

Special Section: Observational Studies

Structural Equation Modeling for Observational Studies

JAMES B. GRACE,¹ United States Geological Survey National Wetlands Research Center, 700 Cajundome Boulevard, Lafayette, LA 70506, USA

ABSTRACT Structural equation modeling (SEM) represents a framework for developing and evaluating complex hypotheses about systems. This method of data analysis differs from conventional univariate and multivariate approaches familiar to most biologists in several ways. First, SEMs are multiequational and capable of representing a wide array of complex hypotheses about how system components interrelate. Second, models are typically developed based on theoretical knowledge and designed to represent competing hypotheses about the processes responsible for data structure. Third, SEM is conceptually based on the analysis of covariance relations. Most commonly, solutions are obtained using maximum-likelihood solution procedures, although a variety of solution procedures are used, including Bayesian estimation. Numerous extensions give SEM a very high degree of flexibility in dealing with nonnormal data, categorical responses, latent variables, hierarchical structure, multigroup comparisons, nonlinearities, and other complicating factors. Structural equation modeling allows researchers to address a variety of questions about systems, such as how different processes work in concert, how the influences of perturbations cascade through systems, and about the relative importance of different influences. I present 2 example applications of SEM, one involving interactions among lynx (*Lynx pardinus*), mongooses (*Herpestes ichneumon*), and rabbits (*Oryctolagus cuniculus*), and the second involving anuran species richness. Many wildlife ecologists may find SEM useful for understanding how populations function within their environments. Along with the capability of the methodology comes a need for care in the proper application of SEM. (JOURNAL OF WILDLIFE MANAGEMENT 72(1):14–22; 2008)

DOI: 10.2193/2007-307

KEY WORDS anurans, Bayesian analysis, complex data, lynx, maximum likelihood, mongoose, path analysis, rabbits, statistics, structural equation modeling.

WHAT IS STRUCTURAL EQUATION MODELING?

Structural equation modeling (SEM) is a collection of procedures whereby complex hypotheses, particularly those involving networks of path relations, are evaluated against multivariate data (Bollen 1989, Grace 2006). Structural equation modeling can be described in terms of its history, purpose, statistical and mathematical framework, and philosophy of application. I describe aspects of SEM that are most relevant to practicing wildlife biologists. I focus on why and how one would use SEM, rather than the variety of mathematical and statistical procedures encompassed by SEM, although I give some information on statistical procedures in association with examples.

Using multiple equations to represent complex path relationships opens a new set of scientific questions for quantitative investigation. This is important because, to a great degree, univariate statistical models are best suited for the study of single processes or responses. In contrast, structural equation models are better suited to study the multiple processes that control the behavior of systems. Researchers commonly want to know the interplay between processes, their relative importance, and how effects of perturbations cascade through systems. Of ultimate importance for many is the desire to understand and predict how properties of the system will behave in natural settings with multiple controlling factors. Conventional multivariate procedures are useful in summarizing complex data but are largely exploratory and not well-suited for representing or evaluating network hypotheses (McCune and Grace 2002).

¹ E-mail: Jim_Grace@usgs.gov

EXAMPLE 1: THE LYNX AND THE MONGOOSE

A simple example for illustrating one application of SEM to wildlife populations comes from Palomares et al. (1998). They were interested in the relationships between Iberian lynx (*Lynx pardinus*) and Egyptian mongoose (*Herpestes ichneumon*) in southwestern Spain and how this relationship was influenced by the protection afforded by the Doñana National Park.

The construction of an SEM depends on what is known or suspected about the elements of the system being studied. The authors were aided by an extensive history of studies of this system (see references in Palomares et al. 1998). Lynx are known to feed primarily on European rabbits (*Oryctolagus cuniculus*), although they are not believed to regulate rabbit populations in the study region. They also are known to kill mongooses and other smaller carnivores. Mongooses, on the other hand, are considered opportunistic feeders, and they feed on rabbits, small mammals, birds, reptiles, amphibians, and carrion. Both lynx and mongooses depend on shrub-land areas and only rarely use open habitats. Lynx are a protected species in Spain, but they are generally more abundant in Doñana National Park, which is buffered from human activities, than in the surrounding landscape. In contrast, mongooses seem quite adaptable to human activities and are abundant outside as well as inside the park.

Palomares et al. (1998) considered 3 alternative a priori models (Fig. 1). The authors believed that the pathways (solid lines in Fig. 1) represented processes operating in their system. Thus, their a priori beliefs were that 1) rabbits would be more abundant where there were more shrubs; 2) the abundance of lynx would be greater with abundant rabbits, dense shrubs, and the protection afforded by the

national park; and 3) mongoose numbers would be reduced by lynx and promoted by shrubs. Finally, Palomares et al. (1998) presumed that shrub densities might differ between protected and unprotected areas (represented by a double-headed arrow between these 2 variables in Fig. 1). If the authors thought that the protection of the park necessarily caused an increase in shrub density, the model structure would have included a unidirectional path from protection to shrubs instead of a double-headed arrow between the two.

