COENOSES 2(2): 67-74, 1987

REGRESSION MODELLING OF PERTURBATION IN SOME VEGETATION TYPES

P.H. Fewster, Department of Plant Sciences,
University of Western Ontario, London, Ontario, Canada N6A 5B7

Keywords: Vegetation, Environment, Regression, Perturbation, Response

Abstract: Where sampling constraints do not allow for the time element to be measured or where it is simply impossible to recognize specific time relationships, regression modelling can facilitate dynamic vegetation studies. The present paper describes such a modelling and presents the results of computer experiments with real data from Boreal and Taiga vegetation in the Yukon Territory. The experiments focus on environmental perturbation and the relationship of this to compositional changes in different vegetation types for which general tendencies and trends are revealed. The mesic types appear the most resistant to perturbation and resistance declines as the types become more hydric or xeric.

Introduction

The relationship among communities is of major concern in vegetation science. The choice of how to measure this relationship often varies, yet the results often lead to a fuller understanding of community structure and dynamics. One view is to consider how vegetation types respond to various levels of disturbance, from very slight disturbance through to very excessive disturbance for which the term perturbation would be appropriate. Here we will not concern ourselves with the time element, but rather, attempt to predict the eventual state of the system following perturbation. The contribution of this is that it allows one to generate hypotheses on the response to perturbation.

To explore this idea, a regression model can be selected. The steps are as follows: (1) Maximizing predictability is the aim. Hence we seek to maximize the fit as measured by a correlation coefficient. In this case, a number of environmental variables can be used so that the model is that of multiple regression. The fact that the precise model is not chosen *a priori*, implying that regression will not be used in the statistical sense, is important. (2) The model to be used will predict vegetation response to environmental factors. One of these factors (and hence, indirectly, all of them) can be perturbed at increasing levels to see what the effects on the vegetation might be. We will take the general view that the vegetation types will be affected differently. (3) To assist the interpretation, this effect can be illustrated graphically by the use of ordination. Here the various vegetation types within and between various perturbation levels can be compared by plotting them on a coordinate plane.

The Regression Model

General characteristics

Since the model was not chosen *a priori*, as it would be in the case of a statistical approach, here primary interest lies not in testing but in maximizing correlations by finding the best fitting regression model. Goodall (1972) treated regression this way. He was interested in prediction and so used squared, cubed, exponential, and even dependent variables on the right side of his equation. In a regression model, the assumption that the right side variables are independent is an arbitrary requirement serving the limitations inherent in the use of the standard distributions for significance tests. Regarding the testing of hypotheses, high powered alternatives to simplistic tests based on these standard distributions can be readily found, but this is not a topic of immediate concern here. Monte Carlo methods, however, allow one to bypass this difficulty in general, and to tailor the test to the actual conditions of the sample (Orlóci 1978). Furthermore, if one is interested in going beyond statistical curve fitting, to representing some degree of the system mechanisms in the model, Goodall's approach is logical.

Multicollinearity

The use of squared and cubed terms along with the linear terms often results in a high degree of multicollinearity. This does not, however, invalidate the use of regression as a curve fitting technique since its absence is not an underlying assumption. Nevertheless one might choose to reduce multicollinearity, as outlined in Freund and Minton (1979), by selectively eliminating some of the variables. Yet this would tend to reduce predictability, our primary aim. Goodall (1972) speaks of a tendency towards model expansion and not reduction for this very reason. The use of regression here is essentially a curve fitting method. By relaxing the statistical bounds and including variables which interact, we develop a model more like the real system. Hence we move towards a mechanistic model, thus increasing predictability, precision, and generality with the view towards application and inference.

Model building

The analysis, being exploratory, followed somewhat convoluted paths before a particular model was decided upon. Only the steps which were actually used in model building are presented here:

Data. The preliminary framework of the data is represented by twenty three vegetation types described by one hundred one species and four environmental variables. The data were derived from a survey done in the summer of 1978 along the Alaska highway in the Yukon Territory, beginning west of Beaver Creek (62.5°N, 141°W) and ending just west of Watson Lake (60°N, 128.5°W). The purpose of the survey was to typify the vegetation and to search for compositional gradients. Sampling was carried out on a stratified random basis where nested strata included ecoregions (Oswald and Senyk 1977) and terrain types (Foothills 1978). A total of three hundred twenty three plots were established and described by three hundred thirty five species and sixteen environmental variables. As explained by Orlóci and Stanek (1979), twenty five vegetation types were found by a cluster analysis. This resulted in a $25\times(335+16)$ data set.

The derivation of the $23\times(101+4)$ data set used for the model is as follows: Two of the twenty five vegetation types were discarded because they had few species in common with the others (Table 1). Of the three hundred thirty five species, one hundred one were deemed characteristic, i.e. present in at least 50% of the plots of at least one type (Table 2). The use of characteristic species seemed to be an appropriate approach as it served to discard much of the randomness in the data and the cosmopolitan species while retaining those species which could predict a type. The original raw data (species by plot) in the form of presence/absence scores rather than cover-abundance were used. This facilitated a transformation to frequency data for species by type. This is analogous to subdividing a plot or quadrat and noting the presence or absence of the species in each subdivision. Here the three hundred twenty three plots are like the subdivisions and the types are like the quadrats.

