An integrated approach to TMDL development for the Neuse River estuary using a Bayesian probability network model (Neu-BERN)

Mark E. Borsuk[1], Craig A. Stow, and Kenneth H. Reckhow

Division of Environmental Science and Policy

Nicholas School of the Environment and Earth Sciences

Duke University, PO Box 90328

Durham, North Carolina 27708-0328

October 21, 2002

[1] Corresponding author. Current address: Department of Systems Analysis, Integrated Assessment, and Modelling (SIAM), Swiss Federal Institute for Environmental Science and Technology (EAWAG), P.O.Box 611, 8600 Dübendorf, Switzerland. Phone: +41 1 823 5082, Fax: +41 1 823 5375, E-mail: mark.borsuk@eawag.ch

Abstract

We develop a Bayesian probability network model to characterize eutrophication in the Neuse River Estuary, North Carolina, and support the estimation of a TMDL for nitrogen. Unlike conventional simulation models, Bayesian network models describe probabilistic dependencies among system variables rather than substance mass balances. Full networks are decomposable into smaller submodels, with structure and quantification that reflect relevant theory, judgment, and/or observation. Model predictions are expressed probabilistically, which supports consideration of frequency-based water quality standards and explicit estimation of the TMDL margin of safety. For the Neuse Estuary TMDL application, the Bayesian network can be used to predict compliance with the dissolved oxygen and chlorophyll a regulatory criteria as a function of riverine nitrogen load. In addition, the model includes ecological endpoints, such as fish kills and shellfish survival, that are typically more meaningful to stakeholders than conventional water quality characteristics. Incorporating these unregulated attributes into TMDL decisions will require explicit consideration of costs, benefits, and relative likelihoods of various possible outcomes under alternate loading scenarios.

Keywords: Neuse Estuary Bayesian Ecological Response Network (Neu-BERN), risk analysis, water quality modeling, Total Maximum Daily Load (TMDL), decision-making

Introduction

The Neuse River estuary, North Carolina (Figure 1), has been experiencing severe consequences of eutrophication in recent years including excessive algal blooms, low levels of dissolved oxygen, declining shellfish populations, large fish kills, and outbreaks of toxic microorganisms. These problems have led to the Neuse River being declared one of the twenty most threatened rivers in the United States (ARF 1997). The Neuse River estuary has also been included on the federal list of impaired waters under section 303(d) of the Clean Water Act. As in many other marine systems, nitrogen has been identified as the pollutant of concern in the estuary because it is believed to stimulate the excessive algal growth that is at the root of other ecological problems. Therefore, the USEPA has required that a Total Maximum Daily Load (TMDL) for nitrogen be developed by the State of North Carolina. TMDLs establish the maximum pollutant loading to a water body that will allow it to meet water quality standards and attain its designated uses (Office of Water 1999). These assessments then provide the basis for states to require watershed-based pollutant controls to achieve the TMDL. The impaired condition of waters across the nation underlies the requirement that thousands of TMDLs for pollutants must be developed in the next ten years (NRC 2001).

To develop a TMDL, a linkage must be defined between the pollutant load and the symptoms of water quality impairment. In many cases, water quality models provide the scientific basis for this pollutant-effect relationship and therefore play a critical role in pollutant load decisions. Most receiving water models used for TMDL development are of the deterministic, mechanistic variety (Lung 2001; Office of Water 1997a). That is, they reflect the belief that the values of water quality endpoints are determined by a finite set of processes that can be represented by mathematical expressions. Once calibrated to a system of interest, these models are assumed to adequately represent reality, and various pollutant reduction strategies are simulated with the model to anticipate water quality effects.

Simulation models of increasing mechanistic complexity have been developed and applied in recent years (Thomann 1998), yet there is little evidence that much confidence can be attached to the predictions of such models (NRC 2001). While calibration studies may sometimes show a close fit between predictions and the observations to which models are calibrated, verification studies against different sets of data suggest that prediction errors may be large, particularly for models of higher resolution and greater mechanistic detail (Reckhow 1994). This result should not be surprising considering the complexity of natural systems relative to even the most sophisticated simulation models. It is not reasonable to expect that all of the mechanisms of natural systems can ever be fully understood and assembled into accurate predictive models (Pace 2001). Nature is simply too complex.

The difficulty with exact representation of nature is even more problematic when attempting to extend water quality models to ecological endpoints, such as fish kills, shellfish mortality, or fish health. At the scale employed by most simulation models, the ecological processes associated with these attributes are too complex or stochastic to be characterized mathematically. For this reason, most mechanistic simulation models are only used to predict biochemical variables such as chlorophyll a or dissolved oxygen concentration. Decision-makers are then left in the difficult position of having to extend model results to the water body attributes of true concern to the public.

An alternative approach to modeling is one that has been adopted by physicists who use probabilistic expressions to characterize the aggregate effects of small-scale molecular motion. In a similar manner, water quality modelers can summarize small-scale, unpredictable, or unmanageable processes with probabilistic expressions and then focus model development on describing the large-scale effects of the most important controlling factors. Models of this type may be more useful for decision support because they can provide direct answers to questions about large-scale ecosystem response. Such policy-relevant questions broadly concern the relationship between a management option and an attribute of concern to the public, such as, "Will the frequency (probability) of fish kills in the Neuse estuary be reduced to a tolerable level by a 30% nitrogen load reduction?" Replication of detailed environmental processes is not usually required to answer such questions.

These observations about the type of models useful for supporting TMDL decisions suggest the general utility of Bayesian networks (Pearl 1988). Also known as probability networks, belief nets, Bayes nets, or influence diagrams, Bayesian networks are graphical models that depict the nature of relationships among a number of uncertain variables. These relationships are quantified using mathematical models, data, or expert opinion that capture the aggregate effect of the controlling processes. The effects of secondary processes are then summarized with probabilistic expressions.

