Chapter 10. The Systematic Application of a Multivariate Perspective to Understanding Plant Diversity Patterns in Ecological Communities

This chapter illustrates the systematic application of a multivariate perspective using SEM to explore a topic. In this presentation, the statistical details of the analyses will be ignored; these have been presented in earlier chapters or can be found in the various publications referenced throughout. Here, the emphasis is on illustrating the broad enterprise of developing, evaluating, refining, and expanding multivariate models in order to understand system behavior and regulation. Throughout, the focus will be on the research enterprise rather than the analytical details. Thus, the philosophy and practice of SEM will be in the forefront, while the analysis of covariances, maximum likelihood, and mathematical details will be deemphasized.

Background Studies and Findings

In 1992, Laura Gough and I conducted a study designed to examine the relationship between plant community biomass and species richness. This work was conducted in coastal marsh communities. The purpose of this study was to first characterize the relationship between biomass and richness. Then we planned to determine the role of competition in controlling the relationship. We expected that we would find a unimodal relationship between biomass and richness, primarily because of several key papers that had been published previously (Al-Mufti et al. 1977, Huston 1980, Wheeler and Giller 1982, Moore and Keddy 1989, Wisheu and Keddy 1989, Shipley et al. 1991, and Wheeler and Shaw 1991). We also expected this relationship because there were several competing theories attempting to explain this phenomenon (Grime 1979, Huston 1979, Tilman 1982, Taylor et al. 1990, and Keddy 1990). Fig. 10.1 presents one of the early examples that inspired much of the subsequent work on this topic. A further influence on us at that time was work by Paul Keddy and his colleagues (see above references) who were attempting to establish general, quantitative, predictive relationships between community biomass and richness.
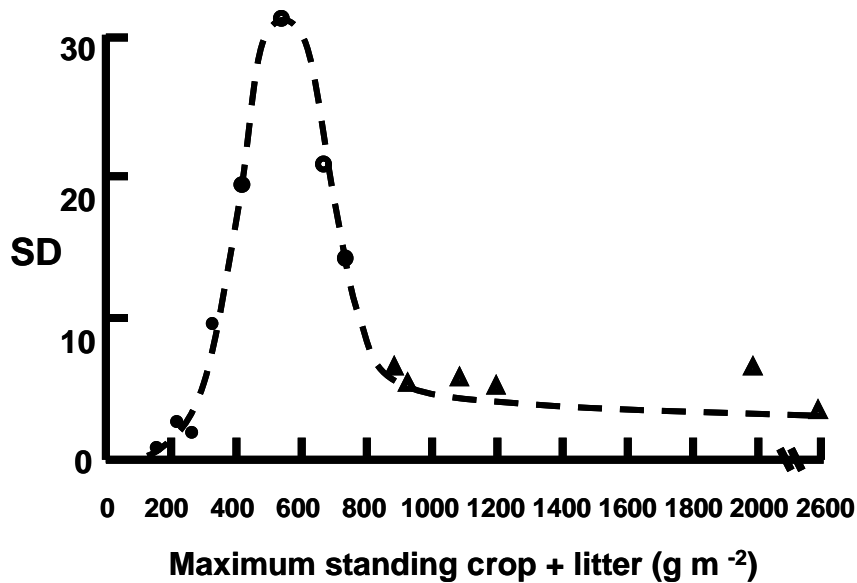
Figure 10.1. Example of the relationship between biomass, measured as the maximum standing crop including litter) and species richness (SD, or species density) presented by Al-Mufti et al. in 1977. Solid circles represent woodland herbs, open circles represent grasslands, and triangles represent tall herbs of open areas.