Of the sixteen environmental variables used by Orlóci and Stanek (1979), four were selected for use based on considerations of their correlations with changes in vegetation structure. The environmental variables were scored at the level of the plots. These scores were grouped by vegetation type and then averaged. In the case of drainage, scoring was based on seven classes ranging from 1-excessive to 7-wet. Soil texture, the second variable chosen, was a combination of four particle sizes such as gravel/stone, sand, silt, and clay. Each plot was scored as to percent content of the soil mass in the four particle size classes. The raw scores were placed into one of eleven classes and means were calculated for each vegetation type. The third variable, organic content and thickness, was a combination of scores for organic content placed into seven classes and scores for thickness placed into six classes. Again means for vegetation types were calculated based on the classes. The fourth variable, erosion potential, was scored as either

Table 1. Description of twenty three vegetation types. The type code is based on constant genera (Orlóci & Stanek 1979). The numerical values match the raw scores, referring to the amount of moisture: 1 - dry; 2 - moderate to dry; 3 - moderate; 4 - moderate to wet; 5 - wet; 6 - very wet.

Type code	Type nomenclature	Type number	Drainage category
AAC	*Artemisia-Agropyron-Calamagrostis*	11	1
CT	*Calamagrostis-Tortula*	12	1
PVFP	*Pinus-Vaccinium-Festuca*	23	1
PASE	*Populus-Arctostaphylos-Shepherdia*	14	2
PAFP	*Pinus-Arctostaphylos-Festuca*	12	2
PVHC	*Pinus-Vaccinium-Hylocomium*	22	2
PVD	*Picea-Viburnum-Drepanocladus*	2	2
PAF	*Picea-Arctostaphylos-Festuca*	9	3
PFP	*Pinus-Festuca-Peltigera*	19	3
PC	*Picea-Carex*	8	3
PHP	*Picea-Hylocomium-Peltigera*	16	3
PHPC	*Picea-Hylocomium-Peltigera*	21	3
PCHP	*Picea-Cornus-Hylocomium*	17	3
PLH	*Picea-Ledum-Hylocomium*	1	4
PAA	*Picea-Arctostaphylos*	7	4
PSAH	*Picea-Salix-Aulacomnium*	18	5
PARA	*Picea-Rhododendron-Aulacomnium*	5	5
PAAP	*Picea-Arctostaphylos-Aulacomnium*	15	6
S	*Salix*	6	6
PLEA	*Picea-Ledum-Aulacomnium*	4	6
SCA	*Salix-Carex-Aulacomnium*	20	6
SRA	*Salix-Rubus-Aulacomnium*	24	6
SC	*Salix-Carex*	3	6

1-low, 2-medium or 3-high, and means for each type were found.

Distance matrices. The vegetational and environmental information was transformed into the Euclidean distance matrices SPF, DR, ST, OR and ER. There were three reasons for this transformation. First and foremost, distances are scalar descriptions of complex multidimensional relationships. For a large data set such as that for the vegetation in most surveys, the matrix SPF is a succinct representation of the information. Second, the number of descriptors varied among the vegetation and four environmental sets. The vegetational data had scores for the twenty three types on one hundred one species. Two of the environmental variables, drainage and erosion potential, had scores for the types on one descriptor while organic content and thickness was scored by two descriptors and soil texture by four. By calculating distances (separately for each of the five data sets) among the types, the five data sets could be matched up for regression purposes. Third, the vegetational raw data had one hundred one descriptors (species) of the twenty three types. However, a parsimonious representation with respect to the types requires a maximum of twenty two descriptors or hyperdimensional axes; the maximum being reduced if any species are correlated. Again a distance matrix was useful for this. The data matrix for species frequencies, composed of one hundred one species by twenty three types, was transformed into a Euclidean distance matrix. A preliminary geometric model was derived and so an ordination (cf. program PCAD; Orlóci 1978) was done. The first three axes accounted for 58.3% of the variation. The scores from these three axes were transformed into the distance matrix SPF which was subsequently used in the regression model. This represents one of the convolutions. Although the original distances may have been used instead of distance matrix SPF, this latter matrix is expected to contain less random variation and therefore be more amenable to revealing trends. Three of the four derived environmental variables were independently transformed into the distance matrices DR (drainage), OR (organic content and thickness) and ER (erosion potential) and then scaled to a (0,1) range. This scaling was done to make them commensurable with SPF which had been scaled via ordination $(1/\sqrt{(N-1)}; N=23)$. The raw data for soil texture was transformed into a distance matrix and then ordinated. This produced 99.9% of the variation on the first three principle axes. The three sets of scores were then transformed into distance matrix ST without scaling as the ordination had already done so. Because almost all of the information was retained by the first three axes, the effect of this convolution was insignificant.

Table 2. List of the one hundred one characteristic species used in the analysis. Nomenclature was af follows: for the vascular plants — Hultén (1968) and Welsh (1974); for the mosses - Crum *et al.* (1973); and for the lichens - Hale & Culberson (1970).

