Beyond Thresholds: Using Logistic Regression Models to Estimate the Probability of Toxicity from Sediment Chemistry

Jay Field, NOAA, Seattle, WA
Sue Norton, EPA, Washington, D.C.
Don MacDonald, MacDonald Environmental Sciences Ltd., Ladysmith, BC
Corinne Severn, EVS Consultants, Seattle, WA
Chris Ingersoll, USGS, Columbia, MO

National Oceanic and Atmospheric Administration • National Ocean Service • Office of Response and Restoration

Abstracts

Contaminant-specific logistic regression models (LRMs) were used to estimate the percentage of samples expected to be toxic at various chemical concentrations in field-collected samples. LRMs enable users to select the probability of effects corresponding to specific assessment or management objectives. Users can also estimate the probability of observing specific biological effects at selected sediment contaminant concentrations. The models drew on a large database of matching saltwater sediment chemistry and toxicity data compiled from field-collected samples from different geographic areas affected by different contaminant sources. Logistic regression models were developed for marine amphipod 10-day survival test endpoints, combining data for Ampelisca abdita (n=2060) and Rhepoxynius abronius (n=1668). Models for 25 chemicals were evaluated for reliability using the regression goodness-of-fit statistic and their ability to correctly classify samples as toxic. The severity of the effect (decreased control-adjusted survival) increased along with the increase in the probability of toxicity. Statistical comparison of models derived from California and Puget Sound data for Rhepoxynius survival showed little regional difference in the individual models for most chemicals. The individual chemical models were combined by using either the predicted maximum or the mean probability for a sample. For both the maximum and the mean probability, the proportion of toxic samples in the entire (unscreened) database increased linearly with increasing probability.

Introduction

Sediment quality guidelines (SQGs) have been developed using a variety of empirical and theoretical approaches. These SQGs identify threshold concentrations of sediment contaminants that are associated with the presence or absence of sediment toxicity.
Such SQGs have proven very useful for a wide variety of applications and their ability to predict toxicity in independent data sets has been demonstrated (e.g., Long et al. 1998).
In contrast to the threshold approach to SQG development, the logistic regression modeling approach (Field et al. 1999) enables users to select the probability of effects that corresponds to their specific assessment or management objective. The approach also can be used to estimate the probability of observing specific biological effects at a selected sediment contaminant concentration.
This poster describes the methods that were used to develop the logistic regression models (LRMs) for a standard suite of contaminants of concern in marine and estuarine sediments and presents a couple of applications of the LRM approach.
The specific objectives of this poster are:

1) to evaluate logistic regression models for a standard suite of 25 chemicals, using a large database of matching sediment chemistry and toxicity data for the standard marine amphipod 10-day survival toxicity test with Ampelisca abdita and Rhepoxynius abronius.

2) to use the logistic modeling approach as a means of comparing the data for a specific toxicity test endpoint from two geographic regions.

3) to apply the combined models for 25 chemicals to assess the probability of toxicity in a large database of matching sediment chemistry and toxicity data.

Methods

Data Compilation

Matching sediment chemistry and toxicity data were compiled from a number of sources. Candidate data sets were evaluated according to minimum data quality requirements established for:

- Methods for collecting, handling, and storing sediments
- Methods for determining contaminant concentrations
- Methods for conducting toxicity tests
- Methods for interpreting toxicity test results

Classification of Samples as Toxic

Samples were designated as toxic using two approaches:

Significance Only approach

Statistical significance compared to negative control and < 90% survival (upper limit for response value based on acceptable negative control response)

Minimum significant difference (MSD) approach (after Thursby et al. 1997)

Statistical significance and < 80% control-adjusted survival

Data Screening

Rationale: To exclude toxic samples from the analysis where the contaminant of concern is not likely to be associated with the observed toxicity.

Data were screened prior to conducting the logistic regression analysis using the screening approach followed by Ingersoll et al. (1996) and similar to that used by others (Long and Morgan 1990; Long and MacDonald 1992).

