WILEY

International Association for Vegetation Science

On the statistical analysis of vegetation change: a wetland affected by water extraction and soil acidification

Introduction

The causes of vegetation changes are often to be evaluated without precise data on the changes in the environment. In many permanent quadrat studies, there are detailed records of past and present species compositions, limited information on the present environment, but no information on the past environment. We present a case study to show which statistical analyses may assist in interpreting vegetation change. In this study, we use data from only two sampling dates but the analyses can be generalized to longer series.

Lowering of the phreatic water table, due to water extraction and drainage, is a common phenomenon in western Europe. It is one of the major threats to the survival of rare plant species especially in wet, oligotrophic and mesotrophic grasslands (Westhoff 1979; Mountford & Sheail 1989; Sukopp & Trautmann 1981). Agricultural land use, including fertilization, is usually intensified at the same time. Well-documented studies on the effects on species composition of groundwater lowering alone are scarce (Ellenberg 1952; Grootjans & ten Klooster 1980; Grootjans et al. 1986).

Lowering of the water table can cause a shortage of water for certain species. It can also affect the storage and flow pattern of groundwater. Mineral-rich groundwater in the topsoil can be replaced by more acid rainwater (van Wirdum 1979, 1981; Grootjans et al. 1991). In wet soils, it also may often cause changes in physico-chemical processes such as aeration and related redox-processes. Enhanced mineralization, leaching and ultimately acidification may result (van Wirdum 1982; Grootjans et al. 1986; Kemmers 1986). Soil acidification can also be induced by air pollution with NH_x, NO_x and SO_x (Schachtschabel et al. 1989, p.115).

The present study was undertaken in order to evaluate the changes during the last decade in the vegetation of a *Junco-Molinion* grassland influenced by groundwater extraction and by acid deposition. Ideally, we would like to identify which part of the vegetational

change is due to water extraction and which part to acidification. However, for an observational study in which the relevant variables are correlated, this objective cannot be reached. Therefore, we aim more modestly at assessing their joint effect on the vegetation by evaluating the changes in plant species composition on the basis of relevés of 20 plots in both 1977 and 1988. We attempt to interpret the changes in terms of changes in groundwater table and acidity by using Ellenberg's (1979) indicator values for the participating species and by exploiting data on the present (1988) spatial variation in water depth and acidity across the reserve. Univariate analyses of each species' abundance separately and of the mean Ellenberg indicator values are supplemented with multivariate analyses, which serve five purposes: (1) to yield a single, overall statistical test of the significance of change in species composition between 1977 and 1988; (2) to test whether the change was constant across the site against the alternative hypothesis that the change depends on water depth and to display this dependence if present; (3) to display jointly the spatial and temporal variation in the species data; (4) to interpret this variation in terms of changes in water depth and pH and (5) to decompose the total variation into components as in an analysis of variance (Borcard et al. 1992). We take the opportunity to explain carefully the multivariate approach and to compare it with the univariate one.

Material and Methods

Study area

The study area is located on the Pleistocene cover sands in the eastern part of the Netherlands (52° 0' N, 6° 31' E). It is a moist oligotrophic grassland of 1 ha and it is part of a small nature reserve of 6 ha, surrounded by farmland. According to aerial photographs it has been hayfields at least since 1934. In 1964 it became a nature reserve and since then it has never been fertilized.

Both & van Wirdum (1981) described the history of the area with emphasis on the hydrology and its effect on the vegetation since the end of the last century. In about 1870 the site of the reserve was in a transition zone between vast, extensively grazed heathlands and small meadows along brooklets. Between 1870 and 1930 the heathlands were reclaimed into meadows. Drainage and fertilizing in adjacent areas have intensified since then, with a peak between 1950-1975. Acid deposition has substantially increased during the last few decades especially by NH_x-losses from excessive manuring up to more than 5200 mol H^+/ha/year (Buijsman et al. 1990).

Pumping wells for drinking water supply are situated to the south-west of the area, in the first, sandy aquifer of ca. 40 m depth at an average distance of 450 m from our study area. Groundwater extraction started in 1965, and was extended to 1.5×10^6 m^3/yr from 1972 onwards. As a result the water table dropped gradually until 1972 and more drastically during 1972.

The lowering of the water table is calculated with simulation models (Anon. 1984; Pastoors 1985). For the mean lowest water table in summer, the reductions in level between 1950 and 1977 were ca. 0.45 and a further 0.27 m between 1977 and 1988. There is no evidence that the drop in level varied across the study area. In the northwest to southeast direction there is a difference of about 0.35 m due to topographical relief. For water-quality measurements see Farjon & Wiertz (1989).

The soil is classified as 'beekaarde' (de Bakker & Schelling 1966). It is a non-podzolised sandy soil with an upper 20 cm of loamy sand. The pH-H_2O in the wettest part of the grassland is 5.7, loss-on-ignition is 9.9 %, and the C/N ratio is 10 (0 - 5 cm in 1977; for more details, see Bannink & Pape 1979).

The vegetation in 1953 still had distinct features of a *Cirsio-Molinietum* community with some characteristics of a *Calthion* community. *Parnassia palustris* was quite abundant (M.F. Mörzer Bruyns unpubl.). Dominant species in 1988 were *Agrostis* spp., *Anthoxanthum odoratum*, *Molinia caerulea* and *Plantago lanceolata*. Frequent species in the drier part include *Danthonia decumbens*, *Carex panicea* and *Salix repens*. These species, together with *Polygala vulgaris*, *Hypericum maculatum* and *Erica tetralix* are indicative of the alliance *Violion caninae*. In the lower part *Festuca rubra* and *Holcus lanatus* are frequent and species more typical of an *Arrhenatheretum* occur, including *Dactylis glomerata*, *Festuca pratensis*, *Ranunculus acris*, *Cynosurus cristatus* and *Phleum pratense*. In 1988, species which are now rare in the Netherlands still occurred: *Briza media*, *Cirsium dissectum*, *Dactylorhiza maculata*, *Gentiana pneumonanthe* and *Platanthera bifolia*.

Sampling of vegetation and soil

In 1977, 20 plots of 20 m × 20 m were laid out in a regular 4 × 5 grid (Fig. 1). The species composition of the plots was described in June/July 1977 (Both & van Wirdum 1981) and 1988 (Farjon & Wiertz 1988) using a simplified Tansley abundance scale: 0 = absent, 1 = rare, 2 = occasional or frequent, 3 = abundant (Tansley & Chip 1926). For each plot, 16 subsamples of the topsoil (0 - 15 cm) were taken and mixed on 28 July 1988. The pH-KCl was measured 16 h after shaking a mixed sample with 1 M KCl (1:5).

