

CLASSIFICATION OF VASCULAR PLANT COMMUNITIES OF THE NORTH CASCADES USING DISCREET SPACE BOUNDARY ANALYSIS

by George Wooten¹ and Peter Morrison²

Published by Floradora Farms in 2006 as an open file report

Available online at

<http://www.okanogan1.com/ecology/eco-ncas>

Abstract

Plant communities in the North Cascades Mountain Range of Washington respond to a broad range of environmental factors, spanning 11.7 SD of species turnover across the primary axis of variation. Vegetation types encompassing maximum satellite spectral and geographic representation across the North Cascades were sampled and 1433 detailed plots were selected for further analysis. These plots were aggregated into composite sample sets using their mean abundance values. The composite samples were partitioned into discreet regions on color-coded scatter plots of the detrended correspondence analysis ordination space, bounded by minimum surfaces along indicator species gradients. The set partitions were then defined by a rule-based sequential filter that separated plots by the abundances of differential plant species. The inequalities which defined the rules formed a dichotomous classification of 14 major sets of North Cascades plant communities. These data sets and their classification keys were compared with other classification systems used in the Pacific Northwest in a classification concordance. The primary ordination axis spanned 11.7 SD and was aligned along an east-west gradient. The secondary ordination axis spanned 6.9 SD, aligned along a wet-dry gradient. Approximately 4 SD represents a complete species turnover, i.e., the ordination distance in which none of the same species occur in the plots. Thus, the primary axis of this data represents approximately 3 complete community turnovers in going from east to west across the North Cascades.

Keywords. Relevés, ordination, heirarchical classification, detrended correspondence analysis, reciprocal averaging, ground truthing, spectral classes.

1. George Wooten, Floradora Farms, Twisp, WA 98856 (509-997-6010).

2. Peter Morrison, Pacific Biodiversity Institute, Winthrop, WA 98862.

Classification of North Cascades Vascular Plant Communities

<http://www.okanogan1.com/ecology/eco-ncas>

© George Wooten 1996+

Introduction

Classification of natural ecosystems into potential plant communities, or habitat types, is an important tool for the long-term management of natural resources (Daubenmire and Daubenmire 1968; Pfister and Arno 1980). The classification of forested successional communities has enabled the mapping of large areas by their climax habitat types or association (Franklin and Dyrness 1973; Pfister 1976).

Through the classification and mapping of sufficient numbers of vegetative ecology plots, regional resource maps have been and are being constructed for large areas in the world (Kuchler 1973; Rubec 1979; Brown et al. 1980; DeGraff et al. 1988; Klinka et al.; 1991).

Remotely sensed images of formerly inaccessible land areas has necessitated the development of alternative methods of land classification to cope with massive amounts of new data. (Craighead et al. 1982; Agee et al. 1985). The classification of satellite multi-spectral images into vegetative cover types (ground truthing) also requires establishing large numbers of ecology plots which can be correlated with site information such as elevation and slope, as well as the spectral signature.

Recent legislation enacted to protect the environment has involved botany, ecology and soil science in developing credible methods of classifying land and vegetation, (Tiner 1993; Environmental Laboratory 1987).

With sufficiently large numbers of samples and species, the mean species abundances (attributes) of the samples become statistically reliable indicators of community and environmental relationships (Ludwig and Reynolds 1988).

Consequently, both habitat-typed and ground-truthed ecology plots can be analyzed with multivariate techniques that (1) allow summarizing large, complex data sets,

(2) aid in the environmental interpretation of and hypothesis generation about community variation and (3) refine models of community structure (Gauch 1982).

The aim of this study is to present the results of and methods of vegetative classification using multivariate statistical analysis on a large, complex vegetation database from the North Cascades of Washington, and to describe specific methods useful in the development of a regional classification (Küchler 1973; Klinka et al. 1991). This is a necessary prerequisite in discerning environmental patterns within plant communities, and contributes to knowledge about how to conduct ecosystem management of those parameters.

The study area is in the North Cascades of the Pacific Northwest, from the western border of the Mt. Baker-Snoqualmie National Forest east to the Okanogan and Columbia Rivers and from the Canadian border south to Interstate 90 (fig. 1). This 2,600,000 hectare study area is equivalent to the North Cascades Grizzly Bear Ecosystem (Almack et al. 1993), in which approximately 150 cover types were mapped and quantified.

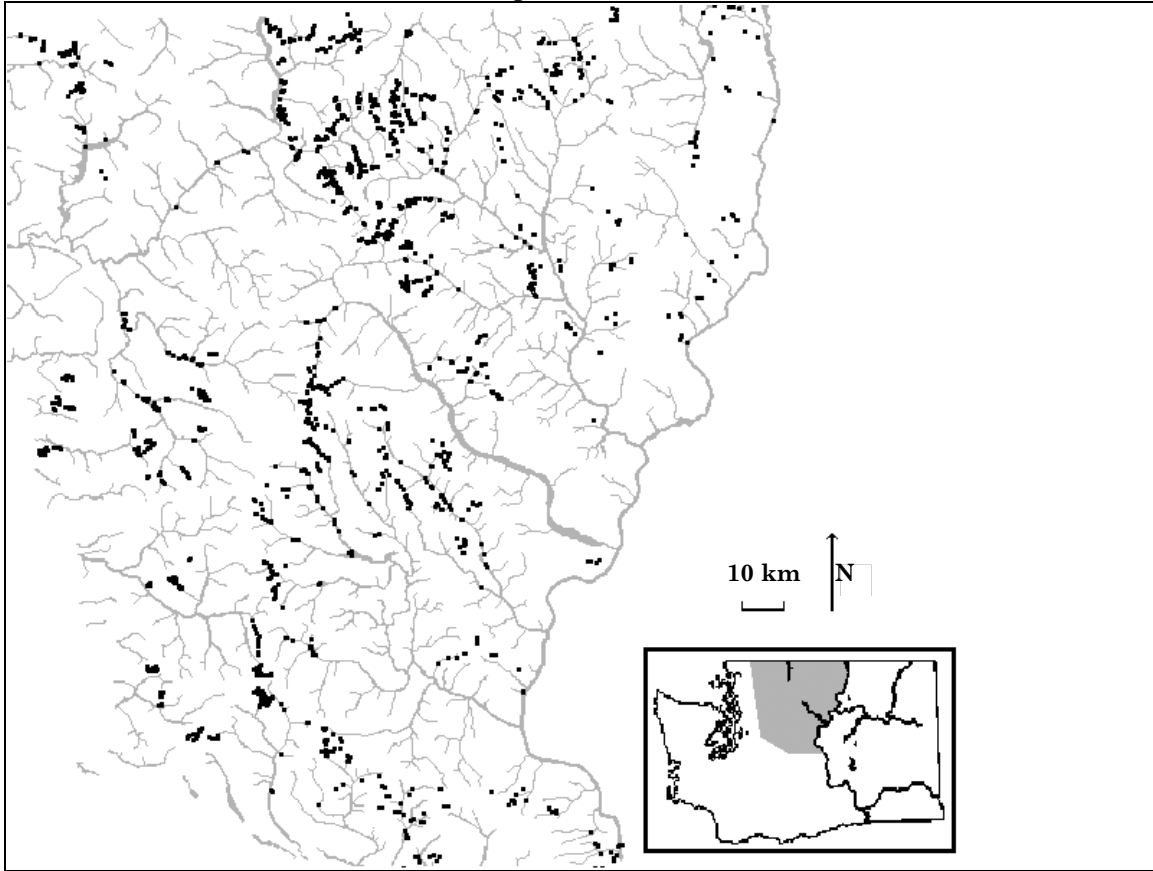


Fig. 1. Watershed map of the North Cascades showing locations of 1741 ecology plots (small black squares). Inset shows boundary of the study area in Washington state (shaded).

Methods

Sampling design. Two Landsat multi-spectral scanner scenes covering the North Cascades of northern Washington and taken in the summer of 1986, were mosaiced together and installed into an EPPL7 GIS (a geographic information system from Minnesota Land Management Information Center), with a 57 m² pixel resolution and georeferenced to UTM coordinates projected onto state zone 10. The western scene, comprising approximately one-third of the area, had 87 spectral classes and the eastern scene had 101. Vegetation "stands" were selected for sampling based on (1) samples having broad, even distribution across the study area, (2) representation of as many spectral classes as possible from among the 188 spectral classes, (3) homogenous vegetation and (4) minimum stand sizes of 3 X 3 pixels.

Since data on 3751 ecology plots dealing primarily with forested, upland plant associations were already available as part of the National Forest and National Park ecology programs, selection of stands in this study were prioritized to augment that data with sampling both above and below the continuous forest zones. To obtain the widest range of geographic and community representation, plots were placed preferentially in areas where previous ecology data was unavailable. This meant that non-forest types, and even non-vegetative types, such as bare ground, roads, open water, rock, snow, and ice, were also sampled. Stand selection was also prioritized by spectral representation, since only 5% aerial coverage was represented by the lowest 30% of spectral frequencies out of the total 101 spectral classes for the eastern satellite scene alone. For this statistical reason, high frequency occurrences of spectral classes were given higher sampling priority, and this proved to be an efficient method of obtaining data, since many of the less represented spectral classes were also located in inaccessible terrain, or did not

Site information. An area near the center of the vegetative stand with typical vegetation was used as the center of a fixed radius plot of 11.35 m radius (0.10 ac). Fixed plots were measured using a standardized protocol for forested stands (Pfister and Arno 1980; Williams and Lillybridge 1983). Fixed plots were used for basal area, stem counts, by diameter size classes, of overstory (> 10 m) and understory (<1 m, 1-3 m, 3-10 m) trees, tree area coverages, and snag and log counts. The remainder of the stand area was used to complete site data, vegetation cover values, noted disturbances and remarks. After completion of the first 237 plots in 1986, the fixed radius part of the plots was replaced by variable-perimeter stands matching the boundary of the spectral class polygons. Exact tree, snag and log counts were categorized with counts of 10 m or greater tree stems in 3.0 dm dia classes (estimated as percentages of the total stem counts), counts of large and small snags and logs in count classes of 0-1, 2-6 or >6, and percentage tree cover in 3 classes of >10 m, 1-10 m, and <1 m, with height of the average tallest cohort recorded.

Trees were defined as species capable of reaching heights of >10 m with a single, upright, woody stem. Canopy cover, or crown density, was determined for all plots, as it is one of the strongest determinants of forest structure, and is inversely correlated with several of the Landsat spectral bands. Therefore it was important to have accurate information on the coniferous canopy density in all plots. Through the year 1989, all plots included 5 measurements of canopy density taken with a spherical densiometer, each measurement being the average of 4 directions and taken at least 36 m distant from each other (or 73 m for canopies less than 50%) in a random direction. Tree canopy density was measured for the sum of all overstory coniferous trees, and again estimated for the sum of the deciduous overstory.

During autumn, leaf coverage was estimated as the outline of the crown for deciduous trees. Species defined as shrubs were not counted as part of the overstory canopy, even if they were overhead.

A separate statistical study of this densiometer data revealed that measurements of crown coverage are least reliable between 20 and 70% cover, but reasonable below 20% or above 70%. Comparison of this method with that of canopy estimation by comparing stands on aerial photographs with graded randomized-dot scattergrams, revealed discrepancies of up to 20% canopy cover between methods. While photographic methods allow visual averaging over an entire stand, and are extremely rapid, they are not species-specific, they are poor at estimating canopy heights, they do not account for interstitial openings in individual trees and there is a tendency to overestimate middle coverages, possibly due to the addition of tree shadows to the dark crown coverages. However, the use of these two methods together, along with the estimates of tree cover, provided important checks on the validity of the data.

Vegetation abundances were quantified by estimates of percent cover (Daubenmire 1959), the area of the vertical projection of the plant outline onto a horizontal plane. Standardization of cover estimation was achieved through training exercises until all field personnel were able to consistently estimate percent cover within approximately 10% of each other. Vegetation coverages between 0% and 1% were recorded as 0%, except for tree species, which were rounded up to 1%. The percent cover of the remainder of the vegetation was recorded using USDA standardized 4- and 5- letter alphacodes (Garrison and Skovlin 1976; see appendix 7), with some modifications to accommodate newer taxa. Because of the overlap of vegetation

layers, total vegetation coverages greater than 100% are normal.

An attempt was made to identify all species within the plot, up to approximately 50 species. Specimens were identified in the field using Hitchcock and Cronquist (1973), or collected and identified later in the laboratory, using Hitchcock et al. (1969). Specimens of difficult species were sent to Western Washington University for identification or verification in the laboratory of Dr. R. J. Taylor.

Site information collected included elevation, aspect, slope, satellite spectral class, stand area, and the estimated percent (from 0% to 100%) of the plot surface covered with bare soil (<1 mm dia), gravel (1-10 mm dia), rock (>10 mm dia), bedrock, moss, litter, soil lichens, rock lichens and surface water (measured on the day of the plot). These percentages plus the area occupied by the basal portion of all plants (where plants contact the soil) total 100%. When logs above the ground accounted for a measurable area within the plot, the area of the vertical projection of the visible log surfaces onto the ground surface was counted as litter if the logs were within 1.5 m of the surface of the ground. Dead plant matter remaining attached to live plants but lying on the ground, as in bunchgrasses, was not counted as litter unless it was separate from the plant.

Miscellaneous information included general remarks, location, plant community type, aerial photo number, notes on adjacent vegetation, landform, patchiness, number of canopy layers, density (0-3) of arboreal lichens, frequency and intensity of all disturbances (including logging), names of observers and time of day (to correspond with time- and date-stamped photographs). Plots were indelibly marked on orthophotos and topographic maps, and pin-pricked on aerial photographs, with

plot numbers marked on the back of the photo. A photograph depicting the appearance of the plot was taken and an aluminum flasher was fixed nearby, except in designated wilderness.