We describe the development and application of a Bayesian network model for TMDL evaluation in the Neuse River estuary (the Neuse Estuary Bayesian Ecological Response Network, or Neu-BERN). The network combines relevant information expressed in a variety of forms into one cohesive structure linking riverine nitrogen loading to the ecological consequences of importance to the public. Most of the individual model relationships are described in detail in our previous publications. Our focus here is on the integration of these relationships into a Bayesian network useful for TMDL decision support.

Modeling Method

Fundamental to the utility of Bayesian networks is their graphical depiction. In such a graph, round cells represent important system variables and connecting arrows represent dependent relationships among these variables. Relationships may reflect direct causal dependencies or the aggregate effect of more complex associations. The conditional independence implied by the *absence* of a connecting arrow between two nodes greatly simplifies the modeling process by allowing separate submodels to be developed for each relationship indicated by the *presence* of an arrow. These submodels characterize conditional probability distributions that reflect the aggregate response of each variable to changes in its up-arrow "parents", together with the uncertainty in that response (Pearl 2000).

Submodels may be based on either (1) mathematical representation of dominant processes, (2) statistical associations derived from historical data, or (3) probabilistic quantities elicited from scientific experts. Any model representation or level of mechanistic detail is appropriate as long as the uncertainty associated with each relationship can be quantified in the form of a conditional probability distribution. We believe that models that are based on mechanistic understanding yet remain within the bounds of data-based parameter estimation will be particularly useful tools in this regard. The incorporation of mechanism will improve confidence in predictions made for changed conditions, while the statistical methods will provide an empirical basis for parameter selection and allow for estimates of predictive uncertainty.

Unfortunately, appropriate and sufficient observational data may not always exist to estimate the parameters of even simple mechanistic models. As a consequence, the elicited judgment of scientific experts may be required to quantify some of the probabilistic relationships. Of course, the use of subjective judgment is not unusual in TMDL modeling. Even the most process-based computer simulations rely on subjective judgment as the basis for the mathematical formulations and the choice of parameter values. Thus, the explicit use of scientific judgment in Bayesian networks should be an acceptable practice. In fact, by formalizing the use of judgment through well-established techniques for expert assessment (see Morgan and Henrion 1990), the Bayesian network method may improve the chances of accurate and honest predictions.

Once all significant system variables are linked in a single network using conditional probabilistic relationships, predictive distributions of model endpoints can be generated for any set of values for input variables. These predicted endpoint probabilities, and the relative change in probabilities between alternative scenarios, convey the expected system response to management while fully accounting for predictive uncertainties. Such uncertainties, which arise from both model uncertainty and natural variability, give decision makers and stakeholders an explicit characterization of the risk of non-attainment of TMDL management objectives.

Identification and Development of Model Relationships

A graphical model representing the variables and relationships important to eutrophication in the Neuse Estuary has been developed through a joint process of stakeholder involvement and scientific characterization (Borsuk et al. 2001a). Predictive endpoints include algal density, as measured by chlorophyll a concentration, abundance of the toxic microorganism *Pfiesteria piscicida*, fish population health, frequency of fish kills, and shellfish survival (Figure 2). These are related to their immediate causal variables which are then related back to their causes, and so on, back to variables that can either be considered marginal variables representing natural variability, or those that will be influenced by TMDL management decisions. This qualitative diagram then serves as the framework for developing quantitative submodels to relate the selected variables. Because intermediate variables and relationships are included in the model only if they contribute to our ability to predict model endpoints, the model can be best explained by starting with the endpoints and proceeding in the "up-arrow" direction.

Fish kills

As revealed by the stakeholder study, the frequency of fish kills is an attribute of significant concern to the public and decision-makers of the Neuse basin. The current scientific belief is that large fish kills are predominantly caused by a combination of low oxygen bottom water (hypoxia) and wind conditions that force that bottom water to the surface, trapping fish along the shores where they suffocate (Crowder 1998). Fish are generally more susceptible if they are already in poor health. Therefore, a probabilistic prediction of fish kills depends on the health of the fish population, the temporal extent to which the estuary experiences hypoxic conditions, and the frequency of cross-channel, "trapping" wind conditions (see Fig. 2).

Of course, a fish kill requires the presence of fish in the area of the upwelling, concurrent with the trapping winds and the presence of hypoxic bottom water. Even with this combination, fish may be able to react and swim away from the upwelling, making mechanistic prediction of the exact timing of fish kills impossible. Therefore, we relied upon the elicited judgment of two experienced estuarine fisheries researchers to characterize the probability of fish kills conditioned on a given state of fish population health, the occurrence of a strong cross-channel wind, and varying bottom water concentrations (Borsuk et al. 2002a). Asking for a probability conditioned on a number of circumstances allowed the scientists to focus on the likelihood of a fish kill only under certain given situations (upon the coincidence of a number of causative factors), rather than having to simultaneously consider the background frequency of the causative variables. The frequency of cross-channel winds can be considered to be a marginal node, without parents, since historical data and observation exist on their occurrence but they cannot be controlled by management.

Prediction of the temporal extent of hypoxia, however, is conditional on the pattern of bottom water oxygen concentrations.

Hypoxia

Oxygen concentration is determined by both the rate of sediment oxygen consumption by bacterial respiration and the duration that the bottom waters are separated from the surface due to salinity stratification (Paerl et al. 1998; Stanley and Nixon 1992). This relationship was quantified using a process-based model of oxygen depletion that is consistent with established theory yet is simple enough to be empirically parameterized from available monitoring data (Borsuk et al. 2001c). The model represents the processes of microbial oxygen consumption and physical reoxygenation, including the effects of temperature and vertical stratification. Nonlinear regression allowed for the direct estimation of rate constants from field data. The resulting model can be used to probabilistically predict the frequency of bottom water hypoxia, conditional on the annual average rate of benthic oxygen demand and duration of stratification (see Fig. 2). It is generally believed that stratification begins to set up whenever cross-channel winds are calm enough to avoid mixing for more than one day (Luettich 1998). Therefore, a variable describing the number of consecutive days between winds of sufficient strength to mix the system is the only variable relevant to stratification. This variable, like fish kills, is dependent on the frequency of strong cross-channel winds.

