Human and Ecological Risk Assessment, 10: 415–434, 2004

ISSN: 1080-7039 print
DOI: 10.1080/10807030490438454

Predictive Assessment of Fish Health and Fish Kills in the Neuse River Estuary Using Elicited Expert Judgment

Mark E. Borsuk
Division of Environmental Science and Policy, Nicholas School of the Environment and Earth Sciences, Duke University, Durham, North Carolina, USA

ABSTRACT

Declining fish health and the occurrence of large fish kills are some of the more publicly meaningful indicators of water quality in the impaired Neuse River Estuary, North Carolina. It is generally believed that such problems are caused by the widespread depletion of dissolved oxygen—an indirect result of anthropogenic nutrient pollution. However, the development of scientific simulation models to predict how improvements in oxygen conditions will improve the health of fish and reduce the frequency of fish kills has proven elusive. As a pragmatic solution to this problem, the expert opinion of estuarine fisheries scientists in possession of relevant data and experience was elicited. The relations between joint and conditional probabilities were exploited to translate quantities that are normally hard to assess into quantities that can be drawn more directly from the experiential knowledge of the experts. A combined model of expert opinion was constructed as an influence diagram, and Monte Carlo simulation was used to generate predictions of fish health and fish kills in the Neuse River Estuary under current and improved oxygen conditions. Full model results are expressed as probability distributions, capturing the effects of natural variability and knowledge uncertainty—both contributors to total ecological risk.

Key Words: expert elicitation, ecological risk assessment, coastal eutrophication, logistic regression, probability network.

INTRODUCTION

Public concern over water quality in the Neuse River Estuary, North Carolina (Figure 1), has been heightened by recent fish kill events involving tens, or even hundreds, of thousands of fish (NCDWQ 2000). These fish, primarily Atlantic

Received 17 January 2003; revised manuscript accepted 23 June 2003.
Address correspondence to Mark E. Borsuk, 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. E-mail: mark.borsuk@eawag.ch

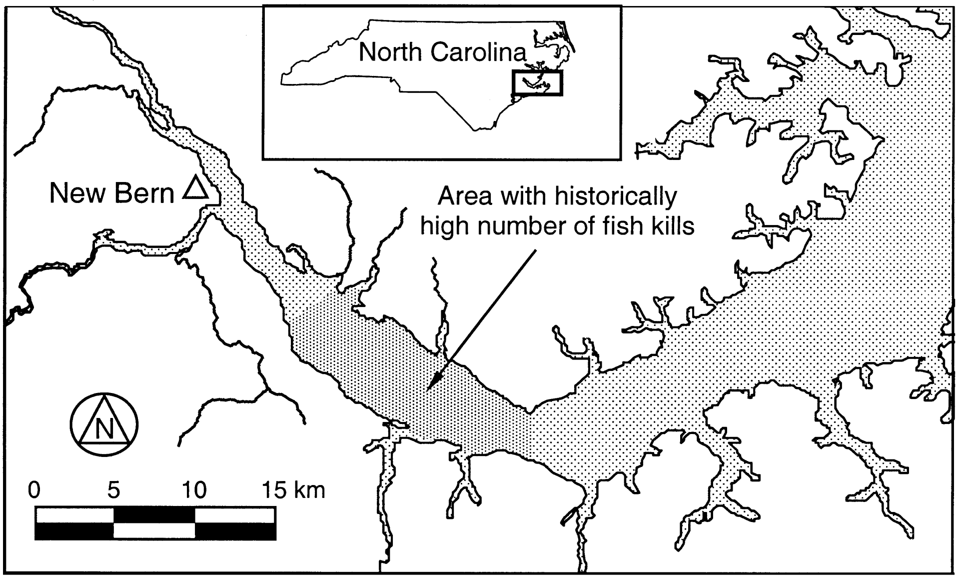

Figure 1. The Neuse River Estuary, North Carolina.

menhaden (*Brevoortia tyrannus*), are found dead along the shore in such large amounts that they often must be cleared away with a bulldozer. The sight is troubling for local residents and attracts substantial media attention, making the occurrence of fish kills one of the more publicly meaningful indicators of water quality (Borsuk *et al.* 2001).

A close association between the occurrence of fish kills and the presence of a potentially toxic microorganism, *Pfiesteria piscicida*, has been used to implicate *Pfiesteria* as a causative agent of widespread fish death in the Neuse River estuary (Burkholder and Glasgow 1997). However, the mechanism for this causal link has been questioned (Dykstra and Kane 2000; Litaker *et al.* 2002), and a recent statistical analysis has indicated that toxic *Pfiesteria* are more likely to be a result than a cause of large fish kills (Stow and Borsuk 2003).

A second, and more widely held, theory is that fish kills in the Neuse Estuary are caused by a combination of low oxygen bottom water and a unique set of wind conditions (Paerl *et al.* 1999). In this scenario, wind blowing across the estuary pushes surface water toward the downwind shore. A compensating flow of bottom water occurs in the opposite direction, causing upwelling along the upwind shore (McNinch and Luettich 2000). If the upwelled bottom water is depleted in oxygen, there is the potential for fish to be trapped without an escape route, leading to mass suffocation. A fish kill requires the presence of fish in the area of the upwelling, concurrent with the strong cross-channel winds and the presence of oxygen-depleted bottom water. Even with this combination, fish may be able to react and swim away from the upwelling, making predictions of the timing of fish kills nearly impossible.

Despite the difficulty in making specific temporal predictions of fish kills, some type of quantitative assessment is necessary to evaluate policy options aimed at reducing their occurrence. Management actions under consideration include efforts to improve oxygen conditions by limiting nutrient loading and resultant algal production (NCDWQ 2001). Nutrient-enhanced algal growth provides excess organic matter to benthic bacteria, fueling decomposition and associated oxygen consumption. This depletes dissolved oxygen in the lower water column, providing the background conditions for a fish kill, should appropriate wind patterns occur. Assessing the magnitude of nutrient reductions necessary to reduce fish kills to a tolerable level requires an assessment relating the oxygen depletion process and the occurrence of cross-channel winds to the frequency of large fish kill events. The interacting effects of anthropogenic changes in water quality and unpredictable natural events are a common source of difficulty in developing predictions for fisheries management (Rose 2000).

Although fish kills are a significant public concern, the numbers of fish involved actually represent a very small proportion of stocks in the estuary at a given time. While the larger fish kills have involved on the order of 100,000 individuals (NCDWQ 2001), commercial landings of menhaden have averaged 67 million pounds per year from 1991 to 2000, or approximately 90 million individuals (NCDMF 2001). It is likely that the sublethal effects of low oxygen water have a more profound population-level effect on a wider variety of species than losses due to direct kills. Sublethal effects include reduced feeding and growth rates (McNatt *et al.* 2000) and increased predation from larger fish and invertebrates (Breitburg *et al.* 1994). Extensive areas of low oxygen can also reduce usable habitat, altering fish distribution and increasing competition (Pihl *et al.* 1991). Even upon successful recolonization of previously abandoned areas, fish may be adversely affected by extensive defaunation of their prey resources (Breitburg 1992). All of these impacts serve to diminish the health and productivity of the fish population and make them more vulnerable to disease and episodic fish kill events of the type described above.

