26 Monitoring for change: Using generalised least squares, non-metric multidimensional scaling, and the Mantel test on western Montana grasslands

Sikkink, P.G., Zuur, A.F., Ieno, E.N. and Smith, G.M.

26.1 Introduction

Monitoring programs are vital to assess how plant community succession is affected by environmental change. Each plant community has many biologic, climatic, and abiotic interactions that affect its species differently over time. In temperate grasslands, plant community composition and species dominance can change rapidly in response to changes in the timing and amount of precipitation (Fay et al. 2002; Knapp and Smith 2001). In fact, some of these grasslands are so sensitive to variations in precipitation that they have been dubbed "early warning systems" for global climate change (Kaiser 2001). However, these ecosystems are also sensitive to temperature fluctuations (Alward et al. 1999), the timing and intensity of grazing and fire (Fuhlendorf et al. 2001; Geiger and McPherson 2005; Jacobs and Schloeder 2002), fire exclusion (Leach and Givnish 1996), and invasion of non-native species (Abbott et al. 2000). Monitoring programs that span several decades are critical to determining which of these environmental stresses are important to compositional change within a particular grassland and how rapidly that change occurs.

Our goal in this case study chapter was to analyse monitoring data from two temperate grassland communities in Montana, USA, to determine whether their composition changed over three to five decades and, if so, whether any of these environmental factors can be related to the changes. Our two study areas are protected reserves that have been sites for long-term natural experiments since the late 1950s and 1960s. Both are dominated by native bunchgrasses, including *Pseudoroegneria spicata* (Pursh) A. Love, *Festuca idahoensis* Elmer, and *F. altaica* Trin.; and both have been affected by climatic fluctuations, grazing, and invasion of non-native species throughout their monitoring history. These particular grasslands are unique because they exist in cool, semi-arid northern landscapes that have not been highly fragmented by increases in human population, converted to agricultural fields, or manipulated in experimental studies.

Our primary questions in this case study address the following: Does the biodiversity of these bunchgrass communities change over time? If so, do changes in biodiversity relate with any particular environmental factor? Does species' membership within the bunchgrass community change significantly over 30–50 years? If so, what environmental factors promote changes in the community over time?

The underlying statistical question for these data is simple: 'Is there a relationship between species present in the community at a particular point in time and the environmental variables?' Answering this question is far from simple, however, because the number of environmental variables is relatively large, the plant data have a large percentage of observations with zero species abundance, and the data have an underlying time component. To address how the environmental variables correlate with (i) biodiversity change and (ii) community change, we adopt a multi-step strategy that ultimately provides two different perspectives on the relationship between species and environmental (i.e., explanatory) variables. First, we will reduce the large number of species into an index that represents biodiversity and relate the diversity index to the explanatory variables. Second, we will reduce the number of explanatory variables using variance inflation factor (VIF) analysis. VIF eliminates environmental variables that are collinear and leaves only the variables that contain unique information. Linear regression is then applied to the remaining explanatory variables. Within the linear regression analysis, we must address a crucial question: 'Do the residuals from this linear regression model show any temporal patterns?' If they do, we will need to address residual temporal structure by comparing the results of the linear regression model with a model created using generalised least squares (GLS). For a second perspective on how environmental variables relate to community change, we apply a multivariate analysis, which incorporates the frequency of abundance of all of the species present at each site. Instead of using a diversity index that represents a site's diversity, we compare similarities among and between communities with time based on similarity values computed for each sample period. Because this dataset has many zeros, we will use classical methods like the Jaccard or Sørensen similarity indices combined with NMDS, the Mantel test, and BVSTEP to look at the relationships between species presence and the environmental variables.

26.2 The data

The sites examined for this case study are located in the northern Rocky Mountains of Montana, USA, within Yellowstone National Park and the National Bison Range (Figure 26.1). Both areas are sites for natural experiments that have recorded the response of bunchgrass communities to fluctuating climate, disturbance, and invasion for the past 30–50 years. The communities of interest are the native *Pseudoroegneria spicata* (Pursh) A. Love, *Festuca idahoensis* Elmer, and *F. altaica* Trin. Yellowstone National Park (YNP) is located at Montana's southern border (Figure 26.1). The area is mountainous and an all-season tourist destination. Although the central and southern areas of the park are famous for

geysers and volcanic hot-water features, much of the northern boundary area is dominated by native grass and shrub communities and is important winter range for several big-game species including wapiti (*Cervus canadensis*) and bison (*Bison bison*). In 1957, YNP initiated a natural experiment to determine how the grass and shrub communities changed with increasing populations of wildlife in the park (Edwards 1957). Park personnel constructed several five-acre exclosures in the northern winter range that eliminated all big-game grazing within the fenced boundaries. They established multiple permanent transects inside each exclosure. Each transect measured 33.3 m (100 ft). Just outside the exclosures, complementary transects were established that remained open to big-game grazing year-round. The transects selected for analysis in this case study lie between 1650 and 2050 m in elevation and have been resampled six to eight times in the last 50 years.

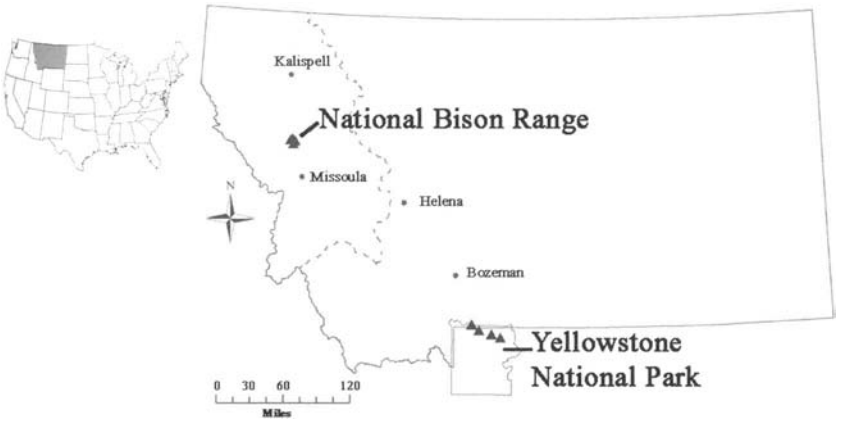

Figure 26.1. Transect locations in Montana are shown by triangles. The dashed line designates the location of the continental divide.

The National Bison Range (NBR) is located near Montana's western boundary (Figure 26.1). It covers approximately 7500 ha. Like YNP, the bunchgrass communities exist at lower elevations within inter-mountain areas. The wildlife refuge is managed to preserve remnants of the once legendary bison population in the western United States and to provide tourists with wildlife viewing and educational opportunities. All areas of the refuge are open to grazing by the same native ungulate species that live in YNP. However, unlike YNP where bison are free to wander outside park boundaries, bison in the NBR are restricted to the refuge and managed within it by several fenced pastures and a rotational grazing system. In the late 1960s, a natural experiment similar to YNP's began in the NBR to monitor range trend and condition. Initially, the experiment began as a way to measure the effects of managed grazing and monitor its trends over very short time frames, but data have continued to be collected for more than 30 years. The transect lines

selected for this study have been resampled an average of 10 times since the late 1960s and range from 875 to 950 m in elevation.