During the third year of the plot collection, when locating plots boundaries along spectral class polygons was begun, a procedure of describing the cover values for microsite or subcommunities within the plot was initiated. However, in this treatment of that data, each numbered plot is given by the mean species abundance of the sum of all the microsities within the plot.

To gather complete geographic and spectral coverage of the study area, 330 abbreviated plots with partial plant information were done that grouped vegetation as overstory, understory, deciduous, coniferous, shrub and herb, recording only the highest three percentages of species cover for those categories. This allowed a more rapid completion of a greater number of plots and enabled the inclusion of data for inaccessible terrain. However, the omission of an unknown number of plants meant many of those plots would be unsuitable for multivariate analysis techniques. Consequently, the abbreviated plots were reviewed and only 75 of them that were judged to have sufficiently complete vegetation information were included in this study.

Taxonomy. All plants were keyed to species level whenever possible and varietal names were avoided, except when appropriate or necessary.

If the species could not be determined, then the lowest division which it was known to be a member of was recorded instead. Difficult cases which often could not be

satisfactorily determined include couplets such as *Arctostaphylos uva-ursi*/*A. nevadensis*, *Aster engelmannii*/*A. ledophyllus*, *Betula X piperi*, *Berberis aquifolium*/*B. repens*, *Castilleja elmeri*/*C. miniata*, *Dodecatheon conjugens*/*D. pulchellum*, *Hackelia arida*/*H. diffusa*, *Potentilla arguta*/*P. glandulosa*, *Rosa nutkana*/*R. woodsii*, and *Stipa occidentalis*/*S. lemmonii*.

Generic or familial names were often necessarily recorded for difficult groups, particularly *Agoseris*, *Antennaria*, *Arabis*, *Arnica*, *Artemisia*, *Balsamorhiza*, *Bromus*, *Campanula*, *Carex*, *Cerastium*, *Crepis occidentalis* group, *Cryptantha*, *Epilobium ciliatum* group, *Equisetum*, *Festuca*, *Hieracium scouleri* group, *Juncus*, *Lomatium*, *Lupinus*, *Microseris*, *Osmorhiza*, *Phlox*, *Ranunculus*, *Ribes*, *Rubus*, *Salix*, *Saxifraga integrifolia* group, *Senecio*, *Stellaria*, *Taraxacum ceratophorum*, *Poa secunda* complex (treated here sensu Hitchcock and Cronquist 1973; but see Kellogg 1985), *Vaccinium* (juvenile forms), *Veronica*, and *Viola*.

Recent changes in the nomenclature affecting indicator species or community names are as follows: *Carex rostrata* Stokes (CARO2), is changed to *Carex utriculata* Boott. (CAUT); *Carex prionophylla* Holm (CAPR8), is changed to *Carex scopulorum* var. *prionophylla* L. A. Standley (CASCP2); *Salix phyllicifolia* L. (SAPH), is changed to *Salix planifolia* Pursh (SAPL).

Species which have since been determined contrary to our initial findings are as follows: *Caltha biflora* DC. var. *biflora* (CABIB) replaces *C. leptosepala* DC. (CALE2), according to Mildred Arnot, University of Washington Herbarium. This may mean that the TSME/CABI plant association of the Mount Baker Ranger District refers to *Caltha biflora* var. *rotundifolia* (Huth) Hitchc. (CABIR), however,

it is notable that the Welcome Pass area of that District supported both varieties. *Poa secunda* Presl. ssp. *secunda* (POSES1) replaces *P. grayana* Vasey (POGR5, *P. arctica* R. Br.), according to R. J. Soreng, L. H. Bailey Hortorium, Cornell University, New York. This is a tall plant of wet habitats, unlike the *Poa secunda* that occurs east of the Cascades in habitats with *Artemisia tridentata* Nutt. ssp. *wyomingensis* Beetle and A. Young.

Classification and ordination. The classification technique described here is an improvement of detrended correspondence analysis space partition (DCASP, Gauch and Whittaker 1981). In this method, termed *discreet space boundary analysis*, ordination space is partitioned into discreet regions bounded by surfaces of minimum area, and constrained to identity with indicator species divisions.

The method takes advantage of the ability of detrended correspondence analysis (DCA, using the computer program DECORANA, Hill 1979a; Hill and Gauch 1980) to extract meaningful community relationships from complex data through dimension reduction. DCA is free from distortions present in RA (reciprocal averaging, or correspondence analysis, CA) and PCA (principal components analysis) axes and its output is scaled in meaningful units of the average standard deviation of species turnover, SD (Gauch 1982). This is a natural outcome of Gaussian frequency distributions of plant populations along environmental gradients (Preston 1962; May 1975). The drawback of multidimensional methods like DCA is that even though the number of dimensions is reduced, the axes are inherently multidimensional. Extraction of environmental gradients and thus DCASP both suffer from a lack of objective criteria about where to make divisions amongst the continua expressed in a data set this large. Furthermore, only two or

three axes at a time can realistically be analyzed with DCASP.

The division of the samples into related groups was accomplished using two-way indicator species analysis (TWINSpan, Hill 1979b, Wheeler 1987), a hierarchical, polythetic, divisive technique for community classification based on RA. Unlike RA, TWINSpan re-computes a new primary axis for each division it produces.

TWINSpan displays the location of hierarchical divisions in a two-way table of samples and species scores. By combining these features with detrended correspondence analysis it was possible to derive an optimum classification using ordination space to display classification sets, and then optimizing the fit between adjacent sets using the TWINSpan table output.

Theory. This procedure is a hybrid ordination-classification. It combines DCA, which is useful for community modeling tasks, with TWINSpan, which excels at divisive classification and generating a two-way table. The linkage of these two techniques was accomplished through *successive refinement* of intermediate results (Gauch 1982, p. 35-37). It is important to be able to examine ordination axes (Whittaker 1987), and in this method this involved using a color-coded scatterplot to examine for goodness of fit between partitions projected onto DCA 2- or 3-dimensional space. The partitions were the boundaries of regions in ordination space that bounded sets of samples belonging to TWINSpan classification divisions. The criteria used for judging goodness of fit was that adjacent ordination partitions should form discreet sets identical with the classification, i. e., samples should not overlap onto the ordination-space of another classification set. Furthermore, the boundaries of adjacent partitions should be flat or convex, and bounded by surfaces of minimum area. By defining plant abundance criteria for

sample set formation based on the fit of those partitions, and then reiterating the procedure until all the sets were judged to be discreet, a classification was developed in conjunction with a set of rules for making the classification dichotomous.

Terms. The term *samples* as used here always refers to groups of data to distinguish them from data about a single vegetation plot, or relevé. The term *plot* is used like *quadrat* when referring to the physical location and shape of the vegetative community, and when referring to the actual vegetation abundances within that particular community. Generally the term *plant community* was used rather than *association*, *habitat type* or *stand*, which in the former two implies a previously described ranked potential category, and in the latter has the connotation that the area is forested. *Relevés* and *raw plot data* are used interchangeably here, typically in the singular and plural senses, respectively.

Preliminary division of the data set. The large number of raw data plots, or relevés, were divided into smaller sample sets defined primarily by geographic distribution. Of 1741 relevés available, 1436 were chosen that had complete species lists and cover values (attributes). These 1436 relevés were divided into 4 sets based on geographic distribution and each set was classified using TWINSpan, with pseudospecies cut levels of 0, 5, 20 and 50. The 4 sets contained from 237 to 479 samples, and 583 to 710 attributes which were reduced in each case to less than 300 attributes, primarily by removal of nonspecific genera and herbaceous species of low constancy or broad distribution.

The 4 classifications resulting from these sets were then regrouped to form 6 new

sets, which after deletion of outliers, contained 1428 relevés. Some of these were duplicated in more than one set, giving a total count of 1433 relevés. These 6 new sets had more consistency than the first sets, except that one set was used for 29 unusual samples. The other sets contained from 193 to 397 samples, and from 427 to 605 attributes, which were reduced and classified with TWINSPAN as before.

Formation of composite samples. In this study, the large number of raw data plots was reduced for computational efficiency as well as for elimination of redundancy. Based on the TWINSPAN divisions, a set of 261 composite samples with 592 species was generated from the 1433 relevés using two Basic programs, MCSORT and MCS2EFZ. MCSORT generated a summary of mean vegetative cover and constancy for each ultimate TWINSPAN division and MCS2EFZ averaged this data into composite samples; all non-generic plants with constancies $\geq 50\%$ were output as mean abundances; plants with $< 50\%$ constancy were excluded unless their mean cover was $\geq 75\%$. For all the TWINSPAN and DECORANA analyses presented here, no species downweighting, or axis transformations were used, since the cut levels were chosen to approximate a logarithmic progression (Preston 1962). Attributes were the TWINSPAN cut levels which represent plant species abundances. For trees, the attributes were grouped into two classes, overstory (> 10 m) and understory (< 10 m).

Initial ordination of the data. The preliminary set of 261 composite samples was reduced to 249 samples by elimination of 12 disjuncts from cultivated and fallow fields and from 553 pseudospecies to the 379 species of highest constancy. The data was then divided into sets based on TWINSPAN classification using cut levels of 0, 0.2, 3.5, 16 and 35% cover, and was then ordinated with DCA. Eigenvalues of the

first 4 axes were 0.917, 0.751, 0.573, and 0.487 (fig. 3).

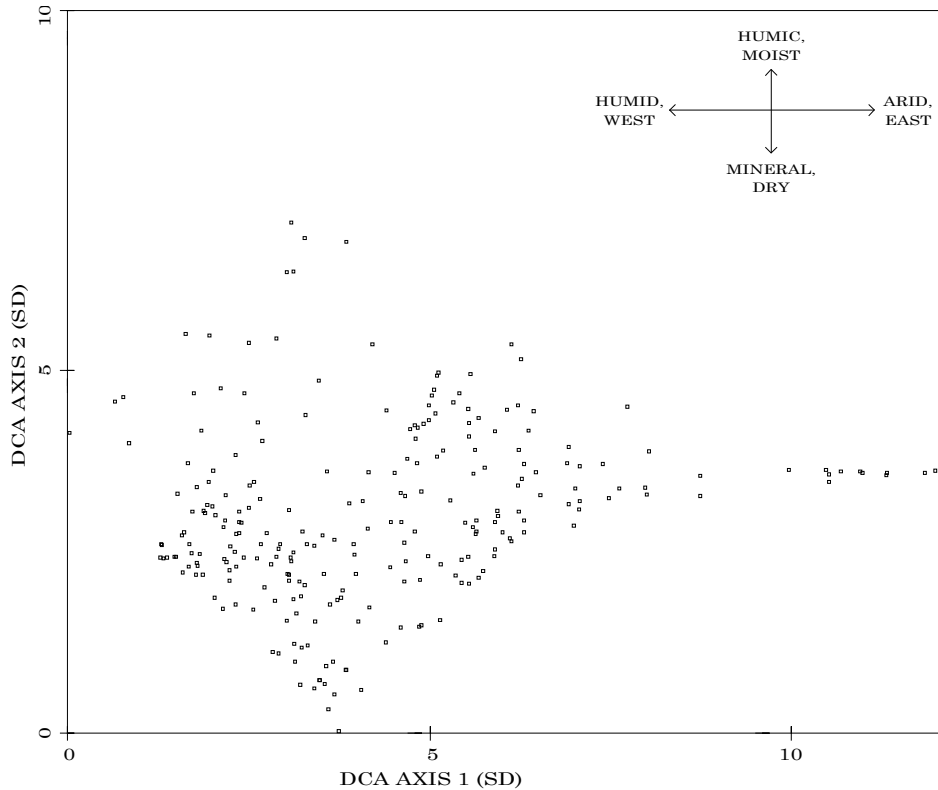


Fig. 3. Sample ordination of the two major DECORANA axes before set optimization. Small squares represent 249 composite samples. Samples representing fallow and cultivated fields were not analysed with this ordination.

Figure 3 tells us that 91.7 percent of the data can be explained by the gradients represented by the first axis. Although the nature of the underlying environmental gradients is a combination of natural processes, in our data, the first and second gradients ultimately correlated with wetness and latitude. Only the first and second axes were used in ordinations.

The species ordination used 261 composite samples, including the 12 samples with fallow and cultivated vegetation, and was reduced from 783 pseudospecies to the 413 of highest constancy. Sets were formed from TWINSPAN divisions of the samples with cut levels of 0.1, 4, 16 and 40% cover, and DECORANA was run on

the sets. Eigenvalues of the first 4 axes were 0.876, 0.595, 0.481, and 0.375 (fig. 4).

