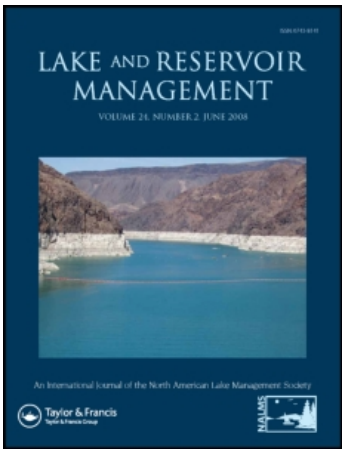

Lake and Reservoir Management
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t905619022

Using structural equation modeling and expert elicitation to select nutrient criteria variables for south-central Florida lakes

Melissa A. Kenney [a]; George B. Arhonditsis [b]; Linda C. Reiter [c]; Matthew Barkley [d]; Kenneth H. Reckhow [a]
[a] Nicholas School of the Environment, Duke University, Durham, NC [b] Department of Physical and Environmental Sciences, University of Toronto Scarborough, Toronto, Ontario, Canada [c] Environmental Resources Management, Mt. Pleasant, SC [d] ARCADIS USA Inc., Golden, CO

First Published on: 01 June 2009

To cite this Article Kenney, Melissa A., Arhonditsis, George B., Reiter, Linda C., Barkley, Matthew and Reckhow, Kenneth H.(2009)'Using structural equation modeling and expert elicitation to select nutrient criteria variables for south-central Florida lakes',Lake and Reservoir Management,25:2,119 — 130
To link to this Article: DOI: 10.1080/02701960902821373
URL: http://dx.doi.org/10.1080/02701960902821373

Lake and Reservoir Management, 25:119–130, 2009

ISSN: 0743-8141 print / 1040-2381 online
DOI: 10.1080/02701960902821373

Using structural equation modeling and expert elicitation to select nutrient criteria variables for south-central Florida lakes

Melissa A. Kenney[1,*], George B. Arhonditsis[2], Linda C. Reiter[3], Matthew Barkley[4], and Kenneth H. Reckhow[1]

[1]Nicholas School of the Environment, Duke University, Box 90328, Durham, NC 27708
[2]Department of Physical and Environmental Sciences, University of Toronto Scarborough, 1265 Military Trail, Toronto, Ontario, Canada M1C 1A4
[3]Environmental Resources Management, 498 Wando Park Boulevard, Suite 100, Mt. Pleasant, SC 29464
[4]ARCADIS USA Inc., 14142 Denver West Parkway, Suite 350, Golden, CO 80401

Abstract

Kenney, M.A., G.B. Arhonditsis, L.C. Reiter, M. Barkley, and K.H. Reckhow. 2009. Using structural equation modeling and expert elicitation to select nutrient criteria variables for south-central Florida lakes. Lake Reserv. Manage. 25:119–130.

To protect the nation's waterbodies from excessive impairments from pollution leading to eutrophication, the Clean Water Act requires states to establish water quality standards. These water quality standards are designed to protect the designated use, or water quality goal; however, they are indirectly measured and assessed using a water quality criterion. An alternative approach to develop nutrient criteria is the predictive approach (Reckhow et al. 2005), which determines the predictive variables by combining water quality data with assessments from multiple experts on the probability of designated use attainment using structural equation modeling (SEM). Our objective was to expand the predictive approach to include a region of waterbodies and to use multiple experts. To demonstrate these extensions, the approach was applied to lakes in south-central Florida using four experts to quantify attainment of a fish and wildlife designated use. Multiple models were built that related eutrophication processes to the designated use. Of the two plausible models, total phosphorus was the most predictive of the designated use followed by chlorophyll *a*. Using the model results, the risk of nonattainment of the designated use for these two predictive variables was calculated; to achieve high attainment (90% or more), total phosphorus should be <0.015 mg/L and chlorophyll $a < 5\mu$g/L. This study provides vital extensions to the previous approach through its use of multiple experts and a region of lakes, making the approach applicable to other regions of waterbodies and conclusions useful to inform policy.

Key words: chlorophyll *a*, eutrophication, expert elicitation, lakes, nutrient criteria, structural equation modeling, total phosphorus, water quality modeling

The United States Environmental Protection Agency (USEPA) named nutrients as the number one cause of water quality pollution for lakes, reservoirs, and ponds, causing the eutrophication of 3.8 million acres of waters (USEPA 2002). Eutrophication is a condition fueled by excess nitrogen and phosphorus that causes problems such as anoxia, noxious algal blooms, and fish kills (Novotny and Olem 1994, Chapra 1997). To protect the nation's waterbodies from excessive impairments, Sections 101(a) and 303(c) of the Clean Water Act require states and tribes to establish water quality standards.

Water quality standards contain an antidegradation clause, a qualitative designated use statement, and a qualitative or

*Current contact information: Department of Geography and Environmental Engineering, National Center for Earth-surface Dynamics, Johns Hopkins University, Baltimore, MD 21218. Email: M.A.KenneyPHD@gmail.com

quantitative criterion. The antidegradation clause is a narrative statement that the water quality standards must prevent additional degradation of a waterbody's use(s). The designated use is a narrative statement that articulates the water quality goal. The designated uses are set by states and describe the desired uses of the waters, such as public drinking water supply, primary contact recreation, and support of aquatic life. Because the designated use cannot be directly measured, a criterion serves as the measurable surrogate for the designated use. The criterion is either a numeric indicator or narrative statement, and is intended to indicate attainment (or nonattainment) of the designated use. Typically, the criterion is a combination of an easily measurable water quality variable and the critical level for that variable; this level for the criterion provides a minimum threshold that must be maintained or attained to support the designated uses. Though states have set criteria for toxic chemicals (e.g., metals or chlorine) and for water quality characteristics (e.g., dissolved oxygen and temperature), they have not widely addressed criteria to protect against cultural eutrophication impairments.

Currently the USEPA is encouraging states to adopt nutrient criteria (USEPA 2000b), identified as any measurable water quality variable or variables that can be used to detect eutrophication impairments (e.g., excessive algae, reduced water clarity) and their associated criteria levels. Selecting criteria is not a trivial issue, and it has led to much debate about what method should be used to establish these criteria.