Throughout the historic record, sampling intervals have been irregular in both areas (Figure 26.2), but all of the transect lines have been sampled using exactly the same methods. Along each line, vegetation or substrate encountered at each 0.33 m mark was recorded. Vegetation hits were identified to species and recorded as either overstory or understory in the canopy. Substrate hits were recorded as bare ground, rock, pavement, litter, or moss/lichen. Each line had a total of 100 hits per line, so all species and substrate data in this case study represent frequency of occurrence in each sample year. The timing for each re-sampling was matched as closely as possible to the timing of historic samplings so changes in species' frequency over the monitoring period were not confused with seasonal physiologic changes.

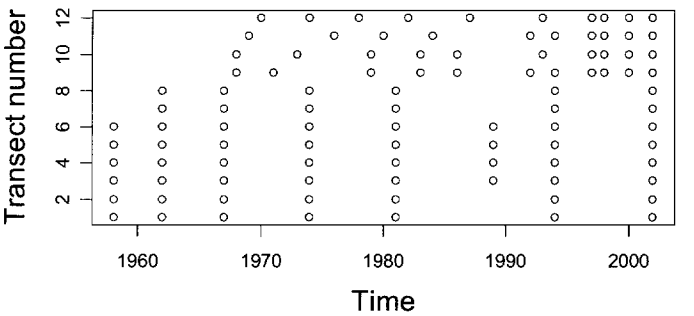

Figure 26.2. Years each transect was sampled (indicated by a circle).

Twenty-one environmental variables are used in this study. They include substrate characteristics, soil texture, and seasonal weather conditions during sampling. Substrate variables include the following surface characteristics at the time of sampling: rock that is greater than 1.9 cm (0.75 in) in diameter (ROCK); exposed soil (BARESOIL); litter (LITTER); and moss or lichen (M/L). Soil samples were analysed for the proportion of sand (PCTSAND), silt (PCTSILT), and clay (PCTCLAY) in their matrix using a combination of sieving and standard hydrometer methods (Bouyoucos 1936). Organic carbon content of soil samples (PCTOrgC) was measured using Loss-On-Ignition procedures (Nelson and Sommers 1982). Soil and organic carbon samples were only collected in 2002, not at each historic sampling.

In general, the climate of both sample areas is semi-arid. YNP has a 38 cm mean annual precipitation and a mean annual temperature of 4.4°C. The NBR is slightly warmer and moister with a 40 cm mean annual precipitation and a mean annual temperature of 7.5°C (Western Regional Climatic Center 2002). To examine how environmental variables correlated with species changes, however, tem-

perature and precipitation values had to be specific to transect site. Specific seasonal values were assigned to each transect for each year from 1958 to 2002 using a surface observation gridding system (SOGS). In the SOGS process, each of the monitoring sites received seasonal climate values that were interpolated from the nearest climate station and adjusted for each site's unique elevation, slope, aspect and location on the landscape (Jolly et al. 2004). The seasonal divisions included fall (September–October prior to sampling year), winter (November–March), spring (April–May), and summer (June–August). The resulting environmental variables include mean minimum temperatures for each season (FallTmin, WinTmin, SprTmin, SumTmin), mean maximum temperatures for each season (FallTmax, WinTmax, SprTmax, SumTmax) and total precipitation for each season (FallPrec, WinPrec, SprPrec, and SumPrec).

For the following data exploration and analysis, transect lines are numerically coded to make interpretations easier. The numeric codes, their corresponding locations, and current management regimes are shown in Table 26.1.

Table 26.1. Transect characteristics in Yellowstone National Park (YNP) and the National Bison Range (NBR), USA.

Site	Transect Number	Location	Management
YNP	1,3,5,7	Inside exclosures	No grazing
YNP	2,4,6,8	Outside exclosures	Open grazing
NBR	9,10,11, 12	North, South, West, East boundaries, respectively	Confined bison grazing, Open wildlife grazing

26.3 Data exploration

The 12 locations from YNP and the NBR contained 93 species, 99 sets of transect data, and 21 environmental variables. As we explore these datasets to find relationships between the species data and the explanatory variables, we are faced with an immediate problem. Many of the explanatory variables are collinear. In previous chapters, we encountered similar collinearity problems, but in each case a simple pairplot gave clear indications of which variables could be dropped. In this study, there are too many explanatory variables to easily use a pairplot. When pairplots become impractical, the PCA biplot is often more appropriate to identify collinearity among variables (Figure 26.3).

Recall from Chapter 12 that lines pointing in the same direction in a PCA biplot indicate that the corresponding variables are correlated with each other, lines pointing in opposite directions are negatively correlated, and lines with an angle of 90^0 are uncorrelated. The directions of lines in Figure 26.3 indicate that many of the climate variables are highly (either positively or negatively) correlated. All seasonal temperatures are highly correlated with each other, and all seasonal precipitation is highly correlated. However, precipitation variables are uncorre-

lated with temperature variables because they are at 90° angles. Soil texture variables are, for the most part, uncorrelated with climate variables.

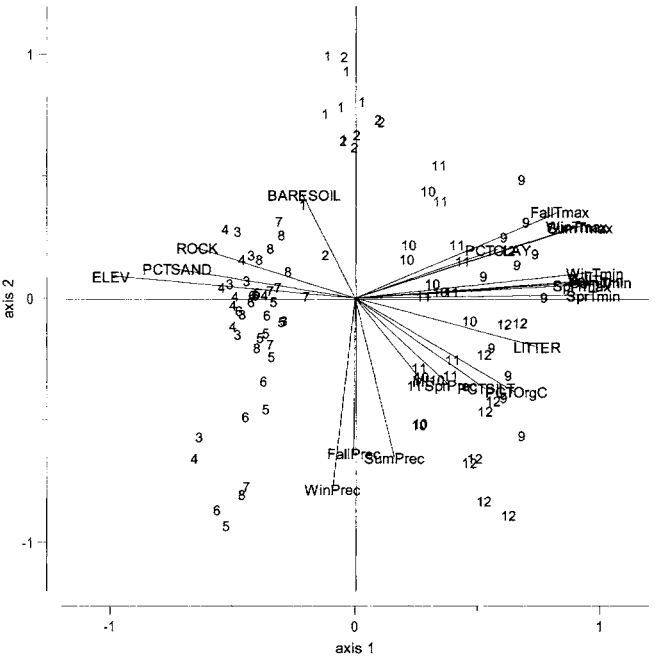

Figure 26.3. PCA correlation biplot for all explanatory variables. The first two eigenvalues are 0.49 and 0.12, which means that these two axes explain 61% of the variation in the data. The numbers in the graph are the PCA scores and represent the transects. Transects 1–8 are from Yellowstone National Park and 9–12 are from the National Bison Range.