Agropyron yukonense	*Populus balsamifera*
Alnus crispa	*Populus tremuloides*
Anemone multifida	*Potentilla fruiticosa*
Antennaria rosea-nitida	*Potentilla hookeriana*
Arctostaphylos rubra	*Potentilla pensylvanica*
Arctostaphylos uva-ursi	*Pulsatilla patens*
Artemisia frigida	*Pyrola secunda*
Betula glandulifera	*Rhododendron lapponicum*
Betulla nana	*Rosa acicularis*
Calamagrostis canadensis	*Rubus arcticus*
Calamagrostis purpurascens	*Rubus chamaemorus*
Carex aquatilis	*Salix alaxensis*
Carex concinna	*Salix glauca*
Carex filifolia	*Salix myrtillifolia*
Carex lasiocarpa	*Salix planifolia*
Chamaerhodos erecta	*Saxifraga tricuspidata*
Cornus canadensis	*Shepherdia canadensis*
Crepis elegans	*Vaccinium uliginosum*
Dryas drummondii	*Vaccinium vitus-idaea*
Dryas integrifolia	*Viburnum edule*
Empetrum nigrum	*Abietinella abietina*
Epilobium angustifolium	*Aulacomnium palustre*
Epilobium latifolium	*Barbilophozia hatcheri*
Equisetum palustre	*Caloplaca cirrochroa*
Equisetum scirpoides	*Ceratodon purpureus*
Erigeron caespitosus	*Cetraria nivalis*
Erigeron compositus	*Cetraria pinastri*
Erigeron purpuratus	*Cladonia arbuscula*
Eriophorum vaginatum	*Cladonia chlorophaea*
Erysimum inconspicuum	*Cladonia ecmocyna*
Festuca altaica	*Cladonia gracilis-dilatata*
Festuca brachyphylla	*Cladonia gracilis-elongata*
Gentiana propinqua	*Cladonia rangiferina*
Geocaulon lividum	*Dicranum undulatum*
Hedysarum alpinum-americanum	*Drepanocladus uncinatus*
Ledum groenlandicum	*Hylocomium splendens*
Ledum palustre	*Hypnum procerrimum*
Linnaea borealis	*Hypogymnia physodes*
Linum perenne	*Lecidea rubiformis*
Lupinus arcticus	*Nephroma arcticum*
Lycopodium annotinum	*Peltigera aphthosa*
Mertensia paniculata	*Peltigera canina*
Mitella nuda	*Pleurozium schreberi*
Oxycoccus microcarpus	*Polytrichum piliferum*
Oxytropis campestris	*Ptilium crista-castrensis*
Pedicularis labradorica	*Sphagum magellanicum*
Penstemon procerus	*Sphagnum sp.*
Picea glauca	*Stereocaulon tomentosum*
Picea mariana	*Thudium abietinum*
Pinus contorta	*Tortula ruralis*
Polemonium pulcherrimum	

Trying regression models. A number of regression models were examined. The number of environmental variables to be used was limited to reduce complexity. This was one reason for choosing the four derived variables. Other reasons included high correlation via an analysis of concentration (Feoli and Orlóci 1979) by Orlóci and Stanek (1979) and a high r as described below. Thus a multiple regression model was settled upon. The dependent variable used was species frequency (SPF), symbolized by Y. The four simple factor variables describing the environment were **X1** - DR, **X2** - ST, **X3** - OR and **X4** - ER. Non-linear and interactive relationships were derived to give other complex factor variables. The particular multiple regression model which gave the highest correlation coefficient without an overwhelming number of variables produced an r of 0.715. A total of fourteen variables were used so that **Y** was a function of **X1, X2, X3, X4,** $\mathbf{X1}^{1/2}$, $\mathbf{X2}^{1/2}$, $\mathbf{X3}^{1/2}$, $\mathbf{X4}^{1/2}$, $\mathbf{X1}^{1/3}$, $\mathbf{X2}^{1/3}$, $\mathbf{X3}^{1/3}$, $\mathbf{X4}^{1/3}$, and $(\mathbf{X1 \cdot X2 \cdot X3 \cdot X4})^{1/4}$). Because a number ranging between zero and one decreases in magnitude if its exponent is greater than one and increases in magnitude if its exponent is less than one, exponents less than one were chosen in order not to make very small distances (large similarities) insignificant in the regression. Note that these exponents refer to the individual elements and not the entire matrix. As well, the final term has been formed via element by element multiplication and not by the usual matrix multiplication.

Perturbations

Assuming that species frequencies are a function of the environment and that a regression model has been constructed encompassing a wide variation of vegetation/environment relations found in the area of study, it is not impossible to have the model predict the eventual vegetation configuration as the result of any environmental configuration which is within its bounds. The point of interest here is in probing the data to discover trends of change as a result of perturbation. Because the model variables are distance variables, it was decided for simplicity, to simulate perturbation by proportional changes in the environmental distance matrices, although one night have chosen to perturb the raw environmental scores and then calculate new environmental distances. For example, a perturbation of magnitude p implies increasing the distances by the proportion p or rather multiplying all distances by the factor $1+p$. This implicitly assumes that any perturbation, irrespective of its magnitude, affects each pair of vegetation types (as described by the environment) in proportion to their distances.