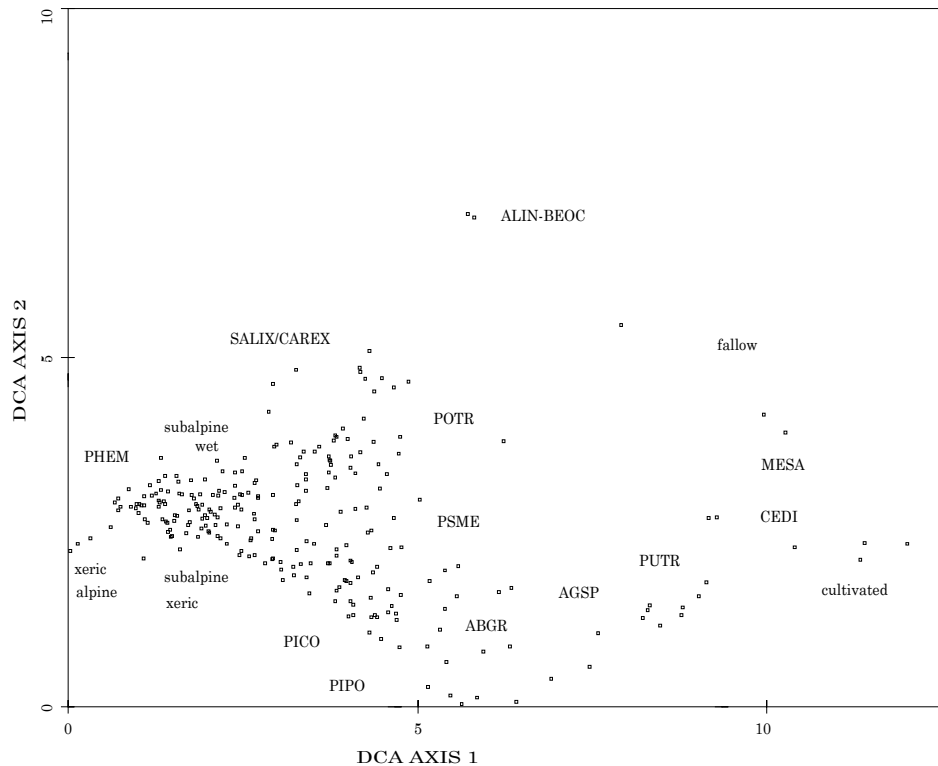


Fig. 4. Partial species ordination of the two major DECORANA axes before set optimization. Small squares represent 413 species from 261 composite samples; approximate locations of important species or species groups are labelled: ABGR = *Abies grandis*, AGSP = *Agropyron spicatum*, ALIN = *Alnus incana*, BEOC = *Betula occidentalis*, CEDI = *Centaurea diffusa*, MESA = *Medicago sativa*, PHEM = *Phyllodoce empetriformis*, PICO = *Pinus contorta*, PIPO = *Pinus ponderosa*, POTR = *Populus tremuloides*, PSME = *Pseudotsuga menziesii*.

Labels for the sets of composite samples. The sample subsets produced by the major divisions of TWINSPAN were labeled to simplify tracking. They can be broadly described as follows: ALP, xeric alpine; ALS, Sitka alder avalanche chutes; BUN, bunchgrass and sagebrush steppe; CUL, cultivated; CUT, clearcuts; EAA, eastern North Cascades mixed conifer of low elevations and dry habitats; EAB, east and central North Cascades mixed conifer of middle elevations, higher and wetter than EAA; HIM, central North Cascades high elevation meadows; POB, western North Cascades high elevation meadows; RIC, eastern North Cascades mountain

alder-water birch swamp, RIX, riparian; RUS, western North Cascades salmonberry shrubfields; TSA, west central North Cascades upper elevation forested; TSH, west central North Cascades low elevation forested; WET, wetlands. These labels retained these meanings throughout the study. Alphacodes, or labels for the most significant plant names are given in appendix 7.

Hybrid ordination-classification of the composite samples to divide the data into classification sets. Data was prepared using the TWINSPAN companion program COMPOSE to remove composite samples with less than 2 species and then the samples were processed with TWINSPAN and DCA. Outliers and disjuncts were identified and removed before running DECORANA, since that program does not handle them well. The results were displayed as scatter plots on the first 2 or 3 DCA axes with the major TWINSPAN divisions displayed in different colors.

Initial development of a key to the differential species. Subdivision of the composite sample set along TWINSPAN divisions. The two-way matrix of samples and species from the TWINSPAN output was examined to determine the names of differential species (species with abundances capable of differentiating between the sets). Those species and the cover values that differentiated the TWINSPAN divisions of samples were noted. The species and cover values were entered as parameters into a Basic program, FKEY.BAS, which operated on the composite sample set. The program separated out a subset of samples from the entire set based on the cover value selection criteria. Samples were keyed for the presence of a given cover level of differential species, stated as a Basic inequality statement of the form, "IF PIPO \geq 20 + VAME \leq 0 THEN ...". Sets of these

inequalities representing each division were logically *ord* and run serially on the composite samples. Samples meeting the key criteria were segregated into subsets and the entire data set was thus reduced to sets of TWINSPAN divisions, with outliers removed.

This routine was reiterated on the main body of composite samples from the beginning a total of 7 times, each time operating on a smaller data set with fewer species attributes (table 1).

Table 1. Input parameters for 7 successive hybrid ordination-classifications of the composite samples and the resultant DECORANA eigenvalues.

Iteration No.	No. samples	No. attributes	TWINSPAN cut levels	DECORANA eigenvalues
1	261	413	0.1, 2, 5, 20, 50	.876, .595, .481, .373
2	238	353	0.2, 3.5, 16, 35	.820, .736, .555, .486
3	223	325	0, 2.5, 9.5, 34.5	.820, .635, .557, .476
4	206	300	0, 2.5, 9.5, 34.5	.811, .646, .519, .436
5	167	283	0, 2.5, 9.5, 30	.815, .631, .535, .413
6	139	265	0, 0.1, 3.5, 16, 35	.822, .603, .506, .446
7	98	218	0, 2, 10, 20, 50	.810, .591, .480, .344

It was not always possible to completely key all members of the subsets in one iteration, but subsequent iterations on reduced sample sets revealed which differential species could then be used to key the data. Based on this hybrid ordination-classification procedure, the 7 ordinations were partitioned among 13 labeled subsets of the 261 composite samples, including one, OUT, that was composed of outliers (table 2).

Table 2. Major subdivisions of 261 composite samples after hybrid ordination-classification.

	Iteration No.	Subdivision label	No. of samples	Total samples
OUT	5	SASI2 (Salix sitchensis)	1	4
	6	EASTSIDE (mesic mixed conifer)	3	

Classification of North Cascades Vascular Plant Communities

<http://www.okanogan1.com/ecology/eco-ncas>

© George Wooten 1996+

CUL	1	CULT (cultivated)	8	9
	3	PYMA (apple orchard)	1	
CUT	4	CC (clearcut)	1	6
	6	CC (clearcut)	3	
	6	ALRU (red alder)	2	
RIC + RIX	3	BEOC (western birch)	2	16
	3	ALRU (red alder)	5	
	3	ALIN (mountain alder)	1	
	3	POTR2 (cottonwood)	2	
	3	COST (red-osier dogwood)	3	
	4	SPDOSOCA (Spiraea/Canada goldenrod)	2	
WET	6	RIP (other riparian)	1	
	2	BOG (bogs)	10	15
	6	RIP (other riparian)	4	
BUN	1	BUN (bunchgrass-shrub-steppe)	11	11
ALP	2	ALPDRY (alpine-dry)	5	5
TSA	4	TSME (mountain hemlock)	25	26
	5	CHNO (yellow cedar)	1	
POB	6	POBIXAS (alpine bistort, low in aster)	17	17
HIM	2	ANLA (xeric fellfields)	3	72
	4	LALY2XT (dry alpine larch)	4	
	5	PHCAVA (heather meadows)	7	
	5	PIALXPI (whitebark pine, low in spruce)	4	
	6	FEVI (green fescue)	3	
	6	HIELMEAD (high elevation meadows)	5	
	7	SUBALP (subalpine parkland)	46	
ALS	5	ALSI (Sitka alder)	8	8
TSH	7	TSHE (western hemlock)	25	29
	3	ALRU (red alder)	3	
	6	PICOGASH (lodgepole/salal)	1	
EAA + EAB	4	PIPO (ponderosa pine)	6	43
	4	SASC (Scouler willow)	5	
	5	PSABCAGE (Douglas fir-alpine fir/elk sedge)	3	
	5	PIALXPI (whitebark pine, low in spruce)	2	
	7	ABLA2 (subalpine fir)	27	

Formation of new sets of composite samples from TWINSPAN divisions.

The raw plot data from each of the 12 groups of composite samples was then aggregated into 12 groups of the same name. The groups of relevés were each separately classified using TWINSPAN with cut levels of 0, 2, 10, 20, and 50. Using these classifications and the Basic programs SELEF, MCSORT and MCS2EFZ, 424 new composite samples with 666 attributes were generated. The MCS2EFZ program formed composite samples from the raw plot data using only those nongeneric species with a mean cover of $\geq 2\%$, or a constancy of $\geq 50\%$. By

judiciously reducing the number of low constancy herbaceous species to 447, it was possible to run TWINSpan and DECORANA on this new matrix.

Keying the raw plot data using differential species. Of the 1433 relevés, 1389 were made into composite samples, 39 became outliers and 5 with likely errors were dropped from the study. Of these, 1432 samples were retained through the final classification. The extension of the classification to the raw data was performed with the goal of retaining consistency between the composite sample ordination and the raw data classification. To do this, the hybrid ordination-classification of the composite samples was used to help define the keyed raw data sets, while minimizing the overlap of the partitions in the ordination diagram.

By adjusting the keying selection criteria in the set-forming program, FKEY.BAS, it was possible to create data set membership based on the goodness of fit of the composite sample partitions in ordination space. To improve the goodness of fit, keying criteria were adjusted to change the set membership of the raw data, which was then used to reform the composite sample set, followed by reiteration of the procedure.

As the refinement of the classification progressed, both the ordination and the two-way table of samples and species were inspected for outliers and misclassified composite samples (samples which overlapped into the domain of other TWINSpan divisions). Outliers were deleted from the set, while misclassifications were allowed to remain in the data to guide refinements of the classification. Some misclassified samples were regrouped or merged to match TWINSpan divisions. Sometimes, the fit was improved by merging TWINSpan divisions or subdividing one division and

grouping one of the subdivisions within an adjacent TWINSPAN division.

Initially, each primary division was keyed using FKEY.BAS on the entire raw data set. Because the same relevé could be keyed into more than one division, it was necessary to develop a dichotomous keying order, so that relevés that keyed out were unavailable for subsequent keying operations. Another Basic program, DESMAT, was used to determine the number of mutual matches between the keyed divisions, hoping to reveal the optimum order of the FKEY.BAS leads that would bring about stepwise division of the data (see Gauch 1982, p. 207, for a discussion of percentage of mutual matches). It was found that, on the average for the 6 largest sets, most of the sets contained a high degree of overlap, ranging from 2% to 74%, and averaging 42%. Thus it was necessary to use a dichotomous key that eliminated subsets matching TWINSPAN divisions from the body of the main before proceeding to the next key test (successive logical classification with elimination).

The sequential set of inequalities in FKEY.BAS formed the basis for rule-based dichotomous key leading exclusively to each TWINSPAN division. In practice, to get 100% fidelity between the key and the desired divisions, it was often necessary to qualify leads with *AND*ed values from several different species' cover values. Sometimes the specificity of these leads keyed out only a few relevés, and often these same relevés were also keyed by some of the other inequalities. Thus, the key to each division was comprised of a set of somewhat redundant statements, joined together by logical *ors*, any one of which would lead to some of the relevés in that division but never to any of the other divisions in the data. This type of key is dichotomous and exclusive between divisions and redundant within divisions. Although its specificity necessitated a greater number of statements, it was simple

to understand, its redundancy insured 100% fidelity between the key and the desired sets, and it allowed for the separation of discrete sets from the remainder of the incomplete classifications.

Coincident formation of a new set of composite samples from both the keyed raw data and matching TWINSpan divisions. Using the resultant keyed divisions of the raw data, the composite samples were then reformed from the raw plot data using the Basic programs SELEF, MCSORT and MCS2EFZ.

In order to obtain a key classification that simultaneously matched the TWINSpan divisions, the sets of keyed relevés were constrained to identity with the divisions of the composite samples produced by TWINSpan. Since the changes in formation of composite samples altered the outcome subsequent TWINSpan classifications, it was necessary to adjust discrepancies between the key and the TWINSpan classification before proceeding to the ordination. Test runs of TWINSpan were examined after each keying operation to assure that the keyed sets exactly matched the TWINSpan divisions.

Ordination and reiteration. The fit between the partitions became obvious when the examined for goodness of fit in DCA ordination-space. Once a set was judged to be bounded by a discrete partition, contained within minimum surface boundaries between adjacent sets, that set was separated from the rest of the data. The remaining sets were then reclassified with TWINSpan, and the new divisions were made into new composite samples. The process was then repeated until all the sets fit together without overlap along linear boundaries.

Initially, many of the borderline composite samples contained relevés that belonged in other composite samples. This was largely due to incomplete subdivision of the data by TWINSPAN, which for simplicity and computational expediency was only run to 6 hierarchical levels. As the set boundaries were successively improved, concave sets were fixed by either splitting the set or by moving some of the relevés to other sets that they resembled more closely. Since the samples are comprised of discreet, or probabilistic, species abundances, whereas the placement of sample points in the ordination was multidimensional, it was impossible to achieve 100% exclusive, non-overlapping fit between the partitions in the ordination.

The TWINSPAN divisions were sequentially separated from the data set before ordinations so that new boundaries could be examined free from effects of other divisions. Consistency was eventually achieved between the key and the classification, through splitting or joining partitions, followed by recreation of the composite samples, running TWINSPAN and DECORANA again, and keying out new subsets. The changes in the classification are detailed in table 3, which shows changes in numbers of relevés in labeled sets and number of composite samples in the sets as the key and classification gradually converged.

Table 3. Changes in the division of sets of 1433 relevés during successive refinement of the classification key. The total of 307 composite samples refers only to the final set of 3/11/92.

Set Label	Date					Composite Samples
	2/5/92	2/11/92	2/16/92	2/20/92	3/11/92	
OUT	181	39	39	39	39	0
CUT	14	14	14	14	14	6
CUL	11	11	11	11	11	5
RIC	0	0	0	3	3	1
RIX	62	67	73	70	81	23
WET	58	92	87	87	89	22
BUN	71	114	114	114	114	25
ALS	34	34	34	34	34	12

Classification of North Cascades Vascular Plant Communities

<http://www.okanogan1.com/ecology/eco-ncas>

© George Wooten 1996+

TSA	106	120	72	72	74	22
RUS	5	6	6	6	6	2
POB	91	90	297	297	287	45
ALP	7	7	7	7	7	2
HIM	391	378	219	219	219	39
TSH	210	218	218	218	218	39
EAA	0	0	0	122	121	33
EAB	191	241	241	119	116	31
Total	1432	1431	1432	1432	1433	307

The 4 products that result from this process of discreet space boundary analysis are a vegetative classification by sets of samples as well as composite samples, an ordination that serves as a model of plant community relationships, a set of composite samples which can be reused as standards of classification, and a key suitable for use on raw plot data.

