for identifying groups at each extreme of one or a small number of environmental gradients. We consider, however, that floristic composition within 200 m^2 plots in seminatural woodland is likely to be governed by very complex interactions among a large number of environmental and recent historical factors: for data of this type, a cluster analysis may be more effective at identifying the 'natural' group structure. In addition, our primary aim was to identify plots similar to those containing the endangered ferns, not to identify gross group structure in the whole data set: this argues for the use of a single-linkage cluster analysis. On more pragmatic grounds, however, the TWINSPAN classification proved more effective. When high-value plots were taken to be those containing one or more of the fern species with a rarity value of 4 (i.e. plots 3, 24 and 31), and high-value groups to be those groups containing those plots, the three-species criterion derived from the TWINSPAN classification (*Luzula sylvatica*, *Viola riviniana*, *Rubus* sp.) proved effective in preliminary field trials.

We consider that simple discriminant analyses of the type used in this study are of potentially wide application. The functions of the indicator species identified by TWINSPAN are very clearly defined by Hill et al. (1975) but, in our opinion, are often misinterpreted in the literature. TWINSPAN indicator species are designed to allow assignment of a full plot inventory to its corresponding group by going through each preceding stage in the decision hierarchy: they are not necessarily those species which best discriminate between one group of plots and all other groups (which is very often what is required). In addition, since the TWINSPAN classification is successively readjusted to maximize the predictive power of the indicator species, it is not possible to carry out a classification in which all species are considered but in which only some species are allowed to be indicators: again, this is very often what is required. Simple discriminant analyses of the type described here can be used to determine which combination of conspicuous (and/or easily identified and/or wintergreen) species best discriminates between a target group and either all other groups or a subset of all other groups.

Our results suggest that the presence of two or all of *Luzula sylvatica*, *Viola riviniana* and *Rubus* is an effective and feasible criterion for identifying sites of value for rare ferns in the Eume woods. This conclusion is based on fairly crude evaluation of the physical characteristics of sites, and the method would certainly be improved by a more detailed analysis: in the case of ferns, for example, the presence of moist rock-formed microhabitats suitable for establishment of the gametophyte generation may be of particular importance. Whether techniques of this type are of value for the study of other endangered species in other areas will depend very much on the characteristics of that species (or that group of species) and on the degree and complexity of community structure in the locations in which it occurs. It is also important to note that we considered a number of rare fern species which tend to co-occur; in many cases it may be a single species which is of interest. Despite these reservations, we consider that techniques of this type may be of value for large-scale surveys of endangered species which are difficult to identify in the field. Floristic criteria derived in this way might also be of value for identifying sites of importance for animal (e.g. insect) species.

Acknowledgments

We are grateful to Carlos Vales for the use of his map of wooded areas of the Eume Valley, and to two anonymous reviewers whose comments provoked careful reassessment of this work.

References

Braun-Blanquet, J. 1979. Fitosociologia: Bases para el estudio de las comunidades vegetales. Blume, Madrid.

Hill, M. O., Bunce, R. G. H. & Shaw, M. W. 1975. Indicator species analysis: a divisive polythetic method of classification and its application to a survey of natural pinewoods in Scotland. J. Ecol., 63: 597–613.

Hill, M. O. 1979a. TWINSPAN – a FORTRAN program for arranging multivariate data in an ordered two-way table by classification of the individuals and attributes. Cornell University, Ithaca, New York.

Hill, M. O. 1979b. DECORANA – a FORTRAN program for detrended correspondence analysis and reciprocal averaging. Cornell University, Ithaca, New York.

Horjales, M., Redodo, N. & Pérez-Prego, J. M. 1988. Notas corológicas sobre pteridófitos gallegos. II. Anal. Jard. Bot. Madrid, 45(1): 338–341.

Izco, J., Amigo, J. & Guitián, J. 1990. Los robledales Galaico-Septentrionales. Acta Bot. Malacitana, 15: 267–276.

Jalas, J. & Suominen, J. 1972. Atlas Florae Europaeae, Vol. 1, Pteridophyta. The Committee for Mapping the Flora of Europe. Helsinki.

Kent, M. & Coker, P. 1992. Vegetation analysis. CRC Press, Boca Raton.

Klopatek, J. M., Kitchings, J. T., Olson, R. J., Kumar, K. D. & Mann, L. K. 1981. A hierarchical system for evaluating regional ecological resources. Biol. Conserv. 20: 271–290.

Losa Quintana, J. M. 1974. Vegetación y flora de la baja cuenca del rio Eume. Estudio botánico y fitosociológico del bosque natural de *Quercus robur* L. Bol. Soc. Brot., 2ª serie, 47: 243–267.

Nilsson, C. 1986. Methods of selecting lake shorelines as nature reserves. Biol. Conserv. 35: 269–291.

Retuerto, R. & Carballeira, A. 1991. Defining phytoclimatic units in Galicia, Spain, by means of multivariate methods. J. Veg. Sci. 2: 699–710.

Rodriguez-Gracia, V., Gómez-Vigide, F., Valdés-Bermejo, E., Garcia-Martinez, X. R. & Silva-Pando, F. J. 1990. Catálogo de la flora vascular gallega. I. Pteridophyta, Gymnospermae y Monocotoledineas. In 'Sobre Flora y Vegetación de Galicia', Silva-Pando (Ed.): 71–89. Consellería de Agricultura, Xunta de Galicia.

Salvo, E. 1990. Guia de helechos de la Peninsula Ibérica y Baleares. Pirámide, Madrid.

Soñora, F. X., Ortiz, S. & Rodriguez-Oubiña, J. 1992. *Hymenophyllum wilsoni* Hooker (Hymenophyllaceae) in the Iberian Peninsula. Rhodora, 94, 879: 316–318.

Tutin, T. G. et al 1964–1981. Flora Europaea. Cambridge University Press, Cambridge.

Vales, C. 1992 Plano de Ordenación dos Recursos Naturais do Val do Eume. I Xornadas Internacionais sobre os Bosques Atlánticos. June 1992, Pazo de Mariñan, Coruña.

Zar, J. H. 1984. Biostatistical Analysis (2nd ed.). Prentice-Hall International, New Jersey.