Some substrate variables correlate with climate variables (LITTER AND ML), but others do not (ROCK, BARESOIL). In the biplot, notice that most observations from the same transect tend to be close to each other, indicating that the composition of each transect over time is more similar to itself than to other transects. The biplot also shows that some of these transects are more variable over time than others (e.g., transects 7 and 9). Finally, the biplot shows distinct separation between transects 1–8 from YNP and transects 9–12 from the NBR, indicating that environmental conditions differ dramatically between these areas. This interpretation must be used with reservation, however, because some of the explanatory variables (especially soil texture) were assigned the same value for all sampling years. Assigning each site's transects the same value eliminated temporal and spatial variation. Nevertheless, the separations between areas are distinct even considering the few identical soil variables.

Although the biplot suggests that many of the climate variables are correlated with each other, using Figure 26.3 alone makes it difficult to decide which variables are highly correlated and should be eliminated from further analysis. To make the decision more objective, we use another tool called the VIF to assess which variables are highly related. In VIF, one explanatory variable is selected as response variable and all others are set as explanatory variables within a linear regression (Montgomery and Peck 1992). The VIF value for the selected variable is given by $1/(1 - R^2)$ where R^2 is the usual R-squared from a linear regression that gives the amount of variation explained by the regression model (Chapter 5). For each analysis, a different explanatory variable is set as response variable in the regression. If there are 20 explanatory variables, this process is carried out 20 times. The underlying idea is that if three explanatory variables Z_1, Z_2 and Z_3 are highly correlated, and one of them is set as the response variable (say Z_1) in the regression model while all others (Z_2 and Z_3) are set as explanatory variables, then the R^2 of the linear regression model will be high and the VIF value from the model will be large. We do the same calculations with Z_2 set as a response model and Z_1 and Z_3 as explanatory variables to get a VIF value. For Z_3, we repeat the process again. The VIF values obtained when Z_1, Z_2 or Z_3 are used as response variables are compared in a table. A high VIF value is a serious indication of collinearity because it means that the variation in the response variable is explained well by the other variables. For the 21 explanatory variables in this study, 9 have VIF values that make them candidates for elimination (Table 26.2). Because there is no real cut off level for the VIF, this decision is subjective; but some statisticians suggest that values higher than 5 or 10 are too high (Montgomery and Peck 1992). Values greater than 50 definitely require elimination from further analysis.

So, where do we start? If the variables can be combined into related groups, we could subjectively remove a variable (or variables) from each group. For example, the variables here are grouped by substrate, temperature, precipitation, and soil characteristics so we could arbitrarily decide to remove one or more variables from each group. Alternatively, we could use a backward selection method to remove one variable at a time, and recalculate VIF after each iteration until we obtain a set of variables that are not collinear anymore. We should keep in mind that if we remove one variable, and repeat the analysis, all VIF values will change. With this dataset, we decided to use the backward selection process. The variable with the highest VIF value was eliminated first. VIF values were then recalculated for the remaining variables, and the process was repeated until all VIF values were smaller than five. The following variables were omitted from further analysis in order from first to last: FallTmin, ELEV, SumTmin, SprTmin, WinTmin, SumTmax and WinTmax. The VIF values of all remaining variables are reasonably low (Table 26.3) indicating that there is not strong collinearity among them. We will work with these variables in the remaining analyses using the richness.

In addition to collinearity, we need to check whether this dataset needs any transformations (Chapter 4). Cleveland dotplots (Chapter 4) of the selected explanatory variables in (Table 26.3) showed that none of the explanatory variables had extreme observations. Therefore, we do not need to transform any of the explanatory variables.

Table 26.2. VIF values for the explanatory variables. The VIF value for PCTCLAY could not be calculated due to perfect collinearity.

Variable	VIF	Variable	VIF
ROCK	2.929	SumTmax	24.978
LITTER	3.924	WinTmax	46.746
ML	1.832	FallTmin	149.896
BARESOIL	1.953	SprTmin	77.184
FallPrec	7.322	SumTmin	135.605
SprPrec	2.519	WinTmin	56.734
SumPrec	4.570	PCTSAND	4.773
WinPrec	4.052	PCTSILT	8.752
FallTmax	43.045	PCTOrgC	2.438
SprTmax	40.199	ELEV	91.398

Table 26.3. Final VIF values for the selected explanatory variables.

Variable	VIF	Variable	VIF
ROCK	2.522	WinPrec	2.087
LITTER	2.943	FallTmax	3.233
ML	1.599	SprTmax	3.064
BARESOIL	1.526	PCTSAND	4.046
FallPrec	2.171	PCTSILT	3.311
SprPrec	1.502	PCTOrgC	2.210
SumPrec	1.560		

With collinearity and transformation issues addressed, the next step is to develop a diversity index that can be used to model general relationships between species in the bunchgrass community and the environmental variables. To begin this process, we need to address a very common problem with plant data; namely, a species dataset with a high percentage of zeros. The bunchgrasss communities in these two areas have 93 species over time and 99 observations per species. Each transect has relatively few species compared with the large number of species in the overall study, so over 90% of the species observations are equal to zero (Figure 26.4). A large number of zeros creates both statistical and ecological problems. Double zeros, or species absent from two transects, does not necessarily mean that the communities are more similar than two others with and without the species. Zeros during sampling also do not indicate whether an environment is unfavourable for a certain species or whether it was just not found at a particular location. This is the classic "zero truncation" problem in ecology. With only 10% of our total observations having a species present, multivariate methods like principal component analysis, redundancy analysis, (canonical) correspondence analysis, and discriminant analysis are not very useful because each of these analyses is sensitive to double zeros (Chapters 12 and 13). For the data that actually do have values in this study, over 32% have very small values (i.e., frequency <1.0%). With data that have many small values and a few large values, a diversity index like the Shannon-Weaver or total abundance might not be appropriate because each is highly

influenced by a few observations. A more appropriate measure of association for these data is a measure that converts all values to presence–absence. Both the Jaccard and the Sørensen indices convert data to presence–absence. We tested both measures on our data and found very little difference in the results, which we expected because these indices have similar definitions (see Chapter 10). Therefore, we decided to use the Jaccard index. For our later multivariate analyses, the Jaccard or Sørensen indices are also more suitable measures of association than either the Euclidean or the Chi-square distance function because they handle data with many zeros better.

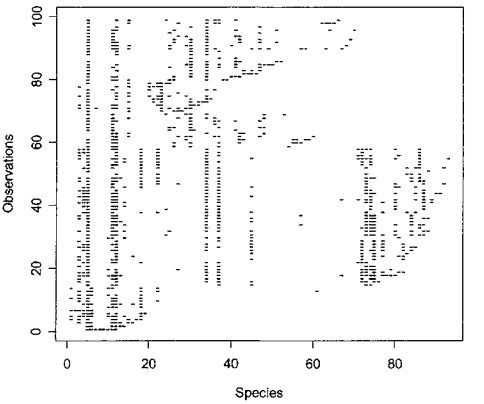

Figure 26.4. Species presence in the dataset. The horizontal axis shows the 93 species in their spreadsheet order and the vertical axis the 99 observations within the transects. A '-' indicates a non-zero value for the species. If all observations had values greater than zero, the graph would be completely filled with dashes.

