COENOSES 2(2): 61-65, 1987

FLEXIBLE GRADIENT ANALYSIS: A NOTE ON IDEAS AND AN APPLICATION

O. Wildi, Swiss Federal Institute of Forestry Research, CH-8903 Birmensdorf, Switzerland
and
L. Orlóci, Department of Botany, University of Hawaii at Manoa,
3190 Maile Way, Honolulu, HI 96822 U.S.A.

Keywords: Vegetation, Theory, Sampling, Data, Gradient, Analysis

Abstract: The standard approach to the study of vegetation is discussed in which sampling represents a first and data analysis a second step. An alternative is suggested where sampling and analysis run concurrently. An example is given to demonstrate possible gains in efficiency and flexibility.

Introduction

The quantitative methods of data analysis had a long tradition of service in Vegetation Science. Often they were used to support classifications, mainly data sorting («Tabellenarbeit»), and also tasks involved in gradient analysis. Since they represented technical innovations to allow precision and to speed up the work, such as in the structuring of phytosociological tables and the preparation of ordinations, these methods have received an early acceptance. Substantial texts cover the field, giving overview and many details (Greig-Smith 1957, Goodall 1962, Sneath and Sokal 1973, Mueller-Dombois and Ellenberg 1974, Pielou 1977, 1984, Orlóci 1978, Gauch 1982, Legendre and Legendre 1983, Feoli *et al.* 1984, and others).

It is interesting to note that many of the methods were not primarily developed for vegetation studies. However, they were adopted since they neatly fitted into the world of the existing theories about the nature of vegetation. For example, the clustering methods gave support to the idea of distinct types (Braun-Blanquet 1964, Whittaker 1962, Westhoff and van der Maarel 1978), ordinations helped to promote the individualistic hypothesis (Gleason 1926, 1939, Curtis 1959, MacIntosh 1967), and the modelling techniques lent strength to reasoning about community development (Clements 1916, van Hulst 1978, 1979, 1980).

The methods long existing have undergone substantial reforms (Orlóci and Kenkel 1985), but the approach itself did not change much: The methods continue to be applied sequentially with sampling first and data analysis next. Although radical consequences are not to be expected from continued refinements, there are signs that the availability of high-powered portable computers may have broadened the possibilities in applications. Simultaneous data analysis, as the sampling progress, and modification of the sampling as the partial results appear, and in this way, optimization of the data for the objectives appear to be the next stage to be attained which acctually may change the approach itself. The present paper offers thoughts and describes possibilities along these lines.

A flexible approach

An approach which allows to change the priorities of the operations is flexible. This might involve individual steps or entire pathways. The result is that the user no longer operates within the confines of a rigidly prescribed method, but accepts the idea that as more information is revealed methodological changes are acceptable. He employs combinations of methods, selects among options and works through steps linked by the flow of data and partial results. Podani's (1984) concept of sampling as a «spatial process» is akin to «flexibility» in that sampling becomes a process in fixed space through allowing sampling unit size to vary continuously, or allowing even the design itself to change while the sampling is in progress. Figure 1 illustrates this. The survey area is subdivided into strata and samples are placed within the individual strata at random. As long as the strata are kept small, sampling is nearly systematic. When the strata size is increased, regularity vanishes. If the strata are allowed to overlap, the design becomes increasingly random. What to gain from these? For one thing, the site is no longer characterized by scalar quantities such as the mean yield, variance, etc., which serve as estimates, but rather by continuous functions reflecting changes in mean yield, variance, etc. as a function of sampling unit size and design. If one is to optimize the information gained, one can examine the functions for optima consistent with one's definition for optimality. This can be carried further, involving data analyses yielding more complex results than just single scalar quantities.

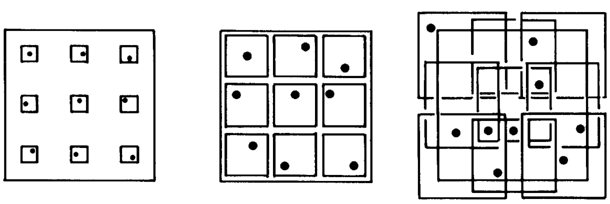

Fig. 1. The flexible sampling design (Podani 1984). Stratified random sampling will result in any arrangement intermediate between quasi systematic (left hand figure, small units) and quasi random (right hand figure, large overlapping units).