Results and Discussion

Sampling. The North Cascades Landsat scenes and the landscapes sampled to define them contain information about many new and undescribed plant communities, many of which are outside of the continuous forest canopy. Those areas for which we have the least ecology information, are at least, if not more, diverse and heterogeneous than in forested areas. Because of the emphasis which was placed on sampling those habitats equitably, the vegetation plots represent data about many previously undescribed plant communities in the North Cascades.

The precision with which GIS plot locations matched field-determined locations was determined to be 71%, or 1234 of 1741 plots that had matching spectral class determinations. Of those, 139 of the 187, or 74% of the spectral classes were represented. The logistics of collecting representative samples from 187 spectral classes spread over 2,600,000 hectares was a major task for 5 years, taking 10 person-seasons to complete. The 26% undersampling rate is reasonable, and was compensated for in the North Cascades Grizzly Bear Evaluation (Almack et al. 1993) by the inclusion of 3306 relevés from other sources (not treated here, see Williams and Lillybridge 1983; Williams and Smith 1991; Henderson and Peter 1982, 1983, 1984, 1985; Agee and Kertis, 1986).

The delineation of variable plot boundaries matching the spectral class polygons has significance with respect to sample composition (Mueller-Dombois and Ellenberg 1974, p. 47). To match vegetative communities in the field to an overlay of spectral class polygons meant that the community boundaries were located with a resolution of about half of a pixel or 27.5 m².

The converse situation, that of being unable to resolve a community boundary when spectral class polygons indicated there was one, resulted in recording a range of spectral classes for each plot, but had no significant effect on the composition of the plot.

For this study, the most significant consequence of matching plot boundaries to spectral class polygons was reflected in the heterogeneity of species included within each given plot. The edges of plot boundaries tended to include species overlap from adjacent communities. The size of the pixels and the minimum 3 X 3 pixel size meant that some plots contained smaller, microcommunities within their boundaries. In forested upland situations, ecology plots of this size and shape were straightforward, but outside of the influence of the forest canopy, or in crossing into a wet areas, distinctive communities unlike the surrounding matrix were found on microsites as small as a square meter.

Not only did this increase the variability, or *noise*, in species abundances between plot boundaries, but it also de-emphasized the presence of subcommunities within a plot. For example, north-facing bedrock cliffs were often grouped in large spectral class polygons representing dark shadows, but there are obviously many communities possible within dark shadows. The homogeneity of vegetation within spectral class polygons is correlated with the percentage of surface area measured by four spectral bands in the ultraviolet to infra-red regions. Although spectral diversity corresponds directly with certain leaf structures, this is complicated by complex interactions with topography, soils, and vegetation structure.

Ecology plots based on forestry stand examinations are generally designed to

minimize such heterogeneity, by locating plots well away from adjacent stand boundaries and selecting areas of homogenous vegetation. This minimizes the noise of the data and increases the statistical value of measured attributes. The drawback of delineating plot boundaries within homogenous communities is that problems arise in classifying this sort of data because few samples exist for the boundary regions, where definitions are most critical with respect to plant abundances.

In contrast, the prioritization of stand selection by spectral class representation reduces human bias in locating plots and increases the sampling continuity between communities. The increased continuity is partly due to noise in the data but also yields information about the boundaries between communities, which are characterized by edge-inhabiting species with affinities to the adjacent habitats. Sampling to the edge of adjacent communities reveals more about vegetation *relationships* than homogenous sampling from the interior of a stand is capable of. This method takes advantage of species overlap between communities to help define divisions of environmental gradients using discreet, or probabilistic, plant occurrences; this forms the basis for a divisive vegetative classification.

The degree of continuity is reflected in the even spread of points in the initial DCA ordinations (figs. 3 and 4). The few outliers were easily recognized and eliminated, thus disjunction of the ordination axes was not a problem. The use of DCASP was not feasible, since the data was spread out too evenly. This contrasts with known patterns of sharp discontinuities common in the North Cascades. The heterogeneity is at least partly due to the sampling scheme, and it indicates that sampling density was indeed broadly and evenly distributed across most of the natural communities in the landscape. This sampling scheme resulted in sampling for maximum stand-

or α -diversity and large scale landscape- or γ -diversity, while reducing the contrasts between communities, β -diversity (Whittaker 1967). The classification presented here takes advantage of the fact that the sum of the plant abundances can be used as multiple variables that define the location of the underlying environmental gradients; the boundary of the communities where the gradients are steepest, is definable in terms of a few of the differential species' abundances. TWINSpan two-way tables were effective in locating and displaying these set divisions as attribute discontinuities, and the program produced divisions of the relevés where expected.

Taxonomic considerations. *Artemisia tridentata* Nutt. ssp. *wyomingensis* Beetle and A. Young is a major component of the sagebrush-bunchgrass steppe at the eastern boundary of the study area, but it was not differentiated in the field from *Artemisia tridentata* Nutt. ssp. *vaseyana* Rydb., which occurs in subalpine openings at elevations within the lodgepole zone. Hitchcock et al. (1969) do not commit to whether these are valid subspecies or simply ecotypes, i.e., identical species responding to variation in habitat. Because they have discontinuous ranges, their segregation appears to be valid.

Another dominant expressed in two ecotypic forms is *Pachistima myrsinites* (Pursh) Raf. It occurs in the PSME/PAMY plant association of Williams and Lillybridge (1983), as a significant understory component of mesic benches and bottomlands. Its presence diminishes with increasing elevation until at elevations above 2100 m, it again becomes an important community component in sandy, eroding scree slopes, with open canopies of krummholz *Pinus albicaulis* and *Abies lasiocarpa* and *Eriogonum pyrolaefolium*.

Amelanchier alnifolia var. *cusickii* overlaps the range of another possibly distinct species or variety of *Amelanchier*, occurring from the Wenatchee area at least north to the Okanogan. That entity, originally proposed as *Amelanchier columbiana* (Jim Barrett, BLM, personal communication), has colonial stems arising from root sprouts, rather than being caespitose; it is low and divaricately branched with a blocky shape, rather than erect and fountain-shaped; it prefers northern exposures with deep soils, rather than dry, rocky talus; it flowers approximately a month later than var. *cusickii*; it has short, rigid petioles less than 1 cm long, rather than petioles longer than a centimeter that are flexuous; its leaves are densely pubescent below, although not so dense to be treated as *Amelanchier utahensis*.

Another difficult genus in the North Cascades is *Lupinus*. In the Methow Valley, there are populations of *L. sericeus* that appear intermediate with *L. leucophyllus*, even though the latter is uncommon or absent from that area. *L. latifolius* was listed as a major component of subalpine habitats in the North Cascades, however it appears to form an intergradient complex with a number of similar species. The differentiation of *L. latifolius* from *L. polyphyllus* was sometimes problematic. *Lupinus* individuals that appear intermediate between *L. latifolius* and *L. sericeus* appear at mid to lower elevations, and some of these were later reassigned to *L. wyethii*. The differentiation of *L. latifolius* from *L. wyethii* was based on whether the longest petioles are basal or cauline, an undependable feature. A densely strigose species of *Lupinus* growing in the Prince Creek drainage of Lake Chelan resembled *L. saxosus*, but this was considered unlikely. It was gradually replaced at lower elevations by increasingly larger-leaved and more glabrate plants, eventually passing through *L. latifolius* to *L. polyphyllus*. Because of its high presence in *Festuca viridula* associations characteristic of that area (Alverson and Arnett, 1986,

p. 19) and the similarity of those communities to other areas (Henderson, 1974, Douglas and Bliss, 1977) further work on *Lupinus* species should be undertaken before descriptions of communities based on its presence in this study area are accepted. For this study, most of the borderline cases were simply recorded as *L. latifolius* (LULA).

West of the Cascade crest, in mossy subalpine headwaters in the *Tsuga mertensiana* zone, several plant communities were identified including CABIR-CANI2-AGTH (*Caltha biflora* var. *rotundifolia*-*Carex nigricans*-*Agrostis thurberiana*) and LEPY2-AGTH (*Leptarrhena pyrolifolia*/*Agrostis thurberiana*). Both of these communities support populations of a diminutive, but occasionally dominant species recorded as *Epilobium alpinum* L. sensu Hitchcock et. al., (1969), and occasionally as *Epilobium minutum* Lind. ex Lehm. in Hook. (which is actually a dryland species), or merely as *Epilobium*. It is about 10 cm tall, not clumped, and has lilac-pink flowers, simple, pink stems, and lacks turions. Up-to-date treatments of that genus were not available to resolve its identity at the time of field work, but it may be *Epilobium oregonense* Hausskn. (*E. alpinum* L. var. *gracillimum* [Trel.] C. L. Hitch.), or a member of the *E. ciliatum* group.

The original field studies misidentified differently sized plants of *Artemisia rigida* as *Artemisia cana* or *Artemisia arbuscula*. In these appendices, all instances of either *Artemisia arbuscula* (ARAB) or *Artemisia cana* have been changed to *Artemisia rigida* (ARRI).

Existing vs. potential vegetation. In contrast with previous landscape-level classifications which grouped series based on potential climax dominants

(Daubenmire and Daubenmire 1968, p. 52), this study and the study from which the data was developed (Almack et al. 1993) classified landscapes on the basis of existing, not potential, vegetation. Unlike the area studied by Daubenmire, many situations in the North Cascades were found with strict limitations on the ability of the same trees to attain canopy dominance, even when successfully reproducing. Such communities as those maintained by snowpack, wind attenuation, fire and avalanche, restrict tree and tall shrub species from attaining canopy dominance. For example in perennial avalanche chutes dominated by the shrub *Alnus sinuata*, the potential overstory climax trees, such as *Abies amabilis*, *Tsuga mertensiana* or *Picea engelmannii*, are unable to attain dominance, even when successfully reproducing. In general, departures from potential climax dominance occurred where environmental gradients were limiting, e. g. topo-edaphic climaxes such as those at high elevation (Arno 1972), on serpentine (Kruckeberg 1969), saturated soils (Tiner 1991), talus, scree, bedrock, etc.

Multivariate methods. These analysis methods gave equal emphasis to all species so that the effects of canopy dominance were not given extra weight. Multivariate analysis techniques can be used to conceptualize the differences between communities, by projecting samples and species onto the most significant axes of regions of multidimensional ecological space, or hyperspace (Williams and Dale 1965; Austin 1976; Whittaker 1967; Whittaker and Gauch 1978), of which complex environmental gradients are axes. The term ordination space, as used here, is the graphical depiction of ecological space, and ordination the derivation and display of meaningful, measurable gradients in that space. Ordination axes are inherently multidimensional, and partition of the set, whether by equal space, equal numbers or along sparse regions, as in DCASP, does not necessarily correspond to discreet,

measurable phenomena. Furthermore, the problem of distinguishing real gradient axes from noise is a problem in evenly spread data.

Gradient analysis methods are one method of determining the relationships between communities. Whittaker (1987) used non-metric multidimensional scaling (NMDS) to derive environmental gradients from ecology data, but unable to undo the multidimensionality of the axes themselves. The use of multiple discriminant analysis (MDA, or canonical variates analysis, CA; Walker 1979; Wiegleb 1981; Gerdol et al. 1985), has been used to derive the significance of determinant species, but like regression techniques, it is based on the assumption that species response curves to underlying gradients are linear.

While it would be desirable to regress environmental gradients such as temperature and moisture directly from measurements of those variables, few such measurements are accurately determined for the study area. Although Zobel et al. (1976) were able to derive temperature and moisture parameters in a diversity study of 16 vegetation communities, their methods would have proved impractical in this study. Available measurements of environmental variables such as slope, aspect, elevation and precipitation are themselves interdependent with other local gradients. It is thus logical that community relationships should be described in terms of the available variables—plant abundances. Such a description comprises this vegetation classification.

Plant community relationships were studied here by developing a classification in concert with DCA ordinations; samples were classified by differential species values along the boundaries of the regions of ordination space where those species were

responding most strongly to underlying environmental gradients. Because the TWINSpan algorithm is recalculated within each subdivision, it was possible to develop a classification consistent with ordination space, as each division is aligned along the most significant gradient. By dividing the ordination space into domains defined by TWINSpan, and then using those domains to define discreet, probabilistic sets of input data for further refinement with TWINSpan, the classification subsets could be defined by sets of inequalities of plant cover values. This is a logical expression of a classification in the form of a key.

Initial ordination of composite samples. The use of composite samples that were the mean cover values of groups of relevés reduces the redundancy of the samples, and decreases computational memory requirements. For large data sets such as this, composite samples simplify visualization of the ordination diagrams; this was an important consideration for manual, as opposed to automatic, separation of the subsets.

The TWINSpan subsets within the data are coded with symbols here, but during the workup of the key, the additional necessity of labeling the samples produced a meaningless jumble of points and symbols. Visibility and membership of sets was improved by the use of color coded TWINSpan partitions, using a commercial graphics package capable of producing colored, labeled scatterplots of 2- or 3-dimensional sample arrays.

Ordination of the data using the two major DCA axes is shown in fig. 3 for 249 composite samples and 379 species, after separation of 12 disjunct samples from cultivated and fallow fields. The eigenvalues for the first 4 axes were 0.917, 0.751,

0.573, and 0.487. The first axis corresponds to a complex precipitation gradient, which is itself partly dependent on elevation, climate, and geography. The eigenvalue of 0.917 indicates that 91.7% of the variation of the attribute scores in the samples is expressed by the primary axis. An eigenvalue of 1.00 corresponds to 100% correlation between the underlying axis gradient and the spread of the samples along the gradient. The eigenvalue for the first axis was only 0.876 when the 12 disjunct samples were included in the ordination.