The method to set nutrient criteria is left to the best judgment of the state; however, the most common method is the USEPA-endorsed ecoregion approach using a reference waterbody strategy (USEPA 2000b, 2000a). The USEPA method recommends setting criteria based on the reference conditions, those that represent pristine or minimally impacted waters. First, to set the criteria variables USEPA recommends that a state or tribe use four variables for each of its ecoregions that reflect both the causal and the response conditions of eutrophication: total phosphorus (TP), total nitrogen (TN), chlorophyll *a*, and Secchi depth. The guidance does not recommend that any variable be weighted as more important than another. Second, the criterion levels were set for each ecoregion by developing seasonal (winter, spring, summer, and fall) frequency distributions using all lakes and reservoirs within an ecoregion for each of the variables. The 25th percentile of each of the four seasons was calculated; the median value of these 25th percentiles was used to set the criterion level (Walker et al. 2007). The USEPA chooses to use the 25th percentile of these distributions because it approximates the 75th percentile of the distribution of reference waterbodies. These reference conditions, USEPA argues, should be used to set the upper bounds for an ecoregion to maintain or achieve natural and attainable conditions for lakes and reservoirs.

This approach to set nutrient criteria has two major flaws. First, it fails to substantively link the criteria to the designated use, meaning that the criterion variable is not necessarily a good predictor of designated use attainment. Second it fails to distinguish between science and societal values, making the implicit risk-based decision incorporated to set the criterion levels seem scientifically driven rather than a combination of science and value judgments (USEPA 2000b, Reckhow et al. 2005).

To address these flaws, an alternative approach was previously developed called the predictive approach to nutrient criteria (Reckhow et al. 2005, Reckhow et al. 2006, Kenney 2007). This procedure includes six steps. First, a dataset is developed that is representative of a given region of lakes (or separately for other types of waterbodies). Second, the probability of designated use attainment is quantified using expert assessments. Third, conceptual models of eutrophication processes are developed that link the water quality characteristics to the designated use. Fourth, given the conceptual models (step 3), structural equation modeling (SEM) is used to evaluate models parameterized with water quality data and designated use attainment assessments (step 1 and 2). Fifth, for those models deemed plausible using statistical tests, determine which variable(s) are most predictive of the designated use and use these variables (step 5) as the nutrient criteria variables. Sixth, using the results of the SEM (step 5), plot the risk of nonattainment of the designated use for the criterion variable(s). This information (step 6) is provided to a decision-maker, who would use the information to choose the criterion level(s). Though this procedure provides an excellent alternative conceptual framework for developing nutrient criteria, the method needs to be expanded beyond a single waterbody and a single expert to fully demonstrate its utility for policy implementation.

Our objective was to extend the methodology of the predictive approach to nutrient criteria to develop models for a region of waterbodies and to use multiple experts to assess the probability of designated use attainment. Using the predictive approach framework, this study determined which measured water quality variables are most predictive of designated use attainment for a region of lakes. These predictive variables were statistically determined, using SEM, by combining water quality data with assessments from multiple experts on the probability of designated use attainment. The approach was applied to lakes in south-central Florida, but the approach is appropriate for application in other regions.

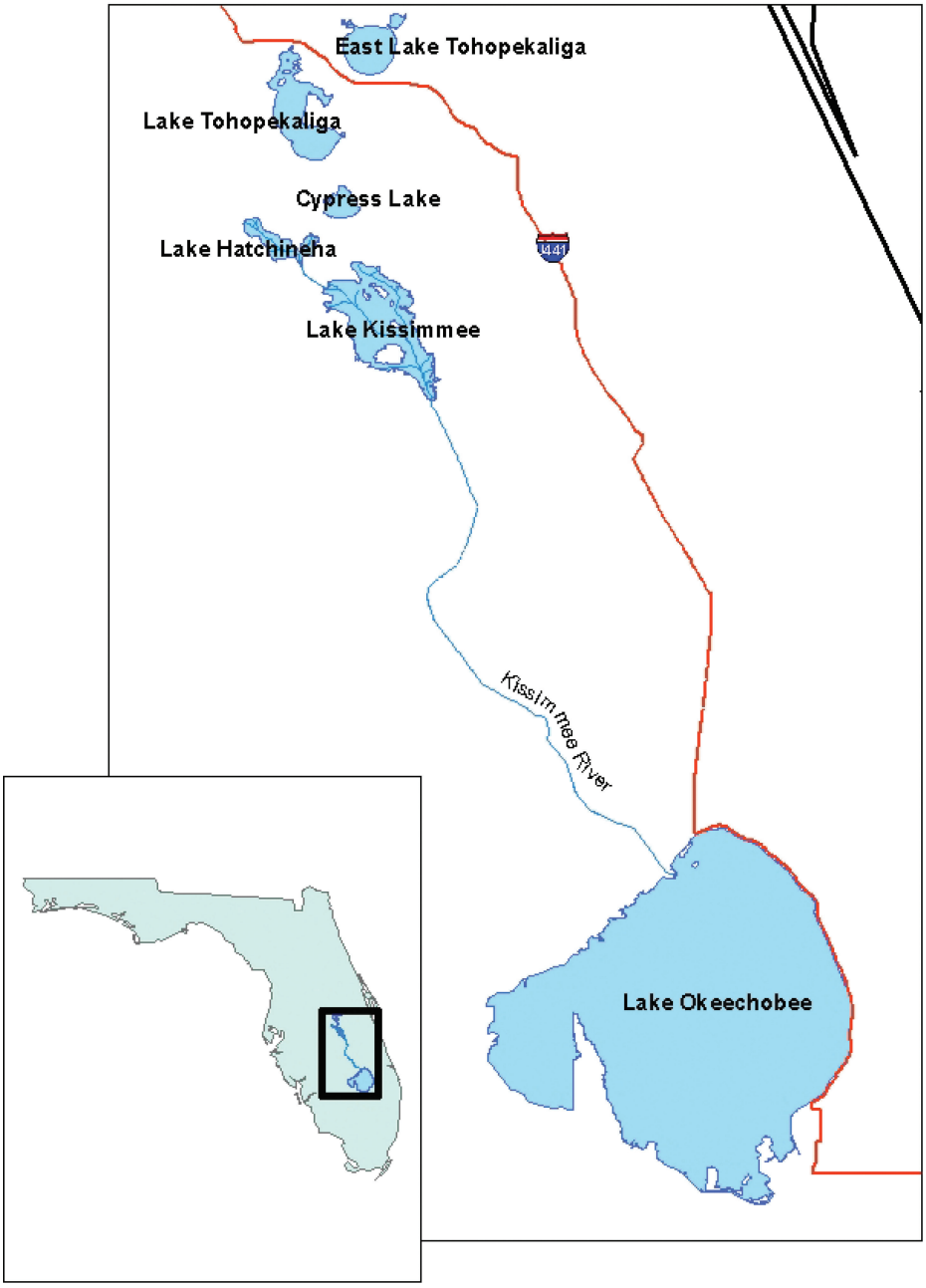

Figure 1.-Kissimmee Chain-of-Lakes in south-central Florida. The lakes in this region from north to south include East Lake Tohopekaliga, Lake Tohopekaliga, Cypress Lake, Lake Hatchineha, Lake Kissimmee, and Lake Okeechobee.