Cluster analysis and ordination may enter the picture. The results will be classifications and ordinations as continuous functions with each point corresponding to a complete dendrogram or ordination graph. The analysis may continue to define the optimal dendrogram or optimal ordination graph. An example is presented in the next section.

Flexible indirect gradient analysis

In indirect gradient analysis, finding the shape and sharpness of a coenocline is a topic of primary interest. The following example describes a simple case, although not an entirely realistic one since it draws on information which were far too well known prior to analysis. The communities involved are oligotrophic wetlands from prealpine locations at Rothenthurm, Switzerland, which were described by Wildi (1977). He distinguished between numerous types which he characterized based on floristic and environmental features. The coenocline of specific interest is tied to the groundwater quality. The exact nature of the coenocline, supposedly unknown prior to the study, is to be revealed. To achieve this objective a flexible approach is applied which involves sampling and resampling, primary data analysis at each stage, and secondary analysis:

Sampling. Eighteen plots were sited within an area of about 20×90 m. The arrangement was systematic using

Table 1. The raw data. Species are arranged in the order of sampling sweeps. The ordering of plots is arbitrary. Some environmental variables are also included. Entries in the table are in units as identified for environmental variables and cover/abundance values for species. Releves are arranged according to increasing pH.

RELEVE GROUP NO.		·	1	1	1	1	1	2	2	2	2	2	2	3	3	4	4	4	4	4
		·	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5	5
		·	2	1	3	0	1	2	2	4	2	0	3	5	6	4	5	5	6	5
RELEVE NO.		·	8	2	7	6	1	2	1	4	7	5	8	0	1	3	1	7	0	6
		·		1	1	1	1	1	1	1		1	1	1			1	1		1
WATERLEVEL, AVERAGE		·	2	5	0	2	4	3	1	8	5	6	5	7	8	4	2	6	9	3
		·																		
		·																		
		·	3	4	3	4	3	3	3	2	4	5	5	3	3	5	4	2	6	3
WATERLEVEL, MIN		·	8	3	1	5	6	2	3	9	6	9	1	6	2	2	1	3	3	4
		·												0				0		
		·	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		·	5	3	5	6	4	4	4	5	5	3	3	0	3	3	3	0	3	3
LOG PEAT LEVEL, CM		·	4	9	4	0	7	7	7	4	4	9	9	0	9	0	9	0	9	0
		·		1		1	1		1			1		1			1	1	1	1
		·	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		·	9	0	9	2	0	9	0	6	7	0	9	0	9	9	0	4	1	4
LOG SLOPE, DEG		·	0	0	5	0	4	5	0	9	7	7	5	0	0	5	0	4	4	7
		·																		
		·	3	4	4	4	4	4	4	5	5	5	5	5	6	7	7	7	7	7
		·	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
PH WATER		·	6	2	3	4	6	7	9	0	1	5	7	8	0	1	1	1	2	6
		·	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2
		·	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		·	4	4	3	4	3	3	4	3	2	3	3	3	2	7	8	3	7	0
LOG OHM/CM WATER		·	1	6	9	6	6	8	6	2	7	4	0	0	7	4	2	9	4	3
		·	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	2
		·	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.	.
		·	3	2	2	3	3	3	4	4	3	4	3	3	4	8	9	7	8	1
LOG CA IN WATER (0.1PPM)		·	9	7	7	8	8	0	3	4	2	4	9	8	7	6	4	9	8	5
1 Vaccinium myrtillus	1	·		+		+						+								
2 Vaccinium uliginosum	2	·		+		+						+								
4 Oxycoccus quadripetalus	3	·	+	+		1		+		+		+								
5 Calluna vulgaris	4	·		1		+		+	+			+								
23 Parnassia palustris	5	·												+	+	+	+	+	+	+
26 Arnica montana	6	·	+	1	1	+	1	+	+	+	+	+								
59 Solidago virgaurea	7	·		+		+														
90 Bellidiastrum michelii	8	·												+	+		+	+	+	+
37 Sphagnum medium	9	·	+	4		2	1	2	+		+	2	+							
57 Sphagnum subsecundum	10	·		+	+					5			+	5	+	+	1			

the intersection points of a 2×9 square grid as plot centers. In the first sweep through the plots the woody species were sampled. These included *Vaccinium myrtillus, V. uliginosum, Oxycoccus quadripetalus* and *Calluna vulgaris.* In two more sweeps, four *(Parnassia palustris, Arnica montana, Solidago virgaurea, Bellidiastrum michelii)* and two more *(Sphagnum medium, S. subsecundum)* species were added to the original four. Finally, a complete species list was compiled. At each step, cover/abundance was estimated for each species by the Braun-Blanquet scale. The raw data are given in Table 1.

