

Development and Evaluation of a Regression-Based Model to Predict Cesium Concentration Ratios for Freshwater Fish

John E. Pinder III

Department of Environmental and Radiological Health Sciences
Colorado State University
305 W. Magnolia PMB 231
Fort Collins, Colorado 80521, USA

David J. Rowan

Atomic Energy of Canada Limited (AECL)
Environmental technologies Branch
Chalk River Laboratories
Chalk River Ontario K0J 1J0, Canada

Joseph B. Rasmussen

Department of Biological Sciences
University of Lethbridge
Lethbridge, Alberta T1K 3M4, Canada

Jim T. Smith

School of Earth and Environmental Science
Science University of Portsmouth
Portsmouth P01 3 QL, UK

Thomas G. Hinton

Star Network of Excellence in Radioecology
Institute for Radiation Protection and Nuclear Safety
Saint-Paul-lez-Durance 13115, France

and

F. Ward Whicker

Department of Environmental and Radiological Health Sciences
Colorado State University
Fort Collins, Colorado 80523-1618, USA

ABSTRACT: Data from published studies and World Wide Web sources were combined to produce and test a regression model to predict Cs concentration ratios for freshwater fish species. The accuracies of predicted concentration ratios which were computed using 1) estimated species trophic levels obtained from random resampling of known food items and 2) K concentrations in the water for 207 fish from 44 species and 43 locations were tested against independent observations of ratios for 57 fish from 17 species from 25 locations. Accuracy was assessed as the percent of observed to predicted ratios within factors of 2 or 3. Conservatism, expressed as the lack of under prediction, was assessed as the percent of observed to predicted ratios that were less than 2 or less than 3. The model's median observed to predicted ratio was 1.26, which was not significantly different from 1, and 50 % of the ratios were between 0.73 and 1.85. The percentages of ratios within factors of 2 or 3 were 67 and 82 %, respectively. The percentages of ratios that were < 2 or < 3 were 79 and 88 %, respectively. An example for *Perca fluviatilis* demonstrated that increased prediction accuracy could be obtained when more detailed knowledge of diet was available to estimate trophic level. The model allows reasonably accurate, species-specific predictions requiring only 1) measures of the K concentration in the water and 2) readily available estimates of trophic level from the FishBase Global Information System (fishbase.org).

1. Introduction

An important parameter used to assess the mobility of the radionuclides ^{134}Cs and ^{137}Cs in freshwater environments and to estimate the potential risks of consuming fish from these systems is the ratio of the mean concentration of the radionuclide in the fish to its mean concentration in the water. This ratio is alternatively termed the concentration ratio or the bioaccumulation factor and has units of L kg^{-1} . Efforts to compile or predict estimates of concentration ratios (hereafter, C_r) that can be readily employed in accident assessments have led to 1) compilations of previously observed C_r (Vanderploeg et al., 1975; Blaylock, 1982; Hosseini et al., 2008; Fesenko et al., 2010; Yankovich et al., 2012) and 2) efforts to develop predictive models based on aspects of fish biology, such as diet, and water quality parameters such as K concentrations. Rowan and Rasmussen (1994) developed a predictive model (Rowan and Rasmussen, 1994; equation 5) applicable to both freshwater and marine systems based on whether fish were piscivorous or non-piscivorous, and measures of K and suspended sediment concentrations in the water column. The model predicts 1) greater C_r for piscivorous fish, 2) smaller C_r in waters with greater K concentrations and 3) smaller C_r in waters with greater suspended sediment concentrations. The model has been shown to predict C_r within a factor of 2 for a majority of cases (Smith et al., 2000).

Although accurate, the model contains two limiting aspects. One limitation is that in its current form it requires input data for both K concentrations and suspended sediment concentrations. The other limitation is that the model requires a judgment of whether a fish species is piscivorous or nonpiscivorous. Although some fish, at least as adults, are clearly piscivorous (*e.g.*, the northern pike *Esox lucius*; nomenclature for common and scientific names follows *fishbase.org*), others are clearly non-piscivorous (*e.g.*, the white sucker *Catostomus commersoni*). Many species show varying proportions of fish and other food items in their diet, and even among nonpiscivorous fish there are

differences in trophic levels and potential concentrations of Cs in their diets. These differences occur between fish that are primarily herbivorous and those that are carnivorous but feed on invertebrates.