One approach to predicting the population consequences of sublethal 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, site-specific information of sufficient detail to parameterize such a model rarely exists (Grimm 1999), and the uncertainty associated with predictions of the cumulative effect of so many poorly characterized processes is likely to be large (Reckhow 1994). In general, it is not reasonable to expect that all the mechanisms of natural systems can be fully understood and used to build accurate predictive models (Pace 2001). An option is to rely on the elicited judgment of experienced estuarine fisheries scientists to characterize the aggregate relationship between fish population health and the annual extent of low oxygen bottom water. Such assessments are meant to focus attention on the most important influence factors, with the effect of smaller scale processes captured by probabilistic depictions of uncertainty (Reckhow 1999). This is a pragmatic approach to the interpretation and integration of diverse knowledge and data, and well-developed methods exist for eliciting such judgments (*e.g.*, Spetzler and Stael von Holstein 1975; Morgan and Henrion 1990; Meyer and Booker 1991).

Once a population health model is established, another probability assessment is then required to link kills of menhaden to the temporal extent of low oxygen conditions and the assessed state of fish population health. The probabilistic approach, which yields annual frequencies, rather than exact times and locations of kills, is a method well suited for describing the interacting roles of natural and anthropogenic causes in generating rare events. The relationship between joint and conditional probabilities can be exploited to translate small probabilistic quantities that are hard to assess into quantities that can be drawn more directly from the experiential knowledge of the experts.

The resulting combined model of expert opinion is used to generate predictions of fish health and fish kills under current oxygen conditions in the Neuse River Estuary. To predict the effects of proposed nutrient management actions, current oxygen conditions are replaced with values which may result from a long-term reduction in nutrient loading and algal growth. While the chosen scenario is only intended to serve as an example, the model described here has been incorporated into a larger set of linked eutrophication models integrated as a Bayesian probability network (Borsuk *et al.* 2003). This model is being used to generate more specific predictions of ecological response to nutrient reductions to support watershed management decisions.

FISH POPULATION HEALTH ASSESSMENTS

Assessment Method

It has previously been shown that, because of the effect of temperature on oxygen saturation and bacterial processes, low oxygen is only a concern in the Neuse River Estuary during the summer season (Borsuk *et al.* 2001). Therefore, probabilities of various categories of fish population health were conditioned on the number of low oxygen days during the summer (July–September). These probabilities were elicited from Drs. Larry Crowder and Lisa Eby of the Duke University Marine Laboratory, Beaufort, NC. These fisheries scientists have conducted a regular fish trawling and water quality sampling program in the Neuse estuary since 1997, as well as a set of *in situ* caging experiments (Eby 2001; Eby and Crowder 2002; Eby *et al.* in review). Their work has been part of a coordinated modeling and monitoring effort of the Neuse (ModMon, Reckhow and Gray 2000) and represents the major fisheries component (Eby *et al.* 2000). Therefore, they were selected as the two most qualified experts on the topic of low oxygen-induced fish disturbance in the Neuse. On the one hand, these two scientists can be expected to have opinions that coincide due their history of collaboration and their equal familiarity with the Neuse Estuary research results. On the other hand, the two surely have unique professional experiences that make their joint responses to elicitation questions more robust than if either were to answer alone. Consequently, the judgments of the two experts were combined using a form of the "nominal group" technique (Clemen and Winkler 1999). Each was first asked to answer the elicitation questions individually. This was followed by a group discussion of the reasoning and evidence supporting their responses. Finally, the experts were asked to reach a consensus answer to each question that accounted for their combined knowledge base.

The first set of elicitation questions posed to the experts was aimed at establishing the probability that fish population health would be categorized as "excellent," "good," or "poor," for multiple numbers of days of summertime low oxygen. The assessment began by asking the experts to define the categories excellent, good, and poor based on their own judgment of important indicators of fish health. The resulting definitions were relatively straightforward and measurable and were stated as:

> Excellent: High average growth rates (>0.6 mm/d); low incidence of visible disease (<1%) on all fish species but menhaden;
> Good: Medium average growth rates ($\leq$0.6 and $\geq$0.2 mm/d); low incidence of visible disease (<1%) on all fish species but menhaden;
> Poor: Poor average growth rates (<0.2 mm/d); medium/high incidence of visible disease ($\geq$1%) on all fish species but menhaden;

where growth rate is measured in the field as described by Eby (2001). 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, conditional probability distributions were assessed by asking the experts for probabilities for each health category at specified values for the number of hypoxic days. This is referred to as the *fixed value* approach (Spetzler and Stael von Holstein 1975). A typical question was,

> Given a summer in which bottom water oxygen concentration (depth greater than 1.5 meters) in the mid-channel of the Neuse Estuary averages less than 2.0 mg/L for 10 out of 92 days in July, August, and September, what is the probability that fish population health at the end of the summer can be characterized as "excellent"? "good"? "poor"?

This question was repeated for multiple oxygen concentration values and multiple numbers of days. The experts were reminded that the sum of the probabilities in the three categories should add to 100%, but were allowed to state a range representing the uncertainty in their assessments.

Fish Health Assessment Results

In developing their responses, the experts stated that they were primarily considering the effect of low dissolved oxygen on susceptibility to disease, the abundance of prey resources, and the effect of crowding due to diminished habitat on growth rate (see Eby *et al.* 2001 for relevant research results). They each stated their belief that a bottom water oxygen concentration of 2.0 mg/L can effectively be considered a tolerance breakpoint for fish, with the probabilities of the various levels of fish health depending primarily on the number of days with concentrations less than 2.0 mg/L. This belief was based upon behavioral avoidance thresholds statistically derived from trawling data for 10 Neuse Estuarine fish species (Eby and Crowder 2002). For convenience, such oxygen conditions will be referred to as "hypoxic." Although minor effects on feeding and growth may be present at higher oxygen concentrations, these are relatively unknown at this time and are not included in the present analysis.

Consensus probabilities (Table 1) indicate that the experts believe the health of the fish population declines with increasing temporal extent of hypoxia. However, this relationship is by no means certain. A number of other factors are believed to affect fish health in addition to oxygen, resulting in some likelihood being attributed to more than one category for almost all assessments. Additionally, most probabilities

Table 1. Assessed probabilities of fish population health for a given number of days of hypoxia during the 92 days of July–September.

Days	Poor (%)	Good (%)	Excellent (%)
10	10	40–50	50
20	10–15	40–60	25–40
30	10–20	50–70	10–20
40	20–30	40–60	5
50	35–65	30–40	5
60	70–80	20–30	0
80	100	0	0

were expressed as ranges, rather than single values. This imprecision in probability specification was indicated by each of experts, and generally increased as a result of their discussion.

To generalize the assessed relationship to numbers of hypoxic days intermediate to the values used in the assessment, a cumulative logit regression model was constructed (Ananth and Kleinbaum 1997). This model is an extension of the familiar logistic regression model for binary responses (*e.g.*, Reckhow *et al.* 1987; Rubin *et al.* 1992) to allow for responses with multiple ordered categories. The basic assumption is that the logarithm of the ratio of each cumulative probability to its complement is a linear function of the covariates. For example, for three response categories and one covariate, the model can be written as,

$$\begin{aligned}\log\left(\frac{\pi_{1i}}{\pi_{2i}+\pi_{3i}}\right) &= \alpha_1+\beta_1 x_i \quad \text{and}\\ \log\left(\frac{\pi_{1i}+\pi_{2i}}{\pi_{3i}}\right) &= \alpha_2+\beta_2 x_i\end{aligned} \tag{1}$$

where π is the probability of each response category, ordered from lowest to highest, i is an index on the value of the covariate x, and α and β are unknown regression parameters. In most applications, it is assumed that the slope parameters of the two expressions, β_1 and β_2, are equal. This is referred to as the "proportional odds" assumption (McCullagh 1980). A property of the proportional odds model is that the parameters are invariant to a reversal in the ordering of the response categories, requiring only a change in their signs. Additionally, the slope parameter β is invariant to deletion or collapsing of the categories of the response; only the intercept parameters, α, will be affected (Ananth and Kleinbaum 1997).

When combined with the requirement that the probabilities of all categories of the response variable sum to one for a given value of the covariate, Equation (1) can be solved for the probability of each category as,

$$\begin{aligned}\pi_{1i} &= \frac{\exp(\alpha_1+\beta x)}{1+\exp(\alpha_1+\beta x)}\\ \pi_{2i} &= \frac{\exp(\alpha_2+\beta x)}{1+\exp(\alpha_2+\beta x)}-\frac{\exp(\alpha_1+\beta x)}{1+\exp(\alpha_1+\beta x)}\\ \pi_{3i} &= 1-\frac{\exp(\alpha_2+\beta x)}{1+\exp(\alpha_2+\beta x)}\end{aligned} \tag{2}$$

The parameters of Equation (2) were estimated for the cumulative logit model of fish health using the assessed values for the categorical probabilities corresponding to each value of the covariate "number of hypoxic days" and a maximum likelihood procedure, as implemented by the *lrm* function of the *Design* library for S-Plus (Harrell 1998). Because the uncertainty in the estimates of model parameters will depend on the number of data values used for estimation, an "equivalent sample size" needed to be determined for the assessment results. This was done by recognizing that the standard error of a multinomial probability estimate is equal to $\sqrt{\pi(1-\pi)/n}$, where n is the sample size (Devore 1991). The ranges given by the experts for the probability estimates (see Table 1) were assumed to represent $\pm$ one standard error from the midpoint, and equivalent sample sizes were calculated for each assessment for which a range was provided. These sample sizes ranged from 11 to 175 for the various assessments with a mean of 63. For simplicity, an equivalent sample size of 60 was used for all assessments, and the midpoint of each assessed range was taken to be the best point estimate.

Results of the model fit (Table 2 and Figure 2) show that the parameters of the cumulative logit model are well determined and that the model predictions and their associated uncertainty appropriately represent the opinions of the experts. The validity of the proportional odds assumption was confirmed using the graphical techniques described by Bender and Benner (2000).

Although the parameterized model is now sufficient for calculating fish health probabilities as a continuous function of the number of days of hypoxia, additional insight can be gained from a small amount of additional analysis. In a multinomial regression model, the categorical response variable can be interpreted as a discretized version of a continuous underlying trait that is divided into intervals by fixed breakpoints (Lunn *et al.* 2001). In practice, it is not necessary that such a latent trait exist, however this conceptualization aids in model understanding and may be helpful in identifying the actual occurrence of such a trait. The probability of a particular category k of the response variable is then the probability that the latent trait Z is in the range (θ_{k-1}, θ_k), where the θ_k are the discretization breakpoints, such that $\theta_k > \theta_{k-1}$ for $k = 1, \ldots, m$, with $\theta_0 = -\infty$ and $\theta_m = \infty$. If Z is assumed to be distributed according to a known distribution, f, then the categorical probabilities

Table 2. Maximum likelihood estimates (MLE), standard errors (SE), and correlation coefficients for the parameters of the cumulative logit model in Equation (2).

Parameter	MLE	SE
α_1	-3.878	0.315
α_2	-1.016	0.230
β	0.084	0.007
	$\text{Corr}(\alpha_1, \alpha_2) = 0.756$	
	$\text{Corr}(\alpha_1, \beta) = -0.916$	
	$\text{Corr}(\alpha_2, \beta) = -0.770$	

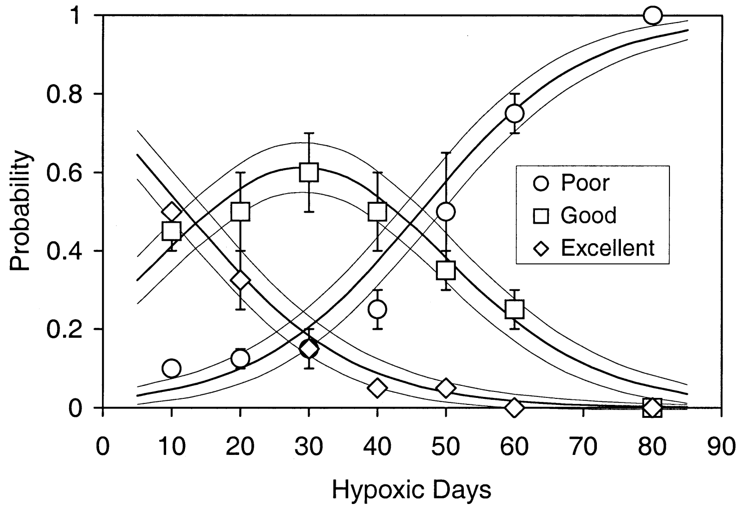

Figure 2. Plot of discrete probabilities for each fish health category as a function of the number of days of summertime hypoxia. Symbols represent the midpoints of the assessed values and vertical error bars indicate the stated ranges. Curves indicate the logistic model predictions ± one standard error.

π_k can be calculated from the cumulative probabilities of Z according to,

$$\pi_k = \Pr(\theta_{k-1} < Z < \theta_k) = \int_{\theta_{k-1}}^{\theta_k} f(s)\,ds = F(\theta_k) - F(\theta_{k-1}) \tag{3}$$

When the distribution of Z is chosen to be normal with a mean that is a linear function of selected covariates, the familiar *probit model* is produced. When the distribution is chosen to be the logistic, the *logit model* results. The logistic distribution is similar in shape to the normal distribution, but with longer tails and a cumulative distribution that can be expressed parametrically as,

$$F(s) = \frac{\exp((s-\mu)/c)}{1+\exp((s-\mu)/c)} \tag{4}$$

where μ is the mean and c is the scale parameter (Evans *et al.* 2000). Equations (3) and (4) lead to the model given in Equation (2) when μ is expressed as a linear function of x, the breakpoints θ are replaced with the unknown constants α, and c is arbitrarily fixed at a value of one.

The experts' definitions of fish health categories included growth rate thresholds of 0.2 and 0.6 mm/d. These suggest that growth rate may serve as the continuous trait, Z, underlying the categorical fish health response variable. Model diagnostics (Bender and Benner 2000) suggested that this conceptualization would be appropriate if growth rates were first subjected to a natural log transformation. The logarithm of growth rates, $\ln(\gamma)$, can then be modeled as a logistic distribution with a mean value linearly related to the number of hypoxic days, x,

$$\ln(\gamma) \sim \text{Logistic}(a + b(x), c) \tag{5}$$

where a and b are linear coefficients and c is a scale parameter that yields a distribution variance of $\pi^2 c^2/3$ (Evans *et al.* 2000).

Equation (5) can also be written as,

$$\ln(\gamma) = a + b(x) + \varepsilon \tag{6}$$

where ε has a logistic distribution with a mean of 0 and scale parameter c. Equation (6) is an example of a *generalized linear model* (Dobson 1990).

Equations (3) and (4) now yield a reparameterized version of the model in Equation (2) for fish health probabilities as a function of days of hypoxia that explicitly incorporates the underlying growth rate variable, γ,

$$\begin{aligned}
\pi_{\text{poor}} &= \Pr(\gamma < 0.2) = F(\ln(0.2)) = \frac{\exp(\ln(0.2) - (a + bx)/c)}{1 + \exp((\ln(0.2) - (a + bx))/c)} \\
\pi_{\text{good}} &= \Pr(0.2 < \gamma < 0.6) = F(\ln(0.6)) - F(\ln(0.2)) \\
&= \frac{\exp((\ln(0.6) - (a + bx))/c)}{1 + \exp((\ln(0.6) - (a + bx))/c)} - \frac{\exp((\ln(0.2) - (a + bx))/c)}{1 + \exp((\ln(0.2) - (a + bx))/c)} \\
\pi_{\text{excellent}} &= \Pr(\gamma > 0.6) = 1 - F(\ln(0.6)) = 1 - \frac{\exp((\ln(0.6) - (a + bx))/c)}{1 + \exp((\ln(0.6) - (a + bx))/c)}
\end{aligned} \tag{7}$$

where $F(\cdot)$ is the cumulative logistic distribution function. Equation (7) can be compared to Equation (2) to solve for the values of a, b, and c, yielding estimates of -0.121, -0.037, and 0.436, respectively. Uncertainty in the values of these parameters can be estimated by Monte Carlo simulation from the uncertainties in the values of α_1, α_2, and β. The generalized linear model for growth rate incorporating these parameter values (Figure 3) now provides a potentially useful setting for model interpretation.

FISH KILL ASSESSMENTS

Assessment Method

Probabilities required to predict the frequency of kills of Atlantic menhaden were also elicited in the same manner from Drs. Larry Crowder and Lisa Eby. Menhaden are more susceptible to low oxygen than other species (Hall *et al.* 1991) and travel in large schools, making their presence in large numbers at a location of upwelling more likely. Other species have historically comprised a relatively insignificant fraction of recorded fish kills and therefore are not considered in the assessments presented here.

To facilitate the assessment process, the probability of a fish kill occurring due to the combination of low oxygen and a strong cross-channel wind under each possible state of fish population health was factored into the product of one conditional and three marginal probabilities as follows:

$$P(K, O, W, H) = P(K|O, W, H)P(O)P(W)P(H) \tag{8}$$

where K indicates a kill event, O represents oxygen concentration, W indicates the occurrence of a strong cross-channel wind, and H indicates the state of fish health. Factoring the background, or marginal, probability of the infrequent event of a fish kill into an expression involving the more likely event of a kill occurring *conditional*

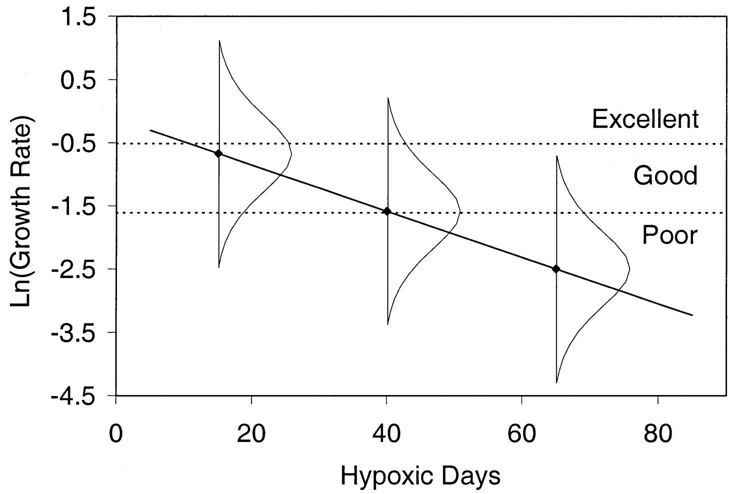

Figure 3. Generalized linear model representation of the fish health model. The solid line indicates the relationship between the mean value of ln(growth rate) and the number of hypoxic days, the curves represent the logistic probability distributions of ln(growth rate) at three representative numbers of hypoxic days (15, 40, and 65 days), and the horizontal dotted lines indicate the growth rate breakpoints (ln(0.2) and ln(0.6)) used to define the three fish health categories. The area under each curve that falls within each health category interval gives the probability of that category, conditional on the corresponding number of hypoxic days.

on the coincidence of a number of causative factors, translates small quantities that are hard to assess into quantities that can be more readily appraised by the experts (Clemen and Winkler 1999). They have now only to focus on the likelihood of a fish kill occurring under certain relevant conditions, rather than having to simultaneously consider the background frequency of cross-channel winds, low oxygen, or a particular state of fish health (all of which are now marginal probabilities to be derived from either historical data or separate models).

Again, the *fixed value* approach was used for probability assessment. A typical question was,

> Given a menhaden population in "poor" health, a day in which bottom water oxygen concentrations average 0.5 mg/L at mid-channel locations, and the strength and direction of winds are such that the bottom water is being brought to the surface along the windward shore, what is the probability of more than 100,000 fish being trapped and dying?

This question was repeated for multiple oxygen concentration values, health states, and numbers of fish. Again, the experts were allowed to state a range representing the imprecision in their assessments.

Fish Kill Assessment Results

The experts stated that their assessments of fish kill probabilities were based on their knowledge of fish movement in response to low oxygen from their monthly fish trawling program (Eby and Crowder 2002), their knowledge of sensitivity to low oxygen from their caging experiments (Eby 2001; Eby *et al.* in review), and their familiarity with the state records on the occurrence of fish kills (NCDWQ 2001). The two experts were in general agreement regarding their assessed values. Minor numerical differences were resolved through discussion and generally resulted in an increase in declared imprecision. This is manifest in assessment results by wider ranges given for each conditional probability.

Fish kills are expected to be relatively rare, even with all the conditions being right (Table 3), with consensus probabilities exceeding 50% only for kills involving less than 1,000 fish and a population in poor health. These probabilities drop substantially for kills involving more fish in better health, with an assessed probability of only 0–2% for a kill involving 100,000 fish in good or excellent health, even at the lowest oxygen concentrations. These low probabilities indicate the high potential for fish to escape and survive low oxygen events. The experts did not believe that there would be a difference in susceptibility between populations with good or excellent health status but did believe that fish of poor health were somewhat more susceptible. Again, a bottom water oxygen concentration of 2.0 mg/L served as a breakpoint,

Table 3. Assessed probabilities of kills involving various numbers of Atlantic menhaden, conditional on the occurrence of strong cross-channel winds, a given bottom water dissolved oxygen concentration, and a given fish population health status.

Oxy. conc. (mg/l)	Kill of >100,000 fish (%)	Kill of >10,000 fish (%)	Kill of >1,000 fish (%)
		Poor health	
0.5	0–5	10–15	50–70
1.0	0–5	10–15	50–70
2.0	0–5	10–15	50–70
4.0	0	0	0
6.0	0	0	0
		Good health	
0.5	0–2	5–10	30–50
1.0	0–2	5–10	30–50
2.0	0–2	5–10	30–50
4.0	0	0	0
6.0	0	0	0
		Excellent health	
0.5	0–2	5–10	30–50
1.0	0–2	5–10	30–50
2.0	0–2	5–10	30–50
4.0	0	0	0
6.0	0	0	0

with an equal chance of a kill occurring at all concentrations below 2.0 mg/L and a negligible chance of a low oxygen induced kill occurring at higher concentrations. The existence of an oxygen tolerance threshold for particular species may be a general phenomenon (Diaz and Rosenberg 1995), and the value of 2.0 mg/L in this case is consistent with behavioral avoidance thresholds observed for menhaden in the Neuse Estuary (Eby and Crowder 2002). The assumption of a threshold significantly simplifies the modeling process by allowing the occurrence of fish kills to be conditioned on the presence or absence of hypoxia (<2.0 mg/L), rather than on continuous oxygen concentration values.

PREDICTIVE MODEL

Model Implementation

Predictions of fish population health and fish kills for various frequencies of hypoxia can be generated from the two sets of assessments described above. To incorporate variability in model inputs and uncertainty in model parameters, Monte Carlo simulation (Rubinstein 1982) was performed using Analytica, a commercially available software program (Lumina 1997). Models constructed in Analytica are depicted graphically as probability networks (Pearl 1988), with uncertain variables represented by round nodes and relationships among variables indicated by connecting arrows. Relationships are quantified by probability distributions for each variable at the head of an arrow, conditional on every possible value of the variables at the tail. These conditional distributions may be described by discrete probabilities or continuous, probabilistic functions (Pearl 2000). Variables without incoming arrows are termed "marginal" nodes and are described by unconditional distributions derived from historical data or expert assessment.

The graphical model of fish health and fish kills (Figure 4) indicates that marginal distributions are required for *Number of Hypoxic Days* and *Frequency of Cross-Channel*

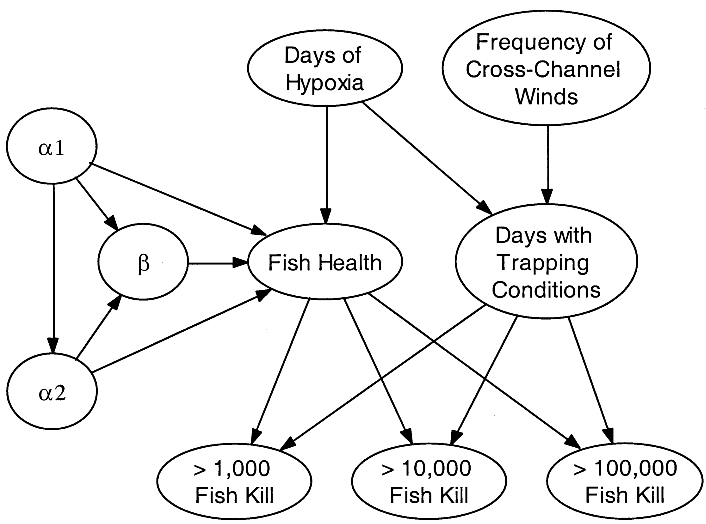

Figure 4. Graphical model of fish health and fish kills.

Winds, as well as for the model parameters, α_1, α_2, and β. The number of days of hypoxia was predicted from an oxygen dynamics model described previously by Borsuk *et al.* (2001) and is approximated by a normal distribution with a mean of 23.8 days and a standard deviation of 4.2. The frequency of cross-channel winds comes from an expert assessment also described by Borsuk *et al.* (2001) and has an exponential distribution with expected frequency of 7 days. The parameters α_1, α_2, and β were assumed to follow a joint multivariate normal distribution described by the means, standard errors, and pairwise correlation coefficients calculated during the maximum likelihood estimation (see Table 2). These were factored into conditional distributions for use in Analytica using the Cholesky decomposition of the variance-covariance matrix (Golub and Van Loan 1983).

Conditional distributions used to quantify the relationships indicated in the graphical model were derived from the expert assessments. Fish population health status was expressed as a function of hypoxic days using Equation (2). It was decided that the categorical description of health ("Excellent," "Good," "Poor") would be more meaningful to the public and more consistent with the aggregate level of detail of assessments than the latent variable, growth rate.

The assessed conditional probabilities of fish kills of varying numbers of menhaden were used directly in the network model. To incorporate the imprecision of the estimates provided by the experts (see Table 3), each probability was represented by a beta distribution with a mean equal to the midpoint of the stated range and a standard deviation equal to one half the range. The beta distribution is a convenient choice for representing the uncertainty in probability values as it is constrained to the [0,1] interval (Evans *et al.* 2000).

The finding that only days with an oxygen concentration less than 2.0 mg/L have the potential for a fish kill suggests the use of a "collector node" to represent the scenarios of concern (Abramson *et al.* 1996). A node labeled "Days with Trapping Conditions" was added to represent the days with the joint occurrence of strong cross-channel winds and hypoxic bottom water. The probability of such conditions is a joint probability, calculated as the product of the distributions describing these two marginal nodes. The assessed fish kill probabilities are then conditioned on the occurrence of "trapping conditions" and fish population health status (Figure 4). Because of the relative infrequency of fish kills of any size, probabilities are expressed as the expected number of fish kills in a 10-year period.

To demonstrate how the effects of nutrient management and associated reductions in algal growth and oxygen demand might be assessed, results are also presented in which the distribution representing the current number of hypoxic days has been replaced with a distribution having the same standard deviation and a mean value reduced by 25%. This reduction is not intended to reflect the results of any particular eutrophication model but only to demonstrate how such predictions of water quality might be extended to evaluate ecological consequences.

Model Results

Results for model endpoints displayed as predictive probability distributions (Figure 5) show the relative likelihood of different values for the endpoints. Under current conditions, the most likely state of fish population health is "good" with

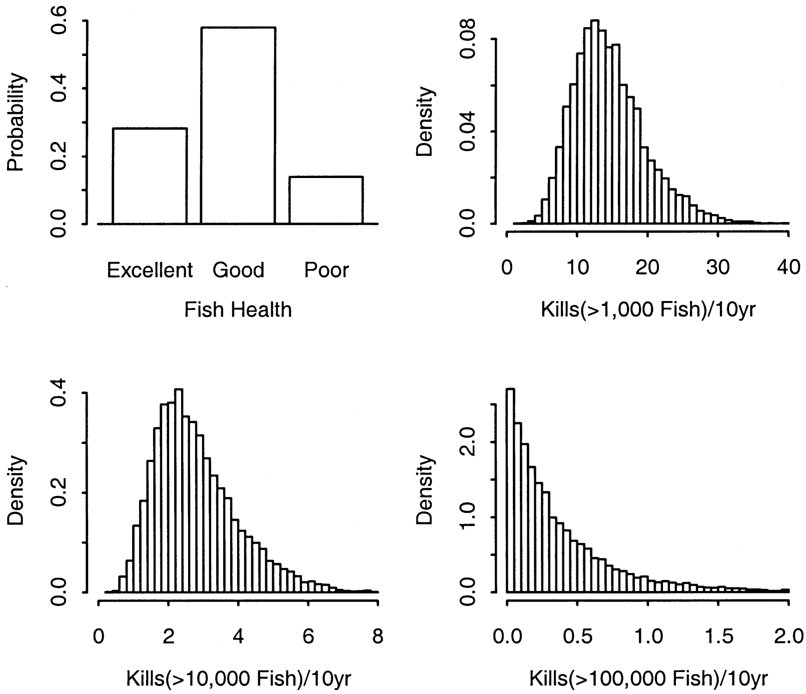

Figure 5. Predictive probability distributions of model endpoints for current conditions, including: a) categorical description of fish health, b) the number of fish kills involving more than 1,000 fish occurring in 10 years, c) the number of fish kills involving more than 10,000 fish occurring in 10 years, and d) the number of fish kills involving more than 100,000 fish occurring in 10 years.

a probability of 58%, while "excellent" has a probability of 28% and "poor" of 14%. Fish kills of any size are predicted to be relatively rare events. Over a period of 10 years, the model predicts between approximately 7.5 and 24.2 kills (the 90% predictive interval) involving more than 1,000 fish, between 1.2 and 5.1 kills involving more than 10,000 fish, and an average frequency of 0.02 to 1.3 per 10 years for kills involving 100,000 fish or more. For reference, there were 8, 5, and 2 documented fish kills of sizes >1,000, >10,000, and >100,000 fish, respectively, between the years 1989 and 1999 in the middle portion of the estuary. Additionally there were 6 kills in which the number of fish involved was not reported (NCDWQ 2001).

If eutrophication management should result in a 25% reduction in the number of days of summertime hypoxia, then the probability of fish health being excellent is estimated to increase to 38%, while the probability of poor health drops to 9% (Figure 6). The 90% predictive interval for the number of small fish kills (>1,000 fish) drops to between 5.1 and 18.3 over a 10-year period, and to 0.9 to 3.8 for kills involving greater than 10,000 fish. Large kills (>100,000 fish) are only expected to occur at a frequency of between 0.01 and 0.9 times per 10 years.

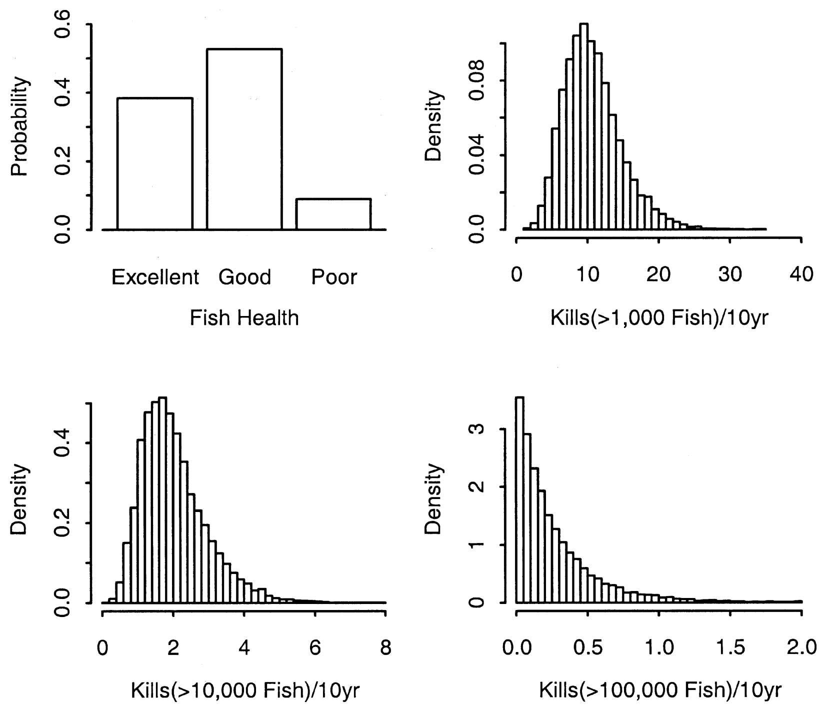

Figure 6. Predictive probability distributions of model endpoints for a 25% reduction in the number of hypoxic summer days, including: a) categorical description of fish health, b) the number of fish kills involving more than 1,000 fish occurring in 10 years, c) the number of fish kills involving more than 10,000 fish occurring in 10 years, and d) the number of fish kills involving more than 100,000 fish occurring in 10 years.

DISCUSSION

Although many factors may potentially control fish growth and survival, in the present study only the effects of hypoxia were explicitly considered. The effects of other factors, such as temperature and salinity changes, recruitment success, commercial harvest, or hurricane effects were considered part of the background "noise," manifest as probabilistic quantities in the expert assessments. These components of variability contribute to overall risk (Hattis and Anderson 1999), but are not currently subject to management control and so were not used as explicit influence factors. This pragmatic choice kept the model consistent with the information content of the available trawling and experimental data upon which the experts based their judgments. While this level of analysis may not satisfy the interests of process-oriented scientists looking to improve basic mechanistic understanding, it is consistent with current policy questions concerned primarily with the degree of eutrophication management necessary to control the harmful effects of hypoxia.

The relatively simple model structure employed in the current study leaves open the possibility of unanticipated feedbacks, non-linearities, and changes in system behavior. These are real concerns which add significant uncertainty to model predictions. However, by definition, these structural errors are currently unknown and cannot be represented in a model that is constrained by the present state of scientific knowledge. Therefore, the model, like all others, must be viewed simply as a tool for exploring the logical implications of a set of evolving assumptions.

The philosophy employed in this study is consistent with the Bayesian approach to statistical inference and decision, which combines the formal properties of probability theory with the conviction that probabilities are a useful way of expressing one's degree of knowledge (Newman and Evans 2002). In the Bayesian framework, results such as those presented here may provide a *prior* distribution for model parameters and predictions that can later be updated using additional data to obtain revised *posterior* distributions (Small and Fischbeck 1999). These can then serve as the prior for subsequent updates. Such an iterative process provides a logical method for recording advancements in scientific knowledge and predictive ability (Ellison 1996).

The literature on obtaining subjective probability distributions for Bayesian analysis generally suggests using three to five experts (see review by Clemen and Winkler 1999). The reasoning is that a larger set of experts will result in more broadly representative, and therefore supposedly more accurate, assessments. However, experimental evidence indicates that groups of experts are only slightly more accurate than the average individual expert, and the most qualified individuals in a group often outperform the group as a whole (Clemen and Winkler 1999). For the present study, we were fairly confident that the combined experience of the selected experts was the most comprehensive available regarding the specific subject of linking hypoxia to fish health and kills in the Neuse Estuary. The unique configuration of the Neuse Estuary requires site specific knowledge that fisheries scientists studying other estuaries would not possess. Additionally, one aim of this effort was to characterize the research results of the Neuse *ModMon* project, and the work of the two selected scientists represents the major fisheries component of that project.

To facilitate the elicitation process, a categorical variable was used to describe fish population health for a discrete number of days of hypoxia. To then interpolate to all possible numbers of days, a multinomial logistic regression model was developed. It was shown that this model could be interpreted as a generalized linear model with fish growth rate as the underlying continuous variable. Although such a conceptualization was not used in the present analysis, it could be used as the basis for future efforts to incorporate mechanistic knowledge about fish response to oxygen stress into ecological risk models.

The occurrence of fish kills, conditional on cross-channel winds, hypoxia, and menhaden health status, was essentially treated as a binomial process, assuming that each event is independent and that the probability of a kill is constant from one event to the next. It seems reasonable to assume independence across events, considering the small fraction of the total population involved in even the largest kills. However the assumption of a fish kill probability that is constant across an entire summer may warrant additional investigation. It is possible that the susceptibility of the fish population to kill events depends on their age or size, so that the probability may change over the course of the season. If this can be supported and quantified, then additional detail can be added to the model.

The predictions of fish kills derived from the expert assessments correspond reasonably well with recorded occurrences over the past decade. However, there is a slight indication of over-prediction of small kills and under-prediction of larger ones. This is somewhat counter to the finding of behavioral scientists who have observed that people tend to overestimate the probability of infrequent, severe events and underestimate the probability of common ones (*e.g.*, Lichtenstein *et al.* 1978). However, research has also shown that people tend to be overconfident in their assessments of probability (Spetzler and Stael von Holstein 1975), which, in this case, would explain the relatively small intervals of imprecision in probability assessments (see Table 3). A higher degree of imprecision would lead to a greater possibility of large kills, which would lead to model results that more closely match the observed frequency. In any case, it is very difficult to draw reliable conclusions about predictive accuracy when only two such events have been recorded. Thus, the actual inaccuracy or uncertainty in model predictions may be greater than that revealed by comparison with historical data.

CONCLUSIONS

Fish kills and fish health are important and tangible indicators of water quality to the general public (Borsuk *et al.* 2001). However, most models used for water quality management only predict the effect of actions on biochemical variables, such as chlorophyll a or dissolved oxygen. As a result, policy-makers are left with the difficult task of extrapolating model endpoints to attributes that matter to the public, even though such a task might be better addressed by scientists. The current study is the first to quantitatively predict the effect of hypoxia on fish population health and the frequency of fish kills in the Neuse River Estuary. Results show that the potential for improvement as a result of eutrophication management is substantial, but so is the predictive uncertainty arising from natural variation and knowledge uncertainty. While the reduction scenario evaluated here is only intended as an example and not to correspond to any particular prediction of hypoxia, related studies (Borsuk *et al.* 2001) suggest that achieving even a 25% reduction in hypoxia will require substantial decreases in estuarine carbon productivity and nutrient loading. This is an important fact for the public and policy-makers to realize so that they can maintain appropriate expectations regarding the benefits of costly management actions.

ACKNOWLEDGMENTS

The author thanks Larry Crowder and Lisa Eby for the time and effort they devoted to the elicitation process and review of the manuscript. Craig Stow also provided helpful comments on an earlier version. The open, "author-directed" peer review of this manuscript was exemplary in its thoroughness, rapidity, and value, and the three reviewers are gratefully acknowledged. The author was supported by an EPA STAR graduate fellowship. Additional support was provided by the Water Resources Research Institute of the University of North Carolina.

REFERENCES

Abramson B, Brown J, Edwards W, *et al.* 1996. Hailfinder: A Bayesian system for forecasting severe weather. Int J Forecasting 12:57–71

Ananth CV and Kleinbaum DG. 1997. Regression models for ordinal responses: A review of methods and applications. Int J Epidemiol 26:1323–33

Bender R and Benner A. 2000. Calculating ordinal regression models in SAS and S-Plus. Biom J 42:677–99

Borsuk ME, Clemen RT, Maguire LA, *et al.* 2001. Stakeholder values and scientific modeling in the Neuse River watershed. Group Decis Negotiation 10:355–73

Borsuk ME, Higdon D, Stow, CA, *et al.* 2001. A Bayesian hierarchical model to predict benthic oxygen demand from organic matter loading in estuaries and coastal zones. Ecol Modell 143:165–81

Borsuk ME, Stow CA, Luettich, RA, *et al.* 2001. Modelling oxygen dynamics in an intermittently stratified estuary: Estimation of process rates using field data. Estuar Coast Shelf Sci 52:33–49

Borsuk ME, Stow CA, and Reckhow KH. 2003. Integrated approach to total maximum daily load development for Neuse River Estuary using Bayesian probability network model (Neu-BERN). J Water Resour Plan Manage 129:271–82

Breitburg DL. 1992. Episodic hypoxia in Chesapeake Bay: Interacting effects of recruitment, behavior, and physical disturbance. Ecol Monogr 62:525–46

Breitburg DL, Rose KA, and Cowen, JH. 1999. Linking water quality to larval survival: Predation mortality of fish larvae in an oxygen-stratified water column. Mar Ecol Prog Ser 178:39–54

Breitburg DL, Steinberg N, and DuBeau S. 1994. Effects of low dissolved oxygen on predation on estuarine fish larvae. Mar Ecol Prog Ser 104:235–46

Burkholder JM and Glasgow HB. 1997. *Pfiesteria piscicida* and other *Pfiesteria*-like dinoflagellates: Behavior, impacts, and environmental controls. Limnol Oceanogr 42:1052–75

Clemen RT and Winkler RL. 1999. Combining probability distributions from experts in risk analysis. Risk Anal 19:187–203

Devore JL. 1991. Probability and Statistics for Engineering and the Sciences. Brooks/Cole Pub. Co., Pacific Grove, CA, USA

Diaz RJ and Rosenberg R. 1995. Marine benthic hypoxia: A review of its ecological effects and the behavioural responses of benthic macrofauna. Oceanogr Mar Biol 33:245–303

Dobson AJ. 1990. An Introduction to Generalized Linear Models. Chapman & Hall/CRC, New Yory, NY, USA

Dykstra MJ and Kane AS. 2000. *Pfiesteria piscicida* and ulcerative mycosis of Atlantic menhaden—Current status of understanding. J Aquat Anim Health 12:18–25

Eby LA. 2001. Response of a Fish Community to Frequent and Infrequent Disturbances in an Estuarine Ecosystem. Ph.D. Thesis. Duke University, Durham, NC, USA

Eby LA and Crowder LB. 2002. Hypoxia-based habitat compression in the Neuse River Estuary: Context-dependent shifts in behavioral avoidance thresholds. Can J Fish Aquat Sci 59:952–65

Eby LA, Crowder LB, and McClellan, C. 2000. Neuse River Estuary Modeling and Monitoring Project Stage 1: Effects of Water Quality on Distribution and Composition of the Fish Community. Water Resource Research Institute of the University of North Carolina, Raleigh, NC, USA