Methods

Study site

The lakes used in this study are those located in south-central Florida: Lake Okeechobee, East Lake Tohopekaliga, Lake Tohopekaliga, Cypress Lake, Lake Hatchineha, and Lake Kissimmee (Fig. 1). All of these lakes are located in an area with similar weather patterns, general lake dynamics, and similar seasonal physical, chemical, and biological trends and attributes (K. Havens, Chair of the Department of Fisheries and Aquatic Sciences, University of Florida, 2005, pers. comm.). These lakes are shallow (2–5 m) and encompass a wide range of trophic states (Havens 2003). The summary statistics for the untransformed variables used in the analysis are provided (Table 1).

Of the six lakes, all are considered eutrophic except East Tohopekaliga, which is considered mesotrophic. East Tohopekaliga has some of the highest color values (>200

Table 1.-Summary statistics for the eutrophication-related water quality variables used in the study. These values are from the data subset of 100 data rows used for the analysis throughout the paper. The data used in the study were log10 transformed, except for Secchi depth; however, the values presented are in their untransformed state. Data were obtained from the South Florida Water Management District database.

	Mean	Standard Deviation	Minimum	Maximum	Skewness	Kurtosis
Chlorophyll a (μg/L)	18.68	18.12	1.00	97.41	1.65	3.50
Total Phosphorus (mg/L)	0.05	0.03	0.01	0.15	1.54	2.60
Total Nitrogen (mg/L)	1.03	0.33	0.25	2.32	0.69	1.34
Total Kjeldahl Nitrogen (mg/L)	1.02	0.34	0.25	2.31	0.65	1.44
Dissolved Inorganic Nitrogen (mg/L)	0.04	0.05	0.01	0.21	1.92	3.04
Total Suspended Solids (mg/L)	5.61	6.91	0.50	36.50	2.18	6.06
Color (PCU)	91.01	62.06	24.00	267.98	1.27	0.91
Secchi Depth (m)	0.91	0.44	0.20	2.60	1.25	1.82

platinum-cobalt unit (PCU) during rain events), but has low chlorophyll *a*, TN, and TP values. All of the lakes in this region are highly regulated and have been periodically drawn down for Hydrilla control and muck removal (Havens 2005).

This system of lakes has several designated uses, but for the purposes of this study, we chose to use "the propagation and maintenance of a healthy, well-balanced population of fish and wildlife (Rule 62-302.400, Florida Administrative Code)," considered to be the most stringent designated use definition. This designated use addresses the Florida Class III classification, with the exclusion of the Class III use of "Recreation."

Step 1: Develop water quality dataset

The South Florida Water Management District (SFWMD) conducts routine monthly water quality sampling and collects data on standard limnological parameters. The dataset for the six lakes in South-central Florida was obtained from the SFWMD database (SFWMD 2007), which ranged from January 1995 to December 2001, consisted of the set of observations from the 12-mo annual data collection cycle for the following water quality variables: chlorophyll *a*, TN, TP, Secchi depth, color, total suspended solids, and total Kjeldahl nitrogen (Table 1). Thus, a data row consists of the data for the previously mentioned variables collected on the same day at the same location.

The dataset was first reduced by omitting any set of observations with a missing value. Next, a subset of 100 data rows was constructed. This subset was determined to be an accurate representation of the full dataset by retaining the sign (positive or negative) and the significance (significant or not) of each of the various bivariate relationships manifested in the original dataset.

Step 2: Expert assessments

Because designated use attainment cannot be directly measured, an expert assessment was conducted to elicit the probability of designated use attainment given correlated water quality data. The expert assessment protocol was similar to the approach used in Reckhow et al. (2005), which was based on best practices (Morgan and Henrion 1990, Keeney and Von Winterfeldt 1991, Meyer and Booker 1991, Clemen and Reilly 2001). Similar to the Reckhow et al. (2005) approach, this assessment had two parts, which were modified as described below.

In Part 1, the experts were asked to translate the given narrative designated use. The experts were additionally asked in this study, given their interpretation of the designated use, to state what they believed to be the ideal measurable variable to assess designated use attainment (for example, fish biodiversity), the change (attainment or nonattainment) point of this variable, and the commonly measured water quality variables that could be used as proxies for the ideal variable.

In Part 2, the experts were asked to go through an exercise to quantify the designated use for 100 correlated water quality data rows. Specifically, the elicitor asked each expert to look at an individual data row, considering all the water quality variables and their associated levels, and answer the question: "Given 100 hypothetical lakes in South-central Florida, all with identical average levels of these variables, and assuming other factors (e.g., morphological, climatic) vary randomly according to the characteristics of these Florida lakes, how many of the 100 lakes would be in attainment of the given designated use?" The image of 100 lakes was used to assist the experts in thinking about the question probabilistically; thus, the response could be directly translated from the assessment to the probability (i.e., an assessment of 50 lakes = 0.5 or 50% probability of designated use attainment). Additionally, the experts were asked to explain a sample of their assessments and were asked questions about their judgments, allowing the elicitors to check for consistency and biases. This check also ensured that experts had a chance to modify their judgments or provide a justification, assuring that the experts' assessments accurately represented their knowledge. These checks were particularly

useful given that experts are not always good at making probability assessments, especially near the extremes (Meyer and Booker 1991, Clemen and Reilly 2001). The final result was an assessment for each data row, provided by each expert, of the probability of designated use attainment (0–100%); this information is necessary to parameterize the models in steps 3–5.

Because the expert assessment included multiple experts, it was essential to combine the assessments into an aggregate designated use value to use the expert judgment as a response variable. After careful consideration of multiple methods, such as Bayesian approaches and behavioral approaches, equal-weighted averaging was used because it is extremely robust (Clemen and Winkler 1999). Furthermore, to employ other techniques, it is essential to have additional information to appropriately weight the relative value of one expert versus another expert (Clemen and Winkler 1999). Because this information was unavailable, the experts were weighted equally.