Primary Analysis. After each sampling sweep, the data set was analysed in principal components analysis (PCA) and ordination graphs were drawn. These are presented in Figure 2. The analysis also included plotting pH, depth to groundwater table, and base saturation within the ordinations to identify those environmental variables which may be responsible for coenoclines. Of these, pH proved to be the most related to variation in species cover/abundance. It also became evident that the shape of the gradient emerges rather early in the ordinations as more and more species are included. Ten species are in fact sufficient to describe the relationship between the vegetation types and pH, and the remaining 94 species could be omitted, saving much of the effort.

Secondary Analysis. Having singled out pH in the primary analysis, its relationship to species cover/abundance values was further examined. For this a contingency table was constructed (Table 2) and on this table concentration analysis (CONA) was performed (Feoli and Orlóci 1984). The results are presented in Table 3 and Figure 3. Clearly, the relationship is very sharp. In fact the canonical correla-

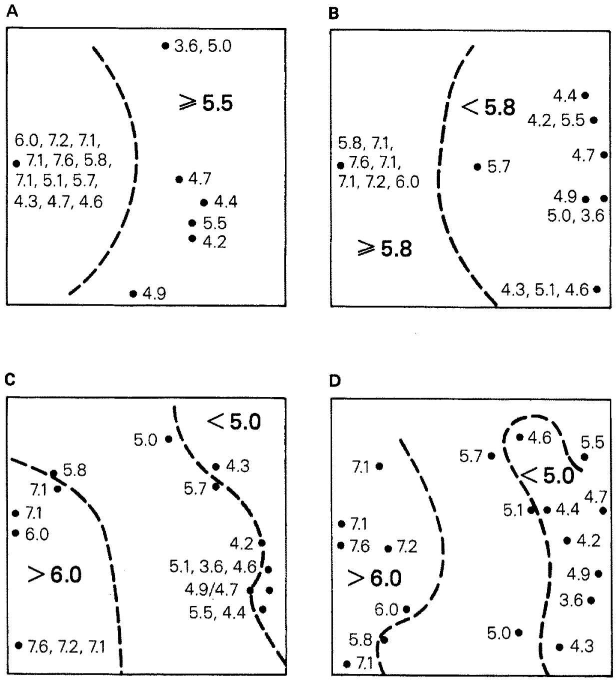

Fig. 2. Ordination diagrams with pH-values shown. The method is PCA based on 4 (graph A), 8 (graph B), 10 (graph C) and 104 species (graph D).

tion of species and pH is almost functional in lattice 1 which accounts for almost 91% of the total chi-squared. The total 55.0 with 207 degrees of freedom corresponds to about 0.001 probability of a more extreme case, suggesting high significance. The mean square contingency coef-

Table 2. Species/pH contingency table. Entries indicate observed frequencies. pH classes have equal widths. Species # identifies species in Table 1.

	pH class				
Species #	3.6-4.6	4.7-5.7	5.8-6.8	6.9-7.9	Totals
1	2	1	0	0	3
2	2	1	0	0	3
3	3	3	0	0	6
4	2	3	0	0	5
5	0	0	2	5	7
6	5	5	0	0	10
7	2	0	0	0	2
8	0	0	2	4	6
9	4	5	0	0	9
10	2	2	2	2	8
Totals	22	20	6	11	59

Table 3. Some statistics of the species/pH relationship obtained in concentration analysis.