The authors were curious about 2 additional processes that might have been operating. Do mongooses control rabbit numbers? Palomares et al. (1998) specified a model that incorporated a direct path from rabbits to mongooses as model 2, and they considered it to be a competing model with model 1 (which omits this path; Fig. 1). Again, it was an important assumption in this example that neither lynx nor mongooses were the primary regulators of rabbit densities, for if they were, one would expect an arrow of negative influence to point from lynx and mongooses to rabbits. A second question of interest to Palomares et al. (1998) was whether mongooses benefit from the protection afforded by the national park. This additional process was the basis for a third competing model, model 3.

As investigators, we also may consider questions that involve networks of variables. Might a positive association between lynx and shrubs be explained entirely by the presence of more rabbits where there are more shrubs? In that case, a positive correlation between shrubs and lynx would be observed, but a direct path would not. If a direct path from shrubs to lynx was found to be needed to explain the data, it would suggest that shrubs were influencing lynx through some other mechanism independent of rabbit abundance. In other words, we are now asking whether we can partition an association between lynx and shrubs into separate processes and how important those processes were.

There are other questions we can ask about the network of relationships (Fig. 1). How mongooses relate to shrubs seems complicated. Although they are observed to use shrubs to avoid detection, the same is true of lynx, their nemesis. What is the net result of such potentially offsetting processes? Is one much stronger than the other so that, for example, mongooses avoid shrubby areas in the presence of lynx? Or, do the 2 processes cancel each other out so that there is no observable relationship between mongooses and shrubs? What about the relationship between rabbits and mongooses? We could imagine several different possibilities. It is possible that the dominant relationship would be negative because of enhanced predation by lynx. It also is possible that the relationship would be positive because both prefer shrubs. Finally, as Palomares et al. (1998) specified in their model 2, mongooses might be more abundant where rabbits are available as a food. Although we can imagine all these processes as true, understanding this system is also about understanding the importance of the individual processes, which is what drives the behavior of the system.

The questions I mentioned thus far are of the sort that

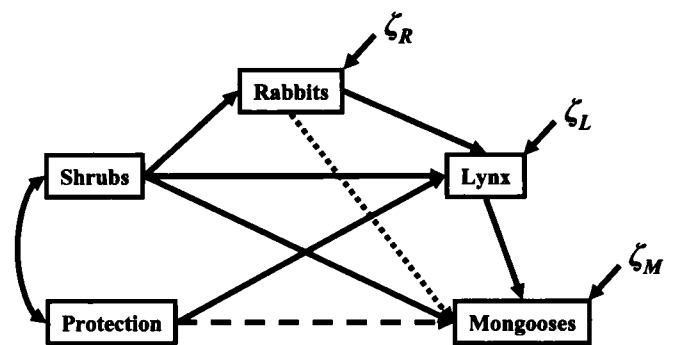


Figure 1. A priori models developed by Palomares et al. (1998) to represent anticipated relationships among mongooses, lynx, rabbits, shrubs, and protection in southwestern Spain. This diagram represents 3 alternative models. Model 1 includes all paths shown as solid lines. Model 2 additionally includes the path from rabbits to mongooses (dotted). Model 3 additionally includes the path from protection to mongooses (dashed). Zeta (ζ) values represent error variances for response variables.

fascinated Sewell Wright and led him to develop the method of path analysis (Wright 1921). When SEM became based on maximum-likelihood estimation procedures in the early 1970s, it became possible to ask additional questions. Are there important forces at work that we have not anticipated? For example, perhaps our methodology contains a bias whereby animal tracks for lynx and mongooses are underreported in some samples but not others. What if some major factor having joint influence on elements of the model was omitted from the study? Would we be able to detect such omissions? In modern SEM we might detect such effects manifested in the form of correlated errors or other missing pathways. Similarly, if it turned out that rabbits were being directly aided by the protection of the park, SEM would lead us to that discovery. Also, we might be able to detect any confounding feature of the park that promoted or reduced rabbit numbers.

Data Collection Methods

Palomares et al. (1998) conducted track surveys during winter and early spring inside and outside Doñana National Park for several years. Along with lynx and mongoose track abundance, they recorded signs of rabbit and shrub coverage was measured. Confidence in the track survey data was based, in part, on previous radiotracking and trapping studies conducted by the authors. Palomares et al. (1998) conducted sampling within squares of a 5 × 5-km grid. Rabbit estimates were based on both sightings of rabbits and their activities (warrens, pellets, and tracks). Rabbit abundance was expressed as a rank-order variable from scarce (= 1) to very abundant (= 4). They collected data at 70 sample sites in the study (for more detail see Palomares et al. [1998]).