In referring to Table 3, specifically the four sets of environment scores DR, ST, OR and ER, it is seen that there are fifteen different combinations of perturbations possible. One may choose to perturb any one of the four, any two of the four, any three, or all four. Since for each choice much work will be done, it would be sensible not to try them all. Not only would there be redundancy due to the interrelationships among the four sets of scores, but the bulk of the resulting graphs and descriptions would confuse the interpretation and detract from the method being developed. How many of the fifteen should we choose? As mentioned, the environmental variables are related so that in choosing one, in a sense, we choose the others. Which one of the fifteen shall we then choose? The literature is helpful here. Wilde (1958) speaks of the importance of soil moisture. Loucks (1962) concurs, saying that soil moisture is a «major influence on the vegetation». Behind much of the work of Whittaker (1967) and Whittaker and Niering (1965) is a «topographic moisture gradient». Knight (1965) describes the importance of soil moisture in influencing other abiotic factors through the fact that water is a

Table 3. Description of the development of the model. Data describing the twenty three vegetation types is separated into vegetation and environment components. Next, the various components are coded as matrices and then as dependent variables. Finally, a functional relationship is formed.

Vegetation types				
Vegetation scores	Environment scores			
species frequencies	drainage	soil texture	organic content and thickness	erosion potential
SPF	DR	ST	OR	ER
Y	**X1**	**X2**	**X3**	**X4**
	X1	**X2**	**X3**	**X4**
Y	$\mathbf{X1}^{1/2}$	$\mathbf{X2}^{1/2}$	$\mathbf{X3}^{1/2}$	$\mathbf{X4}^{1/2}$
	$\mathbf{X1}^{1/3}$	$\mathbf{X2}^{1/3}$	$\mathbf{X3}^{1/3}$	$\mathbf{X4}^{1/3}$
	$(\mathbf{X1\text{-}X2\text{-}X3\text{-}X4})^{1/4}$			

«nearly perfect solvent. It dissolves many substances that would not normally be available to organisms in a solid state». These observations support the choice of drainage (DR, which is variable **X1**) as the candidate for perturbation. Due to the non-linear and interactive regression terms involved, the variables affected are $\mathbf{X1}^{1/2}$, $\mathbf{X1}^{1/3}$ and $(\mathbf{X1 \cdot X2 \cdot X3 \cdot X4})^{1/4}$ (Table 3).

The meaning of a perturbation here is not straight forward because a distance matrix, which we use, is one step removed from the raw data based on the variables. A proportional increase of 10% (p=0.1) for example in the raw data would necessarily have a specific meaning. It would increase a drainage value of say 7 (wet) to a value of 7.7 (wetter). For the moment, consider the types (described only by drainage) as plots occuring in nature. Then, increasing their distances with respect to the distance matrices involvings **X1** will certainly magnify their compositional differences. But in what way? This requires careful examination. A 10% increase of the distances with respect to **X1** does not mean that all the drainage measurements would be increased by 10%. In fact, because any number of raw data sets could produce the same distance matrix, a perturbation in the distance matrix represent an entire family of possible perturbations in the raw data set. Within a particular family, one might consider the possibility of the wet types as getting wetter, the dry types getting drier, and the mesic types hardly changing. As well, one might consider all types to be changing in the same way; for example, all becoming more dry, but with the dryer types changing more than the wetter types. However, it is easiest to devise an approach to interpretation along the lines of sensitivity to perturbation (drainage in this case) and not be overly concerned with the precise meaning with respect to the raw drainage scores.

The strategy to uncovering trends (sensitivity to perturbation) was to gradually increase the amount of perturbation, calculating the new environmental matrices and subsequently the predicted vegetational configurations at each step, and observing the movements of points representing the types in model space. The unperturbed environmental state predicted the vegetational configuration $\hat{\mathbf{Y}}$. This $\hat{\mathbf{Y}}$ was used as a base for making comparisons instead of the raw species matrix **Y**. The various perturbed states predicted various $\hat{\mathbf{Y}}p$. The perturbations, expressed as proportional increases in distances, were p=0.01, 0.049, 0.25, 0.28, 0.5, 0.8, 1.2 and 1.8. A sum of squares clustering was used to create dendrograms for a visual inspection for the type of change resulting from an increasing amount of perturbation.

Graphical Analysis

Ordination

The objective of the analysis is to reveal information about trends in compositional change of the types under perturbation. An ordination technique (Principal Axes (Coordinates) Analysis) will extract these trends from distance matrices while also retaining the original relationships (groups) as they appear in the $\hat{\mathbf{Y}}$ data. It is relevant here to consider two classes of ordination algorithms. According to Gower (1966) one normally considers principal components analysis (PCA), based on a dispersion or correlation matrix, to be an R-algorithm. Orlóci (1966) and Gower (1966) have shown that a Q-algorithm using distances, which they call principal axes analysis and principal coordinates analysis (PCoA) respectively, is a dual to the R-algorithm. The computer program used here is a Q-algorithm which transforms Euclidean distances via q_{jk} as described by Orlóci (1978, p. 115). The program is based on a parsimonious method and so graphing the first two or three sets of scores will retain a large proportion of the variation.