Canonical variate	Canonical correlation	Chi-squared	% of total chi-squared
1	0.920	49.9	90.7
2	0.253	3.8	6.9
3	0.149	1.3	2.4
Totals		55.0	100.0

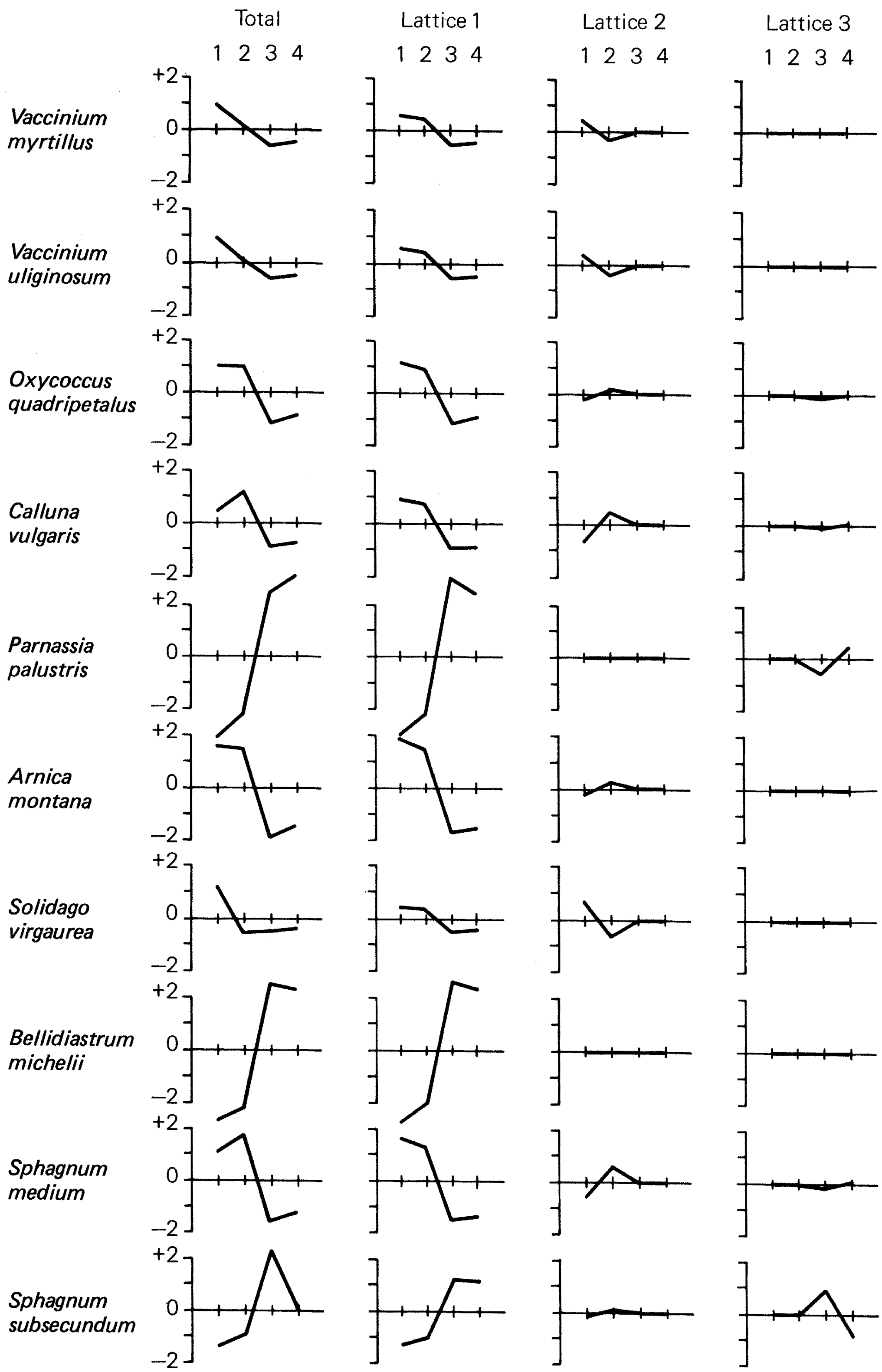

Fig. 3. Deviation from expectation by specimen in the total contingency table and in the lattices for pH-classes 1 through 4 (see Table 2).

ficient is 0.311 which indicates a relatively high sharpness (33%), since its actual maximum in this case is about 0.9.

Examination of the dispersion profiles allows further insight into species behaviour. The most obvious finding is that species change monotonically with little evidence for interference by factors which are non-linearly related to or are statistically independent from pH. This fact is indicated by the monotonic, accentuated relationships in lattice 1 and an almost total lack of trended variation in lattices 2 and 3. The fact that 9% of the total variation is accounted for by the first lattice is a further indication of the importance of pH as a predicted variable.