Steps 3, 4, and 5: Structural equation modeling

Structural equation modeling (SEM) is a multivariate statistical technique that can be used to describe linear relationships among variables (Bollen 1989, Kline 1998, McCune and Grace 2002, Grace 2006). The technique is a general extension of multiple regression where the causal relationships among the variables can be described with multiple linear equations that describe both direct and indirect effects (Bollen 1989). Therefore, SEM is an excellent statistical technique to use when conditions, such as eutrophication, can be described through causal interactions that can be represented by the covariance between variables (Arhonditsis et al. 2006, Grace 2006). The relative strength of these relationships, regardless of the variable's units and scale, are described using standardized path coefficients. These values, similar to the coefficients in regression, can be used to calculate both the direct and indirect effects (the total effects being the sum of the direct and indirect effects).

SEM is an *a priori* method. By this we mean that a researcher develops a model reflective of the background knowledge of how the system works and then tests it with data. Therefore, unlike other approaches, a researcher using SEM hopes to accept the null hypothesis (H_o: data matches model) because this means the model is a plausible representation of the system (McCune and Grace 2002). Rejection of the null hypothesis indicates that the data do not support the model structure. Specifically, an SEM fit is tested by minimizing the difference, or residual, between the model-implied covariance matrix and the data-implied covariance structure (McCune and Grace 2002).

Because no single test statistic can incorporate all the different facets of model fit, multiple test statistics were used to determine the adequacy of the models. Four widely used methods were applied to evaluate our models: χ^2, Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), and Root Mean Square Error of Approximation (RMSEA; Kline 1998).

- The χ^2 statistic indicates whether or not the model, which in SEM is by definition overidentified, differs statistically from a just-identified version of the model; a nonsignificant model (accept the null hypothesis) would not differ statistically as indicated by p-values >0.05.
- The CFI is a test statistic that signifies the overall proportion of variance explained by the model; good fit is indicated by a CFI > 0.9.
- The TLI adjusts for the proportion of explained variance, and a model is considered to have good fit if the TLI > 0.9.
- The RMSEA is a model fit index that considers the model's residuals; an RMSEA < 0.1 indicates a satisfactory model representation.

If one evaluates multiple models, and the test statistics indicate that several models are plausible, it is reasonable to compare the relative fit of the competing models by using two model selection criteria, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC). Both approaches reward models that have both good fit and are parsimonious; thus, models that include additional parameters that do not add additional information are penalized. Smaller values of both of these information criteria indicate preferred models.

The models were structured using the conceptual models provided by the experts in part 1 of the expert elicitation (step 2) and confirmed with literature descriptions of the conditions causing eutrophication in south-central Florida. The final models were a combination of multiple experts' understanding of the relationships. The models were parameterized using the water quality data from the dataset previously described and the designated use attainment data from the expert elicitations. The models were then tested using the SEM test statistics described above. The plausible models were those that that satisfactorily performed on the four model test statistics.

Steps 6: Probability of the risk of non-attainment of the designated use

Using the results of SEM, one can construct plots of the risk of nonattainment of the designated use. These plots were developed by calculating the probability of designated

use attainment given the range of the predictive variable(s), keeping other variables constant. The calculated probability of designated use reflects uncertainty in the elicited judgments as well as natural variability and error in the water quality dataset. Thus, decision-makers can use the plots, and associated values, to determine criterion levels given their risk of nonattainment of the designated use.

Results

Expert assessment

In the first part of the assessment, the experts provided their translation of the designated use. Three of the expert (Experts 1, 3, and 4) interpretations focused on attainment being a full diversity of native aquatic life. The definition of what was included within aquatic life differed, with one expert defining it as those fish and wildlife that live within the water system (e.g., including birds, reptiles), one expert agreeing with the previous expert and expanding it to also include aquatic invertebrates, and one expert agreeing with the previous two experts but additionally expanding the definition to include native plants. The other expert (Expert 2) had a different perspective and stated that the designated use focuses on sport fish population and wildlife associated with a healthy sport fish population. This study did not judge whether one interpretation of the designated use was more "right" than another. The experts also identified what they felt were important indicators of eutrophication-related impairments of the designated use; all experts stated that TP and chlorophyll *a* were important predictors of designated use attainment. The importance of other variables, such as color, TN, and Secchi depth, varied among experts. The experts also supplied conceptual models of the conditions that lead to eutrophication-related impairments in this region; this information provided the starting point for structuring the SEMs.

After learning how the experts interpreted the designated use, in the second part of the elicitation they quantified the probability of designated use attainment. To assess how similarly the experts responded, each expert's responses were correlated against another expert's responses. If the experts responded exactly the same for all the data rows, the correlation was 1; a moderate to high correlation (>0.5) is desirable to demonstrate that the experts are approaching the probability assessment similarly. Because each expert has different expertise and experiences, some variation is expected, but not extreme differences.

The results of the correlation table indicate that the degree of expert correlation varied (Table 2). Specifically, Expert 1 and 3 were strongly correlated (0.74), while Expert 1 and 4 (0.46) and Expert 3 and 4 (0.57) were somewhat correlated. Finally, Expert 2 was negatively correlated with the other three, indicating that this expert's responses were fundamentally different from the others. This difference is largely reflective of Expert 2's notably different definition of the designated use stated earlier.

Table 2.-Correlation of expert responses. Strong or moderate correlations indicate that Experts 1, 3, and 4 were thinking similarly about designated use attainment, as indicated by similar responses to the probability assessment. Extremely low correlations are highlighted in bold; Expert 2's responses were not correlated with any of the other experts.

	Expert 1	Expert 2	Expert 3	Expert 4
Expert 1	1			
Expert 2	**−0.21**	1		
Expert 3	0.74	**−0.22**	1	
Expert 4	0.46	**−0.10**	0.57	1

Because multiple experts were used, it was important to combine the expert judgments. The largest benefit of having multiple experts is the collection of data from a diversity of sources of expertise; therefore, there should be a good reason to weight experts unequally or to exclude an expert from the dataset (Clemen and Winkler 1999). As a result, one dataset was created that used all of the experts' assessments and combined their assessments using equal-weighted averaging (Clemen and Winkler 1999). Because Expert 2's responses were notably different from the other three experts, a second dataset was developed that excluded Expert 2's responses and combined the three experts' responses using equal-weighted averaging. By creating both of these datasets, the authors could assess the sensitivity of the model results.

Structural equation modeling

Of the tested models, two, Model 1 and 2, demonstrated good fit. Model 1 is a simple nutrient criteria model (Fig. 2). In the model, increased TP levels directly affect chlorophyll *a* levels, describing the primary eutrophication process. Designated use attainment is directly affected by TP, chlorophyll *a*, and Secchi depth. In the model, increased levels of TP and chlorophyll *a* decreased the probability of designated use attainment, whereas increased Secchi depth levels increased the probability of designated use attainment. All paths (similar to regression coefficients) in this model, except for Secchi depth, are significant at $p \leq 0.05$. Model 1 was tested both with the dataset that included all experts (Fig. 2b) and with the dataset that excluded Expert 2 (Fig. 2a).