Recently, an extensive online database on the biology and ecology of freshwater and marine fish species (*fishbase.org*) has been developed and made freely available by the FishBase Global Information System (Froese and Pauly, 2011). It includes information on fish diets and food items from numerous references, and uses this information to compute estimates of mean (\pm Standard Error, hereafter SE) trophic levels. Fish trophic levels range from 2, indicating an herbivorous diet, through 3, indicating a primarily carnivorous diet of herbivorous species, and through 4 indicating a primarily carnivorous diet composed of other carnivorous species. For example, mean trophic levels (\pm SE) computed using a random resampling of reported food items for *C. commersoni* and *E. lucius* are 2.46 ± 0.16 and 4.40 ± 1.05 , respectively.

The purposes of this study were 1) to combine these newly available estimates of trophic levels with the data compiled by Rowan and Rasmussen (1994) to produce an alternative predictive model for the concentration ratios of Cs isotopes in freshwater fish and 2) to test the accuracy of this model's predicted C_r using independent data from several sources.

2. Materials and methods

Three data sources were employed in this study to produce and test models to predict C_r for fish. These were: 1) the trophic level estimates from *fishbase.org*; 2) the data from Rowan and Rasmussen (1994) on fish C_r and the concentrations of K and suspended sediment in the water column that were used in model construction; and 3) test data of measured C_r compiled from several literature sources that were used to assess the accuracy of the predicted C_r .

2.1. Trophic level data

The trophic level estimates were obtained from *fishbase.org* where data on fish biology and ecology have been compiled for > 30,000 species from > 45,000 references and made available in various Asian and European languages. Species data may be accessed by searching either common or scientific names. Where taxonomic revisions have been made to scientific names, a search initiated using the previous name leads to the modern synonym and the appropriate data.

Several alternative methods are used in *fishbase.org* to estimate a mean \pm SE trophic level, and the method employed depends on whether the available data from literature sources comprises 1) just lists of food items consumed or 2) more quantitative analyses of diet composition. Where both types of data are available, multiple estimates of trophic level may be made. When using lists of food items consumed, trophic level is estimated using a randomized resampling of the listed items. Where more quantitative data on diet consumption is available, the trophic level is computed from the relative proportion of the food items consumed in a process similar to that used by Vander Zander and Rasmussen (1996) to compute trophic level effects on PCB contamination in fish. These different procedures may produce different estimates of trophic level. For example, the random resampling of 31 food items for *Perca fluviatilis* (European perch) indicated a mean (\pm SE) trophic level of 3.66 ± 0.58 , but an analysis of diet composition for adult fish indicated a mean trophic level of 4.35 ± 0.75 . For all methods, the trophic level of the fish is computed as 1 plus the mean trophic level computed for its diet (*i.e.*, a fish whose diet has a mean trophic level of 2.5 would have a trophic level of 3.5). These estimation methods have been shown to agree with those computed from stable isotopic ratios (Kline and Pauly, 1988). Where data are lacking on diets, a fish's trophic level is inferred from a taxonomically related species of similar size.

Because trophic level estimates computed using the random resampling procedures were available for the majority of the species involved in this study, they have been used in the models to predict C_r . However, the potential effects of using the alternative estimates, such as those for *P. fluviatilis*, will also be discussed. These random resampling trophic levels are computed from a list of n food items, each with its own assigned trophic level, by: 1) randomly choosing one item and assigning it the largest fraction of the diet; 2) randomly selecting each of the remaining $n - 1$ items and assigning each a successively smaller fraction of the diet; 3) computing a trophic level estimate for the diet from the sum of the product of each item's trophic level and assigned proportion in the diet; 4) repeating this random selection process for a total of 100 times; and 5) computing mean and standard error from the 100 replicates.