Results

A summarization of Palomares et al. (1998) results is presented (Fig. 2). Also presented is the matrix of correlations among variables (Table 1). There are several possibilities for how SEM results can be presented. In this case (Fig. 2), the numbers next to pathways are standardized

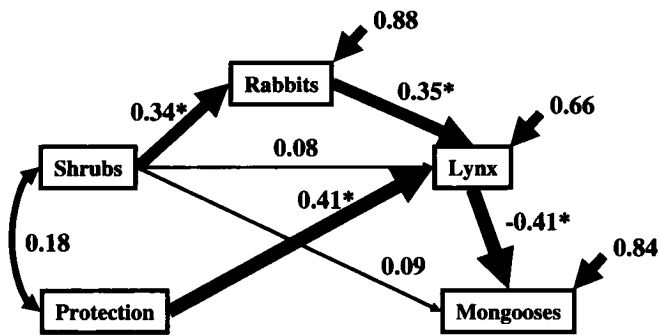


Figure 2. Results from structural equation modeling analyses conducted by Palomares et al. (1998, modified from their fig. 3). Path coefficients without asterisks were judged to be indistinguishable from zero based on model and parameter evaluations.

path coefficients. The coefficient associated with the double-headed arrow between shrubs and protection represents the intercorrelation, which is weak. For the directional pathways, the interpretation of the coefficients requires some explanation. Path coefficients are prediction coefficients. For example, the path coefficient from shrubs to rabbits represents the expected change in rabbits if we were to vary shrub density while holding all other correlated factors constant. In this case, shrubs are the only variable affecting rabbits, so holding other variables constant has no effect, but what about the relationship between shrubs and lynx? According to Palomares et al.'s results (Fig. 2), if we were to vary shrubs while holding rabbits and protection constant, lynx would not vary to a great degree. However, if we were to vary shrubs and allow rabbits to respond, a greater total response in lynx would be expected.

For standardized coefficients, predicted change is measured in standard deviation units. So, the meaning of the coefficient from shrubs to lynx (0.08) is that if we were to vary shrubs by 1.0 standard deviation, rabbits would be expected to vary by 0.08 standard deviations. However, if we were to vary rabbits by 1.0 standard deviation, lynx would increase by 0.35 standard deviations. Standardization based on standard deviations permits us to compare path coefficients in similar units, but it has the problem that standard deviations are a property of the sample. For many purposes, including predictive forecasting and comparisons among populations, we would prefer unstandardized coefficients.

Discussion

The modeling results answer many of the questions I discussed above. They support the expectation that mongooses are strongly and negatively affected by lynx. This effect dominates over any positive influence of shrubs on mongoose habitat use. It also appears that there is no real evidence that mongoose numbers are promoted by rabbit abundance. Although lynx are more common where shrubs are more dense, this association appears to be predominantly due to the indirect benefit that rabbits afford. Lynx also are strongly favored by the protection afforded by Doñana National Park. Rabbits do depend to a moderate degree on

Table 1. Matrix of correlations for Lynx–Mongoose example (Palomares et al. 1998).

	Shrubs	Protection	Rabbits	Lynx	Mongooses
Shrubs	1.00				
Protection	0.18	1.00			
Rabbits	0.34	−0.04	1.00		
Lynx	0.28	0.41	0.37	1.00	
Mongooses	−0.02	−0.10	−0.25	−0.38	1.00

the abundance of shrubs, which is slightly higher in the park, though this does not appear to represent a direct effect on rabbits. There is considerable unexplained variance. The standardized error variances are shown at the end of arrows pointing to the upper right corners of the response (endogenous) variables. These translate to R^2 values of 0.12, 0.34, and 0.16 for rabbits, lynx, and mongooses, respectively. It is unclear how much the low variance explanation depends on 1) the inherently stochastic and patchy nature of the animal abundances, 2) error of measurement, or 3) other important influences that are unmeasured. Additional discussion can be found in Palomares et al. (1998).

Palomares et al. (1998) appropriately kept their analysis based on their a priori questions and their sense of what was important in this case. Structural equation modeling also allows investigators to ask questions beyond those considered by Palomares et al. What if shrubs, rabbits, lynx, and mongooses interacted in an entirely different way in protected areas than in unprotected areas? A multigroup SEM analysis would allow us to explore such a question. We would evaluate 2 models simultaneously, one with and one without protection, determining if relationships inside and outside the park fit a common model. To go a step further, we also could include an assessment of differences among means in our multigroup analysis (note that thus far we have modeled covariances, not means).

One can still ask the question, what have we really gained in this simple example? Is it not possible to arrive at the same place using univariate models? One limitation of using univariate models, such as multiple regression or the general linear model, is that the structure of the model is not conducive to understanding complex relationships among predictors. We can best illustrate this graphically by showing the structure of a univariate model of mongoose numbers as a function of the other elements of the system (Fig. 3). Such models fail to permit much understanding of the system because they treat the suite of predictors as simply being correlated, and they seek only the goal of selecting some minimum set of predictors of y . Missing from univariate models is an explicit representation of the indirect effects of protection, shrubs, and rabbits on mongooses. Although it is possible to ask such questions in a piecemeal fashion using univariate procedures, such an approach is not optimal, and piecewise approaches are not possible for more complex examples (see below). In sum, SEM is useful because it establishes a framework that is amenable to relating the full

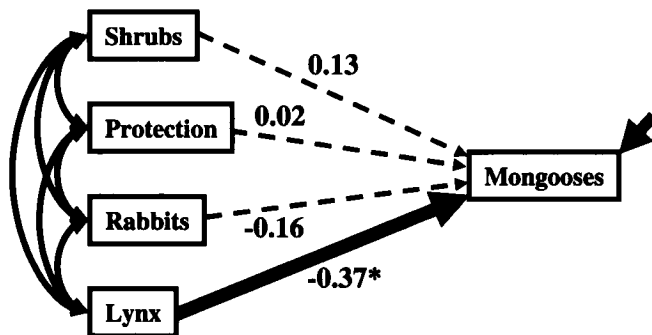


Figure 3. Illustration of structure and results that would be obtained from a multiple regression analysis of mongooses as a function of lynx, rabbits, shrubs, and protection. Dashed lines represent paths deemed nonsignificant using *P*-values.

matrix of covariances to complex hypotheses about the processes operating.

A CLOSER LOOK AT SEM

A Bit of History

The roots of SEM extend back to the pioneering work of Sewall Wright, who published the first application of path analysis in 1921. Parallel to the development of methods designed to isolate effects of individual processes (e.g., analysis of variance [ANOVA]), Wright and others developed the foundations for SEM. Although biologists have used path analysis to some extent for many decades, social scientists have been the ones to push the development of the methodology. Through the 1950s and 1960s, extensive work was conducted on path analysis and factor analysis. In the late 1960s it became clear that a comprehensive system for representing and evaluating complex hypotheses was needed, including both paths and factors. Although many people tried to develop such a system, the LISREL (linear structural relations) system developed by the Swedish mathematician Karl Jöreskog became modern SEM (Jöreskog 1973).

The LISREL system was of revolutionary importance for several reasons (Cudeck et al. 2001, Tomer 2003). First, its sophisticated matrix formulation permits the representation of a large variety of model architectures involving factors (latent variables), paths, correlated errors, reciprocal interactions, and more. Second, a maximum-likelihood solution procedure was developed that permitted 2 important things simultaneously: 1) the ability to estimate many types of complex models, including those containing latent factors and those with reciprocal pathways; and 2) the ability to assess the fit between data and model by comparing expected to observed covariance structures. Jöreskog developed computer software based on the LISREL system, permitting ready access to these procedures for researchers, which contributed to the adoption of his method.

Today's SEM Procedures

Here I give a skeletal presentation of the LISREL system (see Bollen 1989, Lee 2007 for details). First, I operationally define SEM as a multiequational framework. The model in

Table 2. Model fit results for the 3 models evaluated by Palomares et al. (1998; Fig. 2). Results for model 1 also apply to the multiple regression results (Fig. 3). Number of samples was 70.

Model	χ^2	df	<i>P</i>	SBC ^a
Model 1	2.55	3	0.467	-10.20
Model 2	0.83	2	0.661	-7.67
Model 3	2.26	2	0.323	-6.24

^a Schwarz's Bayesian Criterion.

this case (Fig. 1) can be represented (including all paths shown) by the following equations, where x_1 = shrubs, x_2 = protection, y_1 = rabbits, y_2 = lynx, and y_3 = mongooses:

$$y_1 = \alpha_1 + \gamma_{11}x_1 + \zeta_1 \quad (1)$$

$$y_2 = \alpha_2 + \gamma_{21}x_1 + \gamma_{22}x_2 + \beta_{21}y_1 + \zeta_2 \quad (2)$$

$$y_3 = \alpha_3 + \gamma_{31}x_1 + \gamma_{32}x_2 + \beta_{31}y_1 + \beta_{32}y_2 + \zeta_3 \quad (3)$$

Simplifying the LISREL system for the case where there are no latent variables and ignoring the intercept terms, a generalized representation of a system of such equations is

$$Y = BY + \Gamma X + \zeta \quad (4)$$

where Y is a vector of y s, X a vector of x s, B a matrix of coefficients relating the y s to one another, Γ a matrix of coefficients relating y s to x s, and ζ a vector of errors for the y s. Additional parameters are associated with more complex models (e.g., multi-group and multi-level models).

The essential problem in SEM is to resolve the equation $S = \Sigma(\Theta)$, where S is the sample covariance matrix, and $\Sigma(\Theta)$ is the model-implied covariance matrix, expressed in terms of Θ , which is the matrix of estimated model parameters. The maximum-likelihood estimation process involves comparing the observed covariance matrix, S , with the covariance matrix implied by the statistical model $\Sigma(\Theta)$, and the task is to choose values for the coefficients in Θ such that the model-implied Σ is as close to S as possible. This typically involves some fitting function that compares the likelihood for a given model to that of a model with perfect fit. It happens that the values of the fitting functions often follow a chi-square distribution, permitting an assessment of the degree of success of the fitting procedure. Although numerous comparative fit measures are available for use in SEM, the measure of absolute fit provided by the chi-square test is still the most used approach to model evaluation, though it is typically supplemented by other indices. The ability to reject individual models based on the degree of similarity between observed and model-implied covariance matrices using the model chi-square test has been exploited as a core functionality in the hypothesis-testing tradition within SEM. It is important to note that Bayesian approaches to both model evaluation and estimation also exist (Lee 2007).