The distance matrices which produced the dendrograms are analysed separately. Then the results are combined graphically in such a way as to link the successive perturbations for each type. As each data set has been ordinated separately, consideration must be given to its interpretability.

Scatter plots of temporal information

Ordination techniques have been used by others to study vegetation change. The graphical presentation is a set of curves or trajectories which join the points representing the same species or plots (depending on how the analysis is performed) for the different time periods. Goff and Zedler (1972) constructed species succesion vectors where size classes were used for the time element. Carleton and Maycock (1978) did the same using age classes for the time element and the different plots or sites for the dimensions. Of particular methodological relevence are those studies which constructed plot or site vectors of data collected at various sites (q) at different times (t) for various species (s) or descriptors. The raw three-dimensional matrix ($q{\cdot}t{\cdot}s$) was reordered to give a $qt{\cdot}s$ matrix and then ordinated. One of the first studies like this was van der Maarel's (1969) work in phytosociology. Other studies applying this aproach include those by Austin (1977), Bowles (1980) and Bowles and Maun (1982). Ordination of a $qt{\cdot}s$ matrix referred to above implies pooling t sets of order $q{\cdot}s$ prior to analysis. Alternatively one might ordinate the t sets independently and then combine the results. We may refer to the first of these cases as combining *a priori* and to the second as combining *a posteriori.*

Combining 'a priori' or 'a posteriori'

There are two aspects of the ordination method affecting the results which should be considered. They are data centering and multidimensional axes rotation. Data sets which are independently analyzed will be centered and rotated differently. Hence, the results of combining *a priori* will differ from those due to combining *a posteriori.* When does one prefer to combine *a priori* or *a posteriori*? Ordinating a single combined data set is simpler and not surprisingly, is what has normally been done. On the other hand, combining *a posteriori* allows greater control in isolating the components of variation from different sources. Nevertheless, to answer the question of how different the results will be between *a priori* and *a posteriori* combination, three points need to be considered: (1) The trajectories from *a posteriori* combination might look like random walks on first inspection, but upon closer examination trends may appear as the result of sharply different structure within each data set. (2) Axis orientation is usually arbitrary. A PCA type ordination arbitrarily determines the positive/negative orientation of each axis. For two data sets, the ith axis of one may be a mirror image of the other. If indeed the results of the t ordinations are to be compared, this axis flipping must be corrected. (3) A trajectory from the *a priori* approach may be very different from the *a posteriori* approach. This sort of differentiation in trajectories is common when the data structures are not sharp in the sets. Strong structure is a factor in minimizing this. As well, one can show that as the number of trajectoris (t) increases, the possibility of trajectory differentiation drops dramatically. It is noted that valuable information can be expected from inspections of dendrograms, constructed from the results of a clustering algorithm, which can be of assistance when interpreting the observed trajectories.

Results

Type trajectories

A stereogram (Fewster and Orlóci 1978) of type trajectories (up to a maximum perturbation of 0.28) is given in Figure 1. The first three principal axes, represented by the X, Y and Z axes respectively, accounted for at least 77% of the variation. Note that scaling of the ordination results for each data set must be consistent. The scattergram in

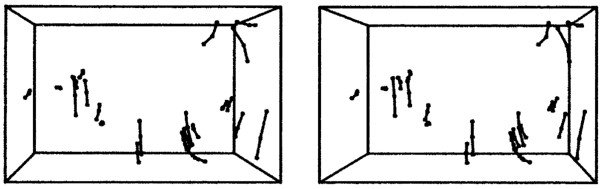

Fig. 1. Stereogram of the trajectories of vegetation types based on a series of increasing perturbations. The trajectories consist of the null state (no perturbation) plus perturbations of 0.049 and 0.28. See Fig. 2 for an orientation of the trajectories.

Figure 2 gives the labels for the types and illustrates their trajectories for the first two principal axes (at least 65% of the variation). This type of graphical presentation gives an overall picture of how the types differentiate along the axes from one degree of perturbation to the next. In order to compare magnitude and direction, the trajectories must be translated so that the tails of earch trajectory (the unperturbed vegetation configuration $\hat{\mathbf{Y}}$) share a common position. A graph (spidergram) of the translated trajectories in the plane of the first two axes is given in Figure 3. (Here the maximum perturbation is 1.8). The first two principal axes accounted for at least 57% of the variation.

Dendrograms and scattergrams

The essential difference between a series of dendrograms from cluster analysis and scattergrams from a combination of ordinations is the way in which they detect or correct for an overall trend as the amount of peturbation is increased. In general, the cluster analyses give no indication as to overall change whereas the ordinations do (cf. Fig. 1).

A protocol for interpreting the trajectories in Figs. 1, 2 and 3 is as follows.

Interpretation

The *first principle* is adapted from Bowles (1980) and is a relative one. The argument used is that if the majority of types show parallel or similar trajectories from one perturbation to the next, then a general trend of type response to perturbation is evident. If the trajectories are otherwise then there could be a) no trend or b) different type-dependent trends. The implication of a) is that the types have no directional response to perturbation or there exists a cyclical response. The implication of b) is that either all trends are different or that there are groups within which there exists parallel trajectories.