Combining our data from two coastal riverine systems with the data of one of our colleagues, Kathy Taylor, we found a pattern rather unlike the one we expected (Fig. 10.2). Instead of a unimodal curve, we found a unimodal envelope (see additional discussion of this relationship in Marrs et al. 1996). Also, we found biomass to be rather unimportant as a predictor, but found that species richness was strongly correlated with microelevation and sediment salinity and, to a lesser degree with soil organic matter. Table 10.1 shows the results of the multiple regression that we performed using all of the measured predictor variables. The fit of this multiple regression model to the data is shown in Figure 10.3.
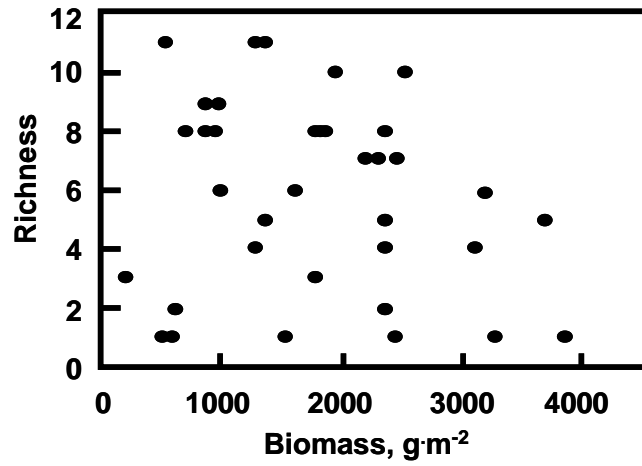
Figure 10.2. Relationship between total aboveground community biomass (live + dead) per m^2 and number of species per m^2 found by Gough et al. (1994) in coastal marsh systems.
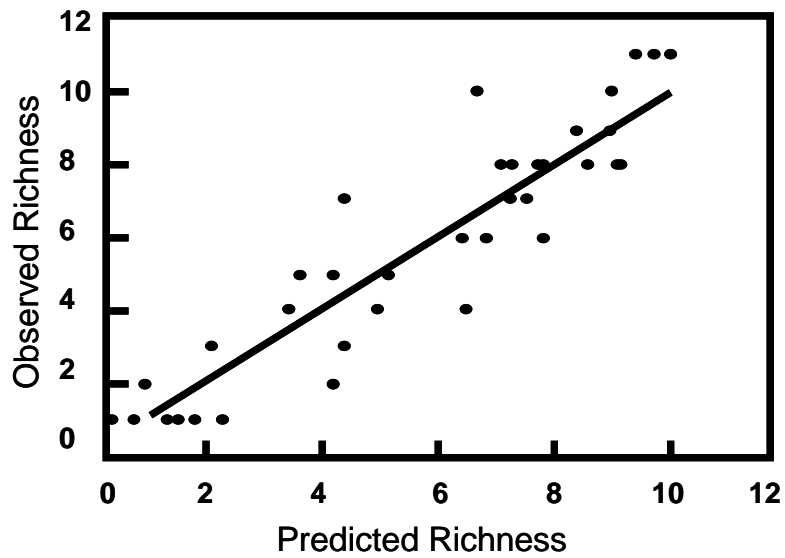

Figure 10.3. Fit of data to the multiple regression model in Table 10.1 (from Gough et al. 1994).

Table 10.1. Multiple regression results for species richness as a function of environmental variables (modified from Gough et al. 1994).

Predictor Variables	Coefficient	Std Error	Cumulative R-square	P <
Constant	-3.90	1.084		0.001
Biomass	-0.0011	0.0003	0.02	0.001
Elevation	3.10	0.377	0.57	0.001
Salinity	0.51	0.137	0.69	0.001
Soil organic	0.052	0.011	0.82	0.001

Based on the results we obtained and our familiarity with wetland systems, we formulated a multivariate conceptual model (Figure 10.4). We further supported this model with additional data about the existing species pools (we used the term "potential richness") and how they varied with salinity. To quote Gough et al. (1994),

In conclusion, in this study we found that biomass was not an adequate predictor of species richness. One reason for this inadequacy appears to be that while stresses such as flooding and salinity may greatly reduce the pool of potential species that can occur at a site, those species that have evolved adaptations to these factors may not have substantially reduced biomass. Thus, we recommend that models developed to predict species richness should incorporate direct measurements of environmental factors as well as community attributes such as biomass in order to increase their applicability.
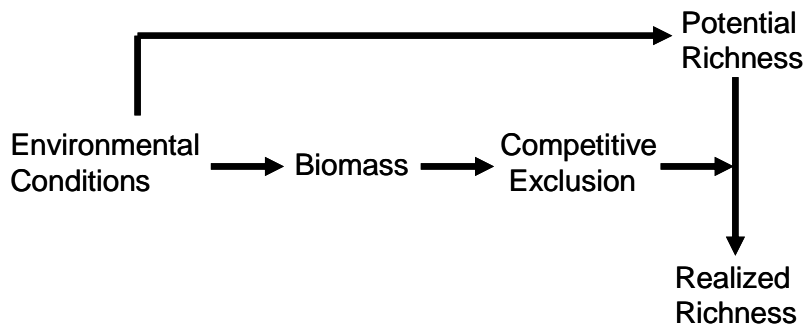

Figure 10.4. Hypothesized conceptual model of factors controlling species richness in plant communities.

Multivariate Hypothesis Formulation and Evaluation

Initial Evaluations of a Multivariate Model at the Pearl River

In 1993, Bruce Pugesek and I initiated a more extensive study designed to evaluate the ideas presented in Figure 10.4 using structural equation modeling. The system to which we wished to apply this model was a coastal riverine marsh landscape located on the shore of the Gulf of Mexico (Fig. 10.5). This landscape contains a number of conspicuous environmental gradients, including gradients in salinity (salt marsh to fresh marsh), microelevation (from plants in deep water to those growing on raised levees), and soil organic content (from sandy soil deposits to organic muck sediments. We also knew from previous studies that natural disturbances were common and resulted from the activities of wild mammal populations as well as from flooding and storm effects. Thus, this system contained a variety of stresses and disturbances of the sort commonly invoked in theories about diversity regulation.