Benthic oxygen demand

Benthic oxygen demand is dependent on the decay rate of organic matter in the sediments, which, in turn, is dependent on the amount of organic matter available (Rizzo and Christian 1996). In a eutrophic estuary such as the Neuse, most of the sediment organic matter is believed to be internally derived via carbon fixation by algae, rather than externally derived via river loading of terrestrial material (Alperin et al. 2000). Because regular measurements are not made of the organic matter decay rate or the sediment organic carbon content, these intermediate steps are not included in the model, and a direct link is shown between algal carbon production and sediment oxygen demand (see Fig. 2).

While abundant water quality monitoring data exist for the Neuse, the historical values of algal carbon production do not span the range that may be expected under a significant anthropogenic change in nutrient inputs. Therefore, we relied on cross-system data from 34 estuaries and coastal zones to parameterize a simple, mechanistic model relating carbon production and sediment oxygen demand, including the effects of water column decay and sediment burial (Borsuk et al. 2001b). To do this, we employed a hierarchical approach which assumes partial, but not complete, commonality in parameter values across different estuarine systems. Both global and system-specific parameters were estimated using Bayesian inference. Using the parameters estimated for the Neuse estuary, annual average sediment oxygen demand can be expressed as a probabilistic function of water depth and annual average carbon production (see Fig. 2).

Algal carbon production

Algal carbon production is primarily determined by algal density, although water temperature also plays an important role (Mallin et al. 1991). Additionally, light intensity and photic depth have been shown to be significant factors (Boyer et al. 1993; Cole and Cloern 1987). However, while these are both observable variables (in that they can be measured), they are neither manageable by nitrogen controls nor predictable from other known factors (as water temperature is from the seasonal cycle). Therefore they are not explicitly included, and the variability they cause becomes part of the model uncertainty (see Fig. 2)

To predict primary productivity from algal density, we used a generalized version of the model proposed by Cole and Cloern (1987) and subsequently modified for the Neuse by Mallin et al. (1991) and, later, by Boyer et al. (1993). The model, which expresses daily algal carbon productivity as a function of biomass and water temperature was fit to approximately five years (mid 1994 through 1999) of biweekly monitoring data at 11 mid-channel sampling locations within the Neuse River estuary (Borsuk et al. 2002c).

Algal density

Among the factors believed to control algal density are nitrogen inputs and water temperature (Pinckney et al. 1997). Additionally, river flow has been shown to be an important factor (Mallin et al. 1993), perhaps because of its influence on estuarine salinity, turbidity, and water residence time. Detailed measurements of water temperature, river flow, and river nitrogen concentration exist, making these suitable marginal nodes (see Fig. 2). Other potential sources of nitrogen to the estuary, including atmospheric sources and groundwater (Paerl et al. 1995), are not considered in this analysis because the TMDL process only regulates nitrogen inputs from the river.

The relationship between algal density, as measured by chlorophyll a concentration, estuarine location, water temperature, and incoming Neuse River flow and total nitrogen concentration was developed using a regression model fit to approximately five years (mid 1994 through1999) of biweekly monitoring data (Borsuk et al. 2002c). Although algal density, itself, may be an important policy variable, of particular concern is the frequency with which chlorophyll a levels exceed the state water quality standard of 40 μg/L. Therefore, a variable representing this exceedance frequency is shown explicitly in the network (see Fig. 2) and its distribution is derived from the distribution of chlorophyll values as described by Borsuk et al. (2002d).

Fish health

Another attribute of policy relevance is fish population health. While a number of factors affect the health of the Neuse estuary fish population, only the effects of hypoxia can be controlled through nitrogen reductions. Harmful effects of low oxygen on fish include reduced feeding and growth rates (McNatt et al. 2000) and increased predation from larger fish and invertebrates (Breitburg et al. 1994). Extensive hypoxia can also reduce usable habitat, altering fish distribution and increasing competition (Pihl et al. 1991). These impacts diminish the health and productivity of the fish population and make them more vulnerable to both disease and episodic fish kill events.

One approach to predicting the population consequences of sublethal oxygen effects has been to develop individual-based models (Huston et al. 1988) linking fish to all the processes and subprocesses associated with the effects (Breitburg et al. 1999). However, information of sufficient detail to parameterize such a model does not exist for the Neuse estuary. Therefore, the relationship between fish population health and the annual extent of bottom water hypoxia was elicited from the same estuarine fisheries scientists questioned for the fish kill model (Borsuk et al. 2002a). Many different definitions of population health are possible, so we asked the researchers to develop a definition that was consistent with their knowledge and experience. They chose to use a categorical variable, with levels defined as,

Excellent: *High average growth rates (> 0.6 mm/d); low incidence of visible disease (<1%) on all fish but menhaden;*

Good: *Medium average growth rates (≤ 0.6 and ≥ 0.2 mm/d); low incidence of visible disease (<1%) on all fish but menhaden;*

Poor: *Poor average growth rates (< 0.2 mm/d); medium/high incidence of visible disease (≥1%) on all fish but menhaden;*

where growth rate is measured in the field as described by Eby (2001). Atlantic menhaden were specifically excluded from measures of the incidence of visible disease because of their high susceptibility to infections and parasites and the seasonal nature of their disease patterns irrespective of oxygen conditions (Goldman et al. in review).

With the health categories defined, questions were next asked regarding the probability of population health being in each of the categories, given a particular temporal extent of low oxygen. Since earlier studies have revealed that low oxygen is only a concern at high water temperatures (Borsuk et al. 2001c), we focused attention on the summer season. The scientists' assessments were based on the results of their monthly fish trawling and water quality sampling program in the Neuse estuary, as well as a set of *in situ* caging experiments (Eby 2001). Such experience-based, probabilistic judgments represent the estimated net result of a number of interacting processes and sources of uncertainty.

Shellfish survival

Shellfish face a similar situation as finfish when subjected to hypoxia. However, because shellfish are sessile, it is not only their health, but also their abundance, that is threatened by long-term exposure to low oxygen conditions. In this regard, both the duration and severity of hypoxia are important considerations, prompting the arrows from nodes representing both duration of stratification and dissolved oxygen concentration (see Fig. 2).

To relate oxygen status to shellfish abundance in the Neuse River estuary, we developed a survival model for the clam species *Macoma balthica* (Borsuk et al. 2002b). The survival rate of *M. balthica* was chosen as an indicator for shellfish abundance because *M. balthica* plays a critical role in the Neuse ecosystem. This later-succession bivalve is the major component of benthic biomass in the estuary as well as a valuable food resource for demersal fish species and blue crabs.

Field studies have shown that the late-summer pattern of abundance of *M. balthica* in the Neuse closely matches the pattern of extended exposure to summertime hypoxia (Powers et al. In review). However, experimental studies have not yet been performed to directly address the sensitivity of this species to low oxygen conditions. Therefore, this sub-model relied upon the expert judgment of two marine biologists to provide the data used in model building. The elicitation method that we used was based on a series of questions to establish points on the cumulative distribution function of times-to-death for multiple dissolved oxygen concentrations. Model parameters were then estimated from the assessed data using Bayesian methods. The resulting model probabilistically relates survival of *M. balthica* to time of exposure (duration of stratification) and dissolved oxygen concentration, as required for the network model (Figure 2).

Pfiesteria abundance

The toxic dinoflagellate, *Pfiesteria piscicida*, is a concern to the public at least in part because of the large amount of media attention it has received in recent years. It has been blamed for having a role in the occurrence of fish kills both by directly attacking the fish and by making them more susceptible to harsh conditions (Burkholder 1999). *Pfiesteria* has also been found to adversely impact the health of laboratory researchers studying the organism by causing respiratory and neurological distress (Glasgow et al. 1995). However, the potential threat to people exposed to *Pfiesteria* under natural conditions is highly controversial (Griffith 1999), and the distinct role the organism plays in fish kills is uncertain (Stow 1999). Many of the scientists we spoke with felt that *Pfiesteria* was just one of many stressors that affect fish, and if *Pfiesteria* were not present in the estuary, other opportunistic organisms would be. Thus, to satisfy the interests of the stakeholders, *Pfiesteria* abundance was included as a variable in the model. However, it was not explicitly linked to fish population health or fish kills, nor was a human health effect included. Perhaps as more laboratory research, fieldwork, and health studies are conducted in the future, the role of *Pfiesteria* in the network can be modified accordingly.

The factors potentially controlling the presence of *Pfiesteria* cells in the water column were recently investigated using a set of mesocosm experiments by Pinckney et al. (2000). These experiments were designed to test the response of *Pfiesteria* zoospores to a range of environmental conditions and potential prey species. Results showed that the density of *Pfiesteria*-like organisms (PLOs) was positively correlated with phytoplankton productivity and total phytoplankton biomass (as measured by chlorophyll a). Apart from the correlation with algal biomass and productivity, PLOs showed no additional significant response to nutrient, sediment, or mixing treatments in any of the experiments. These results suggest that PLOs track the abundance of their prey resources. Fensin (1998) also found a positive correlation between PLOs and phytoplankton biomass (as chlorophyll a) in field samples collected from the Neuse estuary during 1994 and 1995.

We used the data of Pinckney et al. to develop a functional relationship between algal density and PLOs (Borsuk 2001). Data collected by Fensin were not available for our analysis. Our analysis showed that PLO cell counts only reach levels of concern during the summer season. For this reason, the functional equation was quantified using data collected in the summer only. The relationship between algal density and PLOs was found to be approximately linear after a log-transformation of both variables, so parameters were estimated using ordinary least-squares regression.

In expressing concern over *Pfiesteria* abundance, stakeholders were probably particularly concerned about densities that are potentially harmful. A level of 250 cells/ml of toxic zoospores has been cited as a concentration sufficiently high to be lethal to fish (Burkholder et al. 1995). Therefore, the frequency of daily cell densities above 250 cells/ml in the summer season was included as a separate variable in the network. Because the cell counts recorded by Pinckney et al. include all *Pfiesteria*-like organisms, both toxic and nontoxic, the results of our model can be considered an upper estimate of toxic forms.

Integration of Relationships into Bayesian Network

The set of probabilistic relationships described in the previous section can be joined into one integrated network (Figure 3). Each relationship describes the most likely value of the response variable conditional on the values of each of its parents (solid lines in Figure 3). The uncertainty in this relationship, resulting from both model error and parameter uncertainty, is captured by conditional probability distributions (represented by dashed lines in Figure 3). When marginal, or unconditional, distributions are specified for each of the outermost variables (nodes without parents), the resulting predictive distributions of all the remaining variables can be calculated from the network. Marginal distributions for the variables river flow, nitrogen inputs, water temperature, and duration of stratification can be derived from historical data and adjusted, as appropriate, to represent various management alternatives.

We used Analytica, a commercially available software program (Lumina 1997), for implementing the Bayesian network for the Neuse estuary. Analytica allows for the use of continuous or discrete variables related by any functional expression. Uncertainty can be represented by a wide variety of probability distributions and is propagated through the network using Monte Carlo or Latin hypercube sampling.

