COENOSES 3 (3): 135-139, 1988

THE USE OF MULTI-WAY CONTINGENCY TABLES FOR THE STUDY OF EPIPHYTIC LICHEN DISTRIBUTION

T. Herben, Botanical Institute, Czechoslovak Academy of Sciences, CS-252 43 Pruhonice near Praha, Czechoslovakia
and
J. Liška, Institute of Applied Ecology and Ecotechnology, Agricultural University, CS-281 63 Kostelec nad Černymi lesy, Czechoslovakia

Abstract. Data on epiphytic lichen distribution on isolated trees in a moderately polluted area in Czechoslovakia were analyzed using a log-linear model to investigate the interaction of environmental factors. Pollution, bark eutrophication and substrate (tree species) were the factors included in the analysis. Few third-order interactions were identified (only for *Pseudevernia furfuracea, Lecanora carpinea* and *Usnea hirta*), indicating that the effects of the factors can be separated from each other in most cases. The implications for bioindication and similar studies are discussed.

Introduction

One of the common problems in descriptive ecology amounts to relating the occurrences of species (or groups of species) to external ecological variables. This is easy when the response to each individual factor is not dependent on the other factors and can therefore be studied separately. Needless to say, this is not necessarily the case. Different modes of substitution or synergism between factors have been known to exist (*e.g.* Slack 1976, Piipo 1982). While this phenomenon may occur at any level of resolution, complicated cases of factor dependence (response to one factor depending on the actual value of another factor) can obscure strong cases of species response. In many such situations, the interaction of the factors involved may go undetected without further statistical analysis. Then it is necessary:

(1) to test whether the level of one ecological factor influences significantly response to a second factor;

(2) to determine how the two factors are related.

This can be done mathematically in different ways, depending on the type of the available data. For continuous data ANOVA or ANCOVA (Sokal & Rohlf 1981) may be suitable; for categorical data a multi-way contingency table analysis is appropriate. The log-linear model of Brown (Brown 1976, Dixon & Brown 1977) is of this type. This technique may be used both for hypotheses generation and testing.

Identification of interactions is of special interest in bioindication studies. Generally, under this we understand studies which use the presence or quantity of a species to infer the values of some environmental variable (see Ellenberg 1974, Klimeš 1986). Bioindication assumes that the particular species distribution is systematically influenced by infered variable and that this relation is monotonous.

The success of any bioindication study depends on the knowledge of the factor potentially capable of modifying the relation. This problem of modifying factors can be usually dealt with in two ways, both of which require knowledge of the levels of the modifying factor. The first is to select sites for data collection with the same level of the modifying factor. Using this method the significance and the type of interaction need not be known between the inferred and modifying factors. However this manouvre may markedly reduce the available sample numbers. The second way is to calibrate species occurrence - inferred factor relation for all levels of the modifying factor. However, to do this the interaction between the modifying and the inferred factors have to be known, since multiple regression and similar techniques can be used only when this interaction is not significant.

This approach was used for processing a set of lichen floristic data collected for bioindication purposes. As this data set was collected in an area of moderate air pollution, the epiphytic lichen flora composition was determined mainly by factors other than pollution, notably bark eutrophication and tree species. Bark eutrophication is one of the known factors influencing lichen distribution with an effect so dominant that it may occasionally mask the effect of air pollution. Eutrophication is usually regarded as enrichment of the bark with dust containing organic particles, bringing excess nitrogen into the system. For this reason lichen species are often regarded nitrophilous. However eutrophication is a very potent process (*e.g.*, Barkman 1958); it leads to changes in phosphorus, cations and pH, which in turn interact with bark acidification resulting from acid rain. Clearly, the term nitrophilous is not precise. In the following text we use it in the sense of "eutrophy-philous", *i.e.* having preference for nutrient enriched bark. We used the log-linear model to separate the effects of these factors on species occurence and to identify the interactions between them (third-order effects in this design).

Methods

The study area is under low pollution stress. The epiphytic lichen flora is comparatively rich and permit behavioural studies on many species. The locality is in South Bohemia including the town of Tábor (latitude 49° 25'N, longitude 14° 40'E) within a radius of 10-12 km. The most significant polutant is SO_2 emitted from two thermal power plants (about 5,000 tonnes/year): in winter, domestic sources contribute to pollution. A detailed description of the study area is in Herben & Liška (1984). The lichen flora on 217 trees was investigated: 181 trees were analyzed after exclusion rare species.