Summary

We considered an approach to vegetation studies in which the tasks of sampling and data analysis run concurrently. As the sampling progresses, recursive analyses permit the measurement of the success of the sample already taken to reveal ecological information. The example focuses on environmental gradients and sampling stops when a gradient, pH in this case, clearly emerges. Secondary data analysis is performed to clarify the relationships.

Acknowledgement. The authors were recepients of an NSERC operating grant (to L.O.) and institutional support from the Swiss Federal Institute of Forestry Research and the University of Hawaii during tenure of this project.

REFERENCES

ANDERBERG, M. R. 1973. *Cluster analysis for Applications.* Academic Press, N.Y., San Francisco, London. 359 p.

BRAUN-BLANQUET, J. 1964. *Pflanzensoziologie.* 3rd ed. Springer-Verlag, Wien. 845 p.

CLEMENTS, F. E. 1916. Plant succession and analysis of the development of vegetation. Carnegie Inst. Wash. Publ. 242: 1-512.

CURTIS, J. T. 1959. *The Vegetation of Wisconsin: An Ordination of Plant Communities.* Univ. Wisconsin Press, Madison. 143 p.

FEOLI, E., M. LAGONEGRO and L. ORLÓCI. 1984. *Information Analysis of Vegetation Data.* Dr. W. Junk, The Hague. 143 p.

FEOLI, E. and L. ORLÓCI. 1984. Species dispersion profiles of anthropogenic grasslands in the Italian Pre-Alps. Vegetatio 60: 113-118.

GAUCH, H. 1982. *Multivariate Analysis in Community Ecology.* Cambridge University Press, New York.

GLEASON, H. A. 1926. The individualistic concept of the plant association. Bull. Torrey Bot. Club, 53: 7-26.

GLEASON, H. A. 1939. The individualistic concept of the plant association. Amer. Midl. Nat. 21:92-110.

GOODALL, D. W. 1962. Bibliography of statistical plant sociology. Excerpta Botanica, Sect. B 4: 16-322.

GREIG-SMITH, P. 1957. *Quantitative Plant Ecology.* London, Butterworths.

HULST, van R. 1978. On the dynamics of vegetation: patterns of environmental and vegetational change. Vegetatio 38:65-75.

HULST, van R. 1979. On the dynamics of vegetation: Markov chains as models of succession. Vegetatio 40: 3-14.

HULST, van R. 1980. Vegetation dynamics or ecosystem dynamics: Dynamic sufficiency in succession theory. Vegetatio 43: 147-151.

LEGENDRE, L. and P. LEGENDRE. 1983. *Numerical Ecology.* Elsevier, N.Y. 419 p.

MACINTOSH, R. P. 1967. The continuum concept of vegetation. Bot. Rev. 33: 130-187.

MUELLER-DOMBOIS, D. and H. ELLENBERG. 1974. *Aims and Methods of Vegetation Ecology.* John Wiley & Sons, N.Y. 547 p.

ORLÓCI, L. 1978. *Multivariate Analysis in Vegetation Research.* Junk, The Hague. 2nd ed. 451 p.

ORLÓCI, L. and N. KENKEL. 1985. *Introduction to Data Analysis with Applications from Population and Community Ecology.* International Cooperative Publishing House, Fairland, Maryland. 340 p.

PIELOU, E. C. 1977. *Mathematical Ecology.* 2nd ed. John Wiley & Sons, N.Y. 385 p.

PIELOU, E. C. 1984. *The Interpretation of Ecological Data.* John Wiley & Sons, N.Y. 263 p.

PODANI, J. 1984. Spatial processes in the analysis of vegetation: Theory and review. Acta Botanica Hungarica 30: 75-118.

SNEATH, P. H. A. and R. R. SOKAL. 1973. *Numerical Taxonomy.* Freeman, San Francisco. 573 p.

WESTHOFF, V. and E. VAN DER MAAREL. 1978. The Braun-Blanquet approach. 2nd ed. *In*: R. H. WHITTAKER (ed.), *Classification of Plant Community,* pp. 287-399. Junk, The Hague.

WHITTAKER, R. H. 1962. Classification of natural communities. Bot. Rev. 28: 1-239.

WILDI, O. 1977. Beschreibung exzentrisches Hochmoore mit Hilfe quantitativer Methoden. Veröff. Geobot. Inst. ETH, Stiftung Rübel 59. 128 p.

Manuscript received: August 1986