2.2. The modeling data

The development of the predictive model was based on the data for freshwater fishes compiled by Rowan and Rasmussen (1994; Table 1) in conjunction with the newer estimates of trophic levels from the FishBase database. The Rowan and Rasmussen (1994) data included the fish's scientific name and the location of the study, the wet mass Cs concentration in the fish's whole body or muscle, the Cs concentration in the water, the K concentration in the water ($\mu\text{M K L}^{-1}$) and, for some locations, the suspended sediment load (mg L^{-1}) (Table 1). Cesium concentrations were alternatively expressed as mass for ^{133}Cs and Bq for ^{137}Cs , and no distinction was made between the C_r for ^{133}Cs and ^{137}Cs . These data contained no distinction between concentrations for whole fish and those for only muscle tissue and, as a consequence, no distinction with regard to muscle vs. whole-body made in the subsequent use of the model to predict C_r in fish.

Rowan and Rasmussen's (1994) study contained data for 244 observations on more than 60 species of fish from 46 locations in freshwater systems in Europe and North America, but two sets of their data were excluded from the modeling analysis in this study. First, those species which occurred in only a single location were excluded. Second, all data from locations on the U. S. Department of Energy's Savannah River Site (hereafter, SRS) were also excluded. Studies at the SRS (Newman and Brisbin, 1990; Whicker et al. 1990; Pinder et al. 2009) have consistently reported relatively small K concentrations in the water and relatively large C_r for fish species. These data were removed from the modeling analysis and included in the test data to provide tests of model accuracy at the upper extremes of reported C_s concentration ratios. The resulting modeling data set contained 207 measures of C_r and K concentrations in the water from 44 species across 43 locations. A subset of these data also included 84 measures of C_r and suspended sediment concentrations for 22 species across 14 locations.

2.3. Development of the predictive models

Predictive models were developed using simple and multiple regression procedures (Draper and Smith, 1981) that related concentration ratios to the variables 1) trophic levels, 2) K concentrations and 3) suspended sediment concentrations, as well as all possible interactions of these variables. Regression procedures may be used to 1) determine which independent variables from a set of preferably uncorrelated variables have important relationships with the dependent variable (*e.g.*, Rowan and Rasmussen, 1994), 2) to estimate parameter values such as uptake and loss rates (*e.g.*, Smith et al., 2002) or, 3) as in the case of this study, to simply construct a predictive model that relates the values of some predictor variables to the value of an important criterion variable (*e.g.*, Doi et al., 2012).

Regression models which attempt to determine important relationships between independent variables and the dependent variable may be termed explanatory models (Pedhazur, 1997), and these

differ from predictive regression models in that 1) the independent variables in explanatory models are hypothesized to have a causative relationship between their values and the values of the dependent variable and 2) regression is used to identify which of these possible causative agents has measurable effects on the values of the dependent variable. In a predictive model, the predictor variables (*i.e.*, the independent variables) are either readily available or inexpensive to obtain, whereas the values of the criterion variable (*i.e.*, the dependent variable) are often difficult or expensive to obtain. Because prediction, and not explanation, is the only objective in this modeling, the predictor variables need not have a causative effect on the criterion variable. They only need to be correlated with the criterion variable. Such correlations may occur because of a pronounced correlation of the predictor variable with a more difficult to measure or expensive to obtain causative agent. Having some form of causative effect may be preferable, but it is not necessary. The validity of a predictive model is not in the reasonableness of its independent variables but in its ability to accurately predict the criterion variable in an independent data set (Pedhazur, 1997).

The regression models in this study used the logarithm of the concentration ratio as the criterion variable, but the predictive equations will be presented as

$$\text{Predicted } C_r = 10^{[(b_0 + 0.5 \cdot \text{EMS}) + b_i \cdot X_i + \dots]} \quad (\text{Eq. 1})$$

where b_0 = the intercept of the regression equation, b_i = the regression coefficient for the *i*th predictor variable and EMS = the Error Mean Square of the regression Analysis of Variance. The value, $0.5 \cdot \text{EMS}$, was added to b_0 to compensate for a bias after logarithmic transformation that would result in under prediction of the C_r , especially at larger values of C_r (Beauchamp and Olson, 1973).