Figure 10.5. Aerial photograph of the Pearl River marsh complex, which was the site of these investigations.

The next step in the process was to convert the conceptual model (Figure 10.4) into a construct model, as represented in Figure 10.6. The

process of specifying a construct model is an important step forward in theory maturation. The conceptual model presents important ideas, however, it is very general and the meanings of the terms are somewhat vague. This is not to say that the construct model in Figure 10.6 is without ambiguity. Models always have a context, some set of physical circumstances under which the model makes sense. I have argued previously (Grace 1991) that the context for models are often not clearly specified in ecological theories, leading to irresolvable debate. The application of structural equation modeling seeks to make concepts and context tangible in stages, which preserves both general and specific perspectives on the problem. This topic will be discussed in more detail in Chapter 12 as it relates to the concept of theory maturation.
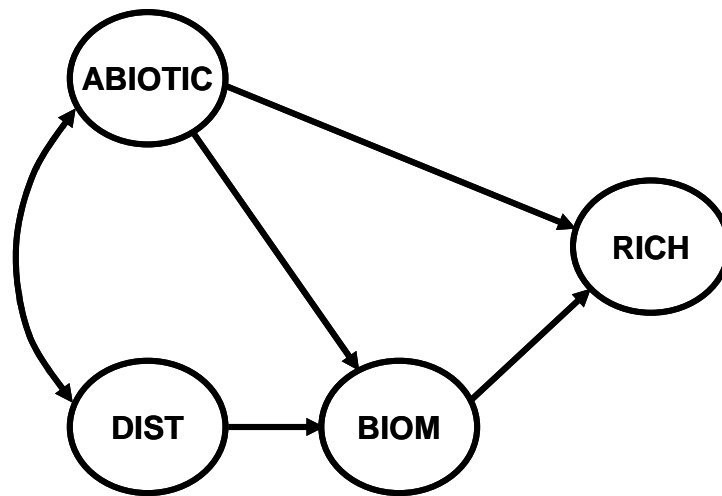

Figure 10.6. Construct model showing major factors presumed to be controlling species richness (taken from Figure 1 in Grace and Pugesek 1997). ABIOTIC refers to abiotic environmental variables, DIST refers to disturbances of vegetation by animals and flooding, BIOM refers to community biomass, and RICH refers to species richness.

One thing that happens when we specify our construct model is that there is an immediate expectation that the concepts represented will have to be made operational. This means it is soon going to be necessary to specify the meaning of the concepts by describing exactly how they will be measured. This reality suggested to us a distinction that we wanted to make that was not specified in the conceptual model, that there are two distinctly different kinds of environmental variables, abiotic conditions and disturbances. Making this distinction was actually not a mandatory

requirement, as the use of composites (Chapter 6) does allow us a method for dealing with highly heterogeneous concepts. However, since much attention has been paid to the effects of disturbance versus the effects of stress on richness, the distinction between abiotic factors and disturbance was one we thought would be valuable.
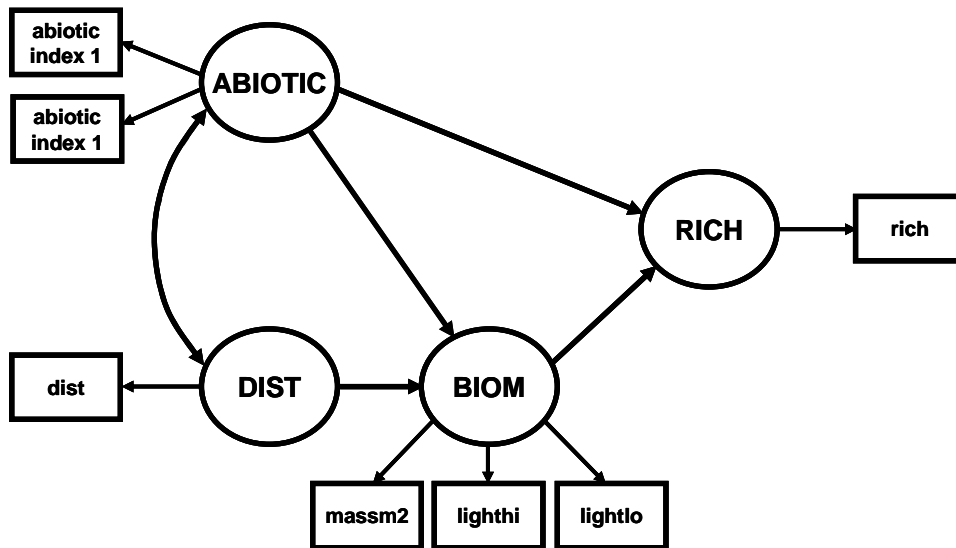

Figure 10.7. Initial structural equation model with latent variables representing concepts and observed variables representing indicators.