Eby LA, Crowder LB, and McClellan, C. Effects of habitat degradation from intermittent hypoxic disturbances on juvenile fishes (*in review*)

Ellison AM. 1996. An introduction to Bayesian inference for ecological research and environmental decision-making. Ecol Appl 6:1036–46

Evans M, Hastings NAJ, and Peacock JB. 2000. Statistical Distributions. Wiley, New York, NY, USA

Goldman RL, Eby LA, Crowder LB, *et al.* The effects of alterations in coastal ecosystems on incidence of disease in Atlantic menhaden (*Brevoortia tyrannus L.*) (*in review*)

Golub GH and Van Loan CF. 1983. Matrix Computations. Johns Hopkins University Press, Baltimore, MD, USA

Grimm V. 1999. Ten years of individual-based modelling in ecology: What have we learned and what could we learn in the future? Ecol Modell 115:129–48

Hall LW, Fisher SA, and Sullivan JA. 1991. A synthesis of water quality and contaminants data for the Atlantic menhaden, *Brevoortia tyrannus*: Implications for Chesapeake Bay. J Environ Sci Health, Part A 26:1513–44

Harrell FEJ. 1998. Design: S functions for biostatistical/epidemiologic modeling, testing, estimation, validation, graphics, and prediction. Available at http://lib.stat.cmu.edu/S/Harrell/

Hattis D and Anderson EL. 1999. What should be the implications of uncertainty, variability, and inherent "biases"/"conservatism" for risk management decision-making? Risk Anal 19:95–107

Huston M, DeAngelis D, and Post W. 1988. New computer models unify ecological theory. Bioscience 38:682–91

Lichtenstein S, Slovic P, Fischhoff B, *et al.* 1978. Judged frequency of lethal events. J Exp Psychol: Hum Learn Memory 4:551–78

Litaker RW, Vandersea MW, Kibler SR, *et al.* 2002. Life cycle of the heterotrophic dinoflagellate *Pfiesteria Piscicida* (Dinophyceae). J Phycol 38:442–63

Lumina. 1997. Analytica for Windows, User's Guide, Version 1.1. Lumina Decision Systems, Denver, CO, USA

Lunn DJ, Wakefield J, and Racine-Poon A. 2001. Cumulative logit models for ordinal data: A case study involving allergic rhinitis severity scores. Stat Med 20:2261–85

McCullagh P. 1980. Regression models for ordinal data (with discussion). J R Stat Soc [Ser B] 42:109–42

McNatt RA, Rice JA, and Noga EG. 2000. Impacts of Sublethal Hypoxia on Feeding, Growth, and Immunity of Juvenile Estuarine-Dependent Fish. N.C. Water Resources Research Conference, Water Resources Research Institute of the University of North Carolina, Raleigh, NC, USA

McNinch JE and Luettich RA. 2000. Hydrography and Circulation. Neuse River Estuary Modeling and Monitoring Project Final Report: Monitoring Phase, Report No. 325-B, pp 7–42. Water Resources Research Institute of the University of North Carolina, Raleigh, NC, USA

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

Morgan MG and Henrion M. 1990. Uncertainty: A Guide to Dealing with Uncertainty in Quantitative Risk and Policy Analysis. Cambridge University Press, Cambridge, UK

NCDMF (NC Division of Marine Fisheries). 2001. Stock Status of Important Coastal Fisheries in North Carolina. Available at http://www.ncfisheries.net/stocks/atlmenhaden.htm

NCDWQ (NC Division of Water Quality). 2000. Estuarine Fish Kills in North Carolina. Available at http://www.esb.enr.state.nc.us/Fishkill/killfact830.pdf

NCDWQ (NC Division of Water Quality). 2001. North Carolina Division of Water Quality Fish Kill Report. Available at http://www.esb.enr.state.nc.us/Fishkill/fishkill00.htm

NCDWQ (NC Division of Water Quality). 2001. 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, USA

Newman MC and Evans DA. 2002. Enhancing belief during causality assessments: Cognitive idols or Bayes's theorem. In: Newman MC, Roberts MHJ, *et al.* (eds), Coastal and Estuarine Risk Assessment, Lewis, Boca Raton, FL, USA

Pace ML. 2001. Prediction and the aquatic sciences. Can J Fish Aquat Sci 58:1–10

Paerl HW, Pinckney JL, Fear JM, *et al.* 1999. Fish kills and bottom-water hypoxia in the Neuse River and Estuary. Mar Ecol Prog Ser 186:307–9

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

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

Pihl L, Baden SP, and Diaz RJ. 1991. Effects of periodic hypoxia on distribution of demersal fish and crustaceans. Mar Biol 108:349–60

Reckhow KH. 1994. Water quality simulation modeling and uncertainty analysis for risk assessment and decision-making. Ecol Modell 72:1–20

Reckhow KH and Gray J. 2000. Neuse River Estuary Modeling and Monitoring Project Stage 1: Executive Summary and Long-Term Modeling Recommendations. Report No. 325-A. Water Resources Research Institute of the University of North Carolina, Raleigh, NC, USA

Reckhow KH, Black RW, Stockton TB, *et al.* 1987. Empirical models of fish response to lake acidification. Can J Fish Aquat Sci 44:1432–42

Rose KA. 2000. Why are quantitative relationships between environmental quality and fish populations so elusive? Ecol Appl 10:367–85

Rubin ES, Small MJ, Bloyd CN, *et al.* 1992. Integrated assessment of acid-deposition effects on lake acidification. J Environ Eng-ASCE 118:120–34

Rubinstein RY. 1982. Simulation and the Monte Carlo Method. Wiley, New York, NY, USA

Small MJ and Fischbeck PS. 1999. False precision in Bayesian updating with incomplete models. Hum Ecol Risk Assess 5:291–304

Spetzler CS and Stael von Holstein C-AS. 1975. Probability encoding in decision analysis. Manage Sci 22:340–58

Stow CA and Borsuk ME. 2003. Enhancing causal assessment of estuarine fishkills using graphical models. Ecosystems 6:11–9

NOTE FROM HERA'S CO-EDITORS-IN-CHIEF

This manuscript was prepared and submitted to *Human and Ecological Risk Assessment* (HERA) under the journal's Author-Directed Peer Review System (HERA 6(1) 2000). The following peer reviewers, submitted by the manuscript's author and approved by HERA's Managing Editor, reviewed drafts of this manuscript, approved the author's revisions, and approved the final manuscript. HERA's Managing Editor approved the final manuscript, as submitted for publication.

- E. Conrad Lamon III, Assistant Professor of Environmental Planning and Management, Department of Environmental Studies, School of the Coast and Environment, Louisiana State University, Baton Rouge, LA
- James Risbey, School of Mathematical Sciences, P.O. Box 28M, Monash University, Clayton, Victoria, Australia
- Mitchell J. Small, H. John Heinz III Professor of Environmental Engineering, Carnegie Mellon University, Civil & Environmental Engineering and Engineering & Public Policy, Pittsburgh, PA

Each reviewer was offered the opportunity to write a 500-word commentary about the reviewed manuscript.