The first 4 ordination axes were 11.7, 6.9, 6.0, and 6.2 SD long. The units of SD represent the average standard deviation of species turnover, and approximately 4 SD represents a complete species turnover, i.e., none of the same species occur in the plots. Thus, the primary axis of this data represents approximately 3 complete community turnovers. Because this axis correlates with precipitation and precipitation correlates with distance from the maritime influence west of the Cascades, this axis approximates a geographic east-west axis. Because of the high degree of landscape- or γ -diversity present in the North Cascades due to the rugged topography, it is important to consider the sample relationships for the second and third axes when analyzing relationships over this long of an axis.

The second axis appears to be a complex gradient of humid low elevation to xeric high elevation. The third axis was also a complex gradient. Both second and third axes were checked against the first during these analyses to assure that the best fit of the classification sets was being achieved in the first 3 dimensions, however diagrams of the third and higher axes are not shown here. Previous studies in this region have found moisture and temperature-elevation to be important gradients (Fonda and Bliss 1969; del Moral and Watson 1978; Agee and Kertis 1987),

however, the strong east-west geographic component of the moisture gradient was only revealed because of sampling over a large geographic area.

The species ordination used all 261 composite samples, including the cultivated field disjuncts (fig. 4 and table 2). Species were reduced from 783 pseudospecies to the 413 most constant; eigenvalues for the first 4 axes of the sample ordinations from this run were 0.876, 0.595, 0.481, and 0.375 and axis lengths were 12.0, 7.1, 5.8, and 4.5. It can be seen that the plants at the top of axis 2 are primarily deciduous and the ones at the bottom coniferous. The weedy and cultivated species are obvious outliers, and do not occupy continuous regions with the rest of the data. They are responding independently of the first two axes to human disturbances and were therefore removed during the sample ordination to reveal more about natural environmental gradients. The continuity of the composite sample ordinations, as well as the length of the axes, indicates that the geographic sampling scheme for the North Cascades was evenly and broadly distributed amongst land types (see table 2).

Initial development of a key to the composite samples. The initial development of a key to the differential species representing TWINSPAN divisions produced a reasonable classification of the composite samples (table 2). The DCA ordinations operated on successively smaller amounts of data (table 1). As classification sets were segregated out, the intersample distances remained similar, resulting in the last sets (HIM, EAA and EAB) being arranged in a dense, even swarm of points near the original center of the ordinations. Removal of 4 outlier samples (SASI2 and EASTSIDE) simplified expanded the swarm and allowed the division of the sample sets. The complexity of the key was greatest during iterations

5 through 7. This is shown in table 2. Early iterations separated out most of the division in one pass (BUN and ALP). Most of the set RIC + RIX were conveniently separated during iteration 3, except that SPDO/SOCA required a further step and RIP could not be separated until the sixth iteration. The most difficult sets to write keys for were HIM, EAA and EAB.

Keying the raw plot data to discreet sets. Given enough computer memory and human time, the development of a regional classification key using differential species could proceed this way by operating on larger and larger data arrays. Alternatively, the classification could be hierarchically divided first by climax dominants, then within those groups, by the remainder of the vegetation. A third method, used here, was to constrain the key to the sets of composite samples produced by TWINSPAN divisions to identity with the key to the raw data sets.

This is an important concept from a statistical point of view, for the composite samples are samples of samples, the mean abundances of abstracted data. The percentage of species abundance agreement between sets of composite samples and the relevés is a measure of the degree of sample abstraction. Conversely, if the keys to samples and to relevés are constrained to identity, then the boundaries produced by TWINSPAN between the composite sample sets in ordination space reflect the degree to which the plant abundances correspond to underlying environmental gradients, *in the local domain of the boundary*. This view of samples in limited regions of ordination space where they are most differential is referred to here as *discreet space*. The space is also discreet in the sense that the sets within them are probabilistic; in fact, the set descriptors of the raw data unmask the integration by DCA of the gradients so well captured by TWINSPAN divisions.

The matching of TWINSPAN divisions of composite samples to key divisions of raw data is accomplished by loading the plant abundance values into the program FKEY.BAS and using it to divide the raw data into subsets. Then the other programs reform the composite samples from the raw data subsets and another TWINSPAN run is used to test if the divisions match the sample sets. Much of the value of TWINSPAN is through the clarity expressed in its two-way table. The composite samples were given aliases according to a rule for assigning the digits that made it easy to determine which TWINSPAN division they belonged to. Sometimes it took several test runs of TWINSPAN and FKEY.BAS before absolute set identity was achieved. Once the two classification sets were identical, DECORANA was run on the composite samples, and the TWINSPAN divisions were viewed in different colors on 2-dimensional scatterplots on DCA axes 1-2 and 1-3. Samples that overlapped the regions of adjacent partitions were noted, and the key was rewritten so that the relevés comprising those samples would be moved into different subsets, then the keying procedure was reiterated. Once a set was removed, however, the key leading to that set remained constant.

Optimization of the fit of the partitions. During the reiteration of the keying procedure, changes occurred to the composition of the subsets in several ways. Outliers and disjuncts were removed. Some of the subsets were merged with adjacent subsets or split into two subsets to improve the fit of the boundaries in ordination space. Minor changes in set membership occurred along the boundaries of the partitions primarily due to incomplete subdivision at the sixth, or default, level of TWINSPAN division that was used. When this occurred, the location of partitions relative to each other was only slightly affected. The main effect this had

was to decrease the convolutedness between the partitions. One reason for this is that moving relevés from one set of composite samples to another has only a small effect on the entire set. The original sample point moves only a small distance toward the rest of the points in its set, while the relocated relevé is placed into another composite sample which moves a small distance away from its set, toward the boundary with the first set. The position of the points that forms the flattest boundary represents the optimum classification for the relevé. Another reason why the location of TWINSpan divisions is relatively stable is that subdivisions are recalculated within each division. Thus, TWINSpan divisions are somewhat insulated from changes in other divisions. An automatic form of this procedure could be developed to quantify the set overlap for borderline relevés in each of two set configurations, and choose the classification with the least set overlap.

Theoretically, this could be extended beyond two dimensions. Both the TWINSpan classification and the goodness of fit of the partitions require criteria about stopping rules for when the routines are complete (Ludwig and Reynolds 1988, p. 183). In the past, these rules have been entirely subjective, however this study indicates that the problem could be overcome using an automatic quantification of set overlap in the lowest 2-4 dimensions of ordination space.

Edge matching versus splitting/merging. The two main types of adjustment to the sets during the classification process can be described as edge matching and splitting/merging. Edge matching represents the minor change in set membership of samples along the boundary of two sets as described above, while splitting/merging represents the major change in set membership associated with joining or splitting of sets along TWINSpan divisions. Although the numerical changes associated with splitting/merging of sets appear significant, in fact they are

not. Merging or splitting of adjacent TWINSPAN divisions does not change the composite sample data content, although it may change the hierarchy, or cladistic structure, of the classification. Contrary to what one might expect, it is the relatively minor changes associated with edge matching that are technically difficult to accomplish. Several reasons account for the difficulty in dealing with minor changes along set boundaries. Often, several edge changes were performed during each classification, some resulting in three- or four-way trades of relevés between composite samples, which resulted in a significant tracking problem. Furthermore each small change affected adjacent points, and these effects needed to be compared with the original classification, and then tested for whichever way produced the best partition fit.

Riparian and wetland communities: overriding the TWINSPAN

classification. As the classification progressed, it became necessary to break some of the above classification rules for the sets RIC, RIX and WET and impose a definition into FKEY.BAS that did not match TWINSPAN divisions (table 3). This was done because those plots were the most likely to encompass several completely different plant subcommunities within the plot boundaries. For example, boulder fields could be mosaics of drought tolerant species atop the rocks with hydrophytes below. Plots containing more than one community are here termed *mosaics*.

The classification was overridden by imposing a definition into the key that would identify plots from the sets RIC, RIX and WET by the presence of even small abundances of important or obligate wetland or riparian species, regardless of whatever else was growing in that plot. The working definition of wetland and riparian species was based on those species that caused plots with obvious riparian

or wetland affinities to be classified based on their drier components, essentially a misclassification for the wetland components. Thus, the definition of wetland and riparian species is only complete for occurrences in plots which were classified unsatisfactorily. Some of the plots in the OUT category are also mosaics, which is where misclassified relevés from the RIP and WET sets would have been placed without the imposition of a wetland-riparian definition. These special sets were formed because of the need to have information about riparian areas and wetlands, although the assumptions regarding the formation of those sets are different from the rest of the sets in this study.

When composite samples were discovered which were misclassified because they contained riparian or wetland species, all relevés having similar or smaller cover values of those species were put into the RIC, RIX or WET categories, even if the preponderance of vegetative cover was upland. This prevented upland associations from becoming mixed with data from wetland or riparian associations, but converse is not true: wetlands and riparian areas were allowed to include parts of upland communities (fig. 6).

Some groups of relevés with borderline wetland or riparian species were still difficult to classify because it was ambiguous whether to use the imposed riparian/wetland rule, or merge the samples into an adjacent set or make them into a separate subset. For instance in a number of areas, SPDO/SOCA (Spiraea/Canada goldenrod) communities occupy large areas, where they are adjacent to riparian areas, upland shrubfields, high elevation meadows and parks. They would have formed a small group on their own, but this was avoided since they seemed to be ecotonal. They had a great deal of similarity to ALS and RUS, except that they

lacked common dominant species. They seemed to belong in the RIX riparian set, since in the field, they were always zonal to riparian areas, however they consistently contained upland species. The situation was complicated by incomplete field identifications for some occurrences of *Calamagrostis* and *Spiraea*; both genera contain upland and wetland species, and *Spiraea* contains hybrids. In the final classification, they were divided among OUT and RIX.

This illustrates the care that needs to be taken in interpreting the wetland and riparian ordinations here, since under this sampling scheme they overlap the ordination space of other communities. Further work in classifying the wetland and riparian samples would first require separating out upland components, although the overlap with upland communities does contain information about the interface between wet or riparian areas and uplands. The desired effect of simplifying the ordination displays for the rest of the classification subsets was successful.

Separation of wetland and riparian sets into separate ordinations, had in fact already been used as a visualization tool during the initial ordination (figs. 3 and 4) because it was noticed then that the wetland and riparian groups were not as consistent as the rest. Because of the clarification that separating these sets out afforded for the ordinations and classification, the key was rewritten with riparian and wetland associations placed near the beginning, which also simplified matching the key to the TWINSpan classification for the rest of the data.

This has important ramifications for land management surveys intended to protect wetlands and provide buffers along riparian corridors. The broad spread of the points representing wetland and riparian areas implies that the use of sampling schemes based on forest stand boundaries or spectral class polygons may be

inappropriate for some wetland and riparian communities, e. g., springs, seeps, sinuous riparian corridors, draws, and zonal associations, which do not occur in the shapes and sizes of upland forested plant communities. Since the factors that govern the composition of upland plant communities are very different than those for riparian areas and wetlands, it follows that riparian and wetland areas should be sampled and classified along different lines, even when they appear to be typical forested communities, (Environmental Laboratory 1987; Kovalchik and Chitwood 1990; Cowardin et al. 1992; Tiner 1993).

Optimizing membership of the subsets of the classification. The progress of this successive refinement of the classification is detailed in table 3, which begins on 2/5/92 with the keying of the raw data using the key developed initially for the composite samples. When used on the raw data, that key failed to identify 182 of the 1433 relevés or approximately 13%. Although 39 of these unclassified relevés eventually turned out to be outliers, this still means that the actual misclassification rate is much higher, since the subsets must be also be misclassified amongst each other. The classifications of the composite samples and the raw data were not expected to agree, since a number of assumptions went into the formation of the composite samples, not the least of which was that species abundances can be summarized as the mean cover of the summed sample abundances.

Not all of the sets went through the classification procedure. In addition to the sets RIC, RIX and WET which were excluded as described above, the set OUT (outliers) was obviously excluded. It is interesting to note that no new outliers were discovered after the second key was written on 2/11/92. This points out the relative

efficiency of TWINSPAN and DECORANA at identifying these entities. The set CUT (clearcuts) did not show very much internal consistency, and formed two widely separate clusters in the ordination. Those samples were treated as a special class of outliers, being excluded from the definitions, while being given a working definition as a set ("clearcut logged"). The set CUL (cultivated) contains a number of highly disjunct relevés (pastures, plowed grain fields and fallow fields), however the key to the set is one of the simplest. Because of the disjunction of the species, they were assigned major divisions by TWINSPAN in initial ordinations.

The sets of ALP (alpine xeric) and ALS (Sitka alder) remained constant till the classification was finished, reflecting for the former the high degree of dissimilarity with the rest of the data and for the latter the great deal of consistency within the set. BUN (sagebrush/bunchgrass), RUS (salmonberry/ladyfern shrubfields) and TSH (western hemlock, red alder and lodgepole/salal) also remained relatively constant once initial keying criteria were adjusted after 2/11/92. BUN was the only set in the original ordination separated far enough from the rest of the sample points to use DCASP on. Both ALS and RUS occupy the central region of the ordination in fig. 5, demonstrating their vegetative similarity to a number of other central Cascades communities. The density of points within their respective regions indicates the relatively high internal consistency within those communities.