The next step in the process was to develop the structural equation model (Figure 10.7). As indicated by the model structure, we specified that abiotic conditions would be measured by combining specific abiotic measurements into indices. We also proposed to characterize the disturbance regime by quantifying the observable signs of disturbance in the vegetation. To measure community biomass, we decided to measure the standing crop of plant material, which was the parameter measured by Al-Mufti et al. (1977), and to also measure the amount of light passing through the vegetation, as an additional measure of the quantity of vegetation. Richness was specified as being measured simply by the number of species we found in 1-m^2 plots.

An important concept associated with SEM is that the observed variables are only required to be indicators of the processes and parameters of interest rather than perfect measures. In order for a measured variable to serve as an adequate indicator, it is only necessary for the relative values of that variable to be correlated with the underlying

process of mechanistic importance. For example, when we specify that the conceptual variable DIST would have an effect on BIOM, what is meant is that the disturbance regime has an effect on community biomass. When a measure of recent disturbance is used as an indicator of DIST, we are proposing that recent indications of disturbance correlate with the disturbance regime over a period of time reflected by the vegetation.

Once the initial structural equation model shown in Figure 10.7 was formulated, we designed a sampling scheme and schedule. Data were then collected over a two-year period to test this model. Only the results from the first year were used in Grace and Pugesek (1997), with the second year's data saved for a subsequent test of whether the model results would hold up over time and how the system changes between years. Data collected included sediment salinity, site microelevation, soil organic and mineral content (components of the abiotic indices), recent disturbance (dist.), measured as the percent of plot disturbed, above-ground biomass per m^2 (massm^2), percent of full sunlight reaching the ground surface in densest part of plot (lightlo), percent of full sunlight reaching the ground surface in sparsest part of plot (lighthi), and the number of species in a plot (rich.).

As is often the case when evaluating multivariate models, the fit between our data and the initial model indicated that it was not adequate. Simply put, our data were not consistent with the expectations implied by the initial model. The part of the hypothesis in Figure 10.7 that failed was the proposition that biomass and light readings can both be used to represent a single conceptual variable. The symptoms of this failure were that lighthi and lightlo were highly correlated, but that massm^2 was not well correlated with either light variable. Because of this, we reformulated our model as shown in Figure 10.8. An important lesson was learned here, regardless of your conceptualization of the problem, if two variables are not consistently and equally well correlated, they will not function as multiple indicators of a single conceptual variable. Of course, we could have used the heterogeneous indicators to represent a composite. However, that did not fit with our objectives in this analysis.

Once the model was reformulated and reanalyzed, another inconsistency between model and data emerged. A large residual correlation between DIST and LIGHT was found to exist. As was discussed in Chapter 8 (see Figure 8.12), this particular residual represents an effect of disturbance on plant morphology that moderates the effect of disturbance on light. In order to proceed further, it was necessary to include a pathway from DIST to LIGHT. Only then was there a consistent match between the relationships in the model and those in the data.
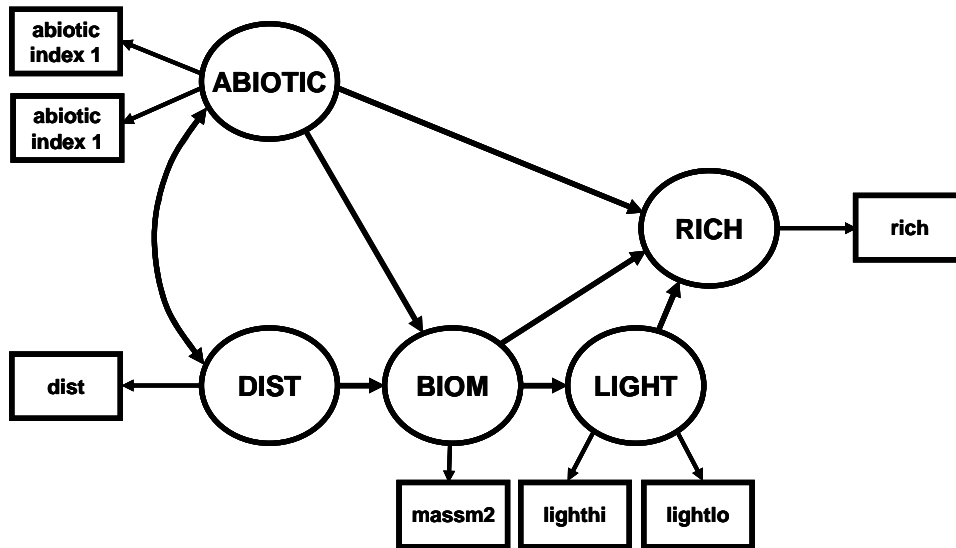
Figure 10.8. Modified full model defining BIOM and LIGHT as separate conceptual variables (from Grace and Pugesek 1997).