Samples were classified into three categories for each selected contaminant:

1) Non-toxic samples

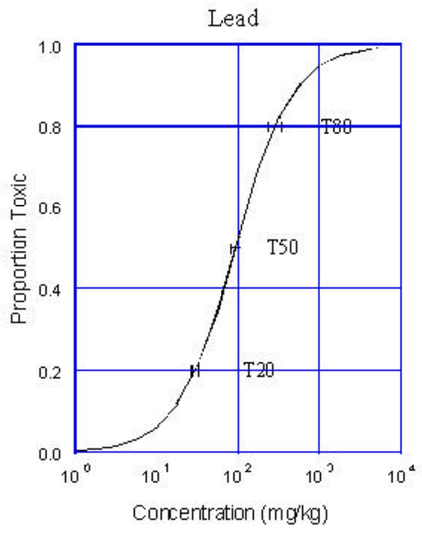
2) Toxic samples with concentration gradient (if contaminant concentration in the toxic sample was greater than the mean concentration in the non-toxic samples for that study/area)

3) Toxic samples with no concentration gradient

Data used to derive individual logistic regression models:
Include all non-toxic samples and toxic samples with a gradient for the selected contaminant.

Exclude toxic samples with no concentration gradient for the selected contaminant.

Logistic Regression Model



- Logistic model estimates the proportion of samples expected to be toxic at a given concentration
- Based on screened data for amphipod survival endpoints
- Logistic regression determines the slope, intercept, and chi-square statistic (-2 log likelihood)
- Chi-square statistics (normalized for sample size) provide a relative measure of the goodness-of-fit for the individual chemical models
- Point estimates (e.g., T20, T50, T80) represent the concentration at which 20, 50, or 80% of the samples would be predicted to be toxic

Concentration Interval Plots

- Concentration interval plots represent a summary of the individual sample data that are used to derive the logistic regression models
- Each point represents the median of a unique concentration interval and the proportion of toxic samples within the interval
- Each interval represents a minimum of 15 samples (may be more if there is more than one sample with the same concentration)

Results

Contents of Marine Database

Data Source	*Ampelisca*	*Rhepoxynius*
NSTP	681	NA
EMAP	1205	NA
MLML	57	768
SEDQUAL	NA	732
Miscellaneous	117	168
Total	2060	1668

NSTP: NOAA National Status and Trends Program

EMAP: EPA Environmental Monitoring and Assessment Program

MLML: Moss Landing Marine Laboratory (California data)

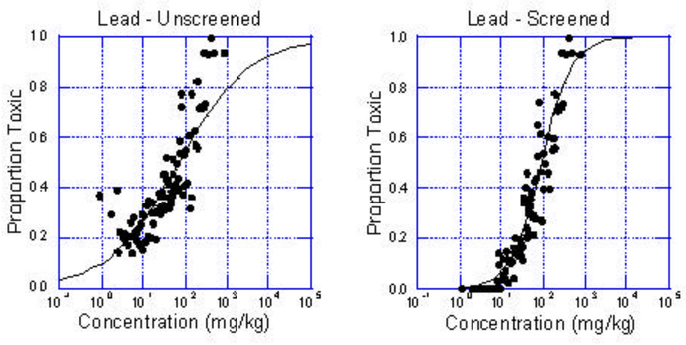
SEDQUAL: Puget Sound Database, State of Washington Dept. of Ecology

Total number of samples in the database broken down by primary data source and species for marine amphipod 10-day survival toxicity tests

Percent of toxic samples in combined amphipod database using two approaches to toxicity classification:

SIG only: 41.1%
MSD approach: 24.7%

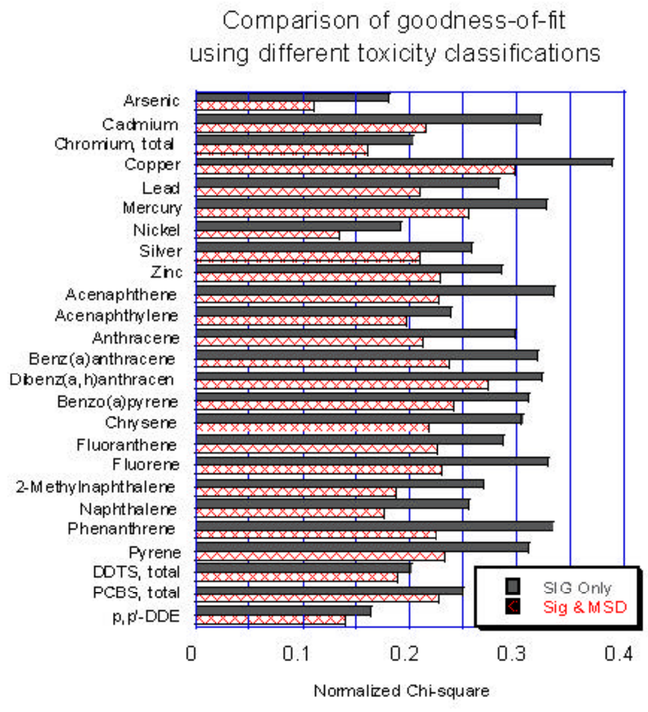
Effect of Data Screening



The plots show the logistic regression curves and the concentration interval plots for the unscreened (n=3131) and screened (n=2598) data for lead.

The plot of the unscreened data shows an elevated proportion of samples toxic at low lead concentrations.

For concentrations greater than 100 ppm, the two concentration interval plots appear similar.

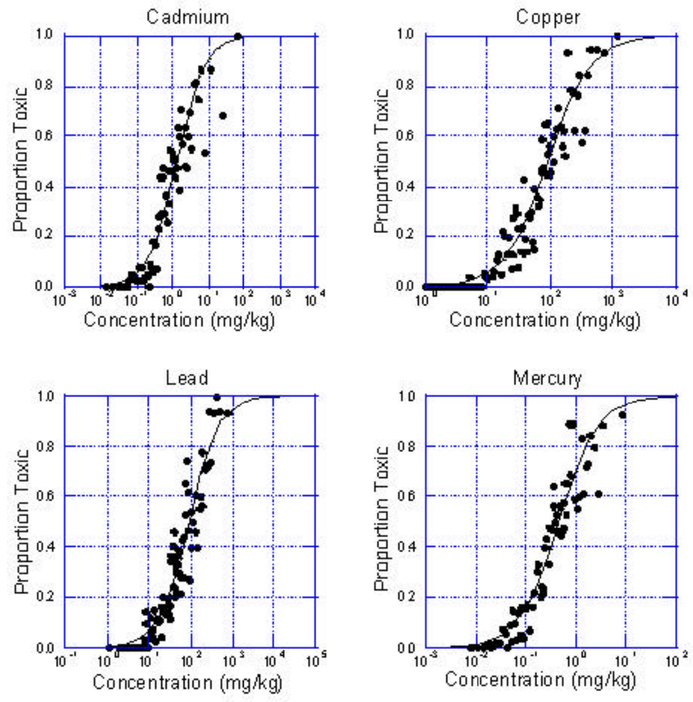


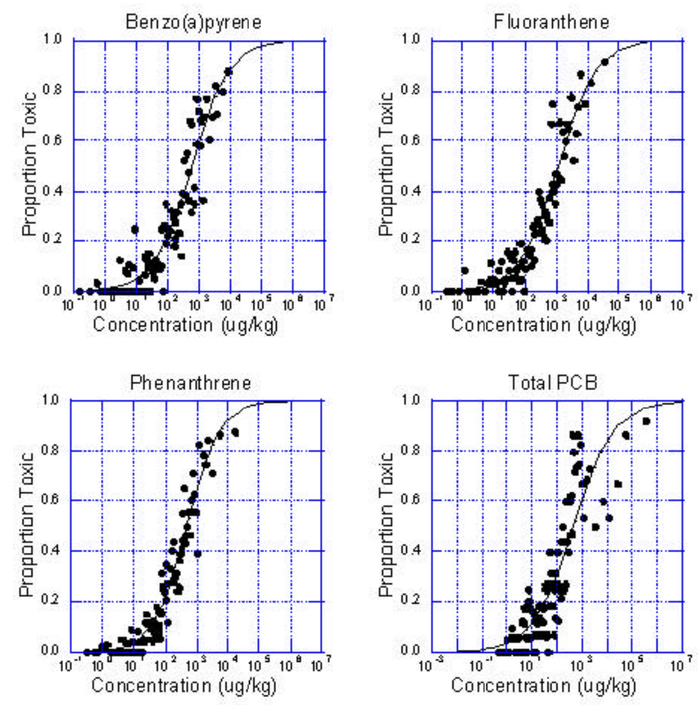
Goodness-of-fit for Individual Chemical Models

- Goodness-of-fit for individual chemical models was evaluated by chi-square values normalized by sample size
- The figure shows the normalized chi-square values for individual chemical models based on two different toxicity classifications: statistical significance only (SIG only) and significance plus minimum significant difference (SIG and MSD).
* Models with lower normalized chi-square values (poorer model fit), such as those for arsenic, nickel, and pp-DDE, are considered less reliable than other models
- Two different approaches to toxicity classification both resulted in quality models, but the goodness-of-fit was consistently better for the “Significance only” approach.

Individual Chemical Models

- Logistic regression models for individual chemicals were derived using the combined amphipod database.
- Examples of the models are shown for eight chemicals, four metals and four organics. The plots show the logistic regression model curve and points representing concentration interval summary plots of the screened data set. Note that the models were derived from the individual sample data, not the summarized data shown in the interval plots.
- The models shown were developed from the screened database, with toxicity determined by statistical significance only.





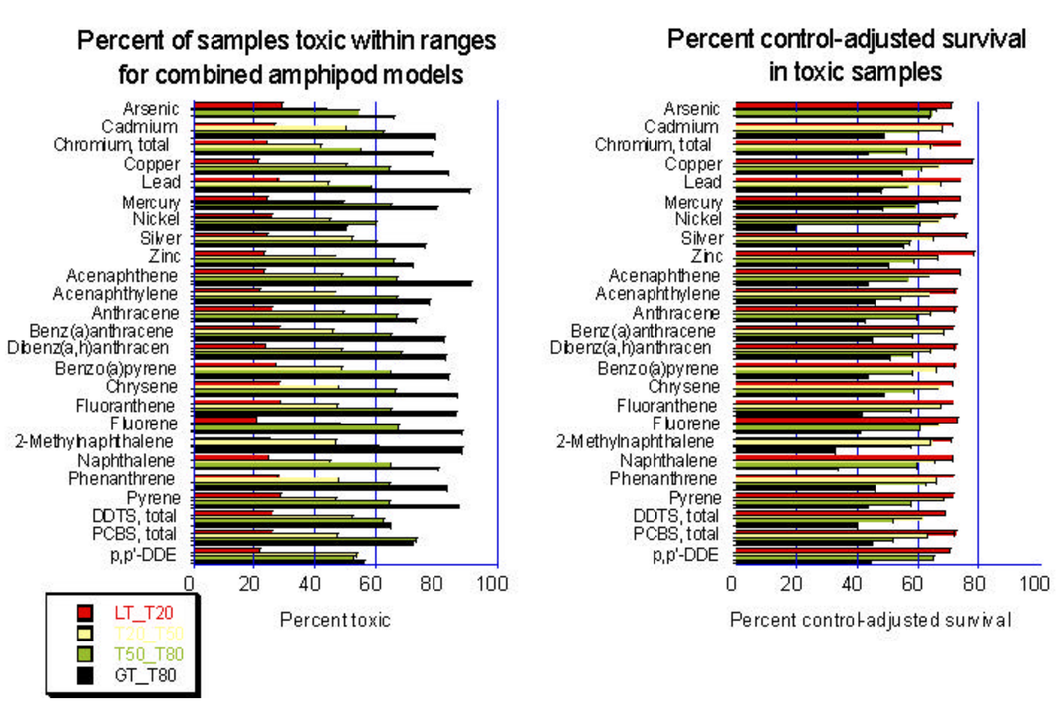
Effect of Different Approaches to Toxicity Classification