We classified trees according to three levels of pollution and two levels of bark eutrophication in the following manner: the area was divided into three zones of air pollution using the overall composition of epiphytic lichen vegetation and flora; each tree was assigned to a pollution level. Pollution zones were recognised following the method of LeBlanc & De Sloover (1970). The level of bark eutrophication was assessed according to the dominant vegetation type on the stand. Then each stand was classified into 6 groups according to the substrate. Only lichen species with a frequency exceeding 15% were included in the analysis.

The observations were arranged into a four-way contingency table. The table cells contain trees counts having the same values of pollution, eutrophication and presence/absence of the particular lichen species. The four-way table has marginal sums, which form 3, 2, and 1-way contingency tables, presenting values summed over 1, 2 or 3 factors. These are called marginal tables. The data analysis used the log-linear modeL(see Brown 1976, Dixon & Brown 1977, Sokal & Rohlf 1981). This method allows to test various deviations from independence in the basic table. The basic idea of the method is that the logarithms of cell frequencies can be written as a linear function of certain parameters, which express the individual main effects and their interactions. Interaction in this context means the departure of independence between factors. These parameters are estimated from the sums in the marginal tables. Including a given parameter in the model means that the relation between given factors, estimated from the appropriate marginal table, is the basis for calculation of expected cell frequencies. The expected frequencies are compared to the observed ones by some goodness-of-fit statistic, usually by a G-test.

The most general model - *i.e.* that includes all the main effects, all interactions between pairs, triplets and quadruplet of variables - can be written as

$$\ln f_{ijkl} = \vartheta + \lambda_i^S + \lambda_j^E + \lambda_k^P + \lambda_l^L + \lambda_{ij}^{SE} + \lambda_{ik}^{SP} + \lambda_{il}^{SL} + \lambda_{jk}^{EP} + \lambda_{jl}^{EL} + \lambda_{kl}^{PL} + \lambda_{ijk}^{SEP} + \lambda_{ijl}^{SEL} + \lambda_{ikl}^{SPL} + \lambda_{jkl}^{EPL} + \lambda_{ijkl}^{SEPL}$$

In this f_{ijkl} - expected count in the all ijkl; ϑ - constant; λ - log-linear parameters; for other abbreviations see Table 1. The model presumes interaction between all factor combinations. The significance of any interaction term may be tested by comparing two models, differing just in the presence of a particular interaction term. Importantly, such an interaction testing (and hence significance) is model-dependent. Brown (1976) proposed two screening methods of searching for interactions likely to be significant. There are:

(1) partial association between factors, testing significance of the interaction of the i-th order in the full model of the i-th order, *i.e.* including all possible interactions of the i-th order;

(2) marginal association between factors, testing significance of a given interactions between *i* factors in a marginal table constructed only from these factors, excluding all the other interactions and main effects.

These tests allow fast searching for significant interactions in the main table, which may be then used for model fitting and computation of residuals. Accordingly, the following procedure is adopted:

1. For each species test the hypothesis that each marginal and partial associations between factors from order 2 to 4 equals zero. Use a significance level of 0.05.
2. If no effect of the third-order is found significant, construct a minimal model including all the second order effects whose deletion significantly impaired the fit of the model. The interaction which expresses the study design (between pollution, eutrophication and tree species) is always included (Sokal & Rohlf 1981).
3. If at least one third-order interaction is significant, test in the particular model, constructed in the same way as in the preceding case. Then conclude that the response to one factor is determined by the level of the second factor and divide the original four-way table into three-way tables. Analyse these separately using the same technique. Divide the table according to the bark eutrophication level.

Results and discussion

There are surprisingly few species which show significant third-order effects. Many species respond to both pollution and bark eutrophication, the response to each can be isolated. Only one species *(Parmelia exasperatula)* showed a significant response to pollution without a significant response to bark eutrophication. There are a few species (see Table 2) which show significant response to both factors. All avoid one level of bark eutrophication.