The results shown in Figure 10.9 represent the partitioning of covariances in the data as specified by the relationships in the model. Since these results were obtained using a maximum likelihood statistical procedure, they satisfy the criterion of being a simultaneous solution for all relationships. It is not my purpose here to describe all the ecological interpretations of these results. The interested reader can consult the paper by Grace and Pugesek (1997). What I would like to point out, however, is that while this model fits the data, we must conclude that our originally formulated model (Fig. 10.7) did not. Thus, further evaluation is still needed using an independent dataset before we can conclude that our accepted model is valid for the system sampled. As this rather weak conclusion reveals, the demands that SEM places on empirical validation are quite stringent. Stated in another way, the model evaluation philosophy pushes the scientist rather hard to demonstrate that their results have consistent validity (i.e., consistent applicability in different places and over time), not just local application.
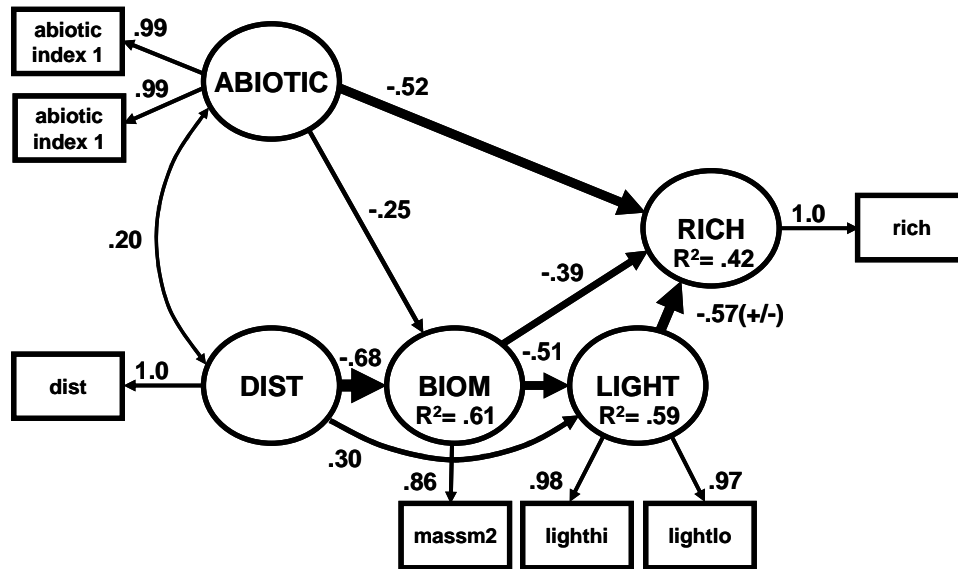
Figure 10.9. Results for accepted model. Standardized partial regression coefficients are given. The (+/-) next to the path from LIGHT to RICH signifies that this path was unimodal.

A More Detailed Examination of the Data

As mentioned earlier, the construct model presented in Figure 10.6 actually represents a family of models that can be represented using a single dataset. To illustrate this point, Figure 10.10 shows the results of a more specific and detailed model in which the abiotic factors were represented individually. The results are based on the exact same data set used to arrive at the more general results in Figure 10.9, except for the fact that the more general model combined the individual abiotic data into indices. Again, no discussion of the ecological interpretations will be presented here. Later in the chapter, I will present an example that used a derivative of this model for the purpose of exploring the role of historical factors in controlling species richness.
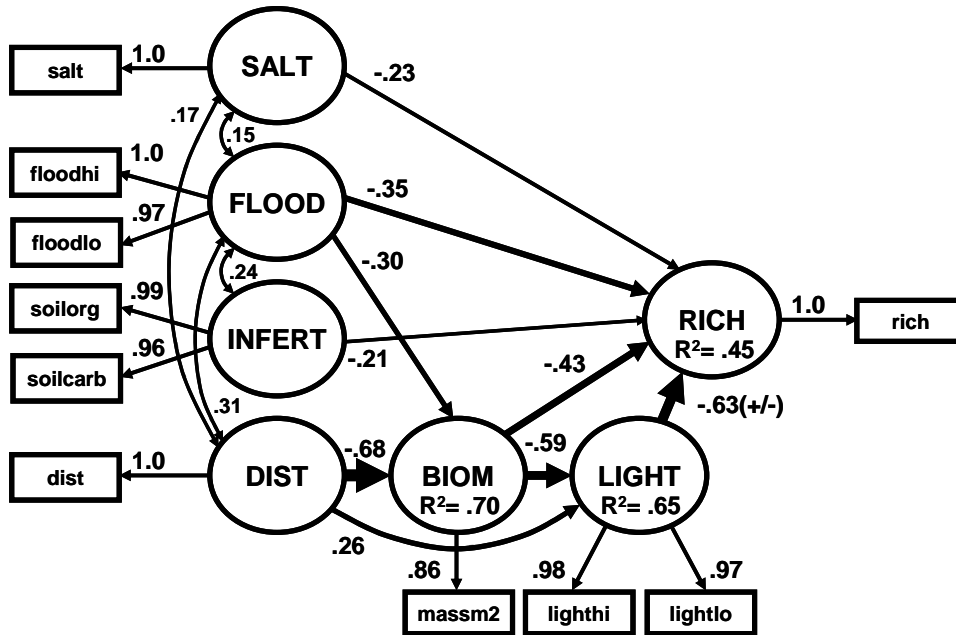
Figure 10.10. Results for a more specific version of the model.