We are now to a point where we can start exploring whether species richness (i.e., our representation of biodiversity) relates to the explanatory variables and whether time factors into the relationship at all. Pairplots are a good tool to visually look for patterns between our diversity index (richness) and the environmental variables that we selected with the VIF (Table 26.3). The patterns in the top row of the pairplot in Figure 26.5 represent the relationship between richness and the explanatory variables. There are no clear patterns. The other panels can be used to visually assess collinearity, and the lower diagonal panels of the pairplot give the corresponding correlation coefficients. A high (absolute) value (>0.8) indicates that two explanatory variables are linearly related. Whereas these correlations and the previously calculated VIF values measure *linear* relationships among explanatory variables, the upper part of this pairplot can be used to assess whether there are *non-linear* relationships among the explanatory variables. In this case, ROCK and LITTER might have a non-linear relationship. Note that a correlation coefficient is only measuring linear relationships.

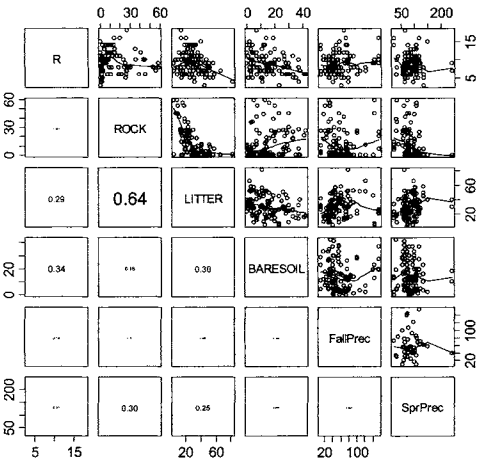

Figure 26.5. Pairplot for species richness and selected explanatory variables. A smoothing curve was added to aid visual interpretation. Other environmental variables are not shown to conserve space. *R* is the species richness. Values on edges represent the units of the variables. The font size of the (absolute) correlation coefficients in the lower diagonal is proportional to their value.

26.4 Linear regression results

In this section, we relate species richness to the explanatory variables using linear regression and try to find the subset of explanatory variables that best explains richness. We start by using the AIC to find the most optimal model. Then we validate the model and explore whether it can explain temporal patterns when 'time' is added as a factor.

Model selection

Recall from Chapter 5 that the AIC measures the goodness of fit but, at the same time, uses the number of explanatory variables as a penalty. The lower the AIC is, the better the model. To find the model with the best AIC, we can either use (i) a backward selection where we start with all explanatory variables and drop one at a time, (ii) a forward selection where we start with models containing only one explanatory variables and add one at a time, or (iii) a combination of forward and backward selections that drops variables at any point in the analysis but can add each back in again at a later stage in the analysis. It could be argued that if the dataset has less than 20 explanatory variables, a fourth option would be to use every possible combination of explanatory variables. This option is not available if analyses are performed using the major statistical software packages, although

programming such an algorithm is not difficult if you choose to use this option. We decided to use option three. Our resulting optimal linear regression model has only five environmental variables that relate significantly to richness at the 5% level. The explained variation in the species richness is 54% (R^2), which is reasonably good. Note that sometimes the AIC comes up with a final model in which some of the explanatory variables are not significantly different from 0 at the 5% level. If this were the case, we would have to further remove explanatory variables using the F-test for comparing nested models (which means that we would have used two different selection criteria to find the optimal model). As all explanatory variables are significant in our analysis (see below), we do not have to do any further model selections.

Variable	Estimate	Standard error	t-value	p-value
ROCK	−0.081	0.017	−4.523	<0.001
LITTER	−0.066	0.022	−2.945	0.004
BARESOIL	−0.102	0.020	−5.022	<0.001
FallPrec	0.015	0.006	2.364	0.020
SprTmax	−0.534	0.089	−5.959	<0.001

Model validation

We now move on to the model validation process, which was outlined in Chapter 5. Here we verify the homogeneity, normality and independence assumptions, and check for model misspecification. The residuals of the optimal linear regression model are plotted against the fitted values (Figure 26.6-A). Recall from Chapter 5 that Figure 26.6-A and Figure 26.6-C can be used to assess homogeneity. In this case, there is no violation of homogeneity because the spread of the residuals is nearly the same everywhere. So there is no need to consider more complicated models like generalised linear modelling with a Poisson distribution. We also have normality (Figure 26.6-B). There are no influential observations as judged by the Cook distance function (Figure 26.6-D).

Next we check for misspecification in our optimal model. Misspecification is indicated when graphs of the residuals versus individual explanatory variables show distinct patterns. For most of the explanatory variables used in our model, there is no obvious pattern. However, the residuals versus ROCK and SprTmax both show pattern. The question is, how serious this is? We fitted a smoother with 95% confidence bands to the residuals (Figure 26.7). The width of the confidence bands indicate that '0' is almost entirely within the 95% bands, so the smoother is not significantly different from 0. Only a few samples are different from 0, and one might argue that this could happen by chance. In this case, the patterns seem weak enough for us to ignore, but others might consider them borderline for taking further action. If the residual patterns were stronger, we would address the problem by (i) adding a quadratic term for ROCK and SprTmax, (ii) using a smoothing function for ROCK to get an additive model (Chapter 7), or (iii) adding interactions (Chapter 5) to our model. Some of these options would require generalised linear modelling or generalised additive modelling procedures, which are covered

in Chapters 6 and 7. As stated, though, we do not consider the residual patterns strong enough to warrant these techniques, so we will proceed to our most important question. Do the residuals show any patterns over time? If so, then we cannot use the linear regression model because standard errors, t-values and p-values might be seriously inflated (Chapter 16).

Time effects

The plant communities at each transect are not only affected by environmental factors, but they are also affected by changes in biotic and abiotic conditions over time. Because the data are taken at irregular intervals, they are not as conducive to time series analysis as other datasets in this book. There are ways, however, to evaluate the effects of time on this regression model. A scatterplot of species richness versus time (Figure 26.8) shows some clear patterns that the linear regression model will hopefully explain. If it does not, we violate the assumption of independence and we may commit a type I error (Chapter 16).