2.4. The test data

Test data were restricted to the results of published studies where there were contemporaneous measures of 1) either ^{133}Cs or ^{137}Cs concentrations in both fish and water, 2) ^{137}Cs concentrations in water, and where available, 3) concentrations of suspended sediment in the water. Data were also limited to those studies where either 1) it was reasonable to believe that the reported C_r were from a system at or near steady-state or equilibrium conditions between the fish and the water or 2) a steady state C_r could be estimated from uptake and loss rate parameters estimated for the fish. To be consistent with the modeling data, all concentrations and C_r are expressed as units of wet mass without corrections between concentrations measured using whole fish or only muscle tissue.

There were four sets of test data. The first was the data of Smith et al. (2000) which had previously indicated the accuracy of the Rowan and Rasmussen (1994) model. These data consisted of C_r for ^{137}Cs measured for 28 instances of 6 species from 10 eastern European lakes and reservoirs and were obtained ≥ 5 years after the Chernobyl release. These were the only test data that contained measures of suspended sediment concentrations. The second set was the temporal analysis of ^{137}Cs for 19 instances of 3 species from 9 European lakes following the deposition of Chernobyl releases (Smith et al., 2002). These temporal analyses estimated an uptake (k_1) and a loss rate parameter (k_2) for ^{137}Cs in fish, and the ratio of these parameters was used to estimate the steady state C_r (Smith et al. 2002; equation 4). The third set was data from the SRS which included C_r measured for ^{137}Cs in 7 species from Pond B in the 1980s, which was 20 years since the last releases of ^{137}Cs into the reservoir (Newman and Brisbin, 1990; Whicker et al., 1990). The fourth set involved C_r estimated from uptake and loss rates in an analogous manner to that of Smith et al. (2002) for 3 fish species following the experimental release of ^{133}Cs into Pond 4 (Pinder et al. 2009).

2.5. Assessing the predictive ability of the regression models using observed to predicted ratios

The predictive ability of a regression model may be assessed as 1) the coefficient of determination, R^2 , which is the proportion of the variation in the criterion variable explained by the predictor variables or 2) the standard error of the estimate, which is the minimum standard error for a prediction (Draper and Smith, 1981; Pedhazur, 1997). These measures of predictive ability are increased by the number of relatively uncorrelated predictive variables, but may be decreased by the use of highly correlated predictor variables which may be providing redundant information at the expense of decreased degrees of freedom in the model (Draper and Smith, 1981; Pedhazur, 1997; Ramsey and Schafer, 2002). Redundancy in predictor variables is expressed as increased values for the Akaike Information Criteria (hereafter, AIC; Ramsey and Schafer, 2002) and the Bayes Information Criteria (hereafter, BIC; Ramsey and Schafer, 2002).

Alternatively, predictive ability may be assessed using statistical summaries of the ratios of observed values of the criterion variable to those predicted by the regression. Such summaries may include the median ratio, the variance of the ratios, the range of the ratios, or the frequency distribution of the ratios. This observed to predicted ratio approach may be more informative where errors of under prediction (*i.e.*, observed values greater than predicted values and ratios > 1) and over prediction (*i.e.*, observed values less than predicted values and ratios < 1) are not equally acceptable. This is the case for predicting concentration ratios which may subsequently be used to estimate radiation exposures. In developing and initially evaluating the models in this study, statistical significance levels and R^2 will be reported as well as the ratios of the observed values in the modeling data to the values fitted by the regression models which are hereafter referred to as observed to fitted ratios. However, in evaluating a model's ability to predict C_r in the test data sets, the ratios of observed to predicted values will be the principal method of evaluating model performance.

In using observed to predicted ratios in evaluating models, two separate aspects of evaluation will be considered. The first is termed accuracy and is measured as the proportion of the ratios that indicate the predicted C_r is within some limits of the observed C_r such as within a factor of 2 (*i.e.*, ratios from 0.5 to 2) or 3 (*i.e.*, ratios of 0.33 to 3). Because under prediction of the C_r , with resulