USGS
science for a changing world

A Statistical Model and National Data Set for Partitioning Fish-Tissue Mercury Concentration Variation Between Spatiotemporal and Sample Characteristic Effects

Scientific Investigation Report 2004-5199

A Statistical Model and National Data Set for Partitioning Fish-Tissue Mercury Concentration Variation Between Spatiotemporal and Sample Characteristic Effects

By Stephen P. Wente

Scientific Investigation Report 2004-5199

U.S. Department of the Interior
Gale A. Norton, Secretary

U.S. Geological Survey
Charles G. Groat, Director

U.S. Geological Survey, Reston, Virginia: 2004

Contents

Figures

Tables

Conversion Factors

Multiply	By	To obtain
Length		
inch (in.)	2.54	centimeter (cm)
mile (mi)	1.609	kilometer (km)
Area		
acre	0.4047	hectare (ha)
Mass		
pound, avoirdupois (lb)	0.4536	kilogram (kg)

NOTE TO USGS USERS: Use of hectare (ha) as an alternative name for square hectometer (hm^2) is restricted to the measurement of small land or water areas.

A Statistical Model and National Data Set for Partitioning Fish-Tissue Mercury Concentration Variation Between Spatiotemporal and Sample Characteristic Effects

By Stephen P. Wente

ABSTRACT

Many Federal, Tribal, State, and local agencies monitor mercury in fish-tissue samples to identify sites with elevated fish-tissue mercury (fish-mercury) concentrations, track changes in fish-mercury concentrations over time, and produce fish-consumption advisories. Interpretation of such monitoring data commonly is impeded by difficulties in separating the effects of sample characteristics (species, tissues sampled, and sizes of fish) from the effects of spatial and temporal trends on fish-mercury concentrations. Without such a separation, variation in fish-mercury concentrations due to differences in the characteristics of samples collected over time or across space can be misattributed to temporal or spatial trends; and/or actual trends in fish-mercury concentration can be misattributed to differences in sample characteristics. This report describes a statistical model and national data set (31,813 samples) for calibrating the aforementioned statistical model that can separate spatiotemporal and sample characteristic effects in fish-mercury concentration data. This model could be useful for evaluating spatial and temporal trends in fish-mercury concentrations and developing fish-consumption advisories. The observed fish-mercury concentration data and model predictions can be accessed, displayed geospatially, and downloaded via the World Wide Web (http://emmma.usgs. gov). This report and the associated web site may assist in the interpretation of large amounts of data from widespread fish-mercury monitoring efforts.

INTRODUCTION

Mercury is a widespread environmental concern. In aquatic ecosystems, mercury can undergo chemical and biological transformations that produce methylmercury, a highly toxic form of mercury that is readily bioaccumulated and bioconcentrated in aquatic communities (Porcella, 1995). The major pathway by which humans are currently exposed to methylmercury is through consumption of fish (U.S. Environmental Protection Agency, 1997a; 1997b). Consumption of fish with high concentrations of methylmercury can be harmful to the health of both humans and wildlife (Wiener and others, 2002). Increased recognition of the prevalence of methylmercury, and its adverse health effects to humans and wildlife, has resulted in considerable monitoring of fish-mercury concentrations. Such monitoring is typically designed to serve three major purposes:

- Identifying sites with high fish-mercury concentrations (U.S. Environmental Protection Agency, 1993; 1995) (hereinafter referred to as measuring spatial variation, because this involves measuring variation in fish-mercury concentrations among multiple locations);

- Measuring temporal trends (hereinafter referred to as measuring temporal variation); and

- Developing fish-consumption advisories (U.S. Environmental Protection Agency, 1993) (fish-consumption advisories recommend limited fish consumption at specific locations where concerns exist about the health effects of consuming fish).

In 2002, 45 states issued mercury-related fish-consumption advisories, which applied to a total of 12,069,319 lake acres and 473,186 river miles (U.S. Environmental Protection Agency, 2003).

The distribution of fish-mercury concentrations within fish communities is complex because fish-mercury concentrations vary with a fish sample's characteristics. Among species, fish-mercury concentrations tend to increase with trophic level (MacCrimmon and others, 1983; Suns and others, 1987; Cope and others, 1990; Kim and Burggraaf, 1999). Within species, fish-mercury concentrations typically vary with length (and other measures of fish size or age) (Wiener and Spry, 1996; Huckabee and others, 1979). Within individual fish, different tissues and organs have different mercury concentrations (Giblin and Massaro, 1973; Boudou and Ribeyre, 1983; Harrison and others, 1990), and therefore different cuts of fish (whole fish, skin-on fillet, skin-off fillet, as examples) are expected to have different fish-mercury concentrations.

A major problem with interpreting fish-mercury monitoring data is partitioning the variation in fish-mercury concen-

trations between spatiotemporal and sample characteristic effects. This problem occurs because it is commonly difficult, and in some cases impossible, to collect samples with consistent sample characteristics over large regions (for example, if no species' range extends over the entire region of interest) or a variety of habitats (if no single species occurs within all habitats). Therefore, changes in fish-mercury concentrations among samples collected over time or across space may be due to either spatiotemporal or sample characteristic effects or some combination. Without a reliable method for partitioning fish-mercury concentration variation between these effects, variability in fish-mercury concentrations due to differences in the sample characteristics can be misattributed to temporal or spatial trends, or actual trends in fish-mercury concentration can be misattributed to differences in sample characteristics. To address the need for a method to partition fish-mercury concentrations between spatiotemporal and sample characteristic effects, the U.S. Geological Survey, in cooperation with the National Institute of Environmental Health Sciences, developed a statistical model using fish-mercury concentration data from the National Listing of Fish and Wildlife Advisories (NLFWA) data set.

PURPOSE AND SCOPE

This report describes a statistical model and national fish-mercury concentration data set that allow fish-mercury concentration variation to be partitioned between spatiotemporal and sample characteristic effects. Because this model and data set provide a method for partitioning fish-mercury concentration variation, this model and data set can be used to predict fish-mercury concentrations across a range of sample characteristics for sites and times when fish were sampled. Such predictions are potentially useful for calculating standardized fish-mercury concentrations that can be used to identify and measure fish-mercury concentration trends over space and time as well as calculating expected fish-mercury concentrations for all species and lengths of fish occurring at individual sites for developing site-specific fish-consumption advisories. The fish-mercury concentration data set contains data collected between June 22, 1977, and March 29, 2001. A rigorous assessment of the accuracy of this model's predictions is beyond the scope of this report. This model is not intended for making predictions for sites or times that were not sampled.

This model's predictions can be accessed, displayed geospatially, and downloaded via the World Wide Web (http://emmma.usgs.gov)[1]. Additionally, this web site provides tools for exploring spatial and temporal trends in fish-mercury con-

[1]The model predictions and data set available on this web site are frequently updated as new observations become available. As additional observations are added to the data set, additional parameters typically must be added to the model to describe those new observations. Although this report refers specifically to the original model (4,910 parameter estimates) and data set (31,813 observations), the conceptual informations provided in this report should aid in the interpretaion of the model predictions available on this web site both now and in the future.

centration data and developing site-specific fish-consumption advisories. This report and the associated web site may assist in the interpretation of large amounts of data from widespread fish-mercury monitoring efforts.

METHODS

There are essentially two ways in which spatiotemporal and sample characteristic effects can be misattributed to one another. In the first case, if a model is incapable of accurately describing the distribution of fish-mercury concentrations over space, time, and sample characteristics, then application of that model to a data set could force misattribution of spatiotemporal and sample characteristic effects. In the second case, if a model is used that can accurately describe this distribution of fish-mercury concentrations, but there are too few observations to properly constrain the parameter estimates, then the poorly constrained parameter estimates could allow misattribution of these effects. The analysis presented in this report, therefore, uses two strategies to better ensure the proper partitioning of these effects. The first strategy is to develop a model that requires as few parameter estimates as possible, while allowing fish-mercury concentrations to vary across space, time, and sample characteristics. The second strategy is to maximize the number of observations available for constraining this model's parameter estimates.

MODEL DESCRIPTION

The model presented in this report is a statistical model related to analysis of covariance and multiple-linear regression and is specifically designed to minimize the number of parameter estimates used to describe variation in fish-mercury concentrations due to sample characteristic effects. This model is similar to the equation for a line:

$$\log_e(C_{ijk} + 1) = \alpha_k \times \log_e(length_{ijk} + 1) + \beta_j + \varepsilon_{ijk} \qquad (1)$$

where α_k and β_j are the slope and intercept, respectively, of the linear relation between the $\log_e (C_{ijk} + 1)$ and $\log_e (length_{ijk} + 1)$ terms;

C_{ijk} is the fish-mercury concentration in μg Hg/kg (micrograms of mercury per kilogram of fish tissue) from the ith sample of the jth sampling event (the term 'sampling event' refers to a collection of samples from a specific site and date) for the kth species and cut combination (A "cut" is the type of tissue sample, such as skin-off fillet.);

α_k is a set of parameters relating variation in fish-mercury concentration to fish length for each of m species and cut combinations of fish;

$length_{ijk}$ is the length of the ith sample of the jth sampling event for the kth species (fish length is used in this model as the measure of fish size because length data is commonly available for fish-mercury results);

β_j is a set of parameters describing variation in fish-mercury concentrations among each of n sampling events; and

ε_{ijk} is an error term for the ith sample of the jth sampling event for the kth species.

Because α_k and β_j can take on m and n values, respectively, equation 1 describes multiple ($m \times n$) lines in log-log space or, after back-transformation into arithmetic space.

Conceptually, these curves describe the mercury concentration of a particular species and cut of fish across the range of possible fish lengths. At a length of 0, all species are assumed to have a concentration of 0. At lengths greater than 0, the fish are predicted to have positive mercury concentrations that vary according to the rate at which that species and cut combination's mercury concentration changes with respect to fish length. Therefore, a constant of 1 is added to both the concentration and length variables in equation 1 to ensure the mercury concentration is near 0 at a length of 0. Additionally, if 1 is not added to length measurement, the curves will have the unrealistic property of crossing at length of 1 in. so that species that typically have higher mercury concentrations relative to other species at lengths greater than 1 in. would be predicted to have lower mercury concentrations relative to those same species at lengths of less that 1 in. Because concentration is expressed in μg Hg/kg, the concentration values typically are 10^2 to 10^3 times the value of this constant, and therefore, the predicted concentrations are relatively unaffected by inclusion of this constant.

This model is a simplified version of the 'non-spatially variable' fish-mercury model presented in Wente (1997), but still retains the ability to describe variation in fish-mercury concentrations due to spatiotemporal and sample characteristic effects for all of the species and cut combinations that occur in a given data set. It is more complex than the version of this model used in Watras and others (1998), which described fish-mercury concentration variation for only one species and cut of fish. All of the aforementioned models and the model presented in this report are specifically designed to partition variation in fish-mercury concentrations into variability due to sample characteristics and variability among sampling events.

MODEL ASSUMPTIONS

Three major assumptions were made in designing this model. First, the fish-mercury concentration is linearly related to the size of a fish for any sample and cut combination and any sampling event in log-log space. Log transformation of the response variable was deemed necessary because the variability of the observed fish-mercury concentrations appeared to be directly related to the fish-mercury concentration. Log transforming the response variable stabilized the variance across the range of the response variable, which produces a data set that better meets the assumptions of regression analysis. Second, the variation in this relation's slope parameters (α_k) among sampling events is assumed to be small enough that each slope can be treated as a constant for all sampling events. Lastly, the variation in this relation's intercept parameters (β_j) also is assumed to be small enough that each intercept can be treated as a constant for all species and cut combinations sampled at each sampling event. These assumptions help to minimize the number of parameter estimates necessary to describe variation in fish-mercury concentrations for sample characteristics, which increases the ratio of the number of observations used to calibrate the model to the number of parameters estimated by the model, and reduces the uncertainty associated with each parameter estimate (assuming these assumptions accurately represent reality). Additionally, making these assumptions reduces the number of samples that need to be analyzed because the slope for a given species and cut combination does not have to be calibrated for each sampling event (this saves on analytical costs and field collection time).

Because most statistical software cannot analyze models in the form of equation 1, the model is rewritten in a different form using the indicator variables, $length_k$ and $event_j$:

$$\log_e(C_{ijk} + 1) = \sum_{k=1}^{m} \alpha_k \times \log_e(length_k + 1)$$

$$+ \sum_{j=1}^{n} (\beta_j \times event_j) + \varepsilon_{ijk} \tag{2}$$

These variables are defined as:

$$length_k = \begin{vmatrix} length_{ijk} & \text{if } k = k_l \\ 0 & \text{if } k \neq k_l \end{vmatrix}$$

$$event_j = \begin{vmatrix} 1 & \text{if } j = j_l \\ 0 & \text{if } j = j_l \end{vmatrix}$$

where k_l is the value of k for the lth observation; and j_l is the value of j for the lth observation.

The $length_k$ term in equation 1 causes only the observations from specific species and cut combinations to be used for calibrating the appropriate slopes, α_k. Similarly, the $event_j$ term causes only the observations from specific sampling events to be used for calibrating the appropriate intercepts, β_j. Although

the notation describing the model (equations 1 and 2) is quite compact, the actual model can require a large number of parameter estimates, because the number of species and cut combinations (m) and number of sampling events (n) can be very large. Because the data set contains left-censored (below laboratory detection limit) values, the model is implemented using a SAS procedure (LIFEREG) that produces unbiased parameter estimates from data sets with censored observations (SAS, 1989).

NATIONAL FISH-MERCURY DATA SET

Maximizing the number of observations available for constraining the model's parameter estimates is implemented by calibrating the model to a large compilation of fish-mercury concentration data from State health and environmental agencies. This compilation is a subset of the fish-mercury concentration data in the NLFWA data set (U.S. Environmental Protection Agency, 2003) as received from the U.S. Environmental Protection Agency (USEPA) on September 28, 2002. This NLFWA data set was comprised of fish-mercury analysis results for 62,314 samples collected during 9,266 sampling events at 7,617 sites. Many samples in the NLFWA data set were missing important sample descriptor information (for example, species, length, or cut) or, less commonly, appeared to contain problematic data (lengths greatly exceeding the species' maximum recorded length or indeterminant or missing sampling dates). Upon removal of these problematic samples, a modified version of the original NLFWA (hereinafter modified NLFWA) data set was obtained containing 31,813 samples from 28 states (table 1). The samples in this modified NLFWA data set predominantly are from freshwater fish species from inland waters of the United States and were sampled as 7 cuts of fish (most frequently as whole fish, skin-on fillet, and skin-off fillet, and less frequently as carcass, eggs, liver, and viscera). A detailed count of the samples in the modified NLFWA data set by species and cut of fish is presented in the appendix. An updated version of the modified NLFWA data set (additional samples added) is available at http://emmma. usgs.gov.

Approximately 22 percent of sites, sampling events, and samples are not geo-referenced with sufficient precision for mapping. Samples that are not geo-referenced are used in the model calibration and help to constrain the model's parameter estimates. Observations and model predictions for sampling events that are not geo-referenced will not appear on, and cannot be downloaded from, the mapping portions of the associated web site. To obtain observations and predictions that are not precisely geo-referenced, the entire data set should be downloaded from the 'Download Data sets' section of the web site (http://emmma.usgs.gov).

The modified NLFWA data set contains 1,398 left-censored observations. Left-censored observations have values that are less than the laboratory detection limit.

MODEL PERFORMANCE ASSESSMENT

The model's performance is assessed using measures of the model's fit to the modified NLFWA data set, magnitude of prediction error, and reliability of the separation of spatiotemporal from sample characteristic effects. Assessing the model's fit to the modified NLFWA data set is complicated by the presence of left-censored, or less-than-detection-limit, data. For observations that have measured concentrations that are above the detection limit, the residual error or lack-of-fit in the model's prediction for that observation can be precisely measured as the difference between the observed and model-predicted values. Left-censored observations, however, are known only to be less than the detection limit value. Because the observed fish-mercury concentration is not precisely known for censored observations, it is impossible to precisely measure residual error or lack-of-fit in the model's predictions for the censored observations. Because the 1,398 censored observations comprise only a small proportion (4.4 percent) of the 31,813 observations in the modified NLFWA data set, model fit was measured as the proportion of variation explained by the model for the 30,415 non-censored observations. This measure will be referred to as the *pseudo-R^2* (pR^2) and is calculated as:

$$pR^2 \;=\; 1 - \frac{SSE}{CSS}$$

where *SSE* is the sum of squared errors for the non-censored log fish-mercury concentration observations; and *CSS* is the sum of squares for the non-censored log fish-mercury concentration observations corrected for the mean of the non-censored log fish-mercury concentration observations [the log fish-mercury concentrations are calculated as $\log_e$ (μg Hg/kg + 1)].

The magnitude of the model's prediction error is estimated from the scale parameter as reported by the SAS LIFEREG procedure. A scale parameter is an unbiased estimate of the standard deviation of the variability left unexplained by the model (error) that can be used with censored data sets. It is analogous to the root mean square error (RMSE) reported by regression analysis software for non-censored data sets. Prediction error expressed as a percentage (*PE*) is calculated as:

$$PE \;=\; 100 \times \left(e^{\left(scale - \frac{scale^2}{2} \right)} - 1 \right)$$

where *scale* is the scale parameter estimated by the SAS LIFEREG procedure.

The reliability of the separation of spatiotemporal from sample characteristic effects is assessed using a comparison of the full and reduced forms of the model. The reasoning behind this assessment is that the parameter estimates of regression analysis, or any type of optimization procedure, will produce reliable parameter estimates only if the parameter estimates are 'well-constrained' by data. The full model includes all variables as in equation 2 and is compared to two reduced forms of the full model. The first reduced model is referred

Table 1. Summary of fish-mercury concentration data contributions to the National Listing of Fish and Wildlife Advisory (NLFWA) data set (September 28, 2002 version of the data set) as modified for calibration of the statistical model used in this report.

States	Sites [1]	Sampling events [1]	Species[2]	Species and cut combinations[2]	Samples[1]
Alabama	141 (141)	175 (175)	22	22	470 (470)
Arkansas	219(218)	249 (248)	20	20	783 (782)
California	61 (61)	84 (84)	23	23	97 (97)
Connecticut	53 (46)	53 (46)	4	4	618 (539)
Delaware	68 (67)	75 (74)	19	25	151 (149)
Georgia	120 (92)	120 (92)	32	32	529 (385)
Illinois	1 (1)	1 (1)	1	1	1 (1)
Indiana	119 (116)	125 (122)	43	59	498 (478)
Kentucky	50 (40)	51 (41)	23	37	114 (96)
Louisiana	68 (68)	72 (72)	30	30	679 (679)
Maryland	21 (21)	22 (22)	7	7	63 (63)
Maine	125 (125)	129 (129)	13	24	354 (354)
Michigan	426 (314)	432 (319)	34	52	5,894 (4,057)
Minnesota	1,190 (1,003)	1,467 (1,246)	42	54	12,246 (10,566)
North Carolina	438 (36)	617 (36)	79	144	2,685 (78)
North Dakota	4 (1)	4 (1)	5	5	76 (3)
Nebraska	7 (6)	26 (25)	2	4	30 (29)
New Hampshire	138 (131)	158 (145)	21	30	750 (643)
New Mexico	37 (30)	44 (33)	27	28	532 (398)
New York	24 (24)	24 (24)	9	9	535 (535)
Ohio	166 (161)	166 (161)	26	33	501 (486)
Oklahoma	46 (45)	54(52)	32	36	252 (241)
Oregon	28 (17)	34 (18)	21	25	456(270)
Texas	27 (27)	39 (39)	14	14	45 (45)
Virginia	16 (16)	16 (16)	10	10	28 (28)
Washington	12 (12)	12 (12)	2	2	41 (41)
Wisconsin	292 (288)	334 (329)	37	54	3,349 (3,309)
West Virginia	14 (8)	14 (8)	9	9	36 (17)
Total	3,911 (3,115)	4,597 (3,570)	[3]163	[3]310	31,813 (24,839)

[1] Values in parentheses indicate the number of precisely geo-referenced sites, sampling events (sampling of a specific site and time), or samples.

[2] Number of unique species or species and cut combinations in a state's data set.

[3] Values indicate a national total for unique species or species and cut combinations (does not represent a cumulative sum of state totals).

to as the "events only" model and omits the variables that describe the sample characteristic effect by omitting the term

$$\sum_{k=1}^{m} (\alpha_k \times log_e(length_k + 1))$$

from equation 2. The second is referred to as the "sample characteristics only" model and omits the variables that describe the spatiotemporal effect by omitting the

$$\sum_{j=1}^{n} (\beta_j \times event_j)$$

term from equation 2. If the full model is only marginally better at describing variability in the modified NLFWA data set than the events only model, then the event term is capable of describing most of the variability that the full model uses both event and sample characteristic terms to describe. Similarly, if the full model fits the NLFWA data only marginally better

than the sample characteristics only model, then the sample characteristic term is capable of describing most of the variability that the full model uses both event and sample characteristic terms to describe. In both cases, the implication is that the full model as calibrated to the modified NLFWA data set is incapable of reliably separating spatiotemporal effect as described by the event term from sample characteristic effect as described by the sample characteristic term. Conversely, if the full model is considerably better at describing variability in the modified NLFWA data set than both reduced forms of the model, the implication is that the event and sample characteristic terms each make unique contributions to describing fish-mercury concentration variation that the other term cannot. This would indicate the modified NLFWA data set does constrain the event and sample characteristic parameter estimates and, therefore, provides evidence that the model is separating spatiotemporal and sample characteristic effects.

The fit of the full and reduced forms of the model to the modified NLFWA data set will be compared by calculating values of pR^2, scale parameter, prediction error, and the Akaike Information Criterion (AIC) for the full and both reduced forms of the model. The AIC is an adjustment of the log likelihood calculated as:

$$AIC = -2\log L + 2p$$

where log L is the log-likelihood and p is the number of parameters estimated for a specific model. However, a better fit is always expected for a linear model that includes additional variables such as the full model in comparison to the reduced models that include fewer variables. Therefore, a statistical test, called a likelihood ratio comparison, is used to determine whether the fit of the full model is significantly better than the fit of the reduced models after accounting for the difference in the number of variables included in each model. In the likelihood ratio comparison tests, a value calculated as 2 times the difference in AIC of the full and one of the reduced models, is compared to a chi-square distribution. The number of degrees of freedom for the test is the difference in number of parameters estimated by the full model and the reduced model being compared.
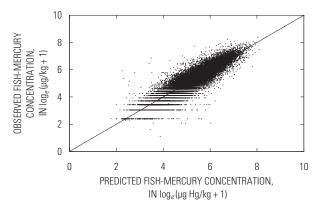
MODEL EVALUATION AND APPLICATIONS

Calibration of the model presented in this report requires one α_k parameter estimate for each of the m species and cut combinations and one β_j parameter estimate for each of the n sampling events in the modified NLFWA data set. The modified NLFWA data set described in this report contains observations from 310 species and cut combinations and 4,597 sampling events (table 1). The statistical model described in this report, therefore, uses 310 species and cut parameter esti-

mates (α_k) and 4,597 sampling event parameter estimates (β_j) for a total of 4,907 parameter estimates.

MODEL FIT

As with all models, the familiar statement, "all models are wrong: some models are useful" (Box, 1979) applies. This model is based on a simplified conceptualization of how fish-mercury concentrations are distributed. The distribution of fish-mercury concentrations will vary with the structure of the aquatic food web during a particular event (Cabana and Rasmussen, 1994). This implies that the model assumption of common slopes (α_k) for all sampling events and common intercepts (β_j) within each sampling event is 'wrong' to some degree. This model, however, does fit the non-censored portions of the modified NLFWA data set well ($pR^2 = 0.82$, fig. 1). The model, therefore, can be thought of as a good approximation of reality and the simplifying assumption of similar food web structure at all sampling events is probably not unwarranted because actual differences in food web structure between sampling events must rarely be large.

Figure 1. Relation of noncensored observations to predicted fish-mercury concentrations in log units.

The fit of the full model and both reduced models to the modified NLFWA data set is compared in table 2. The larger pR^2 value and smaller values for scale parameter, prediction error, and AIC for the full model indicate that the full model fits the modified NLFWA data set much better than either of the reduced models. In table 3, the likelihood ratio comparison tests indicate the better model fit of the full model as compared to that of the reduced models is statistically significant after accounting for the difference in number of variables in each model. These results provide evidence that, in general, the modified NLFWA data set constrains the models' parameter estimates such that neither spatiotemporal nor sample characteristic effect variables alone are able to fit the variability in the data that the other effect's variables are able to fit. Therefore in general, the full model as calibrated to the modi-

Table 2. Comparison of model fit for the full and reduced fish-mercury models.

Models	pR2	Scale parameter	Prediction error (%)	AIC
Full	0.82	0.41	38.2	42,742
Events Only	.60	.61	52.9	67,342
Sample Characteristics Only	.41	.76	60.3	72,922

fied NLFWA data set separates spatiotemporal and sample characteristic effects. However, not all parameter estimates are equally constrained. It is likely that parameter estimates for species and cut combinations that were rarely sampled during events where other more frequently sampled species and cut combinations were collected do not separate spatiotemporal and sample characteristic effects well. Similarly, event parameter estimates that are based solely on data from rare species and cut combinations may also misattribute spatiotemporal and sample characteristic effects. A more rigorous assessment of the accuracy of this model's estimation of fish-mercury concentrations, or predictions, is beyond the scope of this report.

FISH-MERCURY CONCENTRATION PREDICTION

The statistical model presented in this report represents an attempt to provide a realistic description of the distribution of fish-mercury concentrations across space, time, and sample characteristics using as few parameter estimates as possible to describe variation in fish-mercury concentrations across sample characteristics. To predict fish-mercury concentration in log units for a specific species and cut combination and length of fish, equation 1 is simplified as:

$$\log_e(C_{pred} + 1) = \alpha_k \times \log_e(length_{pred} + 1) + \beta_j \quad (3)$$

where C_{pred} is a fish-mercury concentration prediction in units of μg Hg/kg;

α_k is a parameter estimate from equation 2 for a specific species and cut combination for any of the 310 species and cut combinations in the modified NLFWA data set;

$length_{pred}$ is any desired fish length (in inches); and

β_j is a parameter estimate from equation 2 for any specific sampling event of the 4,957 sampling events in the modified NLFWA data set. In log-log space, equation 3 is the equation of a straight line with slope of α_k and an intercept of β_j. It is used to graph the lines in figures 2a and 2c.

To predict fish-mercury concentration in arithmetic units for a specific species and cut combination and length of fish, equation 3 can be transformed into:

$$C_{pred} = e^{(\alpha_k \times \log_e(length_{pred} + 1) + \beta_j)} - 1 \quad (4)$$

Equation 4 is used to graph the curves in figures 2b and 2d. The simplicity of this model's description of fish-mercury variation across sample characteristics is demonstrated in figure 2a, which depicts fish-mercury concentration predictions for a single hypothetical sampling event. Theoretically, one line could be predicted for each of the 310 combinations of species and cut occurring in the modified NLFWA data set. For clarity of presentation, only concentration predictions for skin-off fillet samples of 11 species are depicted throughout figure 2. Each line is depicted for the range of fish lengths sampled from that species in the modified NLFWA data set. The predictions in figure 2a also are shown in concentration and length units in figure 2b. Comparison of figures 2a and 2b shows how small variations of slope in log-log space (figure 2a) can describe variation in the rate of mercury accumulation (per unit of fish length) across multiple species and lengths of fish (figure 2b).

Different sampling events have different intercepts as depicted in the log-log graph of figure 2c. This figure presents predictions for the same species and cut combinations of fish at two sampling events with one-half (blue lines) and twice (red lines) the fish-mercury concentrations depicted in figure 2a. Notice that each line in figure 2a has two corresponding parallel lines (one blue and one red) in figure 2c. Similarly, lines for the same species and cut at any other sampling event also would be parallel. Similar to the relation between figures 2a and 2b, figure 2d shows the predictions of figure 2c transformed into concentration and length units. Figure 2d shows how changes in the intercepts in figure 2c (the corresponding red and blue lines in figure 2c differ by a constant in log-log space) translate into multiplicative changes in fish-mercury concentration between sampling events (the corresponding red and blue curves in figure 2d differ by a multiplication factor in arithmetic space).

Table 3. Likelihood ratio comparisons for the full and reduced fish-mercury models.

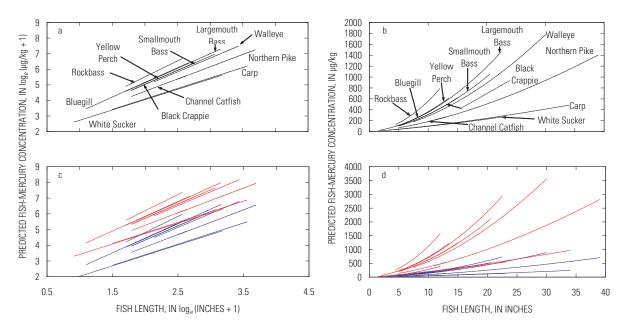| Model comparison	Change in AIC	Degrees of freedom	P-value
Full and events only	24,600	310	< 0.0001
Full and sample characteristics only	30,180	4,596	< .0001

Figure 2. Predicted fish-mercury concentrations as a function of length for skin-off fillet cuts of 11 frequently sampled species in the upper Midwestern United States depicted in (a) log space, and (b) arithmetic space for a hypothetical sampling event; (c) and (d) are similar but compare two hypothetical sampling events with high (red lines) and low (blue lines) fish-mercury concentrations.

PREDICTION ACCURACY

Measures of model fit like the pR^2 (or R^2) do not directly indicate how well this model will perform its intended applications. This is because the applications for which the model is intended commonly require predictions for sample characteristics that were not sampled at the actual sampling events, while measures of model fit only indicate how well the model's predictions match the observations that were actually sampled. For example, measuring spatial and temporal variation requires converting between the samples collected during each sampling event and a standardized sample that may have been collected during few of the actual sampling events. Similarly, development of fish-consumption advisories would be greatly facilitated by the ability to estimate concentrations for fish that likely occur at a site, but were not sampled during the sampling event for that site. Because estimating fish-mercury concentrations for samples that were not included in the calibration data set is less accurate than estimates for samples included in the calibration data set, the ability of the model to predict the fish-mercury concentrations needed for the model's intended applications will be less accurate than the pR^2 measure indicates.

A reduction in prediction accuracy will be especially acute for predictions based on species and cut combination parameter estimates (α_k), or sampling event parameter estimates (β_j), that are poorly constrained. Therefore, users of predictions from this model should determine how well each parameter estimate is constrained. The appendix provides sample counts for individual species and cut combinations[1] , which indicates species and cut parameter estimates (α_k) that are well constrained (high numbers of samples) and poorly constrained (low numbers of samples). A similar table for sampling event parameter estimates (β_j) was judged to be too long for inclusion in this report and would need constant revision as new sampling events are added to the data set. Therefore, users are encouraged to download the current data set from the web site, sort the data set by sampling event code, and judge the quality of the model prediction based on the number of samples collected for an event as well as the number of samples in the data set for each of the species and cut combinations collected during that sampling event.

MODEL APPLICATIONS

Using the model's predictions to assess spatial and temporal variation and develop site-specific fish-consumption advisories is accomplished by setting one or two of the three variables in equation 4 to appropriate constants and varying the remaining variable(s). In this way, spatial and temporal variability can be assessed by setting α_k to a specific or 'standardized' species and cut combination, length of fish (*length $_{pred}$*) to a constant value, and varying the sampling event (β_j). For example, temporal variability is depicted by displaying standardized fish-mercury concentrations for sampling

[1] This data set is correct as of September 28, 2002. The NLFWA data set often is updated. The most recent version of the data set is at http://emmma.usgs.gov

events from the same site across time. Spatial variability is depicted by displaying standardized fish-mercury concentrations for sampling events from a small period of time (the same year, for example) for the sites that were sampled within that time period. Site-specific fish-consumption advisories can be developed by predicting fish-mercury concentrations for multiple species, lengths, and cuts of fish (if necessary) for the most recent sampling event at the site of interest. To accomplish this, β_j is held constant while α_k and $length_{pred}$ are varied for all the species and cut combinations and lengths of fish of interest at that site (similar to figure 2b). Consumption advisories can then be developed for the species, lengths, and cuts of fish that have predicted concentrations that exceed the concentrations of concern for the specific group of fish consumers.

The shape of the curves as depicted in figure 2b can be used to estimate important bio-physical parameters describing the net accumulation of mercury as a fish grows in length. Also, the relative differences between these curves describe which sizes and species of fish typically are safer to eat with respect to fish-mercury ingestion concerns. Even if fish-mercury concentrations have never been measured at a site, or for a particular species at a site, a fish consumer can reduce potential mercury exposure by choosing to consume sizes and species of fish that are expected to have lower fish-mercury concentrations according to the relations described by the model.

FUTURE STUDIES

The model and data set presented in this report could form the basis for many important research products. A brief listing might include:

• *Producing a continuous coverage of local fish-mercury temporal trend estimates.* The model presented in this report is specifically designed to require as few parameter estimates as possible to describe the variation of fish-mercury concentrations with sample characteristics. The model's description of spatial and temporal variation, however, requires a large number of parameter estimates. Future versions of this model could be designed to reduce the number of parameter estimates used to describe temporal and spatial variation using geostatistical methods similar to the 'spatially variable' model of Wente (1997). Such methods could produce local estimates of temporal variation over large regions based on the average temporal variation observed at nearby fish-mercury sampling sites.

• *Prediction uncertainty estimates.* Determining the accuracy of individual model predictions is difficult because the number of fish in a composite fish sample often was not compiled into the NLFWA data set. Compositing a number of nearly identical fish together into a sample will produce a more accurate estimate of the mean fish-mercury concentration than will a sample comprised of a single fish. Obtaining this information in future updates to this data set would allow

better estimates of prediction accuracy to be given for individual model predictions and aid in estimating transformation bias correction factors.

• *Identifying factors that cause spatiotemporal variation in fish-mercury concentrations.* Spatiotemporal variation only is described by the model in this report rather than attributed to any specific cause (for example, changes in methylation efficiency or mercury loadings). The spatiotemporal variation described by this model, however, can be used to formulate hypotheses regarding the factors causing this variation. Additionally, after collecting data on the hypothesized causal factors, a variation of the model and data set could be used to test these hypotheses.

• *Estimating human health effects from fish consumption.* An individual's exposure to fish-mercury is largely determined by that individual's fish-consumption patterns and the concentration of mercury in the fish that the individual consumes. Because fish-consumption patterns are highly variable, better estimates of a local population's exposure risk may be obtained by simulating the local variation in consumption habits (based on fish-consumption survey data) and the local fish-mercury concentrations (as estimated from the model and modified NLFWA data set in this report). Such simulations could be run over large regions to identify which locales have the largest proportions of their populations at risk of adverse affects from methylmercury exposure.

• *Optimization of regulatory strategies for reducing fish-mercury concentrations.* Reducing fish-mercury concentrations over large regions will likely be expensive. It may be beneficial to test different regulatory strategies or combinations of regulatory strategies through models to optimize a final remediation strategy before implementation.

SUMMARY

Many Federal, Tribal, State, and local agencies monitor mercury in fish-tissue samples to identify sites with elevated fish-tissue mercury (fish-mercury) concentrations, track changes in fish-mercury concentrations over time, and produce fish-consumption advisories. The fish-mercury concentration data sets generated by these monitoring programs often contain data from samples with differing characteristics (species, cuts, and sizes of fish) as well as from different locations and dates of sampling. Interpreting such data sets can be difficult because spatiotemporal trends in fish-mercury concentrations can be difficult to reliably separate from the effect of sample characteristics on fish-mercury concentrations. Without such a separation, variation in fish-mercury concentrations due to differences in the characteristics of samples collected over time or across space can be misattributed to temporal or spatial trends; and/or actual trends in fish-mercury concentration can be misattributed to differences in sample characteristics.

This report describes a statistical model capable of partitioning variation in fish-mercury concentrations into

spatiotemporal and sample characteristic effects and a large (31,813 samples) national compilation of fish-mercury data that can be used to calibrate and constrain the parameter estimates of this model. Calibration of this model was performed using a method that produces unbiased parameter estimates from fish-mercury concentration data sets containing left-censored (less than laboratory detection limit) observations. The model's performance was evaluated using measures of model fit, prediction error, Akaike Information Criterion, and likelihood ratio comparisons. These measures provided evidence that, in general, the model as calibrated to a national fish-mercury data set is capable of separating spatiotemporal and sample characteristic effects.

The observed fish-mercury concentration data and model predictions can be accessed, displayed geospatially, and downloaded via the World Wide Web (http://emmma.usgs.gov). This web site also provides tools for exploring spatial and temporal trends in fish-mercury concentration data and developing site-specific fish-consumption advisories for mercury.

REFERENCES CITED

Boudou, A. and Ribeyre, F., 1983, Contamination of aquatic biocenoses by mercury compounds—An experimental ecotoxicological approach, in Nriagu, J.O., ed., Aquatic Toxicology: John Wiley and Sons, New York, p. 73-116.

Box, G.E.P., 1979, Robustness in the strategy of scientific model building, in Launer R.L., and Wilkinson, G.N. , eds., Robustness in statistics: Academic Press, New York, 202 p.

Cabana, G., and Rasmussen, J.B., 1994, Modeling food chain structure and contaminant bioaccumulation using stable nitrogen isotopes: Nature, v. 372, p. 255-257.

Cope, W.G., Wiener, J.G., and Rada, R.G., 1990, Mercury accumulation in yellow perch in Wisconsin seepage lakes—Relation to lake characteristics: Environmental Toxicology and Chemistry, no. 9, p. 931-940.

Giblin F.J., and Massaro, E.J., 1973, Pharmacodynamics of methyl mercury in the rainbow trout (*Salmo gairdneri*)—Tissue uptake, distribution and excretion: Toxicololgy and Applied Pharmacology, no. 24, p 81-91.

Harrison, S.E., Klaverkamp, J.F., and Hesslein, R.H., 1990, Fates of metal radiotracers added to a whole lake—Accumulation in fathead minnow (*Pimephales promelas*) and lake trout (*Salvelinus namaycush*): Water, Air, and Soil Pollution, no. 52, p. 277-293.

Huckabee, J.W., Elwood, J.W., and Hildebrand, S.G., 1979, Accumulation of mercury in freshwater biota, in Nriagu, J.O., ed., Biogeochemistry of Mercury in the Environment: Elsevier/North-Holland Biomedical Press, New York, p. 277-302.

Kim, J.P., and Burggraaf, S., 1999, Mercury bioaccumulation in rainbow trout (*Oncorhynchus mykiss*) and the trout food web in lakes Okareka, Okaro, Tarawera, Rotomahana and Rotorua, New Zealand: Water, Air, and Soil Pollution, no. 115, p. 535-546.

MacCrimmon, H.R., Wren, C.D., and Gots, B.L., 1983, Mercury uptake by lake trout, *Salvelinus namaycush,* relative to age growth, and diet in Tadenac Lake with comparative data from other Precambrian Shield Lakes: Canadian Journal of Fisheries and Aquatic Sciences, no. 40, p. 114-120.

Porcella, D.B., 1995, Aquatic biogeochemistry and mercury cycling model (MCM), in National Forum on Mercury in Fish Proceedings, U.S. Environmental Protection Agency, Office of Water, EPA 823-R-95-002, 211p.

SAS (SAS Institute Inc.), 1989, SAS/STAT User's Guide v. 2, version 6, 4th ed., SAS Institute Inc., Cary, NC. 795 p.

Suns, K., Hitchin, G., Loescher, B., Pastorek, E., and Pearce, R., 1987, Metal accumulations in fishes from Muskoka-Haliburton lakes in Ontario (1978-1984): Ontario Ministry of the Environment, Rexdale, Ontario, and p. 38.

U.S. Environmental Protection Agency, 1993, Guidance for assessing chemical contaminant data for use in fish advisories, vol. I: Fish sampling and analysis, 2nd ed.: U.S. Environmental Protection Agency, Office of Water, EPA 823-R-95-007, Variously paged.

U.S. Environmental Protection Agency, 1995, Guidance for assessing chemical contaminant data for use in fish advisories, vol. II: Risk assessment and fish-consumption limits: U.S Environmental Protection Agency, Office of Water, EPA 823-B-94-004, Variously paged.

U.S. Environmental Protection Agency, 1997a, Mercury study report to congress, vol. III: An assessment of exposure from anthropogenic mercury emissions in the United States: U.S. Environmental Protection Agency, Office of Air Quality Planning & Standards and Office of Research and Development, EPA-452/R-96-001, Variously paged.

U.S. Environmental Protection Agency, 1997b, Mercury study report to congress, vol. VII: Characterization of human health and wildlife risks from mercury exposure in the United States: U.S. Environmental Protection Agency, Office of Air Quality Planning & Standards and Office of Research and Development, EPA-452/R-97-009, Variously paged.

U.S. Environmental Protection Agency, 2003, Update—National listing of fish and wildlife advisories: U.S. Environmental Protection Agency, Office of Water, EPA-823-F-003. 6 p.

Watras, C.J., Back, R.C., Halvorsen, S., Hudson, R.J.M., Morrison, K.A., and Wente, S.P., 1998, Bioaccumulation of mercury in pelagic freshwater food webs: Science of the Total Environment, v. 219, no. 2-3, p. 183-208.

Wente, S.P., 1997, A spatially and temporally variable model of mercury concentrations in aquatic communities with applications to public health protection and water quality assessment, Purdue University: West Lafayette, Ind., Ph.D. dissertation, 132 p.

Wiener, J.G., Krabbenhoft, D.P., Heinz, G.H., and Scheuhammer, A.M., 2002, Ecotoxicology of Mercury, in Hoffman, D.J., Rattner, B.A., Burton, G.A., and Cairns, J., Handbook of Toxicology, 2 ed.: CRC Press, 1344 p.

Wiener, J.G., and Spry, D.J., 1996, Toxicological significance of mercury in freshwater fish, in Beyer, W.N., Heinz, G.H., and Redmon-Norwood, A.W., eds., Environmental Contaminants in Wildlife—Interpreting Tissue Concentrations: Lewis Publ., Boca Raton, Fla., p. 297-339.

Appendix 1. Sample characteristics (species, length, and cut of fish) of fish-mercury analysis results included in the modified National Listing of Fish and Wildlife Advisory (NLFWA) data set (September 28, 2002, version of the data set).

Species and common names	Length range (in.)	Whole fish	Sample counts by cut of fish			Total
			Skin-on fillet	Skin-off fillet	Other cuts	
Achiridae						
Trinectes maculatus (Bloch & Schneider, 1801), Hogchoker	5.5 - 8.9	3	1			4
Acipenseridae						
Acipenser fulvescens Rafinesque, 1817, Lake sturgeon	12.3 - 68.1		5		2	7
A. oxyrinchus oxyrinchus Mitchill, 1815, Atlantic sturgeon	8.2 - 20.0		3			3
Scaphirhynchus platorynchus (Rafinesque, 1820), Shovelnose sturgeon	30.9	1				1
Amiidae						
Amia calva Linnaeus 1766, Bowfin	8.9 - 28.0	44	5	46		95
Anguillidae						
Anguilla rostrata (Lesueur, 1817), American eel	1.5 - 30.5	28	4	3		35
Ariidae						
Bagre marinus (Mitchill, 1815), Gafftopsail Sea Catfish	18.8			1		1
Batrachoididae						
Opsanus tau (Linnaeus, 1766), Oyster toadfish	6.3 - 11.6	1	1			2
Catostomidae						
Carpiodes carpio (Rafinesque, 1820), River carpsucker	10.1 - 21.6		6	22		28
C. cyprinus (Lesueur, 1817), Quillback	7.0 - 17.8	1	17			18
C. velifer (Rafinesque, 1820), Highfin carpsucker	11.5			1		1
Catostomus catostomus catostomus (Forster, 1773), Longnose sucker	13.2 - 22.3		14	12		26
C. commersonnii (Lacepède, 1803), White sucker	1.9 - 23.0	174	1,131	74		1,379
C. macrocheilus Girard, 1856, Largescale sucker	12.1 - 23.3	43				43
C. santaanae (Snyder, 1908), Santa Ana sucker	1.3	1				1
Cycleptus elongatus (Lesueur, 1817), Blue sucker	23.6		1			1
Cyprinella lutrensis (Baird & Girard, 1853), Red shiner	2.1 - 2.5	19				19
Erimyzon oblongus (Mitchill, 1814), Creek chubsucker	5.3 - 14.0	31	1			32
E. sucetta (Lacepède, 1803), Lake chubsucker	7.2		1			1
Gila bicolor (Girard, 1856), Tui chub	1.8 - 11.0	6	4			10
G. orcuttii (Eigenmann & Eigenmann, 1890), Arroyo chub	2.5 - 3.1	6				6
Hypentelium etowanum (Jordan, 1877), Alabama hog sucker	4.9 - 6.6		4			4
H. nigricans (Lesueur, 1817), Northern hog sucker	4.9 - 13.8	21	14			35
Ictiobus bubalus (Rafinesque, 1818), Smallmouth buffalo	13.4 - 32.7	7	33	25		65
I. cyprinellus (Valenciennes, 1844), Bigmouth buffalo	11.4 - 28.7	1	37	25		63
I. niger (Rafinesque, 1819), Black buffalo	20.7			1		1
Minytrema melanops (Rafinesque, 1820), Spotted sucker	11.2 - 23.5	13	19	23		55
Moxostoma anisurum (Rafinesque, 1820), Silver redhorse	9.7 - 23.5	14	19			33
M. carinatum (Cope, 1870), River redhorse	11.3 - 24.6	1	4	6		11
M. duquesnii (Lesueur, 1817), Black redhorse	11.4 - 17.6	10	20			30
M. erythrurum (Rafinesque, 1818), Golden redhorse	10.1 - 22.3	5	29	1		35
M. macrolepidotum (Lesueur, 1817), Shorthead redhorse	8.5 - 24.9	7	42			49
M. poecilurum (Jordan, 1877), Blacktail redhorse	10.1 - 15.4			5		5
M. valenciennesi Jordan, 1885, Greater redhorse	13.2 - 21.5	15	3			18
Scartomyzon rupiscartes (Jordan & Jenkins, 1889), Striped jumprock	6.8 - 7.2	1	1			2
Centrarchidae						
Ambloplites rupestris (Rafinesque, 1817), Rock bass	2.4 - 11.9	18	319	47		384
Centrarchus macropterus (Lacepède, 1801), Flier	3.6 - 7.2	7	3			10
Enneacanthus gloriosus (Holbrook, 1855), Bluespotted sunfish	5.9 - 7.8			4		4
Lepomis auritus (Linnaeus, 1758), Redbreast sunfish	3.0 - 9.0	5	10			15
L. cyanellus Rafinesque, 1819, Green sunfish	2.1 - 9.4	36	30	4		70
L. gibbosus (Linnaeus, 1758), Pumpkinseed	3.3 - 9.5	40	81			121
L. gulosus (Cuvier, 1829), Warmouth	4.2 - 26.0	14	34	4		52
L. macrochirus Rafinesque, 1819, Bluegill	1.9 - 15.0	114	984	98		1,196
L. megalotis (Rafinesque, 1820), Longear sunfish	2.5 - 7.3	25	9			34
L. microlophus (Günther, 1859), Redear sunfish	4.3 - 11.1	14	112	28		154

Species and common names	Length range (in.)	Whole fish	Skin-on fillet	Skin-off fillet	Other cuts	Total
L. punctatus (Valenciennes, 1831), Spotted sunfish	10.8	1				1
Micropterus cataractae Williams & Burgess, 1999, Shoal bass	9.4 - 13.4		3			3
M. coosae Hubbs & Bailey, 1940, Redeye bass	9.4 - 10.6			3		3
M. dolomieu Lacepède, 1802, Smallmouth bass	3.0 - 19.7	36	757	145		938
M. punctulatus (Rafinesque, 1819), Spotted bass	4.3 - 19.1	4	52	99		155
M. salmoides (Lacepède, 1802), Largemouth bass	0.9 - 32.4	214	1,408	1,754	5	3,381
Pomoxis annularis Rafinesque, 1818, White crappie	3.1 - 27.0	68	169	97	2	336
P. nigromaculatus (Lesueur, 1829), Black crappie	3.6 - 14.1	25	733	86		844
Perca flavescens (Mitchill, 1814), Yellow perch	1.9 - 17.7	183	1,230	451		1,864
Sander canadensis (Griffith & Smith, 1834), Sauger	1.4 - 24.8		108			108
S. vitreus (Mitchill, 1818), Walleye	6.8 - 31.4	145	5,733	204	20	6,102
S. vitreus × *S. canadensis*, Saugeye	10.9 - 21.3		9	2		11
Clupeidae						
Alosa chrysochloris (Rafinesque, 1820), Skipjack shad	14.3			1		1
A. mediocris (Mitchill, 1814), Hickory shad	11.0 - 15.9	3				3
A. pseudoharengus (Wilson, 1811), Alewife	3.6 - 11.2	18	1	3		22
A. sapidissima (Wilson, 1811), American shad	8.6 - 16.7	5				5
Brevoortia tyrannus (Latrobe, 1802), Atlantic menhaden	5.1 - 36.3	7	5			12
Dorosoma cepedianum (Lesueur 1818), American gizzard shad	2.0 - 16.1	97	33	34		164
D. petenense (Günther, 1867), Threadfin shad	2.9	1				1
Cottidae						
Cottus asper Richardson, 1836, Prickly sculpin	3.0 - 3.7	3				3
C. cognatus Richardson, 1836, Slimy sculpin	2.4 - 3.0	3				3
C. gulosus (Girard, 1854), Riffle sculpin	3.4 - 3.5	2				2
Leptocottus armatus Girard, 1854, Pacific staghorn sculpin	2.8	1				1
Triglopsis thompsonii Girard, 1851, Deepwater sculpin	2.5	1				1
Cyprinidae						
Acrocheilus alutaceus Agassiz & Pickering, 1855, Chiselmouth	3.4 - 3.6	3				3
Campostoma anomalum (Rafinesque, 1820), Central stoneroller	5.4 - 6.1	3				3
C. pauciradii Burr & Cashner, 1983, Bluefin stoneroller	5.1		1			1
Carassius auratus auratus (Linnaeus, 1758), Goldfish	6.0 - 10.9	6	6	1		13
Clinostomus funduloides Girard, 1856, Rosyside dace	3.0 - 3.3	2	1			3
Ctenopharyngodon idella (Valenciennes, 1844), Grass carp	20.9			1		1
Cyprinus carpio carpio Linnaeus, 1758, Common carp	2.6 - 37.8	340	1,003	927	4	2,274
Hesperoleucus symmetricus (Baird & Girard, 1854), California roach	3.6	1				1
Hybognathus placitus Girard, 1856, Plains minnow	1.0 - 2.5	2				2
Lavinia exilicauda Baird & Girard, 1854, Hitch	3.0 - 3.6	2				2
Nocomis leptocephalus (Girard, 1856), Bluehead chub	2.4 - 8.9	42	4			46
N. micropogon (Cope, 1865), River chub	5.6 - 5.8	2				2
Notemigonus crysoleucas (Mitchill, 1814), Golden shiner	3.3 - 9.7	22	2			24
Phoxinus erythrogaster (Rafinesque, 1820), Southern redbelly dace	5.6		1			1
Pimephales promelas Rafinesque, 1820, Fathead minnow	1.6 - 2.7	14				14
Ptychocheilus oregonensis (Richardson, 1836), Northern pikeminnow	3.8 - 15.6	1	1			2
Rhinichthys cataractae (Valenciennes, 1842), Longnose dace	6.9 - 9.3			5		5
R. osculus (Girard, 1856), Speckled dace	2.8 - 3.7	4				4
Semotilus atromaculatus (Mitchill, 1818), Creek chub	3.6 - 9.0	30				30
S. corporalis (Mitchill, 1817), Fallfish	5.9 - 6.3			2		2
S. lumbee Snelson & Suttkus, 1978, Sandhills chub	6.5 - 6.5	2				2
Embiotocidae						
Rhacochilus toxotes Agassiz, 1854, Rubberlip seaperch	8.2		1			1
Esocidae						
Esox americanus americanus Gmelin, 1789, Redfin pickerel	6.2 - 9.7	7	2			9
E. lucius Linnaeus 1758, Northern pike	1.0 - 44.1	249	4,582	590		5,421
E. masquinongy Mitchill, 1824, Muskellunge	18.2 - 47.9		49			49
E. masquinongy × *E. lucius*, Tiger Muskellunge	27.4 - 30.9			4		4
E. niger Lesueur, 1818, Chain pickerel	3.8 - 26.1	29	44	66		139

Species and common names	Length range (in.)	Whole fish	Sample counts by cut of fish			Total
			Skin-on fillet	Skin-off fillet	Other cuts	
	Fundulidae					
Fundulus parvipinnis Girard, 1854, California killifish	2.0 - 2.5	4				4
F. zebrinus Jordan & Gilbert, 1883, Plains killifish	1.5 - 1.5	2				2
	Gasterosteidae					
Gasterosteus aculeatus aculeatus Linnaeus, 1758, Three-spined stickleback	1.8 - 2.2	12				12
	Gobiidae					
Acanthogobius flavimanus (Temminck & Schlegel, 1845), Yellowfin goby	14.7	1				1
Gillichthys mirabilis Cooper, 1864, Longjaw mudsucker	2.5 - 3.2	6				6
	Hiodontidae					
Hiodon alosoides (Rafinesque, 1819), Goldeye	13.6			1		1
	Ictaluridae					
Ameiurus brunneus (Jordan, 1877), Snail bullhead	5.3 - 11.9	5	2			7
A. catus (Linnaeus, 1758), White Catfish	7.7 - 23.5	3		13		16
A. melas (Rafinesque, 1820), Black bullhead	5.0 - 15.0	9	34	94		137
A. natalis (Lesueur, 1819), Yellow bullhead	4.5 - 13.8	37	35	77		149
A. nebulosus (Lesueur, 1819), Brown bullhead	6.7 - 15.6	32	59	69		160
A. platycephalus (Girard, 1859), Flat bullhead	4.7 - 12.8	14	10			24
Ictalurus furcatus (Valenciennes, 1840), Blue catfish	10.8 - 40.0	2		86		88
I. punctatus (Rafinesque, 1818), Channel catfish	3.6 - 39.4	171	25	964	2	1,162
Noturus insignis (Richardson, 1836), Margined madtom	3.9 - 6.2	3				3
Pylodictis olivaris (Rafinesque, 1818), Flathead catfish	8.9 - 39.3	6		149		155
	Lepisosteidae					
Lepisosteus oculatus Winchell 1864, Spotted gar	19.4 - 28.2	1		8		9
L. osseus (Linnaeus, 1758), Longnose gar	19.3 - 43.5	61	10	2	2	75
L. platostomus Rafinesque, 1820, Shortnose gar	22.2 - 27.3	1		4		5
	Lotidae					
Lota lota (Linnaeus, 1758), Burbot	12.1 - 28.9		22	19		41
	Moronidae					
Morone americana (Gmelin, 1789), White perch	4.3 - 14.8	45	61	38	1	145
M. chrysops (Rafinesque, 1820), White bass	4.7 - 18.2	8	125	62		195
M. saxatilis (Walbaum, 1792), Striped Sea-bass	10.4 - 35.9	24	40	2	17	83
	Mugilidae					
Mugil cephalus Linnaeus, 1758, Flathead mullet	9.9 - 16.0	14	17			31
	Osmeridae					
Osmerus mordax mordax (Mitchill, 1814), Atlantic rainbow smelt	3.0 - 21.3	100	125			225
Spirinchus thaleichthys (Ayres, 1860), Longfin smelt	3.2		1			1
	Paralichthyidae					
Paralichthys dentatus (Linnaeus, 1766), Summer flounder	13.0 - 14.2	1	3			4
P. lethostigma Jordan & Gilbert, 1884, Southern flounder	7.1 - 16.2	2	18	2		22
	Percidae					
Etheostoma whipplii (Girard, 1859), Redfin darter	6.4 - 7.4		6			6
	Pleuronectidae					
Hippoglossoides platessoides (Fabricius, 1780), American plaice	7.0 - 12.0		4			4
	Poeciliidae					
Gambusia affinis (Baird & Girard, 1853), Mosquitofish	1.7 - 2.1	4				4
Poecilia latipinna (Lesueur, 1821), Sailfin molly	2.2 - 3.0	2				2
	Polyodontidae					
Polyodon spathula (Walbaum 1792), Paddlefish	32.0			1		1
	Pomatomidae					
Pomatomus saltatrix (Linnaeus, 1766), Bluefish	9.4 - 19.1	1	8			9
	Salmonidae					
Coregonus artedi Lesueur, 1818, Cisco	7.7 - 19.8	10	130	39		179
C. clupeaformis (Mitchill, 1818), Lake whitefish	10.2 - 26.4		224			224

Species and common names	Length range (in.)	Whole fish	Sample counts by cut of fish			Total
			Skin-on fillet	Skin-off fillet	Other cuts	
C. hoyi (Milner, 1874), Bloater	9.8 - 11.3		2			2
Oncorhynchus clarki clarki (Richardson, 1836), Cutthroat trout	7.1 - 12.0		12			12
O. clarki clarki (Richardson, 1836), Cutthroat trout (listed as Snake river cutthroat trout)	17.5 - 19.5		2			2
O. kisutch (Walbaum, 1792), Coho salmon	4.9 - 26.0		55			55
O. mykiss (Walbaum, 1792), Rainbow trout	2.7 - 31.1	12	248	12		272
O. nerka (Walbaum, 1792), Sockeye salmon (listed as Kokanee salmon)	7.3 - 18.1		21			21
O. nerka (Walbaum, 1792), Sockeye salmon (listed as Sockeye salmon)	10.4 - 15.0		37			37
O. tshawytscha (Walbaum, 1792), Chinook salmon	10.5 - 37.9		72	10		82
Salmo salar Linnaeus, 1758, Atlantic salmon (listed as Atlantic salmon)	6.9 - 21.3	5	3	7		15
S. salar Linnaeus, 1758, Atlantic salmon (listed as Ouananiche)	13.3 - 19.7	12	4	14		30
S. trutta trutta Linnaeus, 1758, Sea trout	5.2 - 28.4	19	259	16		294
Salvelinus confluentus (Suckley, 1859), Bull trout	17.7 - 20.9		3			3
S. fontinalis (Mitchill, 1814), Brook trout	4.3 - 19.9	32	130	52		214
S. namaycush (Walbaum, 1792), Lake trout	7.4 - 37.2	8	820	14		842
Sciaenidae						
Aplodinotus grunniens Rafinesque, 1819, Freshwater Drum	5.1 - 25.6	4	116	94		214
Bairdiella chrysoura (Lacepède, 1802), Silver croaker	7.5 - 7.7	2				2
Cynoscion regalis (Bloch & Schneider, 1801), Gray weakfish	12.3 - 17.3		6			6
Micropogonias undulatus (Linnaeus, 1766), Atlantic croaker	6.1 - 13.2	1	16			17
Pogonias cromis (Linnaeus, 1766), Black drum	18.9 - 21.9			2		2
Sciaenops ocellatus (Linnaeus, 1766), Red drum	12.4 - 26.9	4	4	3		11
Scombridae						
Scomberomorus maculatus (Mitchill, 1815), Spanish mackerel	11.8 - 15.6		4	1		5
Sparidae						
Archosargus probatocephalus (Walbaum, 1792), Sheepshead seabream	11.7 - 18.4			4		4
Lagodon rhomboides (Linnaeus, 1766), Pinfish	5.6 - 8.7	1	5			6
Synodontidae						
Synodus foetens (Linnaeus, 1766), Inshore lizardfish	12.3	1				1
Triakidae						
Mustelus canis (Mitchill, 1815), Dusky smooth-hound	24.0	1				1
Zoarcidae						
Zoarces americanus (Bloch & Schneider, 1801), Ocean pout	11.7 - 14.7		3			3
Species names as recorded in the modified NLFWA that could not be identified to a specific species						
Cusk	5.7 - 18.3	1		3		4
Papio (three species have this common name in the U.S.)	5.1 - 6.8		7			7
Siscowet trout	14.8 - 30.9	19	175			194
Splake trout	4.6 - 28.8	8	60			68
Spotted sea trout	11.4 - 19.7	1	6	7		14
Total		3,016	21,968	6,774	55	31,813