The water depth was derived from 10 piezometers
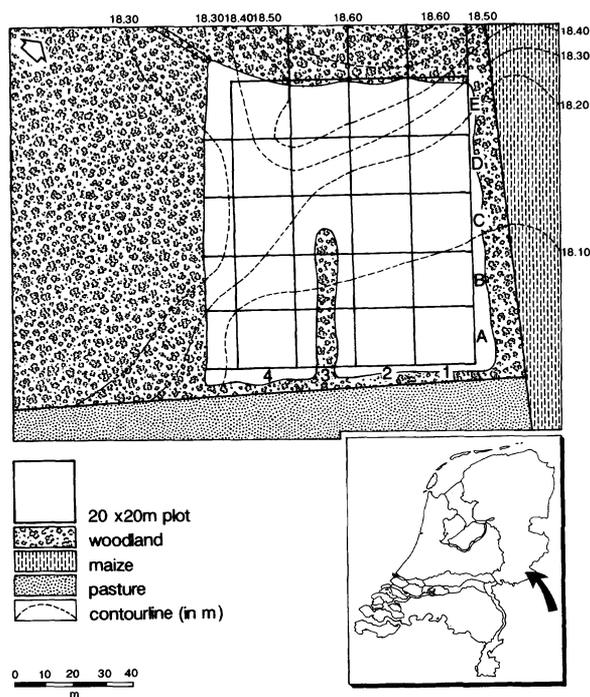
Fig. 1. The layout of the plots in the study area. The inset shows the location of the study area in the Netherlands.

measured every fortnight from 28 May 1988 - 12 May 1989 (Farjon & Wiertz 1989). The depth and filter length of the piezometers was 2.75 m and 1 m, respectively. In the statistical analyses, the mean summer water depth was used which was derived in three steps. First, the fortnightly observations, in the period 28 May - 28 September 1988, of a reference piezometer were averaged. Second, the spatial pattern of the water table was examined using all piezometers. The pattern largely consisted of an average fall in the water table of 0.12 m across the field in the direction of the pumping station. Third, the mean summer depth in 1988 (from now onwards called water depth in 1988) was calculated for each plot by adding to the value of the reference piezometer, the contributions (plus or minus) of the relief and the fall in the water table in the direction of the pumping station. Note that the spatial variation of the water depth counts in the statistical analysis; the reference value has no effect.

The mean pH-KCl was 3.58 (range from 3.89 to 4.28), the mean water depth in 1988 was 1.30 m (range from 1.12 m to 1.55 m). pH and water depth were highly correlated ($r = -0.86$).

Data analysis

A simple univariate statistical approach is to analyze the data for each species separately. In our case, with two recordings for each plot, we assessed the statistical significance of an increase or decrease of each species' abundance by a Wilcoxon signed rank test (Sokal & Rohlf 1981) with deletion of the plots showing no change. This test is not influenced by spatial auto-correlation in the abundance – because it compares the values from 1988 and 1977 within plots – but it is influenced by spatial autocorrelation in the change. The large-scale component thereof will be analyzed later on by ordination. We used a comparison-wise significance level, to avoid loss of power, and thereby accepted that we are likely to mark too many species as showing a significant change.

It may be difficult to interpret the changes on a species-by-species basis. It therefore makes sense to try to reduce the species data to a few variables that are informative on *a priori* grounds. We derived new variables by using Ellenberg's (1979) species indicator values (Böcker et al. 1983; Grootjans et al. 1991) for F (moisture), R (acidity) and N (nitrogen). We calculated the mean indicator values for F, R and N per plot on the basis of species presence/absence values. In cases where Ellenberg indicator values are missing, supplementary values from Wiertz (1992) were used. The changes in the mean indicator values were assessed by the Wilcoxon signed rank test, as was the change in species richness.

As to multivariate analysis, we chose ordination rather than cluster analysis because the variation was expected to be continuous. After a preliminary De-trended Correspondence Analysis on all 40 relevés yielded axes of short length (below 2 SD), we subsequently switched to linear ordination methods, Principal Components Analysis and its constrained form (partial) Redundancy Analysis (RDA, ter Braak & Prentice 1988) using the computer program CANOCO 3.1 (ter Braak 1990b). See further ter Braak (in press). We focus on the results of the redundancy analyses because these directly address the particular questions at hand and thus give more convincing answers than the corresponding indirect gradient analyses (Principal Components Analysis followed by an interpretation by regression analysis). We used the following nominal explanatory variables (with classes between brackets): year (1977, 1988), plot (1 - 20) and strip (A - E in Fig. 1 – roughly indicating increasing terrain height and hence water depth) and the cross-classification strip by year (A77, A88, ... , E88). These factors are coded as series of dummy variables (Jongman et al. 1987, pp. 58-59). These 37 dummy variables are not independent; they take 24 degrees of freedom. In addition, there are two quantitative variables,

pH-KCl of the soil (pH for short) and water depth of the plot in 1988. There are no measurements of pH and water depth for individual plots in 1977. Therefore, relevés from 1977 received the 1988 value of pH and water depth to enable their change to be modelled by the factor year as will be explained below. For each analysis we specify the model in terms of these variables. In CANOCO, explanatory variables can be made covariables so as to adjust for their effect on the species data (ter Braak & Prentice 1988). In the resulting partial ordinations, their effect is eliminated from the ordination. In the analyses that focus on the change with time, spatial differences are accounted for by specifying the plot dummy variables as covariables. Statistical tests of significance are carried out by Monte Carlo simulation (Manly 1990). Monte Carlo tests of significance are distribution-free; they do not require normality of the error distribution. However, they do require independence or exchange ability. In Monte Carlo tests, the permutations are therefore restricted to random interchanges within each plot of the 1977 and 1988 relevés. In the test of interaction effects – strip·year and water depth·year – it is essential that the residuals under the alternative model are permuted rather than the raw species data (ter Braak 1992). The test statistic in CANOCO 3.1 is the usual partial F-statistic, with this difference that the regression and residual sums of squares used to calculate it, are totalled across species. (It could be called a stacked univariate F.) For each test, 999 permutations are used. To display effect sizes, ordination diagrams are in distance scaling (ter Braak 1990b; ter Braak in press). Species scores are adjusted for species variance; this standardization is performed after the RDA is carried out (scaling 1 in CANOCO 3.1). In the diagrams, classes of nominal explanatory variables are indicated by centroids (average scores of the samples belonging to that class).

Extraction can be expected to change the vegetation of wet plots to that typical for drier plots, perhaps after some time-lag. This simple change model (without lag) can be written as:

$$y_k = a_k + b_k (x - \Delta.t) + \text{error} \qquad (1)$$

where y_k is the abundance of species k ($k = 1, ..., m$), x the measured water depth in 1988 and t a dummy variable indicating the relevés from 1977. The parameters of the model are the constant a_k and regression coefficient b_k for species k and Δ the change in water depth between 1977 and 1988. In App. 1 it is proved that the least-squares estimator of Δ is obtained from the first axis of a RDA with model $t + x$, by setting

$$\hat{\Delta} = - c_1 / c_2 \qquad (2)$$

where c_1 and c_2 are the (unstandardized) canonical coefficients of t and x of the first axis, respectively, if $c_2 \neq 0$. The estimate is unstable if the first axis is poorly related to either x or year, or if the first eigenvalue is not well separated from the second. The results of App. 1 can also be used to underpin the former usage of (2) by ter Braak (1987) and Cramer & Hytteborn (1987) in the context of canonical correspondence analysis. App. 1 also gives the extension of (1) to a p-dimensional change model.

A confidence interval for Δ was obtained by bootstrapping the residuals from (1) and estimating c_1 and c_2, and hence Δ, from each bootstrap data set in which each abundance value is the sum of its fitted values under model (1) and a bootstrapped residual. Let Δ_1 and Δ_2 be the 5%- and 95%-percentile of the bootstrap distribution of Δ, then the standard bootstrap 90%-confidence interval (Hall 1988) is $(2 \hat{\Delta} - \Delta_2, 2 \hat{\Delta} - \Delta_1)$. The bootstrap procedure was programmed in GENSTAT 5 (Anon. 1987).

Results

Floristic changes with time: univariate analysis

According to the Wilcoxon signed-rank test, 27 species decreased and only five increased in abundance (Table 1). Among the increasing species are common species of extensively used mesotrophic hayfields (*Plantago lanceolata*, *Anthoxanthum odoratum* and *Holcus lanatus*). Woody species such as *Alnus glutinosa*, *Frangula alnus*, *Rubus idaeus*, *Corylus avellana*, *Populus* spp. appeared for the first time in small quantities in the 1988 relevés, but only *Betula* spp. increased significantly. Decreasing species are mostly species from moist, mesotrophic and neutral hayfields which are not or only slightly manured, such as *Cardamine pratensis*, *Cirsium dissectum*, *Festuca rubra*, *Lotus uliginosus*, *Lychnis flos-cuculi*, *Prunella vulgaris*, *Salix repens*, *Taraxacum officinale* and *Trifolium pratense*. Some species typical of more eutrophic grassland decreased: *Cerastium fontanum*, *Cynosurus cristatus* and *Dactylis glomerata*. Other species that decreased are more representative of drier, oligotrophic and neutral grasslands such as *Achillea millefolium*, *A. ptarmica*, *Centaurea jacea*, *Erica tetralix*, *Hypochaeris radicata*, *Leontodon autumnalis*, *L. saxatilis*, *Leucanthemum vulgare* and *Polygala vulgaris*. Several species disappeared, including some rare taxa, such as *Ophioglossum vulgatum*, *Parnassia palustris* and *Selinum carvifolia*. One rare species showed an unexpected, opposite trend: *Dactylorhiza maculata* which increased.

Concise maps of the abundance in 1977 and 1988 and its changes are given for each species by Farjon & Wiertz (1989).

Table 1. Increase and decrease of species abundance according to the Wilcoxon test and fit in the RDA-analyses of Figs. 2 and 3.

1)	2)	3)		1)	2)	3)		1)	2)	3)	
--	*	**	Achillea millefolium				Filipendula ulmaria				Poa pratensis
--	**	**	A. ptarmica				Frangula alnus				Polygonum amphibium
			Agrostis spec	--	*		Galium palustre †				Polygala vulgaris
	*		Alnus glutinosa				G. uliginosum				Populus spec.
			Angelica sylvestris				Genista anglica				P. tremula
++	*	*	Anthoxanthum odoratum				Gentiana pneumonanthe			**	Potentilla erecta
+	**	**	Betula spec.				Hieracium spec.	--	*	**	Prunella vulgaris
			Briza media			**	H. laevigatum				Quercus robur
			Calamagrostis canescens				H. pilosella	-		**	Ranunculus acris
			Calluna vulgaris	+		**	Holcus lanatus				R. flammula
			Capsella bursa-pastoris †				Hydrocotyle vulgaris				R. repens
-	*	*	Cardamine pratensis				Hypericum maculatum				Rubus fruticosus
			Carex nigra	--	*		Hypochaeris radicata				R. idaeus
			C. panicea			*	Iris pseudacorus	--	*	**	Rumex acetosa
			C. pilulifera				Juncus acutiflorus				Sagina procumbens †
-**	*		Centaurea jacea				J. bulbosus				Salix aurita
--		**	Cerastium fontanum	-	*		J. conglomeratus	-		**	S. repens
			Chamerion angustifolium	--	*	*	Leontodon autumnalis				Selinum carvifolia †
			Cirsium arvense †	-			L. saxatilis				Stellaria graminea
--	**	**	C. dissectum	--	*	*	Leucanthemum vulgare				S. media
--	**	**	C. palustre	--	*	**	Lotus uliginosus				S. uliginosa
		*	Cladium mariscus				Luzula campestris				Succisa pratensis
			Corylus avellana				L. multiflora				Tanacetum vulgare
			Crataegus monogyna	--	**	**	Lychnis flos-cuculi	-		**	Taraxacum officinale
--	**	**	Cynosurus cristatus †				Lysimachia vulgaris				Trifolium dubium †
--	*		Dactylis glomerata	-	*	*	Lythrum salicaria †	--		**	T. pratense
++	*	**	Dactylorhiza maculata				Molinia caerulea				Valeriana officinalis
		**	Danthonia decumbens				Myrica gale				Vicia cracca †
	*		Equisetum arvense				Nardus stricta				Viola canina
			E. palustre				Ophioglossum vulgatum †				V. palustris
-		*	Erica tetralix				Parnassia palustris †				
--	**	**	Festuca arundinacea				Peucedanum palustre †				
			F. ovina				Phleum pratense †				
			F. pratensis	++	**	*	Plantago lanceolata				
--	*	*	F. rubra				Platanthera bifolia				

1) Decrease (⁻, ⁻⁻ with $P < 0.05$ and $P < 0.01$, resp.) and increase (⁺, ⁺⁺ with $P < 0.05$ and $P < 0.01$, resp.) according to the Wilcoxon signed-ranks test; 2) Fit of species in Fig. 3 (RDA with model year + strip.year and plot covariable) (*: fit > 20%, **: fit > 40% in two dimensions); 3) Fit of species in Fig. 2 (RDA with model plot + year + strip.year) (*: fit > 40%, **: fit > 50% in two dimensions); † = disappeared in 1988.

Changes with time of mean indicator value: univariate analysis

The differences between the mean indicator value (Table 2) are rather small in the case of F (moisture) and N (nutrients) but more substantial in the case of R (acidity). Of these, the decrease in N is not significant. The mean species richness per plot decreased considerably ($P < 0.01$). The mean indicator values of F, R, and N show high correlations ($r > 0.7$) across the plots per year, except for F in 1977, N in 1977 and 1988 ($r = 0.5$). The changes in F, R and N show low mutual correlation ($|r| < 0.3$).

Spatial and temporal variation in the floristic data: multivariate analysis

Fig. 2 displays the major variation in the 40 relevés across space and time. Instead of displaying the individual relevés, the centroids of each strip-by-year combination are displayed with lines that connect corre-

sponding strips in 1977 and 1988. This biplot, based on an RDA with model 'plot + year + strip·year', is hardly distinguishable from an unconstrained ordination by PCA and the first four eigenvalues of RDA and PCA are nearly equal (difference < 0.01). Notice that PCA in our case is equivalent to an RDA with model 'plot + year + plot·year'. This model has 15 degrees of freedom more than the RDA model of Fig. 2 (39 vs 24 degrees of freedom). The pattern in the PCA site ordination is well summarized by the reduced RDA model. Apparently, the spatial variation in the vegetation change is most prominent among strips. Fig. 2 shows a spatial trend from the lower part of the terrain (strip A) to the higher part (strip E). The trend is related to pH and water depth in 1988 as judged by their biplot arrows (which are added passively) and the high correlations ($|r| > 0.9$) between the 1988-scores on both axes and both pH and water depth. The temporal change is mainly in the direction 'south-west' to 'north-east' in Fig. 2. The amount of temporal change is largest in strip A and decreases when going to strip E, perhaps indicating a

Table 2. Mean Ellenberg indicator value and species richness in 1977 and 1988 ($n = 20$; * = $0.01 < P < 0.05$; ** = $P < 0.01$ based on Wilcoxons signed-rank test; n.s. = not significant $P > 0.05$).

	Mean indicator value		
	1977	1988	1988-1977
F (moisture)	6.46	6.35	− 0.11 *
R (pH)	3.84	3.57	− 0.27 **
N (nitrogen)	3.47	3.39	− 0.08 n.s.
S (species richness)	43.70	34.90	− 8.8 **

convergence in the development in the upper and driest strips to a vegetation of the *Violion caninae* type. If the individual relevés had been displayed in Fig. 2, an additional spatial trend could have been discernible within strips, which is smaller than that between strips. By connecting the points within each strip (e.g. from A1 to A4), we detected a trend that largely runs in the direction of the temporal change for each strip (except for strip D and E where the within-strip variation is smallest). This spatial trend might be attributed to the drop in the water table in the direction of the pumping station.

Fig. 2 also shows a dashed arrow for water depth. This variable is derived from the depth in 1988 by taking into account the simulation-based drop in the water table since 1977 (0.27 m). This new variable has both a spatial and a temporal component and is more strongly correlated to the first axis than water depth in 1988.

Projection of the species points in Fig. 2 on the spatial gradient, say the arrow for water depth in 1988, shows which species are more abundant in the drier part of the terrain (strip E) and which are more abundant in the wetter part (strip A). A similar projection on the time direction shows that most species displayed in Fig. 2 decrease with time. This can be done for each strip separately. For example, projection of the species points and the origin onto the line A77 - A88, shows that all species (except 4) fall on the 1977-side of the projection point of the origin. In Fig. 2 only those species are displayed of which the first two RDA dimensions, i.e. the figure, explain at least 50 % of the variance of the species abundance (Table 1).

Floristic changes with time: multivariate analysis

Fig. 2 displays only 51% of the variance in the fitted species abundance values in the model 'plot + year + strip·year' (given in CANOCO under the name of percentage variance of the species-environment relation). When looking in detail at the changes with time, it is therefore wise to focus the ordination on year by eliminating the trends in space (differences among plots). This also allows assessment of the statistical significance of the change with time.
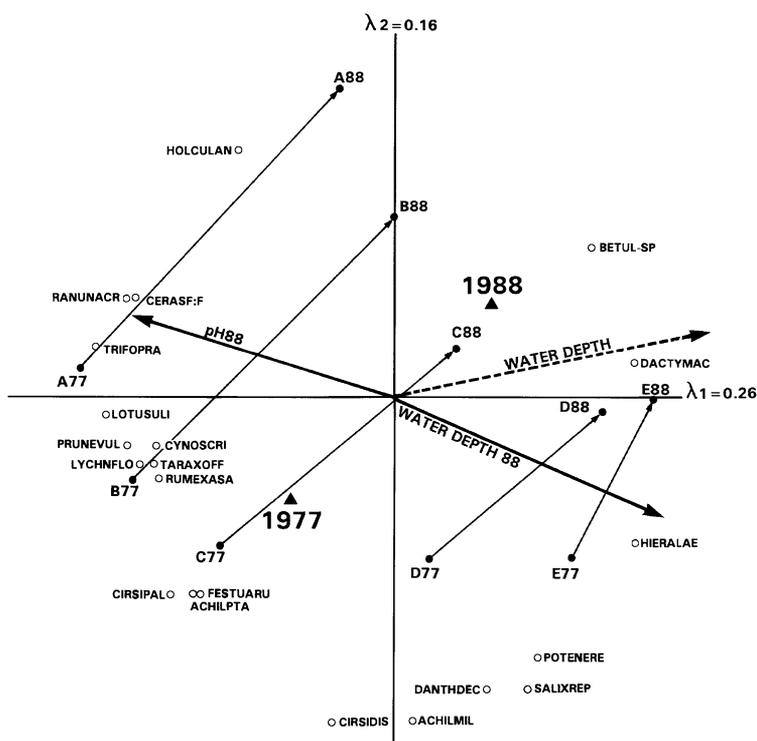
Fig. 2. Ordination diagram showing the spatial and temporal variation in the 40 relevés based on an RDA with model 'plot + year + strip.year'. Shown are centroids of year (1977 and 1988) and strip.year combinations, and species which fit well (more than 50% explained variance). Passively added to the biplot are the arrows for pH and water depth in 1988 (pH88 and WATER DEPTH 88) and water depth (derived from the depth in 1988 by taking into account the simulation-based drop of 0.27 m in the water table since 1977). The change with time is emphasized by connecting, for each strip, the centroid of 1977 with that of 1988. The diagram accounts for 42 % and 51% of the variance of the species data and fitted species data, respectively. The first four eigenvalues are 0.26, 0.16, 0.06 and 0.05. For full species names see Table 1.

The overall statistical test on the change in species composition between 1977 and 1988 (RDA with model 'year' and covariable 'plot') indicated a significant change ($P < 0.01$). To test whether the change was constant against the alternative hypothesis that it depended on water depth, an RDA was carried out with model 'water depth year' and with 'plot' and 'year' as covariables. The Monte Carlo test indicated a significant dependence ($P < 0.01$). The interaction effect of 'strip.year' was also significant ($P \approx 0.01$; covariables 'plot' and 'year').

The change with time and its dependence on strip (using 'strip' as proxy for water depth) is visualized in Fig. 3. The diagram (i.e. the first two axes) displays 85% of the variance accounted for by the model 'year + strip.year'. It is clear from the positions of the centroids for 1977 and 1988 in the diagram that the first axis mainly represents the change with time: many decreasing species on the left, only a few increasing species on the right. The centroids for strip-by-year combinations show a gradual dependence of the change on water depth as indicated by strip A-E. Species that decrease most in strip A are on the top-left side of the diagram, those that decrease most in strip E are on the bottom left-hand side. The formal interpretation of the figure is that a species' change in a strip, say E, can be inferred by projecting its point on the line connecting E77 with E88; on this line the point of no change is the origin. From this, the suggestion arises that species that decrease most in strip A, actually increase a little in strip E, but inspection of the raw data shows that this is not generally true (except for *Centaurea jacea*). The inference from the biplot that *Betula* spp. increased in strip A-C, but decreased in D-E is confirmed in the raw data.

Decomposition of the total variance

On the basis of RDA with a series of models, it is possible to decompose the total sum of squares in the species data as in an analysis of a variance table (Borcard et al. 1992; Økland & Eilertsen 1994). Space, time and their interaction account for 59, 16 and 25 %, respectively (Table 3). About 40 % of the spatial variation can be accounted for by pH and water depth (we did not attempt to subdivide this fraction further because of the high correlation (-0.86) between pH and water depth). The interaction strip.year takes 32 % of the full space·time interaction.

Because sums of squares tend to increase with the associated number of degrees of freedom, the importance of a component is often better judged on the basis of its mean square. The mean squares in Table 3 stress the importance of pH and water depth in the spatial component even more. For judging the significance of the

various components, tabulated critical values for the F-test are useless, because these are meant for univariate analysis. For multivariate analysis, the significance must be assessed by Monte Carlo permutation tests as reported in the previous section for the components time and strip·year.

The component 'pH and water depth' cannot be tested on the basis of random interchanges of the 1977- and 1988-relevés; the test statistic would be constant! To test this component, the spatial and temporal structure of the data was taken into account: (1) to remove the effect of the dependence of the relevés with time on the test, the 1977- and 1988-relevés of a plot were interchanged simultaneously with those of another plot, (2) to remove the effect of spatial dependence of plots, plots were not permuted randomly, but were shifted randomly by wrapping their rectangular spatial layout (Fig. 1) on a torus (Besag & Clifford 1989). This yields 20 different 'permutations', hence, 20 values of the test statistic, of which the observed value is the largest (i.e. $P = 0.05$). Because also the mirror images of the layout can be shifted, as explained more fully in the Update Notes of CANOCO 3.1 (ter Braak 1990b), there are in total 80 valid permutations, hence, 80 values of the test statistic of which the observed one is still the largest, whence $P = 0.012$. In conclusion, there is a statistically significant relation of the vegetation with pH and water depth across the field.

Change in pH and water depth as inferred from the ordination

One would expect extraction to cause the vegetation of wet plots to change to a vegetation that is typical for drier plots, perhaps after some time-lag. A similar simple model for acidification is that it causes plots to change to a vegetation that is typical for the more acid plots. Something of this sort is visible along the first axis in Fig. 2: compared to their position in 1977, the centroids of the strips in 1988 have moved towards the centroid of the drier and more acid strip E in 1977, i.e. in the direction of greater water depth and lower pH. As noted above, the variation within strips also showed a water-table-related trend which largely corresponded to the temporal change. Application of the simple change model (1) resulted in reasonable estimates of change in water depth and pH separately. These are as follows: If pH is assumed constant ($\Delta_{pH} = 0$), the likely change in water depth between 1977 and 1988 is obtained from an RDA with model year+water depth (eigenvalues of the two constrained and first two unconstrained axes: 0.24, 0.12, 0.09, 0.06). The estimated change is 0.19 m, which is to be compared with the value of 0.27 m obtained from the simulation model. The change parameter explained 28%
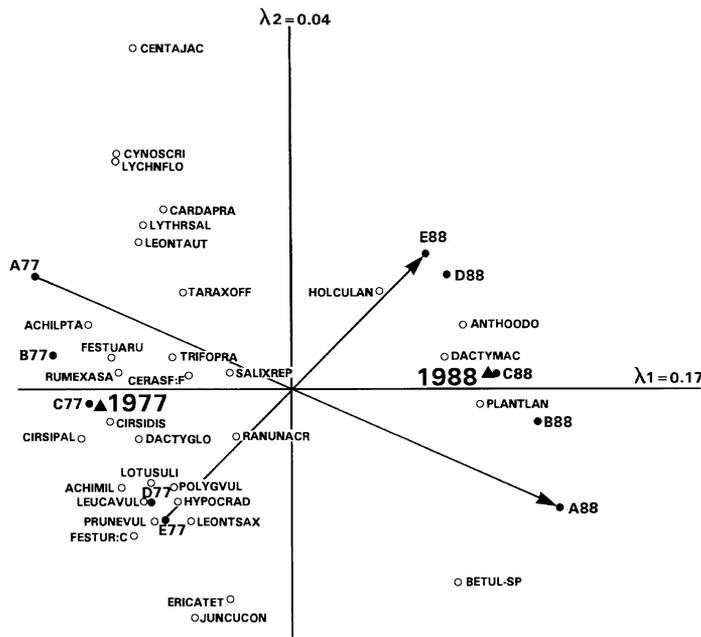
Fig. 3. Ordination diagram showing the change with time and its dependence on strip in more detail than Fig. 2 (RDA-model 'year + strip.year' with covariable 'plot'). Shown are centroids of the strip.year combinations and species which changed significantly in the Wilcoxon signed rank test (Table 1). For visual clarity, only the year-scores for strip A and strip E are connected. The diagram accounts for 50% and 85% of the residual variance (i.e. after fitting 'plot') of the species data and fitted species data, respectively. The first four eigenvalues are 0.167, 0.036, 0.015 and 0.013. For full species names see Table 1.

of the sum of squares associated with time given in Table 3. With $\Delta_{\text{water depth}} = 0$, the inferred pH change is −0.30 unit (with fit and eigenvalues close to those of the analysis with water depth). Joint estimation of the change based on an RDA with model year + pH + water depth (Table 4) resulted, however, in unrealistic estimates: −0.56 m for water depth whereas a positive value is expected and an acidification of 1.3 unit, presumably because of the high correlation between pH and water depth. Further joint analysis was therefore restricted to the first axis of this analysis (which closely resembles the first axis of Fig. 2). For a given change in water depth, the change in pH is estimated by the formula (see App. 1):

$$\hat{\Delta}_{\text{pH}} = - (c_1 + c_3 \, \Delta_{\text{water}})/ \, c_2. \tag{3}$$

Table 5 shows the estimates, with bootstrap 90 %-confidence intervals, for a series of likely changes in water depth. For water depth changes less than ca 0.25 m, the decrease in pH as inferred from the vegetation is significant ($P < 0.05$, one-sided test). At the change in water depth of 0.27 m, the inferred decrease in pH is on average −0.11 unit and just significant. These estimates of change explain 54 % of the sum of squares associated with time.

Discussion

Several authors have studied the effect of lowering of the water table on nutrient-poor, wet grasslands. In a situation with a sudden and large drop in level of more than one metre, the vegetation dies off and is partly replaced by weed species (Walther 1950; Ellenberg 1952). With less of a drop in level, some grass species typical of the *Violion caninae* become dominant. De Vries (1929) mentioned comparable trends in *Junco-Molinion* grasslands with an increase of *Agrostis canina* and *Plantago lanceolata*. Grootjans et al. (1986) found that the most faithful species in a *Junco-Molinion* stand decreased or disappeared, e.g. *Carex hostiana, C. pulicaris* and *C. panicea*, whereas *Festuca ovina, Molinia caerulea* and, locally, *Holcus lanatus, Agrostis canina, Succisa pratensis* and *Mentha arvensis* increased. Our study area was already affected by water table lowering at the time of the observations in 1977. The type of

Table 3. Decomposition of the variance in the species data obtained by RDA. Each sum of squares (ss) is the sum of the corresponding entry in the analysis of variance tables of the 100 individual species and is expressed as a fraction of the total variance in the species data. (df = degrees of freedom, ms = mean square, *: 0.01 < P < 0.05, **: P < 0.01 by Monte Carlo permutation).

Source	df	ss	ms	
Space (plot)	19	0.59	0.04	
pH and water depth	2	0.23	0.11	*
residual	17	0.36	0.02	
Time (year)	1	0.16	0.16	**
Space.Time	19	0.25	0.01	
strip.year	4	0.08	0.02	**
residual	15	0.17	0.01	
Total	39	1.00		

Table 4. Canonical coefficients (between brackets approximate t-ratio's in absolute value) and inter-set correlations of the RDA of the species data with model year + pH + water depth (eigenvalues: 0.25, 0.11, 0.02; first unconstrained eigenvalue: 0.09).

Axis		Canonical coefficients						Inter set correlations		
		1		2		3		1	2	3
Year 1977	(c_1)	− 0.580	(13.1)	− 0.548	(9.1)	− 0.028	(0.8)	0.56	0.71	0.07
pH	(c_2)	1.126	(4.8)	− 0.164	(0.5)	− 1.599	(8.2)	0.76	− 0.47	− 0.27
Water depth	(c_3)	− 1.665	(4.9)	1.365	(3.0)	− 2.250	(7.9)	− 0.76	0.52	− 0.16

changes we observed show, however, some similarity: increase of only a few species such as *Plantago lanceolata*, *Anthoxanthum odoratum* and *Holcus lanatus* and of some woody species; a decrease or even disappearance of many species typical of moist oligotrophic or mesotrophic and neutral hayfields and of several rare species such as *Parnassia palustris*, *Selinum carvifolia* and *Ophioglossum vulgare*. One rare species, *Dactylorhiza maculata*, increased, however. This may be in accordance with the general trend towards the *Violion caninae* or may be due to the often large and unpredictable fluctuations in the population size of orchids (Wells & Willems 1991).

The changes in mean Ellenberg indicator values come close to an interpretation of vegetation change in terms of water level, acidity and nutrients, respectively. However, it cannot be excluded that the large change in the mean indicator value for acidity is caused indirectly by water extraction. There was no increase in nutrient values as indicated by Ellenberg. Enhanced mineralization may have occurred before 1977 and is therefore not detectable any more. A drop in species richness was also found by Grootjans et al. (1986).

The simple change model (1) forces a particular interpretation on the vegetation change with time, namely that the change is caused by an overall change in pH and water depth. Due to the correlation (perhaps causal) between pH and water depth, the model cannot distinguish between 'drying' and acidification and no realistic

joint estimate of the change in water depth and pH could be obtained. Nevertheless, after fixing the change in water depth at the value of 0.27 m from the simulation model (Pastoors 1985), the observed vegetation change points to statistically significant acidification (on average 0.1 pH unit). This estimate is not unreasonable when compared with the few actual measurements of pH change that are available. At three spots near the grid intersections B2, C3 and F3 (Fig.1) the measured changes are − 0.6, − 0.5 and +0.2 pH units, respectively ($\Delta_{pH\text{-KCl}}$ in a mixed sample of 40 subsamples of 0-5 cm topsoil on 50m^2). The measured pH changes suggest a dependence on water depth. In principle, this dependence can be included in the change model, but the present data set seems too small to warrant extension of the model. Finally, note that the change model will underestimate the environmental change if the vegetation lags behind.

Our study area is one of the areas in the Netherlands whose hydrology is studied most intensively. Nevertheless, this case study is presented as one in which precise data on the changes in the environment are lacking. A first reason for this paradox is that, as in many other affected nature reserves, nearby piezometers were only installed after drainage or pumping started. Therefore hydrological models were utilized to determine the extent of the drop. A second reason is that we believe that this vegetation does not react to instantaneous fluctuations in the water table, but only to its long-term pattern, in particular, as far as the pattern reflects moisture deficit. The third and fourth reasons are the variability that is inherent in hydrological measurements and the lack of spatial resolution which together preclude more detailed data on the change. For these reasons we used mean summer water depth and a simple extrapolation thereof across the field. This approach appeared to be successful.

We analysed the floristic changes by both univariate and multivariate methods. The multivariate method is attractive in that it provides an overall significance test of vegetation change. The multivariate approach thus avoids the problem of multiple testing in the univariate analyses. (The problem of multiple testing entails that, if none of the species changed in expected abundance, the observed change of approximately 5% of the species

Table 5. Change of pH between 1977 and 1988 ($\hat{\Delta}_{pH}$) as inferred from the first RDA-axis of Table 4 (with 90% confidence interval) for a range of changes in water depth (m). According to a hydrological model, the change in water depth was 0.27 m.

$\Delta_{\text{water depth}}$	$\hat{\Delta}_{pH}$	
0	− 0.52	(− 0.66, − 0.20)
0.10	− 0.37	(− 0.46, − 0.20)
0.20	− 0.22	(− 0.32, − 0.13)
0.25	− 0.14	(− 0.30, − 0.05)
0.27	− 0.11	(− 0.29, − 0.01)
0.30	− 0.07	(− 0.29, 0.06)
0.35	0.00	(− 0.29, 0.16)

will nevertheless erroneously be judged significant.) Once the overall change is demonstrated to be significant, the detailed non-parametric tests per species in the univariate approach help to single out which species changed. The univariate statistics available with RDA are less attractive: percentage fit per species (standard output in CANOCO 3.1) and the approximate t-tests through the t-value biplot (ter Braak 1990a). The percentage fit is an easy criterion for deciding which species will be displayed in the ordination diagram. In Table 1 we compare the fit per species in the ordinations with the results of the Wilcoxon test. There is a good resemblance between the Wilcoxon test results and the fit in the ordination that focuses on change with time (Fig. 3). Because the Wilcoxon tests are more decisive for change with time, we used a species' significance as a criterion for its inclusion in Fig. 3. This explains why *Cerastium fontanum* and *Trifolium pratense* are displayed in Fig. 3. Fig. 2 includes three species that are not present in Fig. 3; *Danthonia decumbens*, *Hieracium laevigatum* and *Potentilla erecta* all show a marked spatial trend but did not change systematically with time (Table 1).

The number of permutations to be carried out in a Monte Carlo test deserves attention. If at least 19 permutations are carried out for a test at the 5% significance level, the test is exact in the sense that when applied to different data sets, the null hypothesis, when true, is rejected falsely in precisely 5% of the cases. This concept of significance extends the conventional one in which the significance level applies to each particular data set. If closeness to the conventional significance level is important, at least 999 permutations are needed (see Manly 1990). But why not accept the broader concept of significance? Even for very low P, it is not sure whether the null hypothesis is true or false for any particular data set anyway! If the concept of a Monte Carlo significance level is accepted, the incentive for carrying out more permutations is then to increase the power of the test. This increase comes at the cost of computer time. One might argue that this time is insignificant compared to the time involved with field work, data checking and report writing. However, for large data sets it may become prohibitive to do 999 permutations. Also the law of diminishing returns applies. Besag & Diggle (1977) suggested that 99 permutations would be a good compromise. Jöckel (1986) derived a generally applicable formula for the power of the Monte Carlo test on which Jöckel (1991) based the following rule of thumb: at least 199 permutations should be carried out to obtain at least 85% of the power of the test that uses infinitely many permutations. With the relatively small data set we analyzed in this paper, CANOCO could carry out 999 permutations on a 486-machine within a few minutes. Following Jöckel (1991), we

suggest that for larger data sets a reasonable default number is 199.