For the model that used the dataset that excluded Expert 2 (Fig. 2a), the Chi-squared test statistic ($\chi^2 = 1.76$; df = 1;

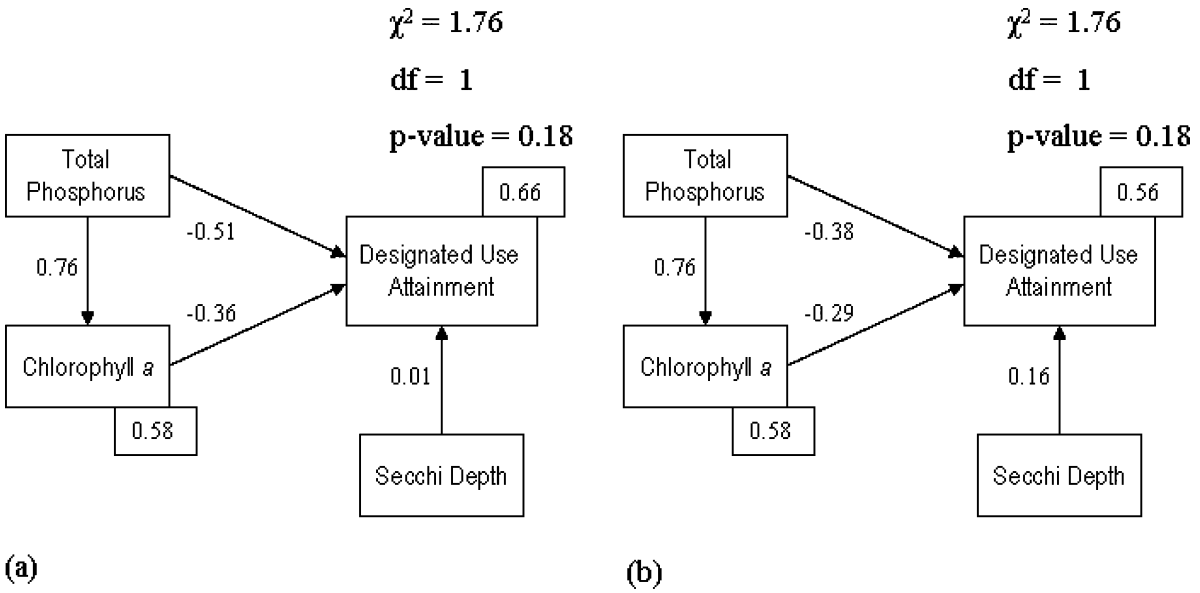

Figure 2.-Structural equation Model 1 for the Kissimmee Chain-of Lakes for (a) dataset without Expert 2, and (b) dataset with all the experts. The values on the arrows are the standardized path coefficients and the values in rectangles are the R^2 values. The χ^2 (Chi-squared test statistic), df (degrees of freedom), and p-value refer to a model fit test statistic; p-values > 0.05 indicate good model fit. In this model, the most predictive variables are total phosphorus and chlorophyll *a*.

p-value = 0.18) indicates that the model is a plausible representation. The other tests of model fit, such as CFI (0.996), TLI (0.98), and RMSEA (0.087), additionally provide support that the model is reasonable.

The standardized path coefficients in this model provide us with information on the relative strength of the relationships, regardless of the variable's units and scale, across the variables. In this model (Fig. 2a), comparing the total effects from the standardized path coefficients, the most predictive variable is TP (direct effect $= -0.51$; indirect effect $= 0.76^* -0.36 = -0.27$; total effect $= -0.51 + -0.27 = -0.78$); the second most predictive variable is chlorophyll *a* (direct effect $= -0.36$; indirect effect $= 0$; total effect $= -0.36$). In comparison, the other paths provide much less explanation of designated use attainment.

The R^2 values for the model can be interpreted similarly to multiple regression. In this model a modest amount of variance is accounted for by chlorophyll *a* (0.58), whereas a considerable proportion of variability is explained for designated use attainment (0.66).

Model 2 is a slightly more complex nutrient criteria model (Fig. 3). In this model, increased levels of TP and Secchi depth cause increased levels of chlorophyll *a*. Increased chlorophyll *a* and color cause a decrease in Secchi depth. Increased levels of TP and chlorophyll *a* decreases the likelihood of designated use attainment; whereas, increased Secchi depth levels increases the likelihood of designated use attainment. All paths in this model, except for the path between Secchi depth and designated use, are significant at the 5% level. Similar to Model 1, Model 2 was tested both with the dataset that included all experts (Fig. 3b) and with the dataset that excluded Expert 2 (Fig. 3a). Color was added to this model because allochthonous organic substances are the primary driver of color and affect the Secchi depth and trophic status (Canfield and Hodgson 1983).

For the model that used the dataset that excluded Expert 2 (Fig. 3a), the Chi-squared test statistic ($\chi^2 = 2.37$; df $= 2$; p-value $= 0.31$) indicates that the model is reasonable. The other tests of model fit, such as CFI (0.999), TLI (0.994), and RMSEA (0.043), additionally provide support that the model is a plausible representation. The interpretation of the predictive variables and R^2 values is similar to Model 1a.

For comparison sake, we also tested the models (Models 1b and 2b) with a designated use attainment dataset that included all of the experts (Fig. 2b and 3b). Using the model fit test statistics and the path coefficients, the results of the models with the dataset of all experts, Model 1b and 2b, were similar to their 1a and 2a counterparts. Of particular note, the most predictive variables and direction of the path coefficients remained the same. The main difference is the value for the standardized path coefficients; because Expert 2's responses were very different from the other experts, the loss in the relative strength in the TP and chlorophyll *a* path

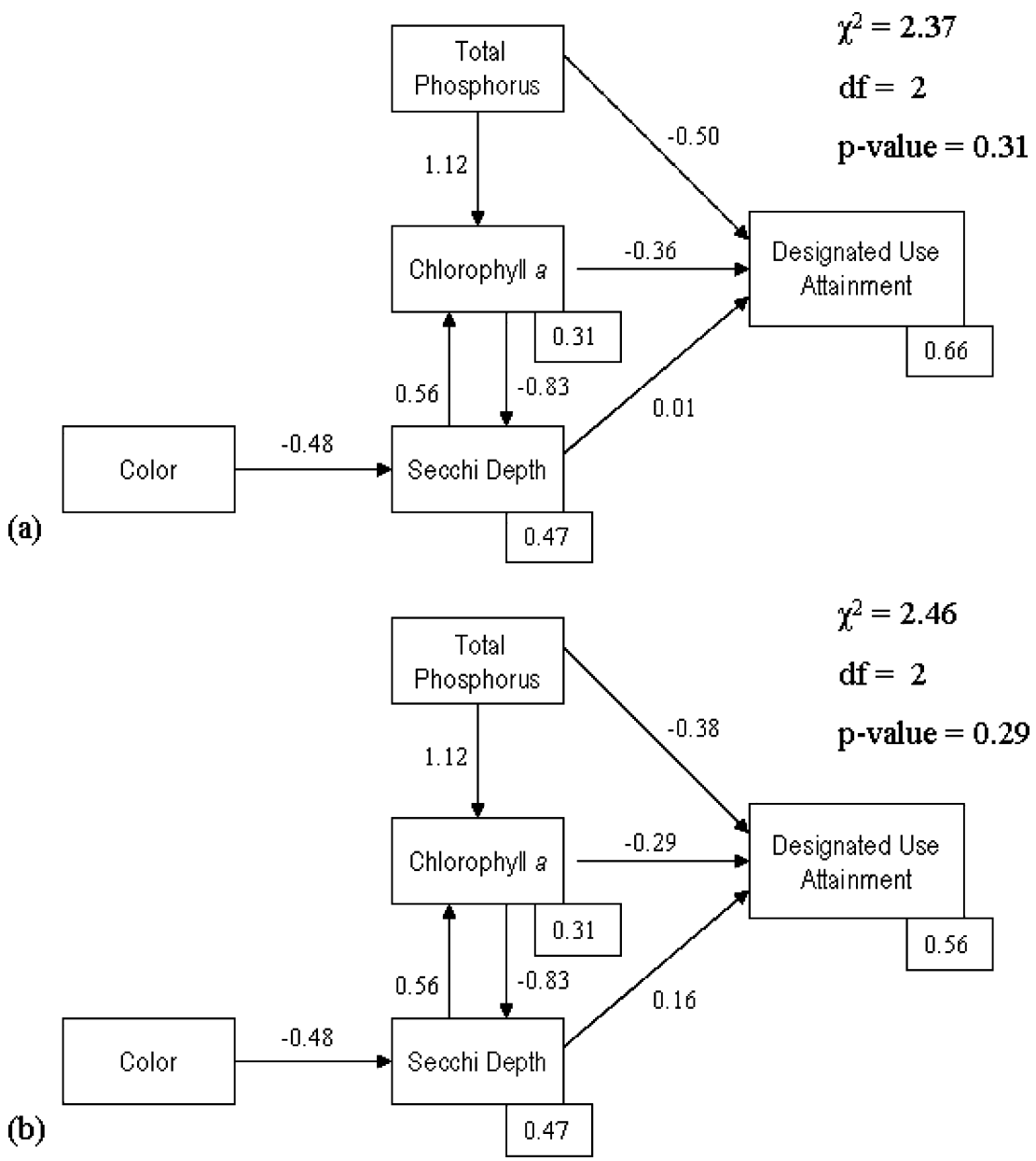

Figure 3.-Structural equation Model 2 for the Kissimmee Chain-of Lakes for (a) dataset without Expert 2, and (b) dataset with all the experts. The values on the arrows are the standardized path coefficients and the values in rectangles are the R^2 values. The χ^2 (Chi-squared test statistic), df (degrees of freedom), and p-value refer to a model fit test statistic; p-values > 0.05 indicate good model fit. In this model, the most predictive variables are total phosphorus and chlorophyll *a*.

coefficients is likely a result of the inclusion of Expert 2's responses.

Comparing the models using AIC and BIC, Model 2a (AIC = 606.4, BIC = 632.5; sample-size adjusted BIC = 600.9) outperforms Model 1a (AIC = 610.5, BIC = 626.2; sample-size adjusted BIC = 607.2). Even though Model 1a has a lower BIC than Model 2a, Model 2a has a lower sample-size adjusted BIC as well as a lower AIC, indicating that Model 2 has better model fit. Overall, the differences between the models are very slight, and both models indicated that TP and chlorophyll *a*, in that order, are the most predictive variables.

The model was applied to consider the probability of designated use attainment for various levels of the two most

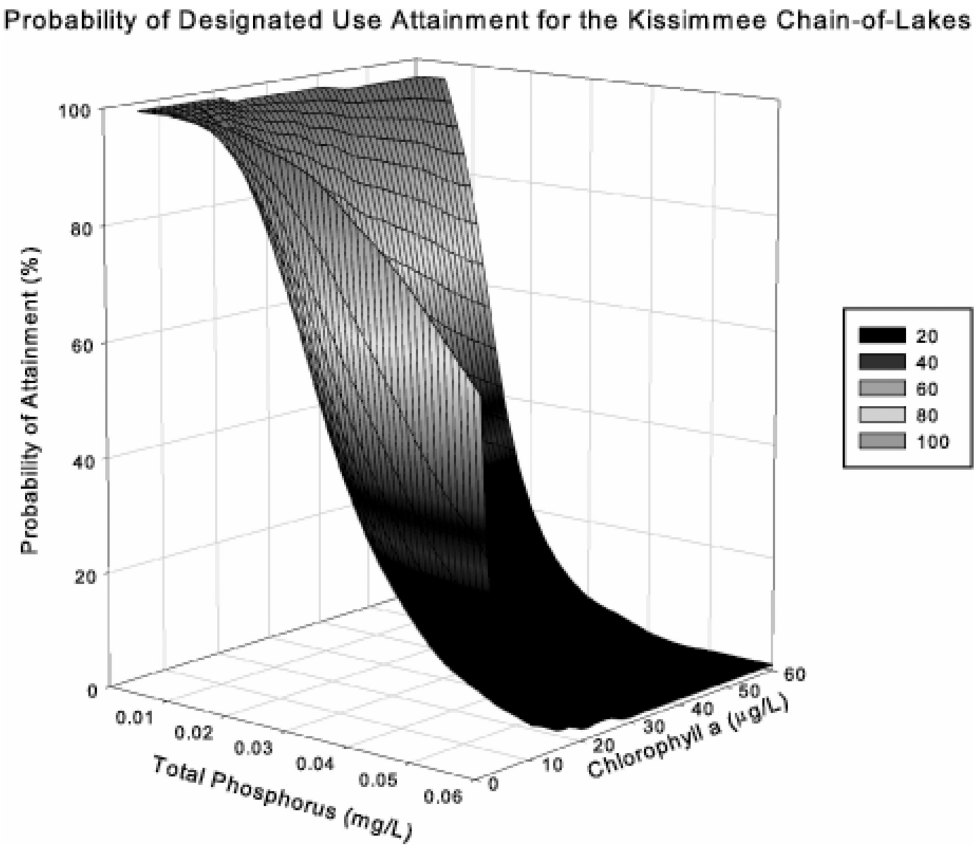

Figure 4.-Three-dimensional surface of probability of designated use attainment for nutrient criteria, total phosphorus and chlorophyll *a*. The average expert response is logit transformed, and the criteria levels are for the most likely range of summer values for both of these variables.