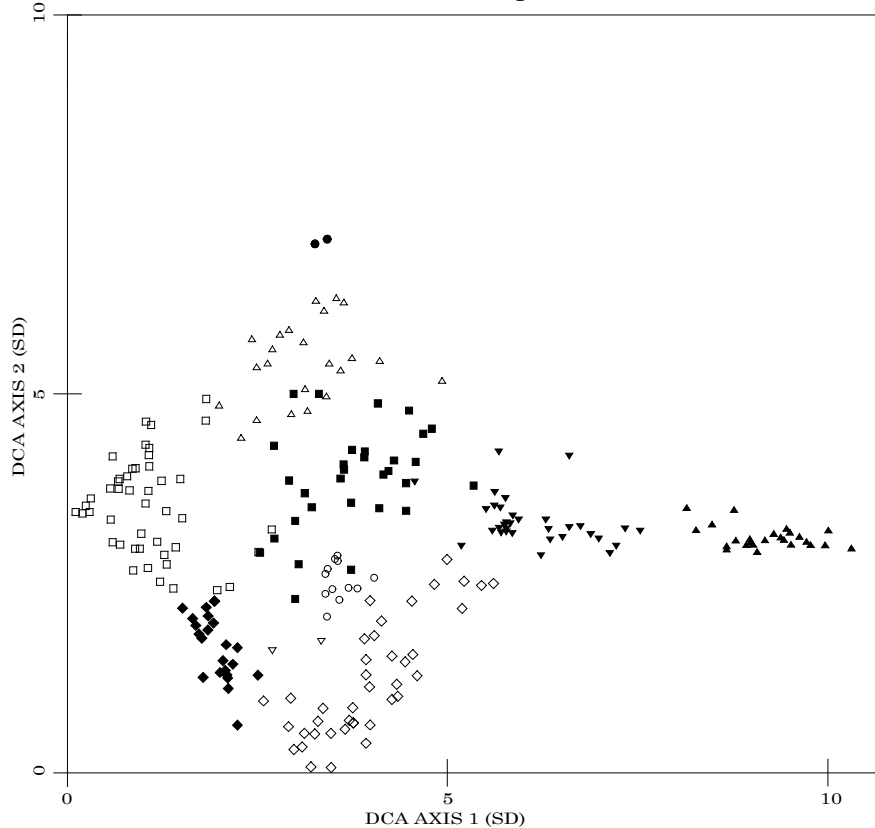


Fig. 5. Ordination of 250 upland composite samples depicting the two major DECORANA axes after optimizing the sets. Fallow and cultivated samples are not shown here. The original form of these plots used color-coded sample sets. Symbols represent the following sets: ●, ALP; ○, ALS; ▲, BUN; ▼, EAA; ■, EAB; △, HIM; □, POB; ▽, RUS; ◆, TSA; ◇, TSH (see text for explanation of labels). Eigenvalues for the first 4 axes were 0.908, 0.660, 0.662, and 0.477.

The classification sets for TSA (mountain hemlock and Alaska yellow cedar) and HIM (high elevation meadows and parklands) stabilized after secondary TWINSpan divisions in those data sets were split or merged among the other subsets. There were a few edge match changes in TSA and HIM during the 2/20/92 classification, the additions to HIM being masked by the deletions, and although these were minor, they required considerable effort to rectify.

The set POB (alpine bistort, heather meadows, wet alpine meadows, western hemlock parklands) went through a major compositional change between 2/11/92

and 2/16/92, due to merging with parts of the HIM and TSA sets. This is not obvious from the counts in HIM (note that no apparent changes occurred in HIM between 2/20/92 and 3/11/92, when in reality the set actually gained 7 relevés and lost 7 relevés from four other sets). Both HIM and TSA sets of points were spread out in the ordinations of 2/11/92, but by moving some of their secondary TWINSPAN divisions into POB, they both became tighter clusters of points in ordination space (fig. 5). These changes were conceptually simple, but took a large number of key leads to accomplish because many of the HIM relevés contained the same species as in POB, and the POB leads preceded them. In contrast, adjusting for minor edge changes in POB required writing only a few key lead inequalities, but they were necessarily complicated statements.

The sets EAA and EAB, which were originally one set, were split into two on 2/16/92. These two sets form diverse assemblages, from the boundary with BUN, the sagebrush/bunchgrass set, to regions of overlap with most of the other sets. The obvious thinning at the boundary of the set EAA with the main body of sample points can be attributed to the effectiveness of this procedure (compare fig. 3 with fig. 5). Most of the relevés represent dry PSME, ABGR, or PIPO mixed conifer, but shrubfields and deciduous trees such as POTR are also represented as are some of the lower elevation ABLA2 stands. In the ordinations, there is a significant thinning of points in the set EAA corresponding to lower timberline in the east Cascades, and another discontinuity took the form of an L-shaped bend between the two sets.

The set EAB has the greatest overlap of all the sets, and covers a greater range of β -diversity than any other set except TSH. This corresponds with a broad

differentiation of vegetative communities in the east Cascades forested zone. del Moral and Watson (1978) explained this in terms of a greater contribution of understory species to community variation where reduced precipitation leads to openings in the forested canopy, which agrees with the results here. Further splitting of the set EAB would continue to improve the goodness of fit of the partitions. The next group of borderline points identified for edge matching to create tighter clusters of sets are relevés representing recent burns, but this has not yet been done.

Ordination of the classification subsets. The ordination of the upland samples is depicted in fig. 5, and that of the wetland and riparian samples are displayed in fig. 6.

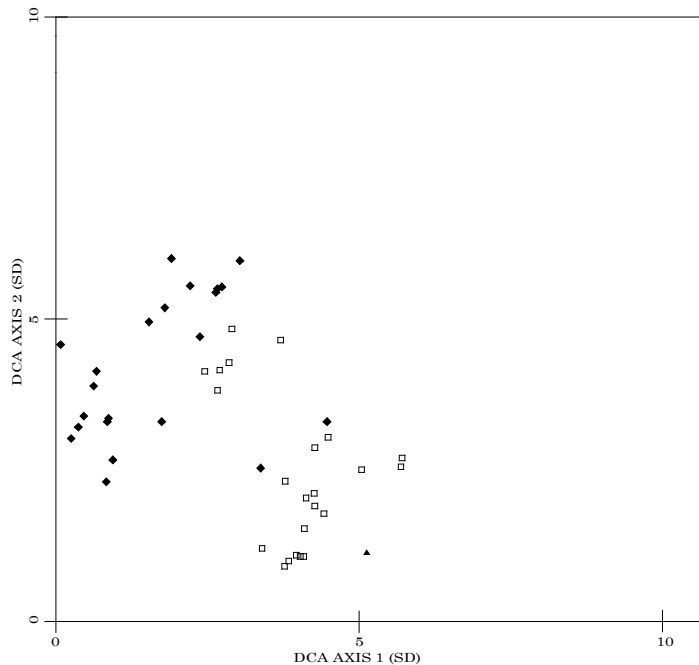


Fig. 6. Ordination of 46 riparian and wetland composite samples on the two major DECORANA axes, before optimizing the sets. Symbols represent the following sets: ▲, RIC; ■, RIX; ◆, WET (see text for explanation of labels. Eigenvalues are 0.908, 0.660, 0.662, and 0.477.

Wetland and riparian sets were separated into two different sets because of the previously explained overlap in sampling riparian and wetland community mosaics from upland plant communities. It is interesting to compare this second ordination in fig. 5 with the initial ordination in fig. 3. The eigenvalues for the first four axes remained high at 0.908, 0.660, 0.622, and 0.477, while the axis lengths were slightly lower at 10.3, 7.0, 5.5, and 6.5. While the gradient expressed by the primary axis is essentially the same, the second axis reversed its positive correlation with humic moisture, and instead corresponds more with a gradient of elevation and climate "harshness", with the xeric alpine cushion communities arrayed at the top of the diagram and lush western hemlock communities put at the bottom.

Ordination interpretation. Considerable structure emerged from the ordination (fig. 5), and the sets can be seen to approximate minimum surface boundary polyhedrons, even in only 2 dimensions. In viewing the first and third axes, the planar polygon boundaries are still apparent, which indicates that the polygons are themselves shadows of higher-order n -dimensional polyhedral regions. This indicates that TWINSpan divisions align themselves perpendicular to underlying gradients. The TWINSpan classification dendrogram is shown in fig. 7.

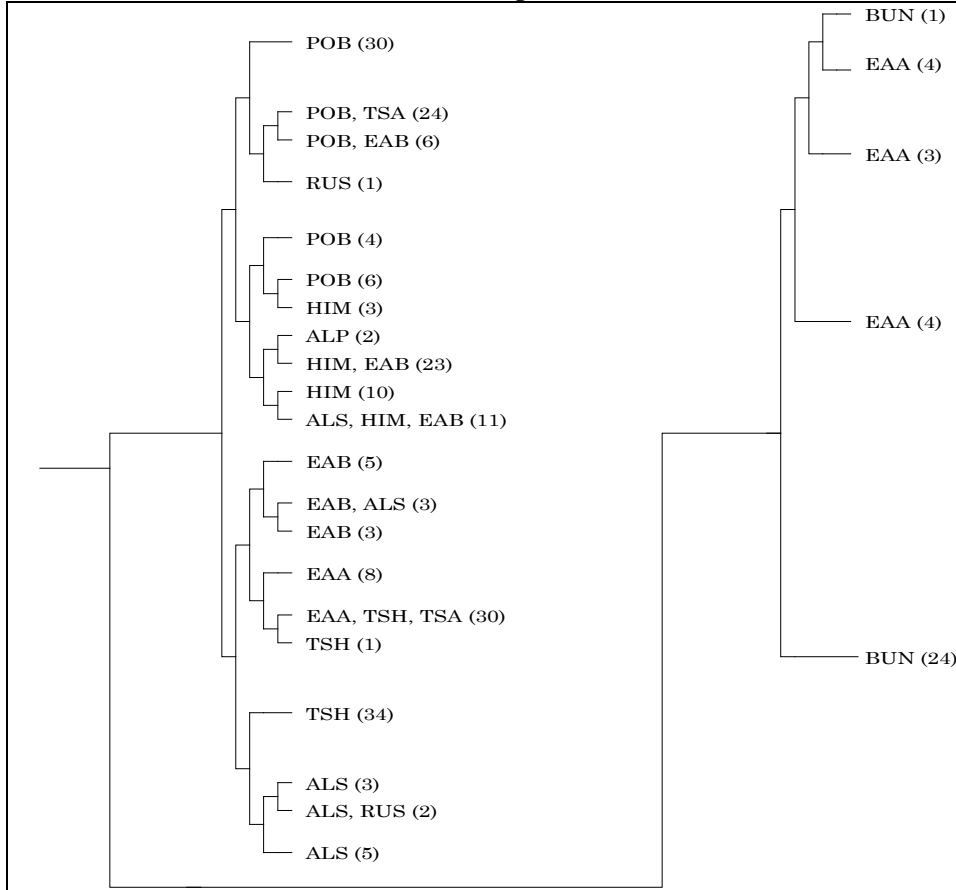


Fig. 7. TWINSpan classification of the final data set of 250 composite samples and 368 attributes, using cut levels of 1, 5, 10, and 25% cover. Results are portrayed by the name of the major sets described in the text. The sets WET, RIX and RIC were excluded from this analysis.

One might question the validity of using ordination for classification of large regions with primary gradients exceeding 4 SD or one complete community turnover.

Although ordination was an effective technique for this study, which spanned more than 10 SD, much of the clarity of the ordination is attributable to the strong sample correlation with a primary axis that is essentially geographic. Ordinations possessing longer axis lengths would be difficult to work with, since tertiary and higher axes may be involuted incomprehensibly with respect to the lower dimension axes. The divisive algorithm of TWINSpan, however, adjusts for new gradient axes at each division, and is not constrained by computational requirements. If a method

of quantifying partition overlap in n dimensions were developed, this method of hybrid ordination classification should be capable of classifying longer ordination axes.

The endpoints of the secondary axis in fig. 5 spread the sample points out in a way that is clearer to visualize than in the initial ordination of fig. 3. In the initial ordination, the secondary axis located points representing east Cascades ALIN-BEOC swamps at the humic, moist endpoint, which failed to show their affinity with the remainder of points representing east Cascades vegetation. The secondary gradient in that initial ordination correlated positively with species going from coniferous to deciduous. It is difficult to visualize deciduousness as an environmental gradient rather than an effect of one. In the final ordination, the points representing those swamps are located between points of moist western hemlock and those of the dry east Cascades, which portrays the community relationships in a clearer fashion.

The overall diamond shape of the data spread in fig. 5 is interesting, with high elevation, mountain hemlock cloud forest communities represented at the west, or left side, and the 1-dimensional sagebrush-steppe points trailing away to the east. The points at the edges of the scattergram are represented by communities at the edge of viable life in talus, flood channels, saturated soils and snowpacks. The flatness of the sides of the scattergram is partly due to elimination of outliers, many of which correspond in nature with less diverse species numbers, but in any case these edges represent the locations where plant growth is limited to all but the most specialized or hardy plants in those environments. The flatness may also be due to linear gradient combinations in the first and second axes, which is supported by the

fact that the third axis also formed a rough diamond pattern against the first axis (not shown).

The dense aggregation of points along the perimeter of the left edges of the data swarm would appear to indicate that a greater diversity of communities exist along those linear edges, i. e., that γ -diversity is higher. The greater number of points along the perimeter of the data can be explained as an artifact of the location of TWINSPAN RA axis endpoints, and the subsequent use of those smaller segments for formation of composite samples from smaller groups of relevés. The *flatness* of the edge, however, is a measure of the degree of abruptness of change from growing to limiting environments, such as at the edge of a cliff, snowpack or flood channel. For instance points along the upper left edge represent mountain heather and dwarf willow communities growing along the edge of snowline and points along the bottom left edge represent communities on north-facing bedrock cliffs and growing in moist to saturated soils.

Even though the number of points along the edge can not be used as a valid estimate of γ -diversity, the continuous point density relative to the greater interpoint distances nearer to the center of the ordination is an indication that β -diversity decreases in going from mesophytic to limiting environments. This is corollary to the findings of del Moral (1972), who measured β -diversity along gradients of moisture and elevation in the North Cascades and found it increased toward milder environments at the minimum richness which allowed canopy dominance to occur. He used the half-change of species turnover as the unit of measurement, which is directly comparable to the units of SD, the average standard deviation of species turnover, in these DCA ordinations. He did not sample to the

limits of plant growth, and located his plots in homogenous stands within the bounds of upper and lower timberline, thus the sharp discontinuity at the limit of plant growth was not observed in his study. In addition to canopy dominance, there are alternative explanations of why plant community β -diversity might vary inversely with limiting factors. The superposition of non-limiting gradients perpendicular to the limit of plant growth results in the creation of many microsites which are constrained in size and extent by irregular variation along the edge of the limiting factors. As these develop into communities, their separation distances can lead to stochastic factors, such as seed dispersal, influencing the development and composition of the community. Consequently, there are numerous sites capable of occupancy by only a limited number of species and individuals from the local flora. Discontinuities such as the presence of a glacier or snowpack also possess a temporal dimension, so that communities can exist in many different seral stages along moving fronts. It is not just canopy dominance that determines greater β -diversity, though. Other limiting factors can become *enabling* to the growth of many different species with just small changes in intensity, for instance in crevices or along the perimeter of a lake.

The monotonicity of the points along the right side of the ordination reflects the relative insensitivity of the secondary axis of harshness to account for the variability of the samples. The far left cluster (BUN) lies below lower timberline, and trends rightward toward even more xeric communities. The furthest right points lie at low elevation, exposed on Columbia flood basalts; short grasslands and shrub grasslands are just left of those. The second set from the right (EAA), occurs at the interface between open, dry, bunchgrass meadows and lower ponderosa pine and Douglas fir timberline. Both of these clusters occur in semi-arid climates, which

is an extension of the interior west sagebrush-bunchgrass steppe. Plant communities adapted to living in the semi-arid west are constrained mainly by the availability of water; other gradients tend to be integrated with that primary precipitation gradient, e. g., elevation or east-west position. Therefore it is reasonable that community variation is so completely explained by the primary axis, with its high correlation with precipitation. This explanation resembles that given by del Moral and Watson (1978) for similar ordination results of plots in the central North Cascades.

The clusters HIM (high elevation meadows), RUS (salmonberry/ladyfern), ALS (Sitka alder) and the combined sets of POB and TSA (upper elevation west Cascades meadows and mountain hemlock) are discreetly partitioned in the first 3 ordination axes. A 3-dimensional quantification of the set overlap would improve the use of discrete space analysis, however 2- and 3-dimensional color scatter plots are effective for 3 dimensions.

The sets EAB, WET and RIX are all incompletely partitioned and lack the ultimate goodness of fit, however these were all large sets with a wide spread of sample points. For the points in these sets lying near the center of the ordination, it is possible that the communities they represent are inhabiting a broad spectrum of mesophytic environments. In light of this, further partitioning of the set EAB into discreet, non-overlapping regions appears necessary to complete the classification. The communities represented in EAB form a large part of the ecosystem east of the Cascade crest. In fact, borderline points representing burned areas have already been identified as outliers which can be put into a separate set for the next stage of the classification.

Community descriptions. Community descriptions for alpine and riparian areas appear to be at least as diverse as for forested types, and the communities described here are only a beginning step toward their understanding. Canopy dominance becomes less important in the order of West Cascades closed forest, East Cascades forest, riparian areas, and then alpine areas. Soils, moisture, weather and topography play a greater role in communities outside of the forest canopy. Forested communities at the limits of their tolerance to environmental stress do not behave as a series of canopy-dominated plant associations. Many tree species have biphasic ecotypes in response to limiting environments, requiring their placement in one of the above regimes before delimiting them in series. Trees which can exist in more than one ecotype in upper timberline environments include LALY, PIAL, ABLA2, PIEN, PICO, PIMO, CHNO, ABAM and TSME. Descriptions of plant associations for these trees should be designated into regimes for uplands, wetlands, riparian areas, parklands, and rocky areas. For instance, the ecological amplitude of ABLA2 allows it to occupy different niches under each of these sets of environments, while still behaving as a dominant species in all of them. Yet within each regime, ABLA2 would have different successional stages, each requiring a separate ABLA2 series for those four types alone. Clearly, it would benefit our understanding of ambiguous association labels if the regime were somehow specified with the association name. An attempt was made to do that for the communities described here; when an additional term such as *talus* or *avalanche chute* clarified the concept of the community, it was appended to the name. In some cases, community names were judiciously chosen to avoid ambiguity, e. g., FEVI-PHDI instead of FEVI-LULA to differentiate the communities here from those of Henderson (1974), even though either name would have adequately described the dominant plants in the

communities here.

Canopied communities. Descriptions are presented here for some riparian areas and wetlands, subalpine meadow parklands (which were treated as mosaics of tree clumps and meadows) and alpine rock-substrate communities (appendix 5). Less attention was given to closed canopy communities since many of these are already documented in the literature. In addition to the types mentioned above, this study recognized the following high elevation closed forest communities: LALY-PIEN/PHEM/LUHI, ABLA2-PIEN/TRLA4-EPGL2, ABLA2-PIMO/PAMY-VASC, PIEN/SASC/ASEN-THOC (SASC treated as a shrub not a tree), ABGR/ROGY, and ABAM/RHAL/depauperate. Subalpine parkland tree communities were generally not differentiated from their meadow matrix using the sampling methods here, however the following open canopy communities and larger tree-shrub complexes were recognized: ABLA2-PIAL/PAMY/VASC, ABLA2-PIEN(-CHNO)/VAME chute, TSME-CHNO/SOSI, ABLA2/VADE, ABLA2/FEVI, ABLA2/VAME, CHNO/VADE/VASI and ABLA2/DAIN. Other open canopy forests that were observed, but not described, here were PIEN-PICO wet meadows at Horseshoe Basin in the Pasayten Wilderness and PICO(-ABLA2)/ARTRV dry meadows on Scaffold Ridge, Chelan-Sawtooth Wilderness. Alpine and subalpine rock communities presented here which support tree species which have the potential to become dominant are: TSME-CHNO boulderfields, (PIAL)/AROB-ERAU fellfields and (PIAL)/SIAC-DROC fellfields. Note that for the latter two, the PIAL is so short (<1 m) and infrequent in this association that its usage merely indicates the position relative to timberline. Other open canopy tree communities observed, but not described here were LALY boulderfields, ABLA2 krummholz, PIAL-ABLA2-PICO serpentine talus slopes, and *Salix nivalis*-lichen communities. The community

descriptions given here include shrubfields, forest burns, wetlands and riparian areas (appendix 5).

Vegetation classification is an integral part of many land classification methods. Land classification maps and assessments of their resource values are only as accurate as the underlying community models. Different ecological methods approach classification differently; statistical routines, in particular, require making a number of assumptions about the utility of sampling methods, the degree to which different species affect the model, and the ability of different numerical methods to accurately predict meaningful, derived, environmental gradients.

A glossary is given in appendix 1, the key that was produced from this process is reproduced in appendix 2, a list of important or indicator species is given in appendix 3, a concordance listing of plant communities in the literature is given in appendix 4, community descriptions are given in appendix 5, species cover data for the named communities are given in appendix 6, and alphacodes used in this report are given in appendix 7.

Conclusions

A regional, dichotomous classification system was developed for 14 major sets of communities of the North Cascades based on sampling spectral class polygons. The effectiveness of multivariate classification techniques is dependent upon sampling plot sizes and shapes that capture the floral composition bounded by the gradients that determine the local environment. Sampling over a broad geographic distribution using variable plot boundaries matching spectral class polygons is an effective way to obtain information about the boundaries between plant

communities, but requires additional interpretation to analyze smaller communities within the plot boundaries. The methods used in this study were effective in relating broad community patterns across moisture and elevation gradients, but work still needs to be done to explain the finer structure of smaller communities.

A discussion of alpine and subalpine tree communities was given to illustrate the use of regime descriptors in association names. In addition to providing some clarity in the naming of associations, codominant tree associations were also described to illustrate the importance of codominance in many subalpine tree associations. Some, like LALY-PIEN/PHEM/LUHI have 4 codominant tree species in older stands (including ABLA2 and PIAL). In parklands and open canopy stands there may be no advantage in canopy dominance, in fact it extracts an exposure penalty (Arno and Hammerly, 1984; Bliss, 1969). Thus, structurally equivalent life forms which don't outcompete each other can coexist for mutual benefit, much as in shrubfields or grasslands, where structural equivalence is also important, but for different reasons. Such a community has also been called a guild.

The method used here to classify the communities of the North Cascades was called *discreet space boundary analysis*, in which the term *boundary analysis* refers to a generalized method which can be further developed. The term *discreet space* was chosen because a defined region, or *discreet space*, was used to obtain each level of the classification, and also because the classification was defined in terms of *discreet* sets of samples. The resulting classification can also be viewed in *discreet space*, which would be those *discreet* regions of ordination space where only differential species axes are considered. It is arguable that this procedure is simply a hybrid ordination-classification, and that is true. However, the successive

refinement of resulting partitions amounts to boundary analysis of ordination space and the particular method used was that of discreet set construction. Sampling techniques gathered sufficient information about community boundaries to make this method possible.

Another concept embodied in discrete space is that samples are abstractions of reality, while the desired solution, a vegetative classification, is probabilistic. Since the sample points are located in ordination space relative to the summed Gaussian plant distributions along environmental gradients, composite samples that are the mean abundances of the relevés are further abstractions from reality. By optimizing the fit between the n -dimensional surfaces that form the boundaries of sets of composite samples, i. e., flattening, the boundary surfaces perpendicular to environmental gradients that determine the regions where the relevés and the composite samples coincide, the classification focuses on the differences rather than the similarities of plant communities. Samples along the boundaries of the classification are brought into focus to develop the classification, enabling the construction of a highly coherent vegetation key as the set becomes defined.

Although this study analyzed ordination space using color visualization of set boundaries, and then exhaustively refined the defined regions, there are more elegant mathematical methods using set theory that do the same thing. For instance, one way would be to mathematically describe the boundary surfaces of the sets in n dimensions and numerically measure the overlap of adjacent sets, although ecologists would still want to be able to inspect 2- and 3-dimensional scatterplots of the data. Other classification methods that are theoretically consistent with ordination space include DCASP and MDA.

The use of successive refinement solves for discreet partition sets in a manner analogous to solutions of thermodynamic equilibria—configurations of successively lower energy are modeled until a solution of minimum energy is obtained.

Partitions reflect statistical changes in plant abundance along underlying gradients within the ecological space of the whole data set. The analogy with thermodynamic equilibria is meant to stimulate thought about the role of entropy in synecology.

This method was supervised in the sense that the stopping rules for the goodness of fit of the partitions was judged subjectively. An automatic method of quantifying set overlap could be developed to specify limiting criteria for the best fit of the partitions.

Current classifications based on potential vegetation are important for regional mapping, however they provide limited or inaccurate information about nonforested, wetland or riparian communities and the nature of community edges. However applicable the assumptions of canopy dominance are in potential vegetative classifications, they may not be applicable when species respond to unknown gradients or communities have limiting factors other than shade tolerance, e. g., open areas or wetlands. This classification is an attempt to define broad groupings of plant communities using whatever species are differential at the division boundary. Further use of both the raw data and the composite samples and improvements in the use of discrete space analysis are hoped for in future ecology studies in this region.

Acknowledgements

Classification of North Cascades Vascular Plant Communities

<http://www.okanogan1.com/ecology/eco-ncas>

© George Wooten 1996+

This data was collected for the North Cascades Grizzly Bear Ecosystem Evaluation under a cooperative agreement between the U. S. Fish and Wildlife Service, Washington Department of Wildlife, Washington State Remote Sensing Lab and U. S. Forest Service, Okanogan National Forest. We would like to give special thanks to Jon Almack, Bill Gaines, Bob Naney, Jim Eby and Ernie Garcia, for their invaluable guidance, and Lanette E. Smith and Aileen Jeffries for their patience and encouragement. Mary Pat Larsen's suggestions were invaluable. Particular thanks are extended to the field botanists who worked so hard to gather and collate this information, Faye Streier, Jerry Thompson, Michael Mancuso and Jody Engle.