The T50 values for models derived from data using the "Significance only" approach were 2 - 3 times lower than the corresponding value for the MSD approach.

Differences were generally greater at low concentrations (e.g., T20 values) and less at high concentrations (e.g., T80 values)

Reliability of Individual Models

- Reliability of individual chemical models was evaluated by determining the percent of samples toxic and average percent survival within four concentration ranges defined by three T-values (T20, T50, and T80) using the entire unscreened database.
- For the combined amphipod models based on statistical significance only, in most cases the percent of toxic samples within the defined concentration ranges increases with increasing concentration and generally corresponds to the predicted probability of toxicity.
- The severity of the effect, as measured by the mean control-adjusted survival for toxic samples, also increases with increasing probability of effects. Percent survival in toxic samples decreases from about 70-75% for samples with concentrations less than the T20 value to 40-50% for samples with concentrations greater than the T80 value.



Comparison of Geographic Areas

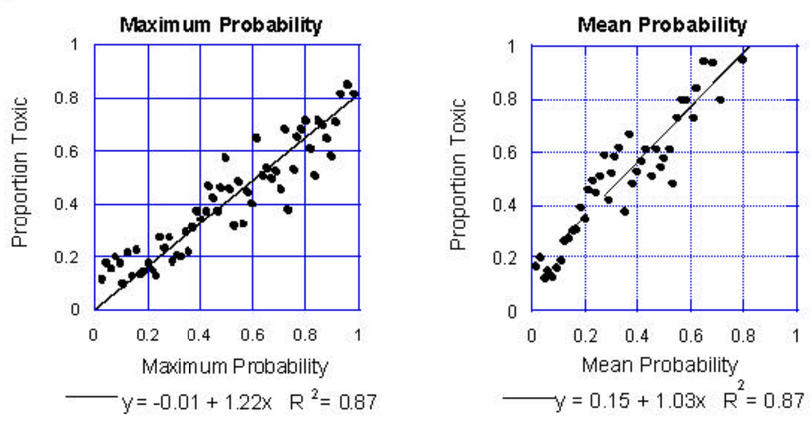
	Preferred Model		
Chemical Name	Combined	Parallel	Different
Arsenic			**
Cadmium	**		
Chromium, total		**	
Copper	**		
Lead	**		
Mercury	**		
Nickel	**		
Silver	**		
Zinc			**
Acenaphthene	**		
Acenaphthylene	**		
Anthracene	**		
Benz(a)anthracene	**		
Dibenz(a,h)anthracene	**		
Benzo(a)pyrene			**
Chrysene	**		
Fluoranthene	**		
Fluorene	**		
2-Methylnaphthalene		**	
Naphthalene		**	
Phenanthrene	**		
Pyrene	**		
DDTS, total of 6 isomers	**		
PCBS, total	**		
p,p'-DDE		**	

Statistical comparison of the individual chemical models for R. abronius 10-day survival tests from two geographic areas: Puget Sound (n=732) and California (n=768).

The results of this analysis show no statistically significant difference (at ~= 0.05) between the models from the two regions for 18 of the 25 chemicals. For 4 of the remaining 7 chemicals, parallel models (common slope, different intercept) were the preferred choice.