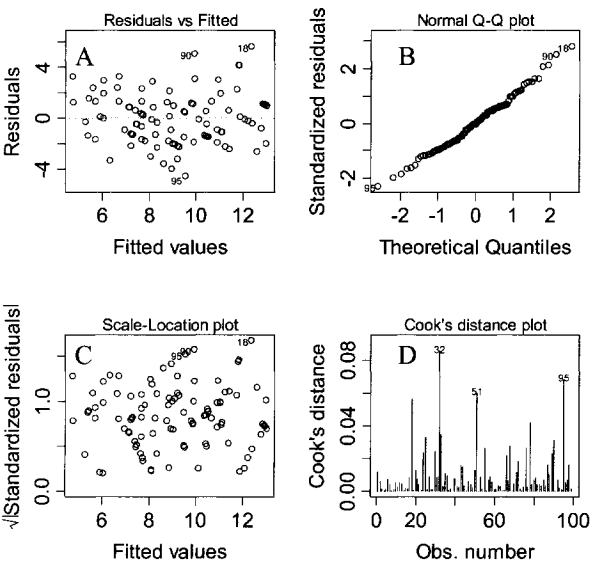

Figure 26.6. Model validation for the optimal linear regression model. Panels A and C indicate that there is no violation of homogeneity. Panel B indicates normal distribution. Panel D shows no outliers using Cook's distance function.

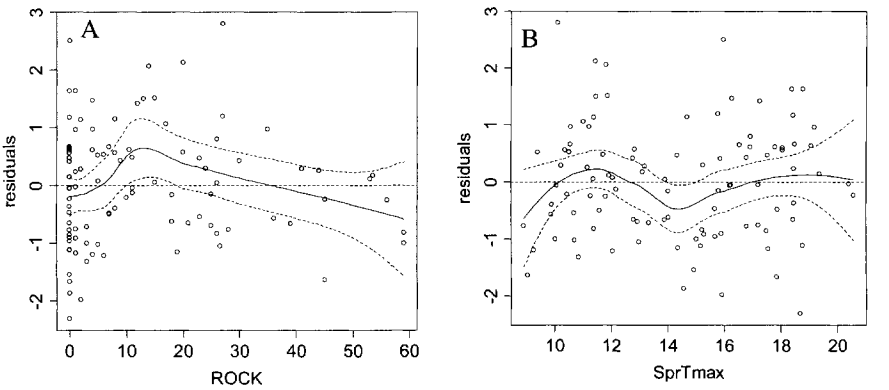

Figure 26.7. Distribution patterns of standardised residuals for the optimal linear regression model versus ROCK (A) and SprTmax (B). The LOESS smoother (solid line) and 95% confidence bands (dotted lines) are shown. All other explanatory variables had smoother lines that were straight horizontal at 0.

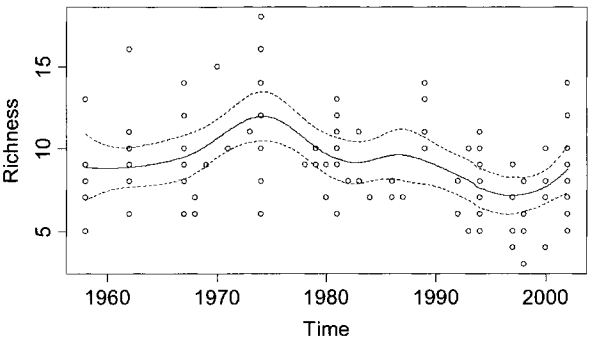

Figure 26.8. Scatterplot for species richness versus time. A LOESS curve with a span width of 0.5 (solid line) was added to aid visual interpretation. The dotted lines represent 95% confidence bands for the smoother.

The easiest way to assess whether the linear regression model explains the temporal pattern in the species richness is to make a plot of the standardised residuals versus time. In our case, this plot does not show a clear temporal pattern (Figure 26.9). A LOESS smoother with a span of 0.75 added to the diagram is almost a straight line at zero indicating that time is not important in the linear regression. Various other tools are available to objectively test whether the residuals contain any temporal pattern. One such tool, GLS, is the subject of the next section.

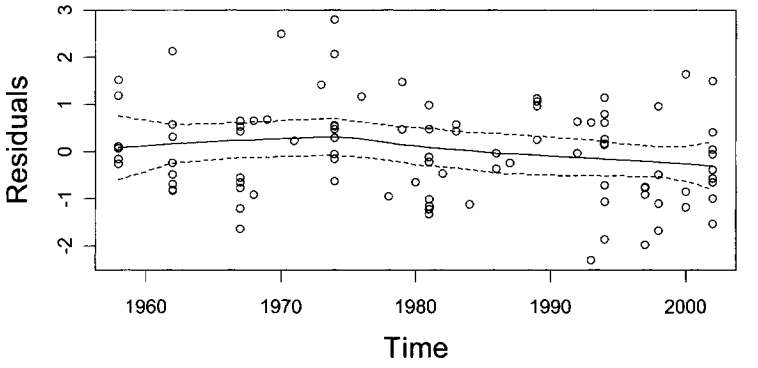

Figure 26.9. Scatterplot of standardised residuals versus year. A LOESS smoother with a span of 0.7 was added. The dotted lines are 95% confidence bands of the smoother.

26.5 Generalised least squares results

In this section, we explore tools that test for a temporal pattern in the residuals. First, we added 'TIME' as a covariate within the linear regression model and tested for its effects. With the YNP and NBR data, 'TIME' was not significant at the 5% level. Alternatively, we could extend the linear regression model so that the residuals are *allowed to* have a temporal pattern, which means that we relax the independence assumption in linear regression. This is done using GLS. When we use GLS, however, there are several ways we can impose a correlation structure on the error terms, and these are discussed below.

ARMA error structure

The auto-regressive moving average (ARMA) adds an auto-correlation structure to the noise component of the linear regression. This technique goes beyond the GLS theory explained in Chapter 16, so we will explain it in more detail here. We start with our optimal linear regression model from the previous section that had five selected explanatory variables:

$$R = \alpha + \beta_1 \text{Rock} + \beta_2 \text{Litter} + \beta_3 \text{Baresoil} + \beta_4 \text{FallPrec} + \beta_5 \text{SprTmax} + N_t \quad (26.1)$$

In linear regression, we assume that the noise component N_t is independent and normally distributed. In GLS, we use the same equation as in (26.1), but the noise component is no longer restricted to be independent. Recall from Chapter 16 that an ARMA model of order (p,q) for a time series Y_t is defined as:

$$Y_t = a_1 Y_{t-1} + a_1 Y_{t-1} + \ldots + a_p Y_{t-p} + \varepsilon_t + b_1 \varepsilon_{t-1} + b_2 \varepsilon_{t-2} + \ldots + b_q \varepsilon_{t-q} \qquad (26.2)$$

where ε_t is independently and normally distributed noise. In this model, a time series Y_t is modelled as a function of its past values, its error component, and its past error components. The p defines the number of auto-regressive (past Y values) components and q the number of past error terms. In GLS, this modelling framework is applied on the error component N_t. Hence, Y_t is replaced by N_t in equation (26.2). So, in linear regression we are imposing the condition that N_t is independent normal noise. In GLS, we are decomposing N_t into small building blocks. These building blocks consist of p past terms of N_t, and q building blocks that are independently normally distributed. The question then becomes 'how many of these building blocks do we need?' Or more formally, how many auto-regressive (p) and moving average (q) terms do we need? The easiest way to determine p and q is to apply the GLS with different values of p and q, and choose the one with the lowest AIC (Table 26.4). The AIC values indicate that the model with $p = 1$ and $q = 0$ has the lowest AIC so it is the 'best' model. However, the AIC of the model with $p = q = 0$, which is the linear regression model, is only marginally larger. The question then becomes whether the decrease in AIC is enough to discard the linear regression model in favour of a more complex GLS model. If the difference in AIC values of two models is smaller than 2 (as is the case in Table 26.4), general statistical consensus dictates using the simpler linear model.

Another more formal way to compare the two models uses a likelihood ratio test (see also Chapter 8 for comparing mixed effects models with different random components). The likelihood ratio test indicates that both the ARMA(1,0) and the model without auto-correlation perform equally well (see below), so there is no need to adopt a complicated GLS model over our original linear regression. Before fully committing to the simpler linear regression model, we will explore two additional GLS error structures to see whether they give different effects. The notation GLS(p,q) refers to the ARMA error structure in the GLS.

Model	df	AIC	logLik	L-Ratio	p-value
GLS(1,0)	8	461.60	−222.80		
GLS(0,0)	7	462.95	−224.47	3.34	0.06

Table 26.4. AIC values obtained by using an ARMA(p,q) structure on the error component in the GLS model ($p = q = 0$ is the linear regression model). Note that the ARMA(1,0) is an auto-regressive model of order 1. The correlation structure in the GLS function within R software can cope with missing values.

	$q = 0$	$q = 1$	$q = 2$	$q = 3$
$p = 0$	462.95	463.69	464.13	466.12
$p = 1$	**461.60**	463.43	465.26	468.16
$p = 2$	463.36	465.24	467.21	469.11
$p = 3$	465.29	466.13	467.66	463.67

Other error structures

In Chapter 35, we examined a salt marsh time series using a special case of GLS error structure, namely the AR(1,0) structure. In the previous section, we examined the effect of time on the residuals using ARMA. In this section, we examine two additional approaches that test the relationships of time lags and covariance between years to the residuals in GLS. The so-called 'compound symmetry approach' is used if it can be assumed that the correlation (of the error) between two different years (within a transect) is equal to ρ whatever the time lag. It is defined as

$$cor(N_t, N_{t+j}) = \rho \text{ for any value of } j \tag{26.3}$$

The indices t and j stand for time and time lags, respectively. So the correlation between N_t and N_{t+1} is ρ, and the same holds for the correlation between N_t and N_{t+2}, N_t and N_{t+3}, etc. For the YNP and NBR data, the estimated value of ρ using the compound symmetry approach is −0.03 and the AIC is 464.66. This AIC value does not improve on the AIC for the linear regression model. The compound symmetry correlation structure tends to be too rigid and simplistic for most time series datasets (Pinheiro and Bates 2000).

GLS error structure can also be defined using a general correlation matrix. A correlation matrix allows for more flexibility than compound symmetry but at a much higher price in terms of estimated parameters. If there are N different years, the general covariance matrix for the error component is of the form:

$$V = \begin{pmatrix} v_{1,1} & v_{2,1} & \cdots & v_{N-1,1} & v_{N,1} \\ v_{2,1} & v_{2,2} & & \vdots & \vdots \\ \vdots & & \ddots & \vdots & \vdots \\ v_{N-1,1} & v_{N-1,2} & \cdots & v_{N-1,N-1} & v_{N-1,N-1} \\ v_{N,1} & v_{N,2} & \cdots & v_{N,N-1} & v_{N,N} \end{pmatrix} \tag{26.4}$$

The component $v_{i,j}$ represents the covariance between two years i and j (see Chapter 16). If we use this structure, we allow the residuals to covary in different ways depending on the time lag between them. To use the correlation matrix, however, the sampling protocol must meet fairly strict guidelines. All sample years should be equidistant apart, and each transect should be measured in each sample year. In this case study, the transects were not sampled regularly in either area nor were the sample intervals the same between the two areas, so the software was not able to converge. All other error structures discussed in this section were able to deal with irregular sampling and missing years.

After trying many options within the GLS model, we still do not have models that relate species richness and environmental variables any better than our original linear regression model. We accept the linear regression model then as the best and simplest model to address which environmental variables best correlate with changes in biodiversity in these plant communities.

26.6 Multivariate analysis results

Analysing the relationship between compositional and environmental changes for bunchgrass communities requires a multivariate analysis approach. In a particular type of multivariate analysis called ordination, the arrangement of these values within a multidimensional graph can help us see whether the community data are structured or contain patterns. The patterns may reflect a community's response to multiple environmental changes over time or more subtle biological interactions. Deciding which multivariate analysis to use, however, depends on the data and focus of the study. We could apply principal component analysis or redundancy analysis on these data (Chapter 12), but these methods are sensitive to double zeros (Chapters 10 and 12). Our data matrix contains over 90% zeros, and therefore, we chose an asymmetric measure of association (i.e., joint absence of species at two sites is not contributing towards similarity), namely the Jaccard index (Chapter 10) in combination with non-metric multidimensional scaling (NMDS).

Using NMDS does make it more difficult to produce triplots and visualise the relationships among observations, species and explanatory variables. These relationships must be informally inferred by comparing the transect positions within the NMDS diagram. Determining how time relates to species gradients is also more difficult using NMDS. For some datasets, it might be more appropriate to use an alternative method, such as db-RDA, to relate distance matrices directly with environmental variables (Legendre and Anderson 1999).

NMDS

NMDS calculates a distance matrix from the species data based on a chosen measure of association. It then uses an iterative process to order the distance matrix in n-dimensional space to find a configuration that matches the distance matrix as closely as possible (Chapter 15). The distance matrix and the position assigned in the n-dimensional space never match perfectly. If they did, a plot of the dissimilarity of each point in the original distance matrix versus the distances between them in the calculated ordination space (also called a Shepard diagram) should be a straight line (or curved for NMDS). To measure the discrepancy of the points from this line, various measures of goodness-of-fit have been proposed. They all use quadratic functions of the original distance and calculated distance in the ordination space (see Legendre and Legendre (1998) or Chapter 15 for details). These measures of goodness of fit are also known as 'STRESS'. High STRESS values indicate a poor fit between the original data structure portrayed by the distance matrix and the ordered positions in n-dimensional ordination space.

In this section, we have two options. We can either apply the multivariate analysis on each of our areas separately or on both areas together. Each option gives different perspectives on community and environmental change. In the previous sections, the univariate analysis clearly revealed a difference in composition and environmental conditions between Yellowstone National Park and the

National Bison Range. Combining the areas into one NMDS analysis will likely have the same result. Analysing the two areas separately, however, has the potential of providing more information on how similar local communities are and enables us to track how composition varies within each area over time. It also gives a different perspective on whether climatic and soil variables are important to species presence in the community locally. Although analysing two areas has double the amount of information, graphs, and numerical output to interpret, the benefit for this study is a clearer perspective on the variations within each transect over time.

If NMDS is used, one has to choose the number of axes. In an initial analysis, we used two axes but the STRESS was 0.214 for the YNP data and 0.254 for the NBR data. These values are rather large (Chapter15), and therefore, we present results obtained by using three axes. The problem with three axes is how to present them. We tried three-dimensional scatterplots and even used spin graphs (these are graphs in which the mouse can be used to rotate the three-dimensional scatterplot) but found three separate two-dimensional scatterplots presented the results most clearly.

When NMDS is applied to data from YNP, the communities of most individual transects form tight clusters and the grazed and ungrazed transects at each area are separated (Figure 26.10). Transects from the warm, dry, low elevations (1-2) also separate from all others. Figures 26.10-A and 26.10-C show that transects 1 and 2 are the most dissimilar from the other transects. Along the third axis we can see a difference between transects 5 and 6 versus the other transects. The STRESS using three axes was 0.15, which is still relatively large. We felt the STRESS was too large to make any statement on changes within transects over time.

For the NBR data (Figure 26.11), transects differ much more among themselves (i.e., same numbers are more spread out) than at YNP over time. Panel C seems to indicate differences among transects 9, 10 and 11, which are all in different pastures of the NBR. Again, the STRESS is relatively large using three axes (0.19), so care is needed with the interpretation. Conclusions about change over time are again probably impractical. Our next problem is to find out whether there is a relationship between the environmental variables and the species that could explain these patterns. This requires a Mantel test (Chapter 10).

The Mantel test

The Mantel test evaluates correlations between two distance matrices of the same size. In this study, we have two data matrices. One matrix, called the '**Y**' matrix, contains the species data. It is (number of sites)-by-(number of species) in size. The second matrix is called the '**X**' matrix, and it contains the explanatory variables. Its size is the (number of sites)-by-(number of environmental variables). When a distance matrix is created from each of these matrices using (for example) the Jaccard index for the similarity measure and Euclidean distance measurements, two equal-sized matrices result ($\mathbf{D}_Y$ and $\mathbf{D}_X$).

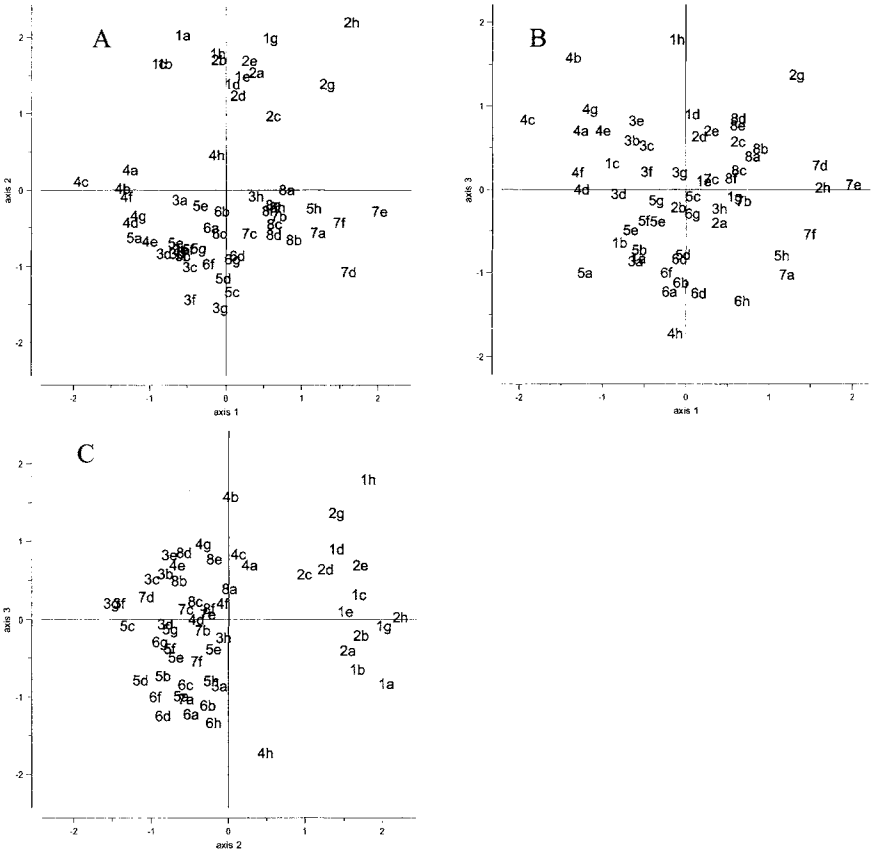

Figure 26.10. NMDS using the Jaccard index for the Yellowstone National Park data. The STRESS value equals 0.15. A: Axis 1 versus 2. B: Axis 1 versus 3. C: Axis 2 versus 3. Alphabetical designations on transect numbers indicate order of sampling (a = first sample).

The ij^{th} element in $\mathbf{D_Y}$ represents the dissimilarity between two observations in terms of species composition, and the ij^{th} element in $\mathbf{D_X}$ the dissimilarity between the same two observations in terms of environmental condition. The Mantel test evaluates how these two distance matrices correlate with each other. To assess the significance of the correlation coefficient, a randomisation process is applied (for details see Chapter 10). It is important to realise what exactly we are testing with the Mantel test. The underlying assumptions are as follows. H_0: distances between the observations in terms of species composition are not linearly related to distances between the observations in terms of environmental conditions. H_1: distances among the observations in terms of species composition are linearly correlated to the distances between the observations in terms of environmental conditions.

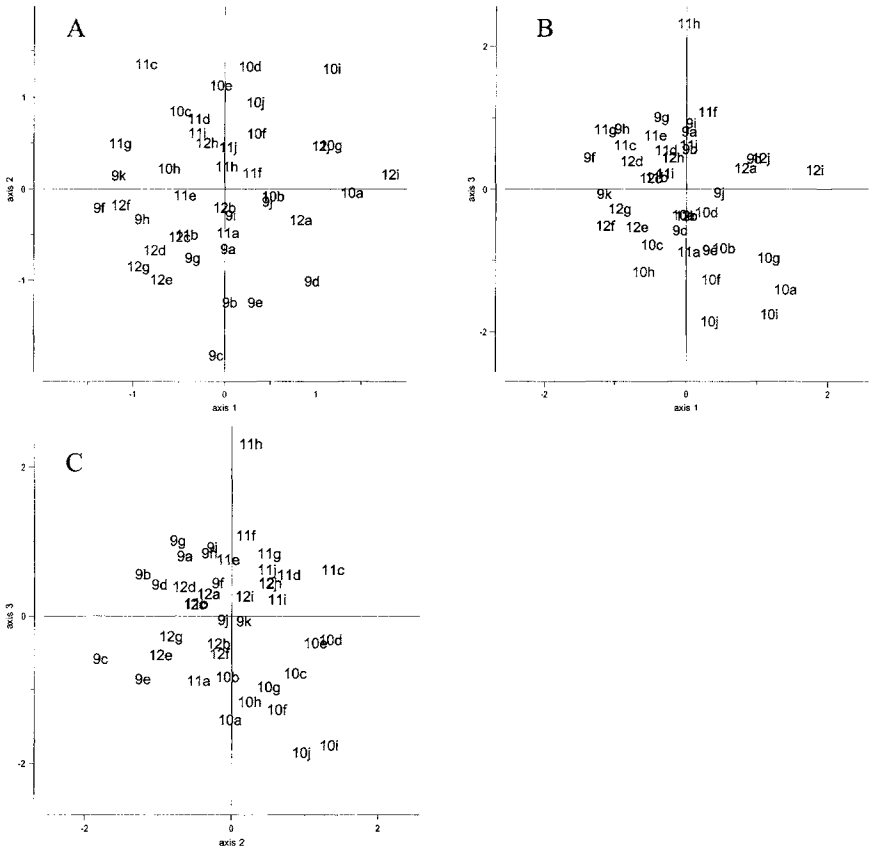

Figure 26.11. NMDS using the Jaccard index for the National Bison Range data. The STRESS value equals 0.19. A: Axis 1 versus 2. B: Axis 1 versus 3. C: Axis 2 versus 3. Alphabetical designations on transect numbers indicate order of sampling (a = first sample).

We applied the Mantel test on the analyses of the separate areas and tested the significance of our correlations between the two distance matrices with a permutation test. The correlation between the elements of $\mathbf{D_Y}$ and $\mathbf{D_X}$ and their significances differed for YNP and the NBR. For YNP, the correlation coefficient was $r_M = 0.378$ ($p < 0.001$). For the NBR, it was $r_M = 0.116$ ($p = 0.06$). For the YNP data, we reject H_0 at the 5% level and conclude that there is a linear relationship between the two distance matrices. For NBR there is no evidence to reject the null hypothesis at the 5% level.

In the previous paragraphs we compared a matrix consisting of distances between observations representing species composition with a matrix reflecting differences in terms of environmental conditions. The question now is 'which explanatory variables should be used when we obtain $\mathbf{D_X}$ so that the correlation

between $\mathbf{D_X}$ and $\mathbf{D_Y}$ is as high as possible?' To answer this question, we first calculate (i) a matrix $\mathbf{D_X}$ and (ii) the correlation with $\mathbf{D_Y}$ for all possible combinations of explanatory variables. We then select the combination of explanatory variables that gives the highest correlation coefficient r_M. This process was called BIOENV by Clarke and Warwick (1994). If there are numerous explanatory variables, the computing time for the BIOENV procedure can be long. Sometimes an automatic forward selection procedure called BVSTEP is preferred to speed the process up (Chapter 10). We applied BVSTEP on data from the two different areas. For YNP, the forward selection procedure first selected WinPrec and then PCTSILT. No other variables were needed to explain the correlations. The correlation coefficient between the elements in $\mathbf{D_Y}$ and $\mathbf{D_X}$, in which $\mathbf{D_X}$ was only determined by WinPrec and PCRSILT, was $r_M = 0.408$. For the NBR, the selected variables, in order of importance, were SprPrec, ROCK, and PCTOrgC. The correlation coefficient using these three variables was $r_M = 0.249$. For the National Bison range, we could not use the explanatory variable PCTSILT because it had the same value in all transects.

26.7 Discussion

Many environmental stresses shape the bunchgrass communities of Montana's temperate grasslands during half a century. In this case study, climatic variation, substrate changes, and soil texture are all important to explaining species composition within these communities. When composition is portrayed as a single biodiversity measure (richness) within linear regression, both substrate and climatic factors relate to its variation. Changes in the frequencies of rock, litter, and bare soil at each transect are significant in the linear regression model and relate inversely to richness. Differences in substrate characteristics between the two study areas are not surprising. Yellowstone National Park and the National Bison Range differ in bedrock and soil textures. The YNP transects are located within lower Tertiary igneous and metamorphic rocks. The NBR exists within fine-grained quartzite, argillite, banded slate, and inter-bedded sandstone of the lower pre-Cambrian. Differences in bedrock affect soil textures and the amount of rock at each site. They also affect soil permeability and moisture retention throughout the growing season. The amounts of litter or bare soil affect moisture retention, surface temperature, and nutrients available for plant growth. The more bare soil, the more rapidly moisture is lost from the surface and the higher the surface temperatures during the growing season. As litter increases, moisture is lost more slowly from the soil surface and nutrients are added to the soil.

Biodiversity, as measured by species richness, is also negatively related to maximum spring temperatures and positively related to the fall precipitation. In this area of the northern Rocky Mountains, fall and winter precipitation are very important to re-charging soil moisture for plant growth to begin in spring.

Although time should be significant in studies on community change, neither adding time as a covariate in the linear regression model nor including different error structures improved the model.

The NMDS shows how similar (or dissimilar) each of the communities are to each other; but because of the relatively large STRESS, it cannot be used in this particular study to assess how much the repeat samplings vary at each location.

The three-dimensional YNP ordination graph has several interesting patterns to explore. The grazed plots form distinct groups among themselves, but the groups are also well inter-mixed with clusters of non-grazed plots. This suggests that management type makes some difference in composition in YNP, but that it is not as important as some other factors, such as elevation and moisture, to park grasslands as a whole. The NBR data, which is all grazed on a rotational schedule, varies much more between individual transects.

A more formal approach of determining which environmental variables affect communities in the NMDS species (BVSTEP) shows that winter precipitation and percentage of silt in the soil are significantly related to species presence in YNP while spring precipitation, frequency of rock, and percentage of organic carbon in the soil are most significant in the NBR.

The univariate and multivariate analyses used in this study give slightly different perspectives on how environmental variables affect these bunchgrass communities. Analysing the long-term monitoring records with univariate analyses shows that substrate conditions and precipitation are very important controls on species presence in these grasslands. Analysing the records with the multivariate NMDS shows that temperature, grazing, and non-native species also affect these communities. Whether the sites are analysed together or separately, the factors that affect these bunchgrass communities over several decades are complex and multiple perspectives are essential to explore that complexity.

Acknowledgement

Research was conducted in Yellowstone National Park under research permit Yell-2002-SCI-5252. Roy Renkin and the Yellowstone archive staff helped access historic transect data and historic photographs. Lindy Garner, Lynn Verlanic, and Bill West of the U.S. Fish & Wildlife Service were instrumental in locating historic records and resampling transects in 2002 on the National Bison Range. Dr. Paul B. Alaback, professor of forest ecology at The University of Montana, provided guidance and advice throughout this project. SOGS interpolations were run by Matt Jolly, Numerical Terradynamic Simulation Group, The University of Montana.