The example from Palomares et al. (1998) represents a very simple application of SEM. The authors presented measures of fit for the 3 models they considered (Fig. 1, Table 2). For model 1, analysis revealed a chi-square of 2.55,

which is a measure of the magnitude of difference between observed and expected covariances. For this model, 3 degrees of freedom exist, representing the difference between the number of known elements in the covariance matrix (15) and the number of model parameters (12 = no. of variances + paths). The *P*-value of 0.467 represents a high probability that the differences between model-implied and observed covariances could be due to chance. Therefore, the authors drew the conclusion that their data were not inconsistent with their model. A very important point is made by this illustration. In SEM, the *a priori* model is the model of scientific interest, not the null model.

The Continuing Evolution of SEM

Many extensions of the original formulation have been made that permit both the relaxation of numerous statistical assumptions as well as an expansion of the variety of models that can be evaluated. Most SEM software packages include procedures for bootstrapping, obtaining robust standard errors, modeling categorical response variables, repeated measures, and hierarchical relations. Methods of estimation used in SEM include several types of maximum-likelihood and least-squares procedures, Monte Carlo methods, and Bayesian estimation (Raftery 1993, Rupp et al. 2004, Palomo et al. 2005, Lee 2007).

Assumptions and Requirements

As with all statistical procedures, the methods incorporated into SEM are based on certain assumptions. The default estimation procedure is maximum likelihood (ML), which provides asymptotically unbiased estimates and has many large-sample properties that make it an attractive solution procedure, as well as reasonably good performance in smaller samples. The ML fitting function typically used in SEM assumes that there is no excessive kurtosis, though this assumption can be relaxed by implementing robust estimation procedures that correct for kurtosis. There is no assumption of uncorrelated error terms using the SEM-ML procedure, which makes the procedure appropriate for nonrecursive models (i.e., ones with reciprocal pathways). A weighted least-squares procedure has been developed for models containing categorical response variables (Muthén 1984). Monte Carlo procedures, including bootstrapping, are now available in most software packages, and Markov chain Monte Carlo methods for Bayesian estimation are included in some software packages.

It is well beyond the scope of my paper to cover the characteristics of software that performs SEM. The package that has been in use the longest is LISREL (Jöreskog and Sörbom 1996). There are numerous other packages, including AMOS, CALIS (part of SAS), EQS, Mplus, R, and several others that have different capabilities (which change regularly as software is updated).

Model Building Versus Hypothesis Testing in SEM

Structural equation modeling not only allows tests of model fit, the results obtained depend on adequate fit. Therefore, if the *a priori* models evaluated by Palomares et al. (1998) did

not show adequate model fit, the authors would be compelled to consider alternative models. Such a requirement means that models are always compared against the saturated model in which all variables are allowed to interact as part of standard SEM practice.

Structural equation modeling can be used either in a model-building exploratory mode or in a hypothesis-testing confirmatory mode. In many situations inadequate theory or insufficient experience with the system (or with SEM) will necessitate a model-building approach. This is a perfectly legitimate use of SEM; however, the potential to misinterpret chance features of the data is increased by model exploration. Models achieved through model-building need further confirmation. For the lynx-mongoose example, the authors already had a good knowledge of their system and a solid theoretical basis for their models. Further, by considering a limited set of alternative models, they were able to minimize the number of comparisons and achieve a test of alternative hypotheses. Bayesian methods for model comparison, such as the Bayesian Information Criterion (BIC) are believed to provide a superior approach to model comparison when many contrasts are to be made. It is typical for studies involving moderate sample sizes that various indices lead to a common set of conclusions about model fit.

Palomares et al. (1998) not only presented an assessment of the absolute fit of their model, the chi-square, they also reported how this model compared to the other models of interest, model 2 and model 3 (Table 2). Comparing models, I found that all had adequate model fit based on the chi-square statistic. I computed a chi-square difference to test whether a model was significantly superior to another. In this case, model 2 possessed a chi-square difference from model 1 of 1.72. For model 3, the difference was 0.29. Neither indicated that the addition of paths led to a significant improvement in model fit. Other indices also suggested that the additional paths associated with models 2 and 3 had small parameter values and were not judged to be significant. Methods that adjust for parsimony, such as Schwarz's Bayesian Criterion, also suggested that models 2 and 3 were not superior to model 1. Thus, the authors felt justified in staying with the results of their original model.

