Available online at www.sciencedirect.com
SCIENCE DIRECT®

Ecological Modelling 192 (2006) 385–409

ECOLOGICAL MODELLING

www.elsevier.com/locate/ecolmodel

Exploring ecological patterns with structural equation modeling and Bayesian analysis

G.B. Arhonditsis [a,*], C.A. Stow [b], L.J. Steinberg [c], M.A. Kenney [a], R.C. Lathrop [d], S.J. McBride [a], K.H. Reckhow [a]

[a] *Nicholas School of the Environment and Earth Sciences, Duke University, Durham, NC 27708, USA*
[b] *Department of Environmental Health Sciences, Arnold School of Public Health, University of South Carolina, Columbia, SC 29208, USA*
[c] *Department of Civil and Environmental Engineering, Tulane University, New Orleans, LA 70118, USA*
[d] *Wisconsin Department of Natural Resources and Center for Limnology, University of Wisconsin, Madison, WI 53706, USA*

Received 7 December 2004; received in revised form 26 May 2005; accepted 18 July 2005
Available online 3 October 2005

Abstract

Structural equation modeling is a multivariate statistical method that allows evaluation of a network of relationships between manifest and latent variables. In this statistical technique, preconceptualizations that reflect research questions or existing knowledge of system structure create the initial framework for model development, while both direct and indirect effects and measurement errors are considered. Given the interesting features of this method, it is quite surprising that the number of applications in ecology is limited, and even less common in aquatic ecosystems. This study presents two examples where structural equation modeling is used for exploring ecological structures; i.e., summer epilimnetic phytoplankton dynamics. Both eutrophic (Lake Mendota) and mesotrophic (Lake Washington) conditions were used to test an initial hypothesized model that considered the regulatory role of abiotic factors and biological interactions on lake phytoplankton dynamics and water clarity during the summer stratification period. Generally, the model gave plausible results, while a higher proportion of the observed variability was accounted for in the eutrophic environment. Most importantly, we show that structural equation modeling provided a convenient means for assessing the relative role of several ecological processes (e.g., vertical mixing, intrusions of the hypolimnetic nutrient stock, herbivory) known to determine the levels of water quality variables of management interest (e.g., water clarity, cyanobacteria). A Bayesian hierarchical methodology is also introduced to relax the classical identifiability restrictions and treat them as stochastic. Additional advantages of the Bayesian approach are the flexible incorporation of prior knowledge on parameters, the ability to get information on multimodality in marginal densities (undetectable by standard procedures), and the fact that the structural equation modeling process does not rely on asymptotic theory which is particularly important when the sample size is small (commonly experienced in environmental studies). Special emphasis is given on how this Bayesian methodological framework can be used for assessing eutrophic conditions and assisting water quality management. Structural

* Corresponding author at: Department of Physical and Environmental Sciences, University of Toronto, Toronto, Ont., Canada M1C 1A4. Tel.: +1 416 208 4858; fax: +1 416 287 7279.
E-mail addresses: georgear@duke.edu, georgea@utsc.utoronto.ca (G.B. Arhonditsis).

doi:10.1016/j.ecolmodel.2005.07.028

equation modeling has several attractive features that can be particularly useful to researchers when exploring ecological patterns or disentangling complex environmental management issues.

Keywords: Structural equation modeling; Ecological patterns; Bayesian analysis; Phytoplankton dynamics; Water quality management

1. Introduction

"...Thinking only in terms of directly observable variables confines our horizons and limits our assessment of complex systems..." (Malaeb et al., 2000, Environmental and Ecological Statistics, pg 95)

Mechanistic understanding and prediction of patterns is a key feature in ecological research (Peters, 1991; Jorgensen, 1997; Pace, 2001; Carpenter, 2002; Arhonditsis and Brett, 2005). Given the hierarchical structure of biological information, observations in a particular study scale are usually associated with upper-level joint behaviors and lower-level processes, thus it is essential for ecologists, when exploring patterns, to be able to shift between different scales of description in space, time, and organizational complexity (Sugihara and May, 1990; Levin, 1992). For example, Vepsalainen and Spence (2000) advocated the development of "general explanatory frameworks" that comprise (i) the focal level, defined by the pattern/process of interest, and (ii) the contiguous lower and upper levels, associated with the initiating conditions and boundary constraints, respectively. Recognizing the intertwined nature of ecological hierarchies, this framework was proposed as a convenient means to select the appropriate level of information for understanding specific observations/events. Similarly, earlier studies by Salthe (1985) and O'Neill et al. (1986) acknowledged the importance of a basic triadic approach, which explicitly considers the effects of slow, larger-scale and fast, smaller-scale processes on the focal level observed ecological patterns. The problem of scale can also be raised from our inability to measure with absolute accuracy characteristics and properties of conceptual interest (McCune and Grace, 2002). Ecologists are frequently interested in processes that cannot be measured effectively by one single variable, and a common way to address this problem is to perceive the ecological concept as a nested hierarchy that can be decomposed into an infinite number of subprocesses and causal interactions. The premise behind this partitioning is that the selective measurement of some of these elements can improve our understanding about the collective behavior/mechanism. The attempt to abstract essential features and reduce the complexity of the real world is ubiquitous in ecological practice. Levin (1992) characterized the study of the transferability of ecological phenomena across scales and the development of laws of simplification and aggregation as a central problem in ecology and evolutionary biology.

Modeling as a tool for elucidating ecological patterns is subject to the same problem of complexity, and the optimal model dimension has been extensively debated in the ecological literature (Levins, 1966; Costanza and Sklar, 1985; Rastetter et al., 1992; Jorgensen, 1999; Reckhow, 1999; Arhonditsis and Brett, 2004). Applied ecologists are inclined to select realism and precision in favor of generality; driven by technical or conceptual limitations, they adopt "intuitively manageable scales" and develop models that aim to provide "faithful descriptions" of the data (Vepsalainen and Spence, 2000). A characteristic example is the application of regression analysis for analyzing data from experimental/observational studies and the use of the best-fit model for inference and hypothesis testing. While useful for investigating causality in nature, regression models have several limitations: (i) the predictor variables are assumed to be free of measurement error or uncontrolled variation, (ii) the assumption of normality is frequently violated by the errors in the resultant models, and (iii) hypotheses are formulated in a way that solely allow for the inclusion of directly observed variables (Malaeb et al., 2000). Therefore, it is increasingly recognized that what is missing from the common ecological practice is a statistical technique with the ability to unravel complex interrelationships and aid generalization and theory testing by relaxing some of these restrictions.