Results suggest that the combined models for the two geographic areas can be applied in either area.

Combined Models



- The results of the 25 individual chemical models were combined by using either the maximum or the mean probability for a sample.
- The individual chemical models were derived from the combined marine database of 10-day survival toxicity tests for *A. abdita and R. abronius*.
- Samples were designated as toxic based on statistical significance compared to a control and less than 90% survival.
- All samples in the unscreened database with at least 10 chemicals measured were included in the analysis (n=2950).
- Data points represent the median of discrete probability intervals and the proportion of samples toxic within each interval. Each interval contains at least 15 samples.
- For both the maximum and the mean probability, the proportion of toxic samples shows a linear increase with increasing probability.
- These results indicate that the combined individual models can be used to estimate the probability of amphipod toxicity from sediment chemistry for this database.

Summary

- Large database (n>3500) used to develop individual chemical logistic regression models for marine amphipod 10-day survival endpoint, with samples from all over coastal North America.
- Screening approach appears to reduce contribution of toxic samples where the selected chemical is not likely a factor in observed toxicity.
- Individual logistic regression models for most chemicals showed high goodness-of-fit to screened data.
- The relative goodness-of-fit was higher for models where toxicity was based on statistical significance only as compared to the MSD approach.
- Individual models demonstrated high reliability based on proportion of samples toxic in the unscreened data set over defined ranges.
- As the probability of toxicity increases, the severity of the effect (decreased survival) also increases.
- Two different approaches to toxicity classification both resulted in quality models, but the T50 values from the “MSD” approach were 2-3 times greater than the values for “significance only.” The differences were greater at lower concentrations and less at higher concentrations.
- Statistical comparison of models derived from California and Puget Sound data for *Rhepoxynius* survival showed that there was either no regional difference in the individual models for most chemicals (18 of 25) or different intercepts only (parallel models) for 4 of the remaining 7.
- The individual chemical models can be combined by using either the predicted maximum or the mean probability for a sample. For both the maximum and the mean probability, the proportion of samples toxic in the entire (unscreened) database increased linearly with increasing probability.

Acknowledgment

Most of the funding for this project was provided by U.S. Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment through an Interagency Agreement with NOAA’s Office of Response and Restoration, Coastal Protection and Restoration Division.

References

Field LJ, MacDonald DD, Norton SB, Severn CG, Ingersoll CG. 1999. Evaluating sediment chemistry and toxicity data using logistic regression modeling. *Environ. Toxicol. Chem.* 18:1311-1322.

Ingersoll CG, Haverland PS, Brunson EL, Canfield TJ, Dwyer FJ, Henke CE, Kemble NE. 1996. Calculation and evaluation of sediment effect concentrations for the amphipod *Hyalella azteca* and the midge *Chironomus riparius.* J Gt Lakes Res 22:602-623.

Long ER, Morgan LG. 1990. The Potential for Biological Effects of Sediment-sorbed Contaminants Tested in the National Status and Trends Program. NOAA Tech Memo NOS OMA 52. National Oceanic and Atmospheric Administration, Seattle, WA, USA.

Long ER, MacDonald DD. 1992. National Status and Trends Program approach. In *Sediment Classification Methods Compendium.* EPA 823-R-92-006. U.S. Environmental Protection Agency, Washington DC, USA.

Long ER, Field LJ, MacDonald DD. 1998. Predicting toxicity in marine sediments with numerical sediment quality guidelines. *Environ Toxicol Chem* 17:714-727.

Thursby GB, Heltshe J, Scott KJ. 1997. Revised approach to toxicity test acceptability criteria using a statistical performance assessment. *Environ Toxicol Chem* 16:1322-1329.

Copies of this poster can be obtained from:
http://response.restoration.noaa.gov/cpr/sediment/sediment.html
or by sending an email request to jay.field@noaa.gov