Users of SEM hypothesis testing need to recognize the philosophy that underlies its application. Most researchers in the natural sciences have been trained with a strong emphasis on variance partitioning and the procedures of experimental design and analysis using ANOVA. Although SEM can be used readily for the analysis of experimental data (Grace 2006, chapter 9), it is not a variance partitioning method but instead emphasizes parameter estimation. One fundamental difference between variance partitioning and parameter estimation is that the former is predicated on having isolated, orthogonal causes, whereas parameter estimation presumes that investigators wish to understand relationships within the context of correlated causal factors. As investigators, we seek models in SEM whose parameters are generally consistent and predictive. Accordant with this,

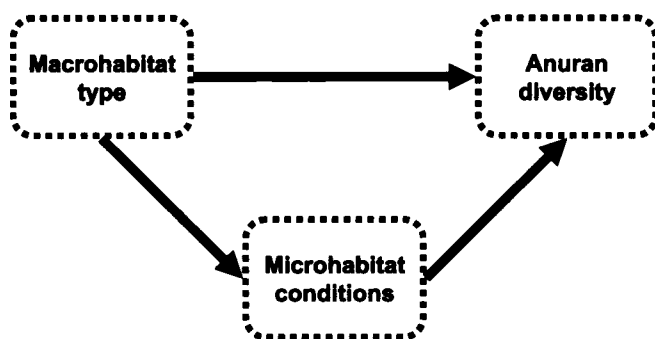


Figure 4. Construct (conceptual) model for anuran diversity study.

emphasis in SEM practice is placed on testing *a priori* models with actual data and on validation using follow-up studies and independent data.

When we use SEM to test hypotheses based on observational data, many factors influence our strength of inference. Model structure and conclusions about cause and effect are based on our total knowledge about the system under observation. The analysis tells us if the data are consistent with the model. However, even when the data are consistent with our models, we must be alert to alternative models that also fit the data. Ultimately, we strive for sequential learning through the use of SEM and by the incorporation of all our knowledge about the system. Perhaps what is most useful to us as researchers is that 1) SEM strongly encourages us to think about the mechanisms operating rather than just the associations, and 2) it provides insights into the relationships most worthy of further study.

A SECOND EXAMPLE: ANURAN DIVERSITY AND HABITAT FACTORS

Background

In this example I use data from Lichtenberg et al. (2006) in which they examined 25 wetlands in the lower Mississippi River alluvial valley, USA, for chorusing anuran species and associated habitat characteristics. One major goal of their study was to understand the habitat characteristics associated with diversity hotspots (i.e., places where a variety of species are abundant). Lichtenberg et al. (2006) examined several types of wetlands. Wetlands were classified as either lakes, impoundments, swales, or riverine areas. They assessed specific microhabitat conditions at each site, including vegetation and topographic features. They assessed anuran diversity using nighttime surveys of chorusing individuals during several seasons of the year. Chorus surveys were used during the year to determine the species using a site. Lichtenberg et al. (2006) measured several characteristics of the vegetation, including herbaceous cover, vegetation density at different vertical positions, woody cover, tree heights, canopy cover, litter cover, and depth (by type). They also measured hydrologic features, including the area of open water and the mean and maximum water depths.

Anuran Model Development and Evaluation

In this example, I provide an overview of model development and evaluation with a brief description of results. See

Lichtenberg et al. (2006) for more detail about the biology and Grace and Bollen (2006) for more detail about the statistical procedures and analyses. I begin with a conceptual model that will guide the development and interpretation of SEM. In this case, I used the construct model (Grace and Bollen 2007), which represents the main theoretical constructs (i.e., ideas) of interest and their expected associations (Fig. 4). The SEM, in contrast, represented the hypothesized statistical relations among variables. Basically, the goal of this study was to understand the degree to which anuran diversity can be predicted by the type of habitat (e.g., lake, impoundment) versus the particular conditions (e.g., vegetation, litter) within that habitat. The construct model makes clear that we, as investigators, expect that the microhabitat conditions may be a function of the macrohabitat type. Although this general question is very simple in structure, relating this question to our data can be anything but simple. When it comes to the specifics, we may have many uncertainties. Which of our measures of a habitat feature is most useful as a representation of the concept of interest? Are there any nonlinearities or interactions? We know that ultimately our SEM will produce results that are highly specific to the particular sample. Our construct model (as well as any more general models we use to establish the foundation for our analysis) helps us to relate the specifics to our more general understanding.

We begin by considering how the available data (relationships summarized in Table 3) relate to the constructs of interest (see Grace and Bollen [2006, 2007] for details). Briefly, we possess 4 measures of macrohabitat type (i.e., lake, impoundment, swale, or riverine). Because our measures are nominal, one possibility is to model this construct using a set of dummy variables representing the possible macrohabitat types and a composite representing their collective influences. We assume that the classification of individual sites as to habitat type was correct. For the construct labeled microhabitat conditions, the specific details of how the measured variables are related was not known *a priori*. For this reason, J. S. Lichtenberg, S. L. King, J. B. Grace, and S. Walls (United States Geological Survey, unpublished analyses) performed an exploratory factor analysis to see if the correlations among the many measured microhabitat variables might suggest the operation of a smaller number of latent factors. The authors identified 2 factors of potential importance to anuran diversity: the abundance of herbaceous vegetation and the abundance of leaf litter. Based on this information, it is possible to represent the construct microhabitat conditions using 2 latent variables and several indicators of each.

An SEM relating the elements of the statistical model (the specific variables and their properties and interrelationships) to the theoretical constructs is presented (Fig. 5). The SEM is based on the hypothesis that the covariances among observed variables can be explained by the relationships between 2 composites (Macrohabitat and Microhabitat) and 3 latent variables (Herbaceous, Litter, and Richness).

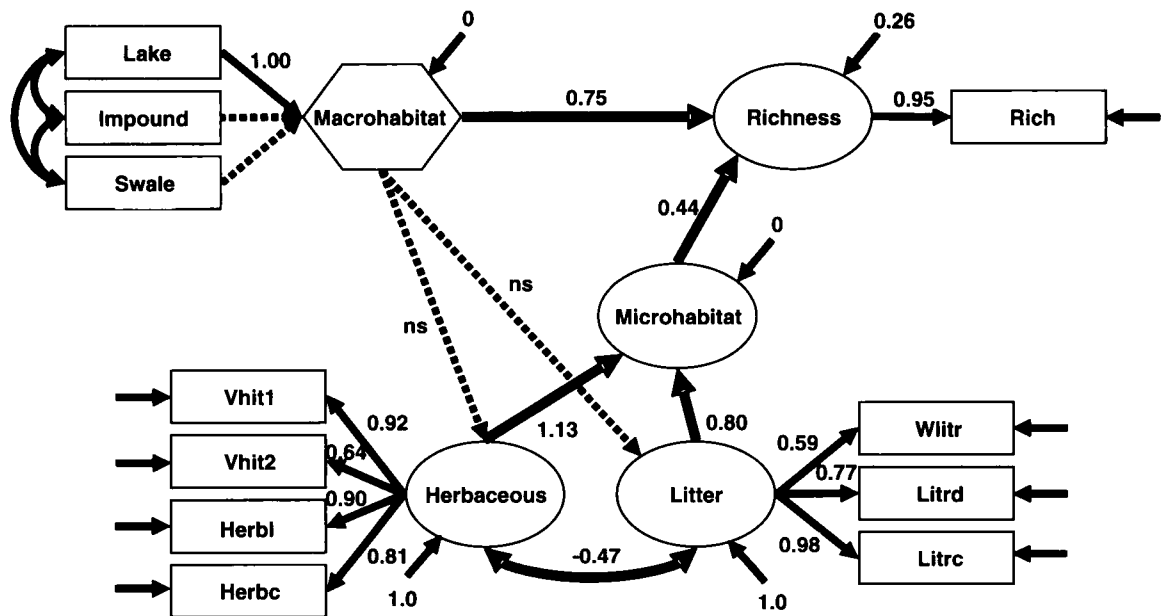


Figure 6. Results obtained from the analysis of a priori model (Fig. 5). Coefficients are standardized values. ns indicates nonsignificant.

nents, Herbaceous and Litter. In the other alternative structure, Grace and Bollen hypothesized that there are distinctly different macrohabitat effects on herbaceous, litter, and richness. In this case, we used multiple measures of model adequacy, including both chi-square tests of overall fit of data to model and measures of residual variation in response variables.

In the presentation of final results (Fig. 6), I show only the standardized coefficients and omit the error variances for the observed variables for simplicity. The model chi-square was 30.03 with 24 degrees of freedom, giving a *P*-value of 0.18. This result indicates that the probability of these data, given the model, was reasonably high. Further, variation in richness was well explained by the model ($R^2 = 0.74$). According to the results, richness in this sample appears to be more sensitive to macrohabitat type than microhabitat conditions (coeff. of 0.75 and 0.44, respectively), although both were important. Unexpectedly, microhabitat components, including both herbaceous plant abundance and litter characteristics, were not consistently related to macrohabitat type; thus, the effects of macrohabitat and microhabitat were found to be largely independent. Results further indicated that the associations among litter and vegetation variables were consistent with the existence of 2 general factors: the abundance of herbaceous vegetation and litter depth.

CAUTIONS AND RECOMMENDATIONS

Although SEM provides a flexible and powerful framework for analyzing complex models, many opportunities exist for missteps. Kline (2005) listed 44 errors that should be avoided when using SEM. These errors can be generally classified into 1) errors of model specification, 2) problematic data, 3) errors of analysis, and 4) errors of interpretation. Model specification often is hampered by either the investigator's knowledge or ability to measure some of the

important factors. Models are, by definition, simplistic representations, so as investigators, we do not expect to include all major influences. However, the omission of certain key influences can lead investigators astray in their interpretations. Practitioners should be careful to appreciate the possibility of alternative models that also fit the data. Ultimately, we must recognize that as we begin using SEM, our first efforts may be poor approximations. Confidence is developed through the process of investigation, modeling, model testing, resampling, retesting, and refinement.

The adequacy of data for addressing particular questions is an ultimate limitation for any study. In SEM, if we have strong confidence in our initial model and find the data are inconsistent with that model, we may conclude that the data are faulty. Usually, however, we have to judge our data by the way in which they were collected, the magnitude of the sample, and whether the information provided by the data measure up to our scientific objectives. Rules of thumb for sample sizes suggest 5–20 observations per model parameter. However, we must recognize that the quality of the data, as well as the behavior of the system components, also influence sample adequacy. Further, SEM is sensitive to the usual set of problems that arise from inadequate inspection of data prior to analysis.