Model-Based TMDL Evaluation

To illustrate the use of probabilistic network predictions for TMDL evaluation, we evaluated five possible scenarios corresponding to nitrogen reductions of 0, 15, 30, 45, and 60% relative to 1991-1995 baseline inputs. The period 1991-1995 was chosen by the N.C. Division of Water Quality as the reference period for the Neuse River TMDL (NCDWQ 2001b). Therefore, daily data from those years served as the basis for the marginal input variables. These variables were represented in the network as a multivariate empirical distribution to maintain any underlying dependencies (indicated by histograms connected by dashed, double-headed arrows in Figure 3). Missing values for the marginal variables were estimated from flow models as described by Borsuk et al. (2002c). The four nitrogen reduction scenarios were evaluated by multiplying all riverine nitrogen concentrations by the complement of the appropriate reduction. All other functions and marginal nodes in the model were left unchanged, and new distributions were computed for the variables of interest. The Latin hypercube sampling method was used to draw 250 samples of all model parameter and error distributions. The median predicted value for each model endpoint as well as the outer limits of the 50% and 90% predictive intervals were then calculated to indicate overall response and predictive uncertainty. Although many of the functional relationships among variables were developed to be applicable to multiple regions of the estuary, we chose the middle region (Figure 1) as the focus of this assessment. This is historically the region with the greatest extent of hypoxia and the most frequent occurrence of fish kills.

Model predictions (Figure 4) show that under the baseline scenario of no nitrogen reduction the annual average chlorophyll a concentration in the middle region of the estuary is expected to be slightly above 20 μg/L, and the state chlorophyll standard of 40 μg/L will most likely be exceeded on more than 10% of the days. As nitrogen inputs are reduced, both the average chlorophyll concentration and the chlorophyll standard exceedance frequency are also expected to decrease. However, only at reductions of 45% relative to the baseline is the median exceedance frequency predicted to approach the 10% EPA guideline (Office of Water 1997b).

Without nitrogen reductions, Pfiesteria-like cell densities at levels of concern are expected to occur between 5 and 17 days during the summer season. This value is expected to decrease in concert with chlorophyll reductions. However, as is the case for chlorophyll, the uncertainty in model predictions increases as conditions depart from overall mean values where the model is most precise. In this case, the scenario with greatest precision is somewhat below baseline because the middle section of the estuary is more impaired relative to the other sections to which the model was also fit. The increase in uncertainty at greater nitrogen reductions implies that the upper range of predicted values is essentially equal for each reduction scenario greater than 30%.

Under the baseline scenario, the summer survival rate of *Macoma* clams is predicted to be low with a median value below 10% but, given model uncertainty and natural variability, is likely to be as low as near 0% or as high as 40%. For comparison, during the summer of 1997, the first year of extensive benthic surveying, the *Macoma* clam community was estimated to be reduced to less than 20% of its spring population (Peterson et al. 2000). The most likely state of fish population health under baseline conditions is "good" with a probability of 0.55, while "excellent" has a probability of 0.32 and "poor" of 0.13. Both summertime shellfish survival and overall fish population health are predicted to increase slightly in response to reduced nitrogen inputs.

In any scenario, fish kills are predicted to be relatively infrequent events. For this reason, probabilities are expressed as the expected number of fish kills in a ten-year period. Without any nitrogen reduction, the model predicts between 5 and 20 kills in ten years involving more than 1,000 fish in the middle portion of the estuary. For reference, there were

8 fish kills of this size in this region during the ten years 1989 through 1999. Additionally there were 6 kills in which the number of fish involved was not reported (NCDWQ 2001a). The frequency of fish kills is not expected to change substantially with nitrogen reductions.

The reason for the relatively minor response of the ecological endpoints can be discovered by looking at the trends in carbon production and days of summertime hypoxia (Figure 5). While carbon production is predicted to decrease relative to baseline values of 350 to 500 gC/m^2y in response to reduced algal stimulation, this effect is dampened out further down the causal chain, so that the change in the number of days of resulting summertime hypoxia is relatively insignificant. The further we move down the probability network and away from the decision variable, the greater the predictive uncertainty. This is due to the uncertainty that is added in every successive relationship, as well as the increasing effects of natural variability.

Using Model Results for TMDL Decision-Making

Given a set of predictions regarding multiple ecological endpoints, the choice of an appropriate load reduction depends on the levels determined by decision-makers to be acceptable for one or more of those endpoints. Such a determination might be based on an analysis of associated costs and benefits (Johansson 1993), a multiattribute utility calculation (Clemen and Reilly 2001), a risk assessment procedure (Suter and Barnthouse 1993), or compliance with a predetermined standard (Barnett and O' Hagan 1997). Choosing appropriate decision criteria is a task for policy-makers, not scientists, because it is a value-based, rather than belief-based, exercise, involving the characterization of societal desires rather than the behavior of a natural system. Unfortunately, such an analysis has not been performed for the Neuse management situation. The only decision criteria that currently exist are the state chlorophyll standard of 40ug/L and the EPA guidance mandating fewer than

10% exceedances of this standard. Thus, our discussion regarding the use of the model results for decision-making will focus on the chlorophyll endpoint, although the ideas will be equally relevant to any of the other endpoints once appropriate decision criteria are determined.

If 10% is considered the maximum acceptable frequency of exceedances of the state chlorophyll standard, then the target nitrogen reduction can be graphically determined from the plot of model results (Figure 6) by drawing a horizontal line at the 10% value and observing where it intersects the curve of predictions. Drawing a vertical line from this intersection to the horizontal axis suggests the necessary reduction. However, given the uncertainty in model predictions, even with a fixed decision criterion the choice of a nitrogen reduction depends on the degree of confidence required by decision makers. Using only the median predictions (or, equivalently, model predictions that do not account for uncertainty) implies 50% confidence that the criterion will be met. If a higher degree of confidence is required, then the outer bound of an appropriate predictive interval must be used in the graphical determination. The difference between the nitrogen reduction necessary to achieve 50% confidence and the reduction necessary to achieve a higher level of confidence can be considered the margin of safety.

Inclusion of a margin of safety in the determination of a TMDL is required under the Clean Water Act (CWA Section 303(d)(1)(c)). Generally this is accomplished through conservative model assumptions (Office of Water 1999). However, this practice confounds values with scientific beliefs and obscures the fact that in making these assumptions the modeller implicitly chooses a particular level of confidence. Choosing the degree of confidence required of a model is a risk management decision that should be made by designated officials, not water-quality modellers. Such a decision should be based on consideration of the potential cost to stakeholders of continued impairment despite the attainment of the chosen nitrogen reduction. A model that assesses the uncertainty associated with predictions provides an explicit basis for choosing a TMDL that includes a margin of safety. Given a quantitative water quality standard and probabilistic model results expressed as the degree of confidence that a criterion will be met for any given loading level (Figure 6), decision-makers simply need to choose the percent reduction that corresponds to their desired level of confidence.

The margin of safety depends on both the risk tolerance of decision-makers and the predictive uncertainty in the water quality model being used to support the decision. Thus, the size of the margin might be reduced in either of two ways: (1) decision-makers and stakeholders must settle for a lower degree of confidence in achieving their objectives, or (2) predictive uncertainty must be reduced. Assuming that the chosen confidence level is based on a rational process that cannot be changed (perhaps an unlikely assumption!), then the margin of safety is wholly reliant on the uncertainty inherent in model predictions. Because the size of the margin of safety has a direct impact on the nitrogen reduction required and therefore on the cost of management, adequate uncertainty analysis of TMDL models should be a high priority.

Discussion

The probability network, Neu-BERN, is one of multiple estuarine response models currently being used to inform the near-term selection of a TMDL for the Neuse River (see Bowen et al. and Wool et al. this issue). Compared to the others, its process-representation is relatively simple (see Roessler et al. this issue). Complex physical, chemical, and biological processes are combined into aggregate components described by measurable, operationally defined variables. The model does not invoke more than is necessary, emphasizing the fact that it should not be considered a representation of reality, but rather a simplification for a limited purpose. In this case, the purpose is to serve as a framework for TMDL decision-making by organizing current scientific understanding and assumptions. This information may exist in a variety of forms, including historical monitoring data, cross-system comparisons, mesocosm results, modelling experiments, or observational experience; a Bayesian network can explicitly accommodate such variety.

Results of the integrated model show that predictive uncertainty, arising from both natural variation and knowledge uncertainty, is high. This is especially true for variables that are not easily measured (such as shellfish survival), infrequent to occur (fish kills), or further down the causal chain (fish health). However, these types of variables are precisely those that are of greatest concern to the public and decision-makers. This suggests that additional data collection is necessary, particularly on some of the more uncertain relationships. These include the effect of nitrogen inputs and water temperature on algal density, the relation between algal density and carbon production, and the connection between hypoxia and fish kills (see Figure 3). It should be kept in mind, however, that there is a limit to predictive precision. Stochastic variability is an inherent property of natural systems and contributes uncertainty that must be considered but, for a given model, cannot be reduced. Recognizing this fact will help water quality stakeholders maintain realistic expectations concerning ecological forecasts.

The presence of significant uncertainty in model-based TMDL predictions should not preclude decisive management action. Model results indicate that nitrogen reductions are likely to lead to ecological improvements however uncertain the magnitude of those improvements may be. Preliminary implementation plans can be made under the condition that additional monitoring and research will occur. In fact, models that quantify uncertainty facilitate the prioritization of future data collection efforts based on their ability to improve predictions. After new information is obtained, new predictions are generated and a revised set of actions and research strategies is developed. This process can be repeated until stakeholder objectives are met or additional management costs exceed the expected benefits.

Decision-oriented methodologies similar to that presented here have been proposed previously for estuarine nutrient assessment. Jaworski and Villa (1981) suggest an integrative approach using multiple, publicly meaningful criteria. They acknowledge that various quantitative methods, from subjective analyses to complex mathematical models, may be suitable depending on available information. They also emphasize the need for communicating uncertainties in model relationships by expressing predictions probabilistically. However, their framework is merely conceptual and is not accompanied by specific quantitative tools or applications. Our present analysis demonstrates Bayesian networks as one possible tool for this type of evaluation and presents results for a specific TMDL application.

More recently, Fitzpatrick and Meyers (2000) reviewed methods for determining estuarine nutrient criteria and highlighted the variety of approaches that may be possible, depending on the situation. However, similar to the general stance of the EPA, they suggest that simple, databased models be used only for initial screening purposes, to be replaced by more realistic simulation models for final analysis. We disagree with this view and believe that simple models focusing on the major processes may be *more* realistic and useful representations of natural systems than complex models that strive to include processes at every scale. Scientific understanding of mechanism is advanced, but only to the point of being able to characterize aggregate relationships, not to quantify all of the small-scale dynamics.

Recognizing the limits of mechanistic knowledge is especially important when attempting to extend water quality attributes to ecological effects. Ecological variables are more reliable indicators of whether a water body is meeting its designated uses, and their importance to future TMDL determination has been emphasized (NRC 2001). By permitting carefully elicited expert knowledge as a practical alternative to "hard" data, the Bayesian approach facilitates extension of models to ecological endpoints. Scientific experts can assess the response of ecological variables to their immediate causes and then summarize remaining variability and uncertainty using probabilistic expressions. Predictions expressed as probabilities then give stakeholders and decision-makers realistic expectations of the chances of achieving desired outcomes. This type of knowledge can be expected to lead to more informed and effective TMDL decisions.