A *second principle* developed here is concerned with type susceptibility. This is an absolute argument based on specific assumptions made by considering Figure 1. It is clear that the ordinations produced a horseshoelike configuration for the types. This is to be expected (Fewster and Orlóci 1984) under the circumstances, since the sample environmental gradient is long. Beginning at the left are the wet types (Table 1). Following the horseshoe trend, progressions is through mesic types ending with dry types in the upper right. Based on the analysis of Orlóci and Stanek (1979, Fig. 18, p. 41), the ordinations appear to have detected a moisture gradient. With respect to the categories in Table 1, a mesic grouping of categories 3, 4 and 5 (types 9, 19, 8, 16, 21, 17, 1, 7, 18 and 5) can be recognized whose trajectories are generally from top to bottom with a slight left to right shift. This combination of category 5 (types 18 & 5) with categories 3 & 4 is supported by the dendrograms (Fewster 1985). A hydric group (category 6; types 15, 6, 4, 20, 24 and 3) can be seen to be on a path directed, in some cases right to left, and in others from bottom to top. A xeric grouping of categories 1 and 2 (types 11, 12, 23, 14, 33, 22 and 2) is moving essentially upwards with the exception of type 11. Again the dendrograms support this combination. It makes sense that any vegetation type which is affected by perturbation would move along the horseshoe rather than going directly from one extreme to another. For this reason the overall trajectory of the mesic grouping quite possibly indicates resistance to change at low perturbation levels. This suggests an assumption to be used for the second principle of the protocol. It appears that types 15, 6, 4, 20, 24 and 3 (very wet category) are susceptible to change whereas types 5 and 18 (wet category) are resisting change more than might be expected. Types 14, 13, 22 and 2 (moderate to dry category) are affected by increasing perturbation in the same way as types 12 and 23 of the dry category. Apparently type 11 may be resisting change. Overall, most of the mesic types are resisting change whereas the hydric and xeric types are not. For the amount of perturbation used in Figure 1, types 11, 18, 5 and possibly 14, 13, 22 and 2 are not following the horseshoe and hence are behaving differently from what is expected.

The analysis continues with a *third principle* which calls

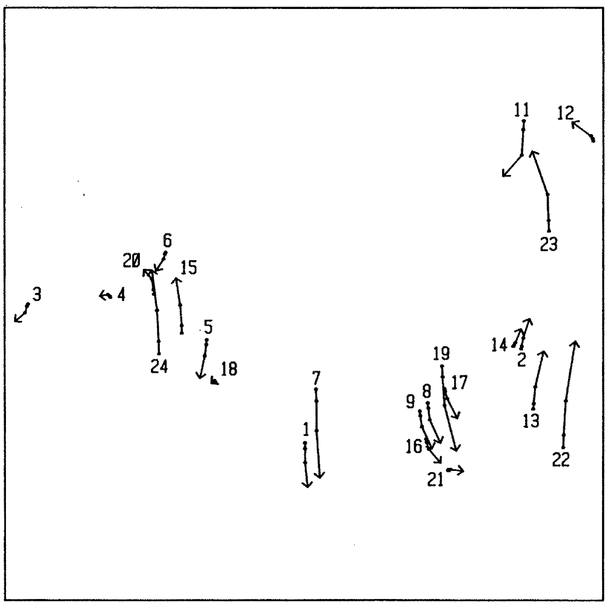

Fig. 2. A two dimensional scattergram of trajectories, similar to Fig. 1, with a maximum perturbation of 0.28. Arrows represent direction of increasing perturbation and labels refer to the types as given in Table 1. Note the direction of types 4 and 18 is from right to left.

for the translation of each trajectory whereby their tails share a common point. Note the amount of perturbation has been increased to 1.8. Figure 3 illustrates that categories 1 and 2 respons similarly although the dry category is more susceptible to change. Type 11 behaves differently than 12 and 23 at first, that is at low perturbations, but as perturbation is increased, it begins to conform to the others. Figure 3 illustrates very well how type 18 behaves differently at first but later on conforms to the others. Figure 3 is very interesting in that what can be interpreted here is not apparent from just the dendrograms (Fewster 1985). Types 15 and 24 are not classified as a pair (lowest fusion level) yet their translated trajectories suggest they behave in a similar manner to perturbation. The same holds for type pair 6 and 3 and type pair 4 and 20. Further, at low perturbations the response of types 4 and 20 is similar to 15 and 24 but at higher levels their response is more like the pair 6 and 3.

Discussion

It is interesting in ecological investigations to consider what the general tendencies of the system might be, following perturbation, even if time is undefined. One might ask what should be the eventual state of the system following perturbation without regard to time? This question leads to the regression approach to modelling based on the assumption that the vegetation is a function of the environment. The value of choosing a regression model is that it can absorb unlimited variation in the data. Here, regression has been used informally in seeking indications for purposes of exploring system relationships (Goodall 1967; also see Tukey 1962, Tukey and Wilk 1966).

It should be realized that when regression predicts the vegetation resultant, no restriction is placed on time. However, the resulting vegetation state after perturbation may be resolved quickly or it may take a long time to develop as in the case where mature vegetation is destroyed and the new is forced to grow back. Hence it is possible that as the vegetation changes, environmental variables could be affected. As well, the vegetation change could involve feedback from the environmental variables.

As explained, distance matrices, rather than the raw data, were used in the regression model. Consequently, an interpretation of the specific effect of any particular perturbation level on the raw data was not possible. Only general trends could be identified. This, plus the fact that only one, although apparently the most influential of the four environmental variables were perturbed, limited the extent to which the results could be interpreted ecologically. Nevertheless, dynamics of various community types emerged. Type 3, with the nomenclature *Salix-Carex* in Table 1, and type 6, *Salix*, are two of the wettest types. Their trajectories (Figs. 2 and 3) indicate that if the 'jolt' due to perturbation is sufficiently high, they will tend towards the mesic condition and allow for the establishment of certain higher arboreal species (*Picea glauca* and *P. mariana* for example). The same could be said for the driest types (11, 12, 23) with type 11 requiring less of a

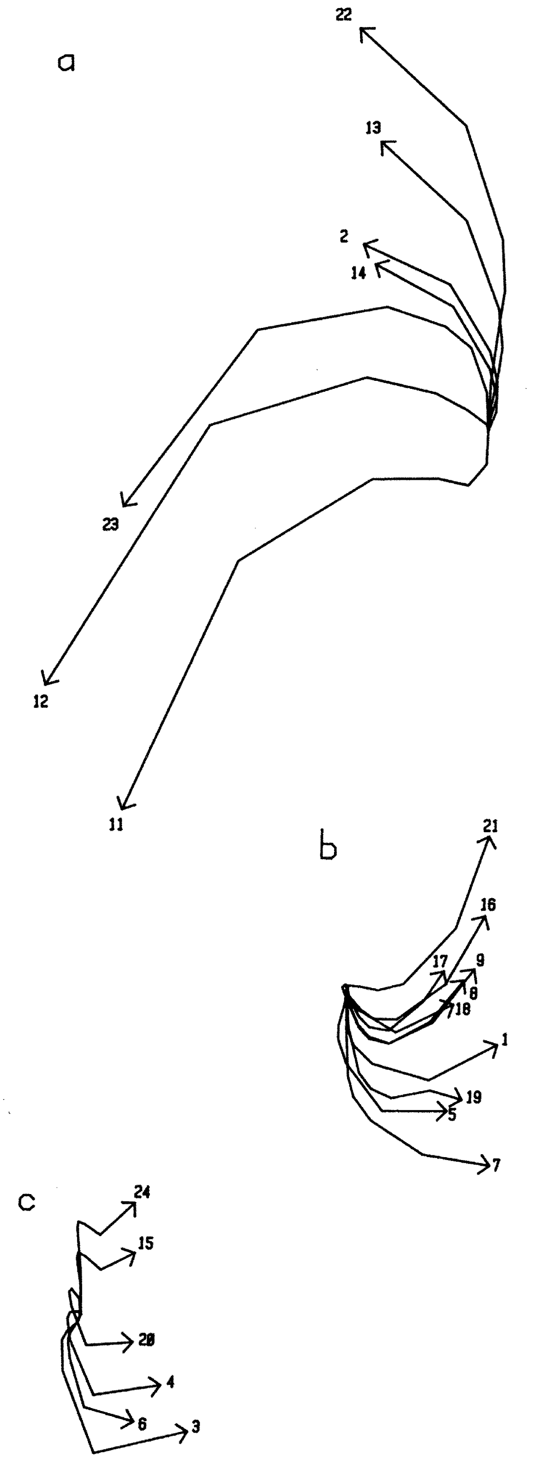

Fig. 3. Translated trajectories of types with a maximum perturbation of 1.8. Subsequent to the ordinations, the trajectories, as in Figs. 1 and 2, have been translated so that the null states ($\hat{Y}$) of each type share a common point allowing for comparisons based on magnitude and direction: a - categories 1 and 2; b - categories 3, 4 and 5; c - category 6.

'jolt' than the other two. By contrast, the trajectories of types 14, 13, 22 and 2 appear to indicate that with increasing perturbation, conditions become less favourable for the arboreal species.

Many (Carleton and Maycock 1978; and others) describe community dynamics in (boreal) areas similar to the Yukon Territory, as not being Clementsian (mono-climax), preferring to suggest the dependance on disturbance and catastrophe for the rejuvenation or perpetuation of certain vegetation types. In general, the normal trends following disturbance, or even in the absence of disturbance, have not been described in detail because little work spanning a long time period has been done. Some have been able to make suggestions about vegetation change. However, their descriptions have been based on very broad vegetation types and not on types which are as refined as those used here (see Table 1). From the description of the results above, we have noted the establishment, or possibly the reestablishment, of certain arboreal species as a consequence of disturbance. The method used here (regression modelling) can help in answering questions such as how large would the disturbance have to be before this reestablishment occurs? In referring to Figs. 2 and 3, we would expect to see more or less circular trajectories. Hence, we would be interested in seeing how large these circular trajectories would be.

Acknowledgement. This work contributes to the ecological survey of the right-of-way of the proposed gas pipeline in the southern Yukon and represents part of a Ph.D. thesis. The author wishes to thank Dr. L. Orlóci for invaluable discussion during the tenure of the project. Financial support from NSERC (to Dr. Orlóci) is appreciated. Acknowledgement is extended to Drs. D. Lausi and P.L. Nimis of the University of Trieste, and W. Stanek of the Department of the Environment, Canadian Forestry Service, for advice, discussions, and data.

REFERENCES

AUSTIN, M.P. 1977. Use of ordination and other multivariate descriptive methods to study succession. Vegetatio 35: 165-175.

BOWLES, J.M. 1980. *Effects of human disturbance on the sand dunes at Pinery Provincial Park.* Ph. D. thesis. University of Western Ontario, London, Canada.

BOWLES, J.M. and M.A. MAUN. 1982. A study of the effects of trampling on the vegetation of Lake Huron sand dunes at Pinery Provincial Park. Biol. Conserv. 24: 273-283.

CARLETON, T.J. and P.F. MAYCOCK. 1978. Dynamics of the boreal forest south of James Bay. Can. J. Bot 56: 1157-1173.

CRUM, H.A., W.C. STEERE and L.E. ANDERSON. 1973. A new list of mosses of North America north of Mexico. The Bryologist 76: 85-130.

FEOLI, E. and L. ORLÓCI. 1979. Analysis of concentration and detection of underlying factors in structured tables. Vegetatio 40:49-54.

FEWSTER, P.H. 1985. *Methodologies of predicting vegetation response to the environment and to perturbation on a regional level: a thesis in three parts treating temporal and time static models.* Ph. D. thesis. University of Western Ontario, London, Canada.

FEWSTER, P. and L. ORLÓCI. 1978. Stereograms to aid group recognition and trend identification in vegetation data. Can. J. Bot. 56: 162-165.

FEWSTER, P.H. and L. ORLÓCI, 1984. On choosing a resemblance measure for non-linear predictive ordination. Vegetatio 54: 27-35.

FOOTHILLS. 1978. *Terrain Evaluation Maps (YK-04-0200-D series). Explanation of Data Categories.* Foothills Ltd., Edmonton, Alberta.

FREUND, R.J. and P.D. MINTON. 1979. *Regression Methods: A Tool for Data Analysis.* M. Dekker Inc., New York.

GOFF, F.G. and P.H. ZEDLER. 1972. Derivation of species succession vectors. Amer. Midl. Nat. 87: 397-412.

GOODALL, D. W. 1967. Computer simulation of changes in vegetation subject to grazing. J. Indian Bot. Soc. 46: 356-362.

GOODALL, D.W. 1972. Building and testing ecosystem models. In: J.N.R. JEFFERS, (ed.). *Symposium of British Ecological Society on Mathematical Models in Ecology.* pg. 173-194. Grange-over-Sands, March 1971.

GOWER, J.C. 1966. Some distance properties of latent root and vector methods used in multivariate analysis. Biometrika 53-325-338.

HALE, M.E. and C.L. CULBERSON. 1970. A fourth checklist of the lichens of the continental United States and Canada. The Bryologist 73: 499-543.

HULTÉN, E. 1968. *Flora of Alaska and Neighbouring Territories.* Stanford Univ. Press. Stanford, California.

KNIGHT, C.B. 1965. *Basic Concepts of Ecology.* Collier-Macmillan Canada Ltd., Toronto.

LOUCKS, O.L. 1962. Ordinating forest communities by means of environmental scalars and phytosociological indices. Ecol. Monogr. 32: 137-166.

MAAREL, E. van der. 1969. On the use of ordination models in phytosociology. Vegetatio: 19: 21-46.

ORLÓCI, L. 1966. Geometric models in Ecology. I. The theory and application of some ordination methods. J. Ecol. 54: 193-215.

ORLÓCI, L. 1978. *Multivariate Analysis in Vegetation Research.* Junk, The Hague.

ORLÓCI, L. and W. STANEK. 1979. Vegetation survey of the Alaska Highway, Yukon Territory: Types and gradients. Vegetatio 41: 1-56.

OSWALD, E.T. and J.P. SENYK. 1977. *Ecoregions of the Yukon Territory.* Canada Department of the Environment, Canadian Forestry Service, Pacific Forest Research Centre. Victoria, British Columbia.

TUKEY, J.W. 1962. The future of data analysis. Ann. Math. Statist. 33: 1-67.

TUKEY, J.W. and M.B. WILK, 1966. Data analysis and statistics: an expository overview. AFIPS Conf. Proc. Fall Joint Comput. Conf. 29: 695-709.

WELSH, S.L. 1974. *Anderson's Flora of Alaska and Adjacent Parts of Canada.* Brigham Young University Press, Provo, Utah.

WHITTAKER, R.H. 1967. Gradient analysis of vegetation. Biol. Rev. 42: 207-264.

WHITTAKER, R.H. and W.A. NIERING, 1965. Vegetation of the Santa Catalina Mountains, Arizona: a gradient analysis of the south slope. Ecology 46: 429-452.

WILDE, S.A. 1958. *Forest Soils.* The Ronald Press Co., New York.

Manuscript received: May 1986