The magnitude of the individual effects can be roughly assessed if the corresponding terms in the model showed a marked difference in significance (as in *Evernia prunastri* or *Parmelia sulcata*). However, the second-order effects should be interpreted with caution. In both pollution and bark eutrophication, the level assignment was done according to lichen vegetation without

Table 1. Models for lichen species. For each term included in the model, significance of the difference in fit between the final model and the model with this particular term deleted is tested; for other terms, significance of the difference in fit between final model and the model with this term added. Abbreviations: S - substrate (phorophyte species), E - bark eutrophication, P - pollution, L - lichen species. Combination of *n* letters means an effect of n-th order including those factors. Nomenclature of the lichens follows Wirth (1980).

Lichen	Frequency (Total = 181)	Final model	Fit of the model	Deleted term	Significance	Addes term	Significance
Buellia punctata	102	SEP, EL	0.917	EL	0.000	SL	0.458
						PL	0.502
Candelaria concolor	40	SEP, L	0.944			SL	0.333
						EL	0.381
						PL	0.107
Candelariella xanthostigma	60	SEP, PL	0.812	EL	0.000	SL	0.271
						PL	0.262
Evernia prunastri	67	SEP, PL, EL, SL	0.999	SL	0.005	SEL	0.886
				PL	0.000	SPL	0.920
				EL	0.046	EPL	0.379
Hypocenomyce scalaris	28	SEP, EL, PL	0.972	EL	0.002	SL	0.280
				PL	0.006	EPL	0.768
Hypogymnia physodes	118	SEP, EL, PL	0.971	EL	0.000	SP	0.792
				PL	0.003	EPL	0.127
Lecanora carpinea	55	SEP, EPL	0.957	EPL	0.011	SL	0.154
Lecanora chlarotera	55	SEP, EL	0.859	EL	0.000	SL	0.352
						PL	0.097
Lecanora conizaeoides	28	SEP, PL, EL	1.000	PL	0.000	SL	0.929
				EL	0.001	EPL	0.851
Lecanora varia	28	SEP, EL	0.861	EL	0.000	SL	0.569
						PL	0.377
Parmelia exasperatula	58	SEP, PL	0.917	PL	0.005	SL	0.176
						EL	0.054
Parmelia sulcata	116	SEP, PL, SL	0.961	PL	0.000	EL	0.531
				SL	0.048	SPL	0.728
Physcia adscendens	53	SEP, SL, EL	0.992	SL	0.002	PL	0.053
				EL	0.000	SEL	0.911
Physcia dubia	47	SEP, EL	0.991	EL	0.000	SL	0.979
						PL	0.123
Physcia orbicularis	38	SEP, EL	0.994	EL	0.000	SL	0.070
						PL	0.336
Physcia stellaris	26	SEP, EL, SL	1.000	EL	0.013	PL	0.692
				SL	0.019	SEL	0.999
Pseudevernia furfuracea	33	SEP, EPL	0.686	EPL	0.024	SL	0.609
Scoliciosporum chlorococcum	43	SEP, EL	0.972	EL	0.000	SL	0.632
						PL	0.362
Usnea hirta	40	SEP, SEL, PL	0.960	SEL	0.030	SPL	0.847
				PL	0.000	EPL	0.149
Xanthoria candelaria	42	SEP, EL	0.959	EL	0.008	SL	0.420
						PL	0.212
Xanthoria parietina	61	SEP, PL, EL	0.971	PL	0.000	SL	0.488
				EL	0.000	EPL	0.385

Table 2. Structure of the final models.

no interaction	Candelaria concolor
with eutrophication (+)	Buellia punctata, Candelariella xanthostigma, Lecanora chlarotera, Physcia orbicularis, P. dubia, Xanthoria candelaria
with eutrophication (—)	Lecanora varia, Scoliciosporum chlorococcum
with pollution (—)	Parmelia exasperatula
with pollution (—) and substrate	Parmelia sulcata
with eutrophication (+) and substrate	Physcia adscendens, P. stellaris
with pollution (—), eutrophication (—) and substrate	Evernia prunastri
with eutrophication (+) and pollution (—)	Xanthoria parietina
with eutrophication (—) and pollution (+)	Hypocenomyce scalaris, Hypogymnia physodes, Lecanora conizaeoides
with eutrophication + + pollution	Lecanora carpinea, Pseudevernia furfuracea
with eutrophication + + pollution and substrate	Usnea hirta

the independent set of data. Thus some hypotheses about species behaviour are incorporated into the input data *a priori*. We believe this does not invalidate the analysis of the resulting multi-way contingency table, since

(1) the level assignment is done using flora attributes (mainly quantitative parameters) rather than the presence of species, which is used in the model;

(2) we are interested mainly in finding third-order effects. Since assignment of both pollution and bark eutrophication levels was done independently, third-order effects are much less likely to be incorporated in the model.

The substrate (= phorophyte species) is insignificant in most cases; only 3 species showed a significant response to it (Table 2). In addition, in the case of *Usnea hirta*, the phorophyte species enters into third-order interaction with bark eutrophication.

The third-order effect including pollution, bark eutrophication and lichen occurrence, was found in two species. Both at least partly, occur on eutrophicated just as on non-eutrophicated bark *(Lecanora carpinea, Pseudevernia furfuracea)*, great caution has to be exercised when the data on occurrence of such species are used for bioindication purposes. However, other species found on both substrate types *(e.g. Candelaria concolor)* do not show the third-order effect to be significant. The low relative number of third-order effects indicates that bioindication disregarding these modifying factors would probably not be seriously impaired. In a similar study using log-linear models, Zobel (1986) found that *Hypogymnia physodes* did not show a significant interaction of the third order between occurrence, pollution and altitude of the stand.

Although the log-linear model is useful in identifying higher-order interactions between factors which are hard to detect by other techniques, the problem of its use in higher dimensional designs lies in the steeply increasing number of cells in the basic table. This requires very high numbers of observations in such studies.

We believe the study of factor-interactions to be important even in descriptive ecology and should be performed, at least in more complicated cases of multifactorial relations. Importantly, the analysis of more than one external factors is common (Whittaker 1973). Study of interactions of factors may be especially important when two studied factors are related to a third factor, which is a primary determinant of the species presence, but for some reasons cannot be measured. For example, a species presence may depend on the annual course of moisture (which is difficult and costly to determine), which can be sufficiently well described by variables such as soil composition, slope, aspect, etc. Then the correct resolution of interactions between these secondary variables may help to identify the basic factor.

REFERENCES

BARKMAN, J.J. 1958. *Phytosociology and Ecology of cryptogamic epiphytes*. van Gorcum, Assen.

BROWN, M.B. 1976. Screening effects in multidimensional contingency tables. Appl. Statist. 25: 37-46.

DIXON, W.J. and BROWN, M.B. 1977. *Biomedical Computer Programs*. P-series. University of California, Berkeley.

ELLENBERG, H. 1974. Zeigerwerte der Gefässpflanzen Mitteleuropas. Scripta Geobot. 9: 1-97.

HERBEN, T. and LIŠKA, J. 1984. The use of average number of neighbours for predicting lichen sensitivity: a case study. Lichenologist 16: 289-296.

KLIMEŠ, L. 1987. Exploitation of tabulated indicator values in the gradient analysis of vegetation. Preslia 59: 15-24.

LEBLANC, F. and DE SLOOVER, J. 1970. Relation between industrialization and growth of epiphytic lichens and mosses in Montreal. Can. J. Bot. 48: 1485-1496.

PIIPO, S. 1982. Epiphytic bryophytes as climatic indicators in eastern Fennoscandia. Acta Bot. Fenn. 119: 1-39.

SLACK, N.G. 1976. Host specificity of epiphytic bryophytes in eastern North America. J. Hattori Bot. Lab. 41: 107-132.

SOKAL, R.R. and ROHLF. J.F. 1981. *Biometry*. 2nd ed. Freeman, San Francisco.

WIRTH, W.V. 1980. *Flechtenflora: Ökologische Kennzeichnung und Bestimmung der Flechten Südwestdeutschlands und angrenzenden Gebiete*. Ulmer, Stuttgart.

ZOBEL, K. 1986. Analysis of the ability of Hypogymnia physodes to indicate atmospheric pollution in mountainous conditions. Folia crypt. Eston. 21: 4-7.