Errors of analysis, as well as errors of interpretation, are largely caused by insufficient training in SEM procedures. I find that the training many receive in experimental statistics courses that emphasize variance partitioning provides inadequate background for using SEM. A much better background is provided from training in regression or statistical modeling. Ultimately, biologists will benefit from courses in SEM, which are reasonably common in programs in sociology, psychology, economics, and health sciences, although a bit of translation is required.

MANAGEMENT IMPLICATIONS

In the application of science to real-world problems, it is typical that managers and decision-makers wish to understand how natural and managed systems will behave under various scenarios. Conventional approaches to the analysis of data based on the univariate model are not well suited either to understanding systems, nor predicting their behavior. Responses to management or to changing conditions are conditional. We cannot successfully understand systems by attempting to factor out correlated environmental effects, but instead, need to understand their contributions to function within our models. As an example, reliance on the univariate model has led to the proliferation of dozens and dozens of models of biodiversity based on individual processes. The use of structural equation modeling has the ability to address the relevant question of which processes dominate in controlling system behavior and move us into the realm where science naturally informs management (e.g., Grace et al. 2007). Thus, SEM would appear to have a particularly important role to play in applied science.

ACKNOWLEDGMENTS

I thank F. Palomares, M. Anderson, B. Vairin, and 2 anonymous reviewers for helpful comments on the manuscript. My special thanks to T. L. Shaffer and D. H. Johnson for organizing this special collection of papers and for extremely helpful comments on the manuscript.

LITERATURE CITED

- Bollen, K. A. 1989. Structural equations with latent variables. John Wiley and Sons, New York, New York, USA.
- Cudeck, R., S. H. C. Du Toit, and D. Sörbom, editors. 2001. Structural equation modeling: present and future. Scientific Software International, Lincolnwood, Illinois, USA.
- Grace, J. B. 2006. Structural equation modeling and natural systems. Cambridge University Press, Cambridge, United Kingdom.
- Grace, J. B., T. M. Anderson, M. D. Smith, E. Seabloom, S. J. Andelman, G. Meche, E. Weiher, L. K. Allain, H. Jutila, M. Sankaran, J. Knops, M.

- Ritchie, and M. R. Willig. 2007. Does species diversity limit productivity in natural grassland communities. *Ecology Letters* 10:680–689.
- Grace, J. B., and K. A. Bollen. 2006. The interface between theory and data in structural equation models. U.S. Geological Survey Open-File Report 2006–1363. <<http://pubs.usgs.gov/of/2006/1363>>. Accessed 25 Oct 2007.
- Grace, J. B., and K. A. Bollen. 2007. Representing general theoretical concepts in structural equation models: the role of composite variables. *Environmental and Ecological Statistics* 14:in press.
- Jöreskog, K. G. 1973. A general method for estimating a linear structural equation system. Pages 85–112 in A. S. Goldberger and O. D. Duncan, editors. *Structural equation models in the social sciences*. Seminar Press, New York, New York, USA.
- Jöreskog, K. G., and D. Sörbom. 1996. LISREL 8: user's reference guide. Scientific Software International, Lincolnwood, Illinois, USA.
- Kline, R. B. 2005. Principles and practice of structural equation modeling. Second edition. The Guilford Press, New York, New York, USA.
- Lee, S.Y. 2007. Structural equation modeling: a Bayesian approach. John Wiley and Sons, New York, New York, USA.
- Lichtenberg, J. S., S. L. King, J. B. Grace, and S. Walls. 2006. Habitat associations of chorusing anurans in the lower Mississippi river alluvial valley. *Wetlands* 26:736–744.
- McCune, B., and J. B. Grace. 2002. Analysis of ecological communities. MJM, Gleneden Beach, Oregon, USA.
- Muthén, B. 1984. A general structural equation model with dichotomous ordered categorical, and continuous latent variable indicators. *Psychometrika* 49:115–132.
- Palomares, F., P. Ferraras, A. Travaini, and M. Delibes. 1998. Co-existence between Iberian lynx and Egyptian mongooses: estimating interaction strength by structural equation modelling and testing by an observational study. *Journal of Animal Ecology* 67:967–978.
- Palomo, J., D. B. Dunson, and K. A. Bollen. 2005. Bayesian structural equation modeling. Pages 163–168 in L. Sik-Yum, editor. *Handbook of structural equation models*. Elsevier, New York, New York, USA.
- Raftery, A. E. 1993. Bayesian model selection in structural equation models. Pages 163–180 in K. A. Bollen and J. S. Long, editors. *Testing structural equation models*. Sage, Beverly Hills, California, USA.
- Rupp, A. A., D. K. Dey, and B. D. Zumbo. 2004. To Bayes or not to Bayes, from whether to when: applications of Bayesian methodology to modeling. *Structural Equation Modeling* 11:424–451.
- Tomer, A. 2003. A short history of structural equation models. Pages 85–124 in B. H. Pugesek, A. Tomer, and A. von Eye, editors. *Structural equation modeling*. Cambridge University Press, Cambridge, United Kingdom.
- Wright, S. 1921. Correlation and causation. *Journal of Agricultural Research* 10:557–585.

Associate Editor: Shaffer and Johnson.