References

Alperin, M. J., Clesceri, E. J., Wells, J. T., Albert, D. B., McNinch, J. E., and Martens, C. S. (2000). "Sedimentary processes and benthic-pelagic coupling." Neuse River Estuary Modeling and Monitoring Project Final Report: Monitoring Phase, Report No. 325-B, Water Resources Research Institute of the University of North Carolina, Raleigh, NC, 63-105.

ARF. (1997). "Report on the 20 most threatened American Rivers.", American Rivers Foundation, New York.

Barnett, V., and O' Hagan, A. (1997). *Setting Environmental Standards*, Chapman and Hall, London.

Borsuk, M. E. (2001). "A Graphical Probability Network Model to Support Water Quality Decision Making for the Neuse River Estuary, North Carolina," Ph.D. Dissertation, Duke University, Durham, NC.

Borsuk, M. E., Clemen, R. T., Maguire, L. A., and Reckhow, K. H. (2001a). "Stakeholder values and scientific modeling in the Neuse River watershed." *Group Decision and Negotiation*, 10, 355-373.

Borsuk, M. E., Eby, L. A., and Crowder, L. B. (2002a). "Probabilistic prediction of fish health and fish kills in the Neuse River estuary using the elicited judgment of scientific experts." *In review.*

Borsuk, M. E., Higdon, D., Stow, C. A., and Reckhow, K. H. (2001b). "A Bayesian hierarchical model to predict benthic oxygen demand from organic matter loading in estuaries and coastal zones." *Ecological Modelling*, 143, 165-181.

Borsuk, M. E., Powers, S. P., and Peterson, C. H. (2002b). "A survival-based model of the effects of bottom-water hypoxia on the density of an estuarine clam population." *In review*.

Borsuk, M. E., Stow, C. A., Luettich, R. A., Paerl, H. W., and Pinckney, J. L. (2001c). "Modelling oxygen dynamics in an intermittently stratified estuary: Estimation of process rates using field data." *Estuarine Coastal and Shelf Science*, 52(1), 33-49.

Borsuk, M. E., Stow, C. A., and Reckhow, K. H. (2002c). "The confounding effects of nitrogen load on eutrophication of the Neuse River estuary, North Carolina." *In review*.

Borsuk, M. E., Stow, C. A., and Reckhow, K. H. (2002d). "Predicting the frequency of water quality standard violations: A probabilistic approach for TMDL development." *Environmental Science & Technology*, 36(10), 2109-2115.

Boyer, J. N., Christian, R. R., and Stanley, D. W. (1993). "Patterns of phytoplankton primary productivity in the Neuse River estuary, North-Carolina, USA." *Marine Ecology-Progress Series*, 97(3), 287-297.

Breitburg, D. L., Rose, K. A., and Cowan, J. H. (1999). "Linking water quality to larval survival: predation mortality of fish larvae in an oxygen-stratified water column." *Marine Ecology-Progress Series*, 178, 39-54.

Breitburg, D. L., Steinberg, N., DuBeau, S., Cooksey, C., and Houde, E. D. (1994). "Effects of low dissolved oxygen on predation on estuarine fish larvae." *Marine Ecology-Progress Series*, 104, 235-246.

Burkholder, J. M. (1999). "The lurking perils of Pfiesteria." *Scientific American*, 281(2), 42-49.

Burkholder, J. M., Glasgow, H. B., and Hobbs, C. W. (1995). "Fish kills linked to a toxic ambush-predator dinoflagellate: distribution and environmental conditions." *Marine Ecology-Progress Series*, 124, 43-61.

Clemen, R. T., and Reilly, T. (2001). *Making Hard Decisions*, Duxbury, Pacific Grove.

Cole, B. E., and Cloern, J. E. (1987). "An empirical model for estimating phytoplankton productivity in estuaries." *Marine Ecology-Progress Series*, 36, 299-305.

Crowder, L. B. (1998). "Personal Communication.".

Eby, L. A. (2001). "Response of a fish community to frequent and infrequent disturbances in an estuarine ecosystem," Ph.D. Thesis, Duke University, Durham, NC.

Fensin, E. (1998). "Seasonal population dynamics of the phytoplankton community and of Pfiesteria piscicida/Pfiesteria-like dinoflagellates in the mesohaline Neuse Estuary, North Carolina," M. Sci. Thesis, North Carolina State University, Raleigh, NC.

Fitzpatrick, J. J., and Meyers, M. B. (2000). "A review of methodologies that determine nutrient criteria for coastal and estuarine systems." *Water Environment and Technology*, 12, 35-39.

Glasgow, H. B., Burkholder, J. M., Schmechel, D. E., Tester, P. A., and Rublee, P. A. (1995). "Insidious Effects of a Toxic Estuarine Dinoflagellate on Fish Survival and Human Health." *Journal of Toxicology and Environmental Health*, 46(4), 501-522.

Goldman, R. L., Eby, L. A., Crowder, L. B., and McClellan, C. (in review). "The effects of alterations in coastal ecosystems on incidence of disease in Atlantic menhaden (Brevoortia tyrannus L.).".

Griffith, D. (1999). "Exaggerating environmental health risk: The case of the toxic dinoflagellate Pfiesteria." *Human Organization*, 58(2), 119-127.

Huston, M., DeAngelis, D., and Post, W. (1988). "New computer models unify ecological theory." *BioScience*, 38(10), 682-691.

Jaworski, N. A., and Villa, O. (1981). "A suggested approach for developing estuarine water quality criteria for management of eutrophication." Estuaries and Nutrients, B. J. Neilson and L. E. Cronin, eds., Humana Press, Clifton, NJ.

Johansson, P.-O. (1993). *Cost-benefit analysis of environmental change*, Cambridge University Press, Cambridge.

Luettich, R. A. (1998). "Personal Communication.".

Lumina. (1997). "Analytica for Windows.",, Lumina Decision Systems, Denver, CO.

Lung, W.-S. (2001). *Water Quality Modeling for Wasteload Allocations and TMDLs*, Wiley & Sons, New York.

Mallin, M. A., Paerl, H. W., and Rudek, J. (1991). "Seasonal phytoplankton composition, productivity and biomass in the Neuse River estuary, North-Carolina." *Estuarine Coastal and Shelf Science*, 32(6), 609-623.

Mallin, M. A., Paerl, H. W., Rudek, J., and Bates, P. W. (1993). "Regulation of estuarine primary production by watershed rainfall and river flow." *Marine Ecology-Progress Series*, 93(1-2), 199-203.

McNatt, R. A., Rice, J. A., and Noga, E. J. "Impacts of sublethal hypoxia on feeding, growth, and immunity of juvenile estuarine-dependent fish." *N.C. Water Resources Research Conference*, Raleigh, NC.

Morgan, M. G., and Henrion, M. (1990). *Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis*, Cambridge University Press, Cambridge.

National Research Council. (2001). *Assessing the TMDL Approach to Water Quality Management*, National Academy Press, Washington, D.C.

NCDWQ. (2001a). "North Carolina Division of Water Quality Fish Kill Report (Online).",, http://www.esb.enr.state.nc.us/Fishkill/fishkill00.htm North Carolina Department of Health and Natural Resources, Raleigh, NC.

NCDWQ. (2001b). "Phase II of the Total Maximum Daily Load for Total Nitrogen to the Neuse River Estuary, North Carolina.", North Carolina Department of Environment and Natural Resources, Raleigh, NC.

Office of Water. (1997a). "Compendium of Tools for Watershed Assessment and TMDL Development." *EPA-841-B-97-006*, U.S. Environmental Protection Agency, Washington, DC.

Office of Water. (1997b). "Guidelines for Preparation of the Comprehensive State Water Quality Assessments." *EPA-841-B-97-002*, U.S. Environmental Protection Agency, Washington, DC.

Office of Water. (1999). "Protocol for Developing Nutrient TMDLs." *EPA 841-B-99-007*, U.S. Environmental Protection Agency, Washington DC.

Pace, M. L. (2001). "Prediction and the aquatic sciences." *Canadian Journal of Fisheries and Aquatic Sciences*, 58, 1-10.

Paerl, H. W., Mallin, M. A., Donahue, C. A., Go, M., and Peierls, B. L. (1995). "Nitrogen loading sources and eutrophication of the Neuse River estuary, North Carolina: direct and indirect roles of atmospheric deposition." *Report No. 291*, Water Resources Research Institute of the University of North Carolina, Raleigh.

Paerl, H. W., Pinckney, J. L., Fear, J. M., and Peierls, B. L. (1998). "Ecosystem responses to internal and watershed organic matter loading: consequences for hypoxia in the eutrophying Neuse river estuary, North Carolina, USA." *Marine Ecology-Progress Series*, 166, 17-25.

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*, Morgan Kaufmann, San Mateo, CA.

Pearl, J. (2000). *Causality: Models, Reasoning, and Inference*, Cambridge University Press, Cambridge, UK.

Peterson, C. H., Dozier, M., Powers, S. P., Johnson, G. A., and Christian, R. R. "Does the value of essential fish habitat in the Neuse River Estuary vary with annual severity of hypoxia?" *N.C. Water Resources Research Conference*, Raleigh, NC.

Pihl, L., Baden, S. P., and Diaz, R. J. (1991). "Effects of periodic hypoxia on distribution of demersal fish and crustaceans." *Marine Biology*, 108, 349-360.

Pinckney, J. L., Millie, D. F., Vinyard, B. T., and Paerl, H. W. (1997). "Environmental controls of phytoplankton bloom dynamics in the Neuse River Estuary, North Carolina, USA." *Canadian Journal of Fisheries and Aquatic Sciences*, 54(11), 2491-2501.

Pinckney, J. L., Paerl, H. W., Haugen, E., and Tester, P. A. (2000). "Responses of phytoplankton and Pfiesteria-like dinoflagellate zoospores to nutrient enrichment in the Neuse River Estuary, North Carolina, USA." *Marine Ecology-Progress Series*, 192, 65-78.

Powers, S. P., Peterson, C. H., Christian, R. R., Sullivan, E., Powers, M. J., Johnson, G., and Buzzelli, C. P. (In review). "Linking water quality to fisheries sustainability: the effects of estuarine eutrophication on seafloor habitat and prey resources of demersal fish.".

Reckhow, K. H. (1994). "Water quality simulation modeling and uncertainty analysis for risk assessment and decision-making." *Ecological Modelling*, 72(1-2), 1-20.

Rizzo, W. M., and Christian, R. R. (1996). "Significance of subtidal sediments to heterotrophically-mediated oxygen and nutrient dynamics in a temperate estuary." *Estuaries*, 19, 475-487.

Stanley, D. W., and Nixon, S. W. (1992). "Stratification and Bottom-Water Hypoxia in the Pamlico River Estuary." *Estuaries*, 15(3), 270-281.

Stow, C. A. (1999). "Assessing the relationship between Pfiesteria and estuarine fishkills." *Ecosystems*, 2(3), 237-241.

Suter, G. W., and Barnthouse, L. W. (1993). *Ecological risk assessment*, Lewis Publishers, Boca Raton.

Thomann, R. V. (1998). "The future "golden age" of predictive models for surface water quality and ecosystem management." *Journal of Environmental Engineering-Asce*, 124(2), 94-103.