An Experimental Test of the Multivariate Model

There exists a certain amount of experimental evidence that supports the dependency assumptions in the above models (for a review of some of this literature, see Grace 1999). This is not to say that there are no feedback processes that have been omitted; for example, a reciprocal effect of species richness on community biomass. What has been assumed is that the relationship between biomass and richness is asymmetric, with the predominant influence being in the direction specified. Regardless, the principle of theory maturation pushes us to ask whether the results from the nonexperimental studies so far described have any predictive power. Testing predictions based on our accepted models (Fig. 10.9 and 10.10) not only addresses the question of whether there is consistent validity but also represents an opportunity to further refine our model and to determine the limits of its applicability.

Between 1993 and 1995, Laura Gough devised and conducted an extensive experimental study designed to manipulate a number of key variables believed to control species richness in the Pearl River system. The results from these experiments were then used in two ways to evaluate multivariate hypotheses (Gough and Grace 1999). First, her results were compared with predictions from a model like the one in Figure 10.10, except that it was based on the second year's data collected by Grace and

Pugesek. Second, the results were represented as structural equation model of its own. This model was presented in Chapter 9 and can be seen in Figure 9.11.

Figure 10.11 presents two graphs showing how observed values of species richness in Gough's experiments compared to those predicted from nonexperimental data in the broader landscape. When all plots were included (Fig 10.11A), there was considerable scatter and only 35% of the variance was explained. Further analyses showed that this was due to the fact that fencing and fertilizing caused effects that were not quantitatively predicted. With plots subjected to either of these treatments removed (Fig. 10.11B), the remaining treatments, which included those subjected to changes in salinity and flooding as well as the controls, demonstrated a stronger correlation between predicted and observed ($R^2 = 0.63$).
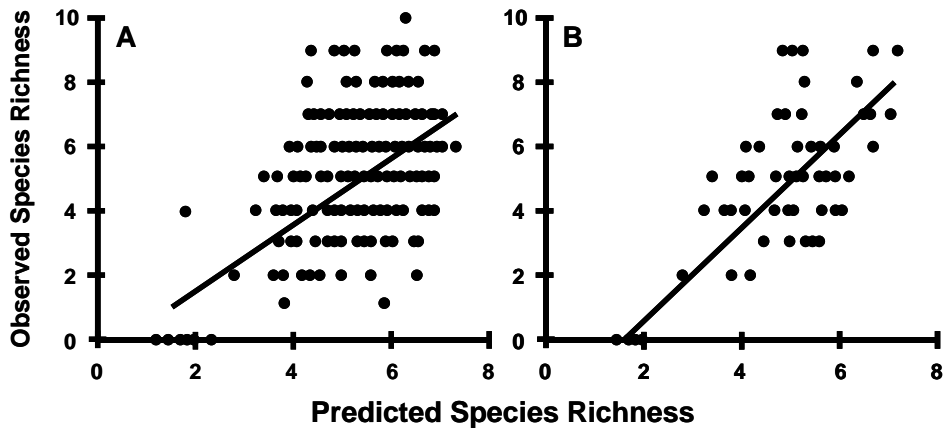

Figure 10.11. Comparisons between predicted and observed species richness taken from Gough and Grace (1999). (A) All plots, (B) excluding fertilized and fenced plots.

Details of the interpretations of the experimental study can be found in Gough and Grace (1999). What should be pointed out here, however, is that many aspects of the model of this system based on nonexperimental data were supported by the results of the experimental treatments. It was our interpretation that where the model based on nonexperimental data failed to predict accurately is for conditions that were experimentally created, but that did not exist naturally in the field. In other words, it appears that the nonexperimental data can be used to reveal how factors relate in the unmanipulated community while experimental treatments permit one to ask the question, "What will happen if we change the

conditions in the system?" Thus, the combination of experimental and nonexperimental model building and comparison is complementary, each with a role to contribute.

A Search for Evidence of Historical Effects on Species Richness in the Pearl River System

In 1994, Glenn Guntenspergen and I set out to see if we could discover the existence of additional factors that might explain variation in plant species richness at the Pearl River study area. Of particular interest to us were factors that could be related to landscape position and that might reflect past events. Prior work (Brewer and Grace 1990) had suggested that periodic tropical storms, which are common in this region, leave a long-lasting effect on plant community zonations. To examine the possibility that this might be important in understanding diversity patterns at the Pearl River, we established a sampling scheme that placed plots relative to the mouth of the Middle Pearl River (downstream to upstream) and relative to the river channel (streamside to interior). The assumptions that we wished to test were (1) that distance from the river's mouth would reflect the effects of past saltwater intrusions from tropical storm events and (2) that distance from the stream channel would reflect past overbank flooding events. It is important to point out, that the specific question we were asking was not if these events happened, but whether they had lingering effects of richness that were not reflected in current environmental conditions.