predictive variables, TP and chlorophyll *a*. Using these variables, we considered appropriate candidate criterion levels by plotting the variables against a logit transformation of the average experts' responses. The values of the other variables were set to their mean values; this assumption means that the uncertainty is understated because the full range of the other variables was not considered. Given the model results, the two candidate criteria (Fig. 4) and a single TP or chlorophyll *a* criterion (Fig. 5a and 5b) were plotted as candidate criterion(a).

Values for TP and chlorophyll *a* were presented for a given probability of designated use attainment, or the surface of the probability of designated use attainment for TP and chlorophyll *a* (over a range of most-likely values of both variables), conditional on Secchi depth remaining constant (Fig. 4). Using this graph, a decision-maker could determine his/her risk of noncompliance of the designated use and assess which levels of TP and chlorophyll *a* would lead to an acceptable level of attainment.

Instead of considering a surface of potential values, one could consider the same problem in two-dimensions for TP (Fig. 5a) and chlorophyll *a* (Fig. 5b). For TP (Fig. 5a), the probability of designated use attainment remains high when the TP levels are <0.015 mg/L and then dramatically decreases until it levels out at 10% or less attainment of designated use when TP values are >0.05 mg/L.

For chlorophyll *a* (Fig. 5b), the figure exponentially decays over the range of values. Therefore, the chlorophyll *a* values must be <5 μg/L to achieve a high (90% or greater) attainment of the designated use; setting a chlorophyll *a* criterion level at a value such as 30 μg/L would yield a 40% attainment of the designated use. The three-dimensional surface of the probability of designated use attainment (Fig. 4) emerges from the combination of the exponentially declining chlorophyll *a* graph (Fig. 5b) and the reversed S-shaped TP graphs (Fig. 5a).

Discussion

Given the variety of expertise, this range was captured by using multiple experts and then aggregating their assessments.

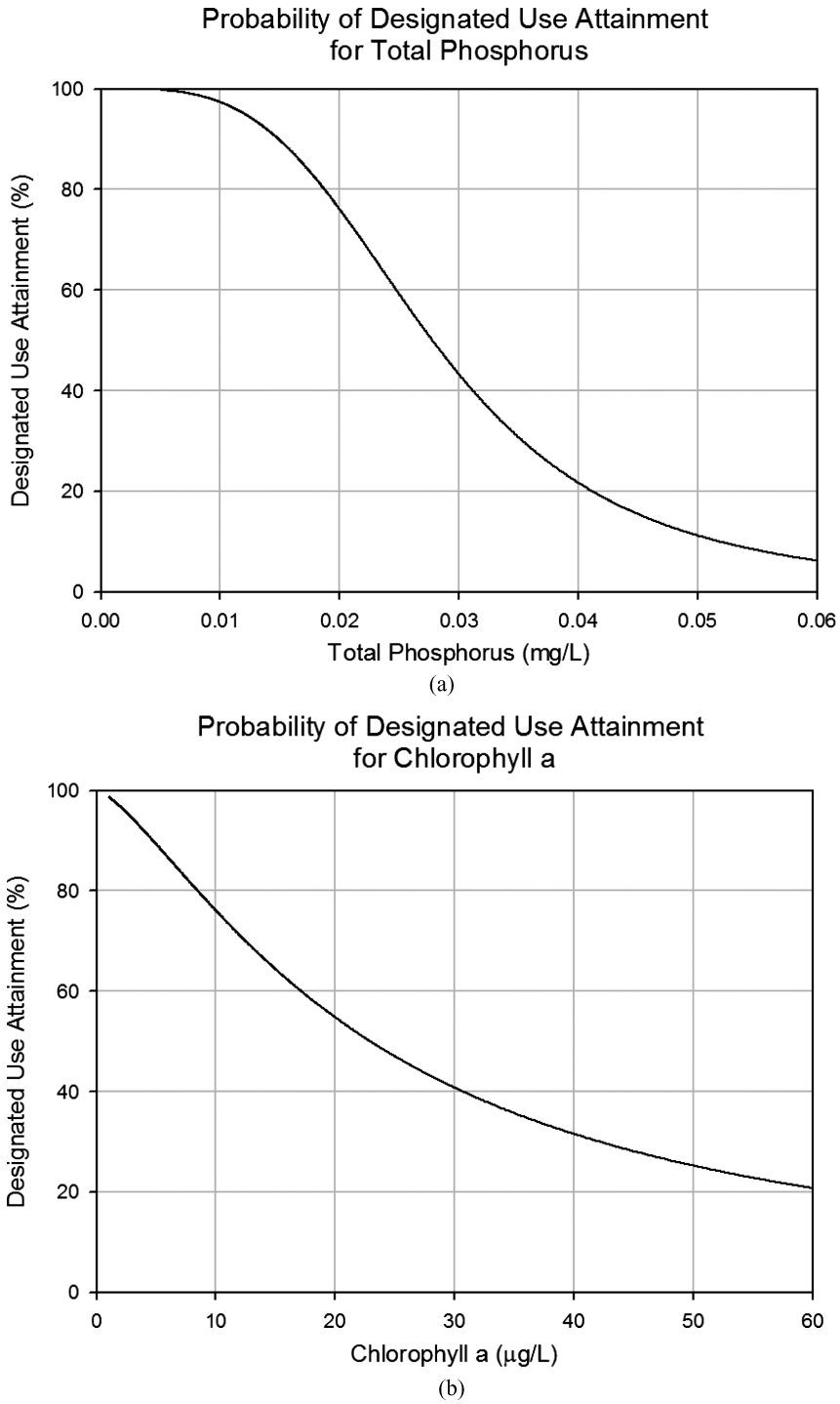

Figure 5.-Two-dimensional graph of a probability of designated use attainment versus the candidate criterion, (a) total phosphorus, and (b) chlorophyll a. Decision-makers can use such a graphic to determine the criterion level that will meet their risk of nonattainment of the designated use.

The inherent variability in the experts' responses, particularly when one seeks and assesses such a diverse group of experts as performed in this study, stresses the need to use multiple experts.

The structural equation modeling results highlight that multiple plausible models can link eutrophication and designated use attainment; our study had two plausible models. Because both of the models indicated that the same variables

were predictive of the designated use, TP and chlorophyll *a*, the authors are confident that the results are appropriately identifying the most predictive variables.

Additionally, using the model results, an analysis of the risk of nonattainment of the designated use was presented (Fig. 4 and 5). A decision-maker can use such information to set a criterion level or criteria levels based on their risk of nonattainment of the designated use. This is the key difference between this method and the USEPA approach. With the predictive approach to nutrient criteria, a decision-maker(s) (advisor, state committee, agency, or legislature) would determine what they believed was an acceptable probability of designated use attainment and then use the values on the graph (or associated results) to set the criterion level.

Research directions underway include an extension of this approach to a statewide assessment, with multiple ecoregions and multiple designated uses. Part of this research team has also been involved in developing methods that could provide a decision-maker with a specific nutrient criterion level using his or her value tradeoff between water quality and cost. This emerging method combines water quality models with decision value models to determine the optimal criterion level, thus maintaining the transparency between scientific assessments and value judgments (Kenney 2007).

Setting nutrient criteria that are truly predictive of eutrophication-related designated use impairments is essential to properly identify, and then maintain or restore, the desired uses of lakes and reservoirs. This study provides important extensions to the predictive nutrient criteria approach, using a region of lakes and assessments from multiple experts. Using the expanded approach presented in this paper, states and tribes can now apply the method to other regions to select nutrient criteria predictive of the designated use.

Acknowledgments

This project was funded by a grant from the Water Environment Research Foundation (WERF) through the University of North Carolina Water Resources Research Institute (NC WRRI). M.A. Kenney was supported by a fellowship from the National Water Research Institute (NWRI) and was partially supported by the STC program of the National Science Foundation via the National Center for Earth-surface Dynamics under the agreement number EAR-0120914. Dan Canfield, Paul Gray, Karl Havens, and E. Conrad Lamon, III provided expert assessments. Water quality data were provided from South Florida Water Management District (SFWMD). Robert T. Clemen, Ralph L. Keeney, Craig A. Stow, and three reviewers provided useful comments to improve this article.

References

Arhonditsis, G.B., C.A. Stow, L.J. Steinberg, M.A. Kenney, R.C. Lathrop, S.J. McBride and K.H. Reckhow. 2006. Exploring ecological patterns with structural equation modeling and Bayesian analysis. Ecol. Model. 192:385-409.

Bollen, K.A. 1989. Structural equations with latent variables. John Wiley & Sons, Inc., New York.

Canfield, D.E. and L M. Hodgson. 1983. Prediction of Secchi disc depths in Florida lakes: Impact of algal biomass and organic color. Hydrobiologia 99:51-60.

Chapra, S.C. 1997. Surface water-quality modeling. McGraw-Hill, Boston, MA.

Clemen, R.T. and R.L. Winkler. 1999. Combining probability distributions from experts in risk analysis. Risk Anal. 19:187-203.

Clemen, R.T. and T. Reilly. 2001. Making hard decisions with decision tools. Duxbury Thomson Learning, Pacific Grove, CA.

Grace, J.B. 2006. Structural equation modeling and natural systems. Cambridge University Press, Cambridge.

Havens, K.E. 2003. Phosphorous-algal bloom relationships in large lakes in South Florida: Implications for establishing nutrient criteria. Lake Reserv. Manage. 19:222-228.

Keeney, R. L. and D. Von Winterfeldt. 1991. Eliciting Probabilities from Experts in Complex Technical Problems. Transactions on Engineering Management 38:191-201.

Kenney, M. A. 2007. Which Nutrient Criteria Should States and Tribes Choose to Determine Water Body Impairment? Using Science and Judgment to Inform Decision-Making. Pages 20-25 *in* National Water Research Institute Graduate Fellowship Research Conference, Orange, CA.

Kline, R. B. 1998. Principles and practice of structural equation modeling. The Guilford Press, New York.

McCune, B. M. and J. B. Grace. 2002. Structural Equation Modeling. Pages 233-256 Analysis of Ecological Communities. MjM Software Design, Gleneden Beach.

Meyer, M. A. and J. M. Booker. 1991. Eliciting and Analyzing Expert Judgment: A Practical Guide. Academic Press Ltd., London.

Morgan, M. G. and M. Henrion. 1990. Uncertainty: A guide to dealing with uncertainty in quantitative risk and policy analysis. Cambridge University Press, Cambridge.

Novotny, V. and H. Olem. 1994. Water Quality: Prevention, Identification, and Management of Diffuse Pollution. 1 edition. Van Nostrand Reinhold, New York.

Reckhow, K. H., G. B. Arhonditsis, M. A. Kenney, L. Hauser, J. Tribo, C. Wu, K. J. Elcock, L. J. Steinberg, C. A. Stow, and S. J. McBride. 2005. A predictive approach to nutrient criteria. Environmental Science & Technology 39:2913-2919.

Reckhow, K. H., G. B. Arhonditsis, M. A. Kenney, S. J. McBride, R. J. Gosnell, C. A. Stow, and H. W. Paerl. 2006. Water Quality Indicators: Nutrient Impacts on Chlorophyll or Algae

Species Composition. Report 02-ECO-1, Water Environment Research Federation.

SFWMD. 2007. Environmental Monitoring: Kissimmee Chain-of-Lakes Dataset 1995-2001. South Florida Water Management District, West Palm Beach, FL.

USEPA. 2000a. Ambient water quality criteria recommendations: Information supporting the development of state and tribal nutrient criteria for lakes and reservoirs in nutrient ecoregion XII. U.S. Environmental Protection Agency, Office of Water, Washington, DC. EPA 822-B-00-013.

USEPA. 2000b. Nutrient criteria technical guidance manual: lakes and reservoirs. U.S Environmental Protection Agency, Office of Water, Washington, DC. EPA-822-B00-001.

USEPA. 2002. National water quality inventory. U.S. Environmental Protection Agency, Office of Water, Washington, DC.

Walker, J.L., T. Younos and C.E. Zipper. 2007. Nutrients in lakes and reservoirs - A literature review for use in nutrient criteria development. Virginia Ploytechnic Institute and State University Virginia Water Resources Research Center, Blacksburg, VA.