References

- Agee, J. K., S. G. Pickford, J. Kertis, M. Finnney, R. de Gouvenain, S. Quinsey, M. Nyquist, R. Root, S. Stitt, G. Waggoner, and B. Titlow (1985). *Vegetation and Fuel Mapping of North Cascades National Park Service Complex*. Final Report. National Park Service contract CX-9000-3-E029, College of Forest Resources, University of Washington, Seattle, WA.
- _____, and J. Kertis (1986). *Vegetation Cover Types of the North Cascades*. Report CPSU/UW 86-2, National Park Service Cooperative Park Studies Unit, College of Forest Resources, University of Washington, Seattle, WA.
- _____, and J. Kertis (1987). Forest types of the North Cascades National Park Service Complex. *Canadian Journal of Botany* **65**:1520-1530.
- Almack, J. A., W. L. Gaines, R. H. Naney, P. H. Morrison, J. R. Eby, G. F. Wooten, M. C. Snyder, S. H. Fitkin, and E. R. Garcia (1993). *North Cascades Grizzly Bear Ecosystem Evaluation*. Final Report. Interagency Grizzly Bear Committee, Denver, CO.
- Alverson, E. and J. Arnett (1986). From the steppe to the alpine: a botanical reconnaissance of the Lake Chelan-Sawtooth Ridge area, Washington. In *Plant life of the North Cascades: Lake Chelan-Sawtooth Ridge, Stehekin Valley, and Glacier Peak*, Douglasia Occasional Papers, Vol. 2, pp. 1-63, R. J. Taylor, (ed.). Washington Native Plant Society, University of Washington, Seattle, WA.
- Arno, S. F. and R. P. Hammerly (1984). *Timberline: Mountain and Arctic Forest Frontiers*. The Mountaineers, Seattle, WA.
- _____, and J. R. Habeck (1972). The ecology of alpine larch (*Larix lyallii*) in the Pacific Northwest. *Ecological Monographs* **42**:417-450.
- Austin, M. P. (1976). On nonlinear species response models in ordination. *Vegetatio* **33**:33-41.
- Bliss, L. C. (1969). Alpine community pattern in relation to environmental parameters. In *Natural Vegetation of Oregon and Washington*, J. F. Franklin and C. T. Dyrness (eds.), pp. 248-90. (1973, repr. 1984), Oregon State University Press, Corvallis, OR.
- Borgias, D. D. (1984). *Subalpine Picea engelmannii-Abies lasiocarpa Forests in the Pasayten Wilderness, Northeastern Cascades, Washington*. M. S. Thesis, Western Washington University, Bellingham, WA.
- Brayshaw, T. C. (1965). The dry forest of southern British Columbia. In *Ecology of Western North America*, Vol. 1, V. J. Krajina (ed.), pp 65-75. University of British Columbia Dept. Botany.
- Brown, D. E., C. H. Lowe, and C. P. Pase (1980). *A Digitized Systematic Classification for Ecosystems with an Illustrated Summary of the Natural Vegetation of North America*. USDA Gen. Tech. Rep. RM-73. Fort Collins, CO.
- Cowardin, L. M., V. Carter, F. C. Golet and E. T. LaRoe (1992). *Classification of Wetlands and Deepwater Habitats of the United States*. Report FWS/OBS-79/31. U. S. Department of the Interior Fish and Wildlife Service, Office of Wildlife Services, Washington D. C. 20240.
- Craighead, J. J., J. S. Sumner, and G. B. Scaggs (1982). *A Definitive System for Analysis of Grizzly Bear Habitat and Other Wilderness Resources*. Wildlife-Wildlands Institute Monograph No. 1, University of Montana Foundation, University of Montana, Missoula, MT.

- Daubenmire, R. (1959). A canopy-coverage method of vegetational analysis. *Northwest Science* **33**:43-64.
- _____, and J. B. Daubenmire (1968). *Forest Vegetation of Eastern Washington and Northern Idaho*. Wash. Agr. Expt. Sta. Tech. Bull. XT-0060. Rrepr. 1984, Cooperative Extension, Washington State University, Pullman, WA.
- _____. (1970). *Steppe Vegetation of Washington*. Wash. Agr. Expt. Sta. Tech. Bull. 62, Washington State University, Pullman, WA.
- DeGraff, R. M., R. G. Bailey and R. T. Brooks (1988). *A Comparison of Five National Land Classification Maps*. USDA Forest Service Handbook 672, Washington, D. C.
- del Moral, R. (1972). Diversity patterns in forest vegetation of the Wenatchee Mountains, Washington. *Bulletin of the Torrey Botanical Club* **99**:57-64.
- _____, A. F. Watson, and R. S. Fleming (1976). Vegetation structure in the Alpine Lakes region of Washington state: classification of vegetation on granitic rocks. *Syesis* **9**:291-316.
- _____, and A. F. Watson (1978). Gradient structure of forest vegetation in the central Washington Cascades. *Vegetatio* **38**:29-48.
- Douglas, G. W. (1969). *A Preliminary Biological Survey of the North Cascades National Park and the Ross Lake and Lake Chelan National Recreation Areas*. National Park Service, Seattle, WA.
- _____. (1972). Subalpine plant communities of the western North Cascades, Washington. *Arctic and Alpine Research* **4**:147-166.
- _____, and L. C. Bliss (1977). Alpine and high subalpine plant communities of the North Cascades range, Washington and British Columbia. *Ecological Monographs* **47**:113-150.
- Environmental Laboratory (1987). *Corps of Engineers Wetlands Delineation Manual*. Final Report, Y-87-1. Department of the Army, U. S. Army Corps of Engineers, Waterways Experiment Station, Vicksburg, MS.
- Fonda, R. W. and L. C. Bliss (1969). Forest vegetation of the montane and subalpine zones, Olympic Mountains, Washington. *Ecological Monographs* **39**:279-301.
- Franklin, J. F., and C. T. Dyrness (1973). *Natural Vegetation of Oregon and Washington*. Repr. 1984, Oregon State University Press, Corvallis, OR.
- Garrison, G. A., and J. M. Skovlin (1976). *Northwest Plant Names and Symbols for Ecosystem Inventory and Analysis*, Fourth ed. USDA Forest Service Gen. Tech. Rep. PNW-46, PNW Forest and Range Experiment Station, Portland, OR.
- Gauch, H. G., R. H. Whittaker, and T. R. Wentworth (1977). A comparative study of reciprocal averaging and other ordination techniques. *Journal of Ecology* **65**:157-174.
- _____, and R. H. Whittaker (1981). Hierarchical classification of community data. *Journal of Ecology* **69**:537-557
- _____, (1982). *Multivariate Analysis in Community Ecology*. Cambridge University Press.
- Gerdol, R., C. Ferrari, and F. Piccoli (1985). Correlation between soil characters and forest types: a study in multiple discriminant analysis. *Vegetatio* **60**:49-56.
- Hamman, M. J. (1972) Vegetation of alpine and subalpine meadows of Mount Ranier National Park, Washington. (Unpub. M. S. Thesis at Washington State University, WA), In *Natural Vegetation of Oregon and Washington*, J. F. Franklin and C. T. Dyrness (eds.), pp. 248-90. (1973, repr. 1984), Oregon State University Press, Corvallis, OR.
- Henderson, J. A. (1974). *Composition, Distribution and Succession of Subalpine Meadows*

- in *Mount Ranier National Park*. Ph.D.dissertation, Oregon State University, Corvallis, OR.
- _____, and D. Peter (1982). *Preliminary Plant Associations and Habitat Types of the Snoqualmie and Adjacent Skykomish River Drainages, Mt. Baker-Snoqualmie National Forest*. USDA Forest Service, PNW Region.
- _____, and D. Peter (1983). *Preliminary Plant Associations and Habitat Types of the Northern Skykomish Ranger District, Mt. Baker-Snoqualmie National Forest*. USDA Forest Service, PNW Region.
- _____, and D. Peter (1984). *Preliminary Plant Associations and Habitat Types of the Darrington Ranger District, Mt. Baker-Snoqualmie National Forest*. USDA Forest Service, PNW Region.
- _____, and D. Peter (1985). *Preliminary Plant Associations and Habitat Types of the Mt. Baker Ranger District, Mt. Baker-Snoqualmie National Forest*. USDA Forest Service, PNW Region.
- Hill, M. O. (1979a). *DECORANA - A FORTRAN Program for Detrended Correspondence Analysis and Reciprocal Averaging*. Cornell University, Ithaca, NY.
- _____. (1979b). *TWINSPAN - A FORTRAN Program for Arranging Multivariate Data in an Ordered Two-way Table by Classification of the Individuals and Attributes*. Cornell University Ithaca, NY.
- _____, and H. G. Gauch, Jr. (1980). Detrended correspondence analysis, an improved ordination technique. *Vegetatio*, **42**:47-58.
- Hitchcock, C. L., A. Cronquist, M. Ownbey, and J. W. Thompson (1969). *Vascular Plants of the Pacific Northwest*, Parts 1-5. University of Washington Press, Seattle, WA.
- _____, and A. Cronquist (1973). *Flora of the Pacific Northwest*. University of Washington Press, Seattle, WA.
- Kellogg, E. A. (1985). A biosystematic study of the *Poa secunda* complex. *Journal of the Arnold Arboretum* **66**:201-242.
- Klinka, K., J. Pojar, and D. V. Meidinger (1991). Revision of biogeoclimatic units of coastal British Columbia. *Northwest Science* **65**:32-47.
- Kovalchik, B. L., and L. A. Chitwood (1990). Use of geomorphology in the classification of riparian plant associations in mountainous landscapes of central Oregon, U. S. A. *Forest Ecology and Management* **33/34**:405-418.
- _____. (1993). *Riparian Plant Associations on the National Forests of Eastern Washington*. Draft Version 1. Colville National Forest, Colville, WA.
- Kruckeberg, A. R. (1969). Plant life on serpentinite and other ferromagnesium rocks in northwestern North America. *Syesis* **2**:15-114.
- Kuchler, A. W (1973). Problems in classifying and mapping vegetation for ecological regionalization. *Ecology* **54**:512-523.
- Larson, J. W. (1972, unpub.). Ecological Role of Lodgepole Pine in the Upper Skagit River Valley, Washington. (Unpub. M. S. Thesis at University of Washington Seattle, WA), In *Natural Vegetation of Oregon and Washington*, J. F. Franklin and C. T. Dyrness (eds.), p 312. (1973, repr. 1984), Oregon State University Press, Corvallis, OR.
- Lillybridge, Terry, B. Kovalchik, C. Williams and B. Smith (1995). *Field Guide for Forested Plant Associations of the Wenatchee National Forest*. USDA-FS PNW Res. Sta. PNW-GTR-359, Portland, OR.
- Lincoln, R. J., G. A. Boxshall, and P. F. Clark (1982). *A Dictionary of Ecology, Evolution and Systematics*. Repr. 1990, Cambridge University Press, NY.
- Ludwig, J. A., and J. F. Reynolds (1988). *Statistical Ecology*. John Wiley and Sons.

- May, R. M. (1975). Patterns of species abundance and diversity. In *Ecology and Evolution of Communities*, M. L. Cody and J. M. Diamond.(eds.) Belnap Press, Cambridge, MA.
- McLean, A. (1970). Plant communities of the Similkameen Valley, British Columbia and their relationships to soils. *Ecological Monographs* **40**:403-424.
- Mueller-Dombois, D., and H. Ellenberg (1974). *Aims and Methods of Vegetation Ecology*. John Wiley and Sons.
- Pfister, R. D. (1976). Land capability assessment by habitat types. In *America's Renewable Resource Potential—1975: the Turning Point*, pp. 312-26. Proc. 1975 Natl. Conv. Soc. Amer. For., Washington, D. C.
- _____, B. Kovalchik, S. Arno and R. Presby (1977 repr.). *Forest Habitat Types of Montana*, USDA-FS GTR-INT-34, Intermountain Forest and Range Experiment Station, Ogden, UT.
- _____, and S. F. Arno (1980). Classifying forest habitat types based on potential climax vegetation. *Forest Science* **26**:52-70.
- Preston, F. W. (1962). The canonical distribution of commonness and rarity, Part I. *Ecology* **43**:185-215.
- Rubec, C. D. A. (1979, ed.). *Applications of Ecological (Biophysical) Land Classification in Canada*. Canadian Government Publishing Centre. Hull, Quebec, Canada.
- State of Minnesota Land Management Information Center (1992). *EPPL7 Environmental Planning and Programming Language*, Version 7, release 2.1. St. Paul, MN.
- Tiner, R. W. (1993). The primary indicators method—a practical approach to wetland recognition and delineation in the United States. *Wetlands* **13**:50-64.
- _____(1991). The concept of a hydrophyte for wetland identification. *Bioscience* **41**:236-247.
- Tisdale, E. W., and A. McLean (1957). The Douglas-fir zone of southern interior British Columbia. *Ecological Monographs* **27**:247-266.
- _____(1947). The grasslands of the southern interior of British Columbia. *Ecology* **28**:346-382.
- Vanbianchi, R., and S. J. Wagstaff (1988). A floristic survey of Big Beaver Valley. In *Plant Life of Washington State*, Douglasia Occasional Papers, Vol. 3, pp. 1-61. Washington Native Plant Society, University of Washington, Seattle, WA.
- Walker, H. A., S. B. Saila, E. L. Anderson (1979). Exploring data structure of New York Bight benthic data using post-collection stratification of samples, and linear discriminant analysis for species composition comparisons. *Estuarine and Coastal Marine Science* **9**:101-120.
- Wheeler, D. L. (1987). *Computer Analysis of Ecological Data, a User's Manual for the Data General MV-Series*. Review Draft. USDA Forest Service, Siskiyou National Forest.
- Whittaker, R. H. (1967). Gradient analysis of vegetation. *Biological Reviews* **42**: 207-264.
- _____, and H. G. Gauch (1978). Evaluation of ordination techniques. In *Ordination of Plant Communities*, R. H. Whittaker (ed.), pp. 277-336. Junk, The Hague.
- _____(1987). An application of detrended correspondence analysis and non-metric multidimensional scaling to the identification and analysis of environmental factor complexes and vegetation structures. *Journal of Ecology* **75**:363-376.
- Wiegleb, G. (1981). Application of multiple discriminant analysis on the analysis of correlation between macrophyte vegetation and water quality in running waters of Central Europe. *Hydrobiologia* **79**:91-100.
- Williams, C. K., and B. G. Smith (1991). *Forested Plant Associations of the Wenatchee National Forest*. Draft. USDA Forest Service, Wenatchee, WA.
- _____, and T. R. Lillybridge (1983). *Forested Plant Associations of the Okanogan National*

Classification of North Cascades Vascular Plant Communities

<http://www.okanogan1.com/ecology/eco-ncas>

© George Wooten 1996-2009

Forest. USDA Forest Service Report R6-ECOL-132b-1983, PNW Region.

Williams, W. T., and Dale, M. B. (1965). Fundamental problems in numerical taxonomy. *Advances in Botanical Research* **2**:35-68.

Zobel, D. B., A. McKee, G. M. Hawk and C. T. Dyrness (1976). Relationships of environment to composition, structure, and diversity of forest communities of the central western Cascades of Oregon. *Ecological Monographs* **46**:135-156.