To address this question, we first developed a multivariate model that included a minimum set of the best predictors of richness (based on the previous experience of Grace and Pugesek 1997). This model included soil salinity, microelevation, disturbance, and light readings (as a measure of plant abundance). Then we asked, if the inclusion of landscape position variables might explain additional variance in richness. This model and its results were presented in Chapter 8 as an illustration of sequential hypothesis testing (see Figure 8.6) and will not be reproduced here. What is important to the current discussion is that we found that distance from the mouth of the river did explain an additional 12% of the observed variance in richness, while distance from the river's edge did not contribute new information. Based on these findings, we concluded that landscape position could reveal effects of past events that influence current diversity patterns.

A Further Examination of Spatial Effects

At a later time, Glenn Guntenspergen and I returned to the question of whether there were hidden controls of richness that could be detected from landscape position. This time, we were joined in the search by an eager group of graduate students at the University of Louisiana who were involved in a multicampus course in Biocomplexity offered by the National Center for Ecological Analysis and Synthesis. Together, we reexamined the earlier data Guntenspergen and I collected to see if the grids of plots at each site along the river held additional clues. This time, the question we wished to address was whether small-scale historical effects might show up as positive correlations in richness among adjacent plots that were not related to known environmental gradients (Mancera et al. 2005).

Starting with an examination of the data, we determined that there was spatial autocorrelation among plots. In other words, we found that plots that were spatially close were similar to one another in richness more often than would be expected by chance. Such spatial autocorrelation has been reported before (Legendre 1993) and is likely very common. It is entirely possible, of course, that such spatial autocorrelation in richness simply reflects spatial autocorrelation in controlling environmental conditions. As seen in Figure 10.12, the relationship between spatial patterns in richness and spatial patterns in environmental variables represents an important problem for interpretations. Do the spatial patterns represent a tight mapping to spatial variations in environmental conditions or do they represent historic effects, such as dispersal? To address this problem we first factored out the variation in species richness that could be ascribed to known environmental factors and then tested that residual richness still showed spatial autocorrelation. This sequential hypothesis testing is represented in Figure 10.13, with the test of neighbor richness represented by a "ghost" variable, indicating that its effects was determined after the effects of the other variables had been considered.

The analyses showed that once environmental factors were considered, spatial autocorrelation in species richness disappeared. This means that we were unable to find any evidence of small-scale historical effects or other unmeasured causes of spatially-controlled variations in richness in this system.
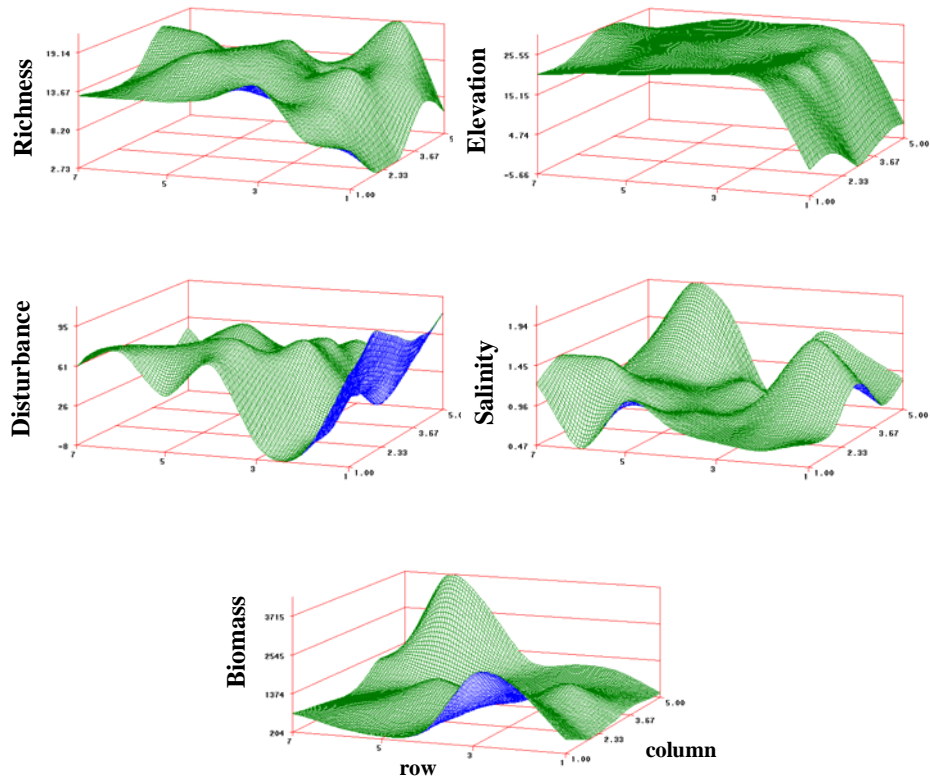
Figure 10.12. Topographic plots of spatial variation in species richness and other variables at one of the five sample sites along the Pearl River, based on sampling in a 5 x 7 grid of 1m^2 plots. In these figures, rows are at different distances from the river's edge (in meters) and columns are at different locations along the river at a site (also in meters).
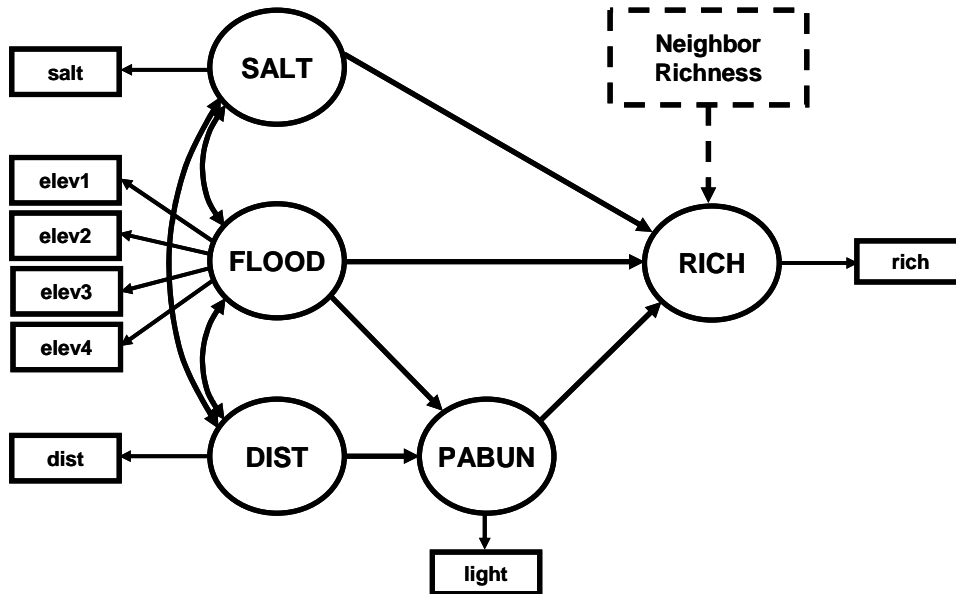
Figure 10.13. Model used to evaluate relationship between species richness and contemporary variables at each of the five grid sample sites at the Pearl River. PABUN represents plant abundance. The "ghost" variable, Neighbor Richness, was evaluated for a relationship to RICH after the effects of all other variables were removed, representing a sequential variance explanation test.

Applicability of Findings to Other Systems

More recently, the questions that have interested me are (1) how plant diversity is regulated in a wide array of systems and (2) whether there are general features that apply broadly across systems. In a study of meadows in Finland, Grace and Jutila (1999) examined the relationships of plant richness to both grazing and environmental gradients (see Figure 7.2). These results generally indicate support for a common construct model such as shown in Figure 10.6. General support for such a construct model was also found by Grace et al. (2000) in coastal tallgrass prairie, though the influence of recent disturbances were minor in this system. In a study of woodlands in Mississippi, Weiher et al. (2004) found that the presence of trees in a prairie grassland moderates soil and biomass effects on herbaceous richness, requiring an alteration of the general construct model for such situations. Investigations of diversity regulation in California chaparral (Keeley and Grace unpublished) revealed both similar construct

relations, along with the importance of landscape features, and the increasing importance of spatial heterogeneity with increasing plot sizes. In other studies, the overwhelming importance of abiotic constraints on diversity patterns of serpentine endemics (Harrison et al. unpublished) suggest a different balance of forces at play for plant occupying extreme environmental conditions. Altogether, these studies suggest support for certain general features across systems, and numerous specific factors of importance in particular situations or contexts. We will return to the question of generality in our final chapter where we consider how SEM methods may evolve to permit the evaluation of very general models that can apply across systems diverging in specific properties.

Summary

This chapter has sought to give the reader an insight into an example of the ecological perspective that can be created through a committed approach to multivariate model development, evaluation and refinement. The point of presenting such an extensive example is not to imply that all these steps are required. Instead, what I hope to have accomplished is to show the reader how the adoption of a multivariate perspective opens up a new way of learning about ecological systems. It has been my personal experience that the pursuit of a multivariate understanding has greatly enhanced the ecological insights I have gained. I am eager to extend these studies to include additional variables and pathways, such as reciprocal interactions between biomass and richness, exploration of the role of habitat variability, and interactions with other trophic levels. In Chapter 12, I followup on this extended example with a more philosophical discussion of how multivariate theories can contribute to the maturation of ecological science.