The validity of the Monte Carlo tests depends critically upon the way the permutations are carried out. Data that are statistically dependent should not be permuted. This principle led in our application to ways of permuting that differed among tests: random interchanges within plots for tests that involved time and a specialized (constrained) type of permutation for the test that involved space. The latter test can also serve as a test of spatial trend that takes account of autocorrelation of residuals. If trend is not of interest but the pure environmental effect is, the geographical coordinates can be used as covariables in the test of the environmental effect (Borcard et al. 1992).

The multivariate approach stimulated us to look at the spatial variation in the change, hence the test on the strip.year interaction and ordination diagrams showing this variation. It seems impractical to do this for each species or to calculate for each species an analysis of variance or analysis of deviance (Anon. 1987) analogous to Table 3. But, separate detailed analyses for the most interesting species may be worthwhile and go beyond simple linear models (Yee & Mitchell 1991). In ordination diagrams the similarity in response among species stands out. In the univariate analysis such similarity must be searched for in the numeric output per species. On the other hand, ordination diagrams often present only a part of the full information, are easily overcrowded and require skill for their correct interpretation. These limitations can be alleviated by presenting partial ordinations that focus on particular aspects (Fig. 3), by careful selection of the species for display and by training, respectively. Yet, at several points in the analysis we checked conclusions derived from the ordination diagrams against the mapped data per species.

It may be useful to perform multivariate methods directly on the output of the univariate analyses for a series of species rather than on the raw data. This is actually what Redundancy Analysis does: it is a Principal Components Analysis of the table of fitted values obtained by multiple regressions. Simultaneously, Redundancy Analysis achieves a dimensional reduction of the table of multiple regression coefficients (ter Braak 1990a; ter Braak & Looman 1994). Multivariate analysis complements rather than replaces univariate analysis.

References

Anon. 1984. *Natuurreservaat Koolmansdijk. Invloed van een verminderde waterwinning op de vegetatie.* Report Land-inrichtingsdienst, Arnhem.

Anon. 1987. *GENSTAT 5: A general statistical program.* Clarendon Press, Oxford.

Bannink, J.F. & Pape, J.C. 1979. *Enige gegevens betreffende de bodemgesteldheid van de natuurterreinen Empese en Tondense Heide, Koolmansdijk en Nijkamps Heide.* Report nr. 1456. STIBOKA, Wageningen.

Besag, J. & Clifford, P. 1989. Generalized Monte Carlo significance tests. *Biometrika* 76: 633-642.

Besag, J. & Diggle, P.J. 1977. Simple Monte Carlo tests for spatial patterns. *Appl. Stat.* 26: 327-333.

Böcker, R., Kowalik, I. & Bornkamm, R. 1983. Untersuchungen zur Anwendung der Zeigerwerte nach Ellenberg. *Verh. Ges. Ökol.* 11: 35-56.

Borcard, D., Legendre, P. & Drapeau, P. 1992. Partialling out the spatial component of ecological variation. *Ecology* 73: 1045-1055.

Both, J.C. & van Wirdum, G. 1981. *Waterhuishouding, bodem en vegetatie van enkele Gelderse natuurgebieden.* RIN-report nr. 81/18 Rijksinstituut voor Natuurbeheer, Leersum.

Buijsman, E., Blom, W.F. & Onderdelinden, D. 1990. *Luchtkwaliteit. Jaarverslag 1989.* RIVM-report nr. 222101006. Rijksinstituut voor Volksgezondheid en Milieuhygiene, Bilthoven.

Cramer, W. & Hytteborn, H. 1987. The separation of fluctuation and long-term change in vegetation dynamics of a rising seashore. *Vegetatio* 69: 157-167.

Davies, P.T. & Tso, M.K.-S. 1982. Procedures for reduced-rank regression. *Appl. Statist.* 31: 244-255.

de Bakker, H. & Schelling, J. 1966. *Systeem van bodem-classificatie voor Nederland. De hogere niveaus* (with English summary). Pudoc, Wageningen.

de Vries, D.M. 1929. Het plantendek van de Krimpenerwaard III. *Ned. Kruidk. Arch.* 39: 245-403.

Ellenberg, H. 1952. Auswirkungen des Grundwassersenkung auf die Wiesengesellschaften am Seitenkanal westlich Braunschweig. *Angew. Pflanzensoziol.* 6: 1-47.

Ellenberg, H. 1979. Zeigerwerte der Gefäszpflanzen Mitteleuropas. 2. Aufl. *Scripta Geobot.* 9: 1-122.

Farjon, A. & Wiertz, J. 1989. *Milieu- en vegetatieveranderingen in het schraalland Koolmansdijk (Lichtenvoorde) 1952-1988.* RIN-report nr. 89/18. Rijksinstituut voor Natuurbeheer, Leersum.

Grootjans, A.P. & ten Klooster, W. P. 1980. Changes of groundwater regime in wet meadows. *Acta Bot. Neerl.* 29: 541-554.

Grootjans, A.P., Hartog, P.S., Fresco, L.F.M. & Esselink, H. 1991. Succession and fluctuation in a wet dune slack in relation to hydrological changes. *J. Veg. Sci.* 2: 545-554.

Grootjans, A.P., Schipper P.C. & van der Windt, H.J. 1986. Influence of drainage on N-mineralisation and vegetation response in wet meadows. II. *Cirsio-Molinietum* stands. *Oecol. Plant.* 7: 3-14.

Hall, P. 1988. Theoretical comparison of bootstrap confidence intervals (with discussion). *Ann. Stat.* 16: 927-981.

Jöckel, K.-H. 1986. Finite sample properties and asymptotic efficiency of Monte Carlo tests. *Ann. Stat.* 14: 336-347.

Jöckel, K.-H. 1991. Monte Carlo techniques and hypothesis testing. In: Nelson, P.R., Dudewicz, E.J., Öztürk, A. & van der Meulen, E.C. (eds.) *The frontiers of statistical computation, simulation & modelling,* pp. 21-42. American Sciences Press, New York, NY.

Jongman, R.H.G., ter Braak, C.J.F. & van Tongeren, O.F.R. 1987. *Data analysis in community and landscape ecology.* Pudoc, Wageningen.

Kemmers, R.H. 1986. Perspectives in modelling of processes in the root zone of spontaneous vegetation at wet and damp sites in relation to regional water management. *Comm. Hydrol. Res. TNO Proc. Inf.* 34: 91-116.

Manly, B.F.J. 1990. *Randomization and Monte Carlo methods in biology.* Chapman and Hall, London.

Mountford, J.O. & Sheail, J.E. 1989. *The effects of agricultural land use change on the flora of three grazing marsh areas.* Focus on nature conservation No. 20. Nature Conservancy Council, Peterborough.

Økland, R.H. & Eilertsen, O. 1994. Canonical correspondence analysis with variation partitioning: some comments and an application. *J. Veg. Sci.* 5: 117-126.

Pastoors, M.J.H. 1985. *Nota inzake berekening van de grondwaterstandsverlagingen ten gevolge van de winning te Lichtenvoorde en de compensatie voor diepinfiltratie in een natuurgebied.* Report. Rijksinstituut voor Volksgezondheid en Milieuhygiene, Bilthoven.

Schachtschabel, P., Blume, H.-P., Hartge, K.-H. & Schwertmann, U. 1989. *Lehrbuch der Bodenkunde.* Enke, Stuttgart.

Sokal, R.R. & Rohlf, F.J. 1981. *Biometry.* Freeman, San Francisco, CA.

Sukopp, H. & Trautmann, W. 1981. Causes of the decline of threatened plants in the Federal Republic of Germany. In: Synge, H. (ed.) *The biological aspects of rare plant conservation,* pp. 113-116. Wiley, London.

Tansley, A.G. & Chip, T.F. (eds.) 1926. *Aims and methods in the study of vegetation.* British Empire Vegetation Committee. Whitefriars Press, London.

ter Braak, C.J.F. 1987. The analysis of vegetation-environment relationships by canonical correspondence analysis. *Vegetatio* 69: 69-77.

ter Braak, C.J.F. 1990a. Interpreting canonical correlation analysis through biplots of structural correlations and weights. *Psychometrika* 55: 519-531.

ter Braak, C.J.F. 1990b. *CANOCO - a FORTRAN program for CANOnical Community Ordination by [partial] [detrended] [canonical] correspondence analysis, principal components analysis and redundancy analysis, version 3.10.* Microcomputer Power, Ithaca, NY.

ter Braak, C.J.F. 1992. Permutation versus bootstrap significance tests in multiple regression and ANOVA. In: Jöckel, K.-H., Rothe, G. & Sendler, W. (eds.) *Bootstrapping and related techniques,* pp. 79-85. Springer Verlag, Berlin.

ter Braak, C.J.F. In press. Canonical community ordination. Part I: Basic theory and linear methods. *Ecoscience* 1.

ter Braak, C.J.F. & Looman, C.W.N. In press. Biplots in reduced-rank regression. *Biom. J.* 36 (7).

ter Braak, C.J.F. & Prentice, I.C. 1988. A theory of gradient

analysis. *Adv. Ecol. Res.* 18: 271-317.

van der Meijden, R., Weeda, E.J., Adema, F.A.C.B. & de Joncheere, G.J. 1983. *Heukels' Flora van Nederland.* Wolters-Noordhoff, Groningen.

van Wirdum, G. 1979. Dynamic aspects of trophic gradients in a mire complex. *Comm. Hydrol. Res. TNO Proc. Inf.* 25: 66-82.

van Wirdum, G. 1981. Linking up the natec subsystem in models for water management. Comm. Hydrol. Res. *TNO Proc. Inf.* 27: 108-128.

van Wirdum, G. 1982. *The ecohydrological approach to nature protection.* In: Annual Report 1981, pp. 60-74. Research Institute for Nature Management, Arnhem.

Walther, K. 1950. Unkrautherden als Zeiger Grundwasser geschädigter Grünlandsgesellschaften auf Niedermoor-

böden. *Mitt. Flor. Soz. Arbeitsgem.* N. F. 2: 43-51.

Wells, T.C.E. & Willems, J.H. (eds.) 1991. *Population ecology of terrestrial orchids.* SPB Academic Publishing, The Hague.

Westhoff, V. 1979. Bedrohung und Erhaltung seltener Pflanzengesellschaften in den Niederländen. In: Willmans, O. & Tüxen, R. (eds.) *Werden und Vergehen von Pflanzengesellschaften.* Cramer, Vaduz.

Westhoff, V. & den Held, A.J. 1969. *Plantengemeenschappen in Nederland.* Thieme, Zutphen.

Wiertz, J. 1992. *Schatting van ontbrekende vocht- en stikstofindicatiegetallen van Ellenberg (1979).* RIN-rapport 92/7. Instituut voor Bos- en Natuuronderzoek, Wageningen.

Yee, T.W. & Mitchell, N.D. 1991. Generalized additive models in plant ecology. *J. Veg. Sci.* 2: 587-603.

App. 1. Proof that Eq. (2) is the least-squares estimator of model (1)

Consider the multivariate regression model

$$y_k = a_k + b_k x + d_k t + \text{error.} \tag{A.1}$$

Let $\mathbf{M}$ be an $m \times 2$ matrix, the columns of which contain the regression coefficients b_k and d_k, respectively. Model (1) can be derived from (A.1) by constraining the regression coefficients for t by $d_k = -\Delta b_k$. This is a linear constraint on the columns of $\mathbf{M}$ and reduces the rank of $\mathbf{M}$ from 2 to 1. Model (1) is thus a special parametrization of the reduced rank regression model (Davies & Tso 1982) with a rank of 1. The least-squares fit thereof is obtained from the first axis of redundancy analysis (ter Braak & Prentice 1988; ter Braak 1990b). The relation between these models can also be seen explicitly as follows: the one-dimensional RDA-model is

$$y_k = a^*_k + b^*_k (c_1 t + c_2 x) + \text{error.} \tag{A.2}$$

If $c_2 \neq 0$, model (1) can be obtained by inserting

$$a^*_k = a_k, \; b^*_k = b_k/c_2, \; c_1 = -\Delta c_2 \tag{A.3}$$

in (A.2). The least-squares fit of Δ thus follows from the estimates of c_1 and c_2 by equation (2). This concludes the proof.

With changes in p environmental variables, $x_1, x_2, ..., x_p$, the change model becomes

$$y_k = a_k + \sum_j b_{kj} (x_j - \Delta_j t) + \text{error} \tag{A.4}$$

Again, the change model is a constrained version of the multivariate regression model

$$y_k = a_k + \sum_j b_{kj} x_j + d_k t + \text{error} \tag{A.5}$$

the linear constraint being

$$d_k = -\sum_j \Delta_j b_{kj} \tag{A.6}$$

The least-squares solution is thus obtained by fitting a reduced rank model with rank p. On using the reduced rank estimates of d_k and b_{jk} the least-squares estimate of the changes $\{\Delta_j\}$ follows from (A.6) by matrix inversion [this is equivalent to a regression through the origin of $-d_k$ on $b_{k1}, b_{k2}, ..., b_{kp}$ ($k = 1,..., m$); this regression has zero residual sum of squares]. Equivalently, we can use the canonical coefficients of the first p axes. By rewriting the p-dimensional RDA-model explicitly as a change model, it can be shown [cf. (A.3)] that a formula analogous to (A.6) holds for the canonical coefficients with axis-numbers replacing the species index k in (A.6). Therefore, the $\{\hat{\Delta}_j\}$ [$j = 1, ..., p$] are obtained from the canonical coefficients by a regression through the origin of the canonical coefficients for $-t$ on to those of $x_1, x_2, ..., x_p$.

The change model (A.4) can also be used if the vegetation was sampled on more than one previous occasion. Then, t should not be a dummy variable, but a quantitative variable that measures 'time before the last sampling date' (assuming that environmental measurements are available from the last sampling date). This modification of the model assumes linear changes with time of the environmental variables with rates $\{\Delta_j\}$. If this assumption is not warranted, the model can be extended with a series of dummy variables, as we hope to demonstrate elsewhere.

If the environmental variables are multicollinear, then the change estimates are unstable. To alleviate this problem, a single axis is used for estimation in the main text. Equation (3) follows from the equation

$$c_1 = -(c_2 \Delta_{pH} + c_3 \Delta_{water}) \tag{A.7}$$

which is analogous to (A.6).

Canonical correspondence analysis can be seen as a form of RDA applied to transformed data. The above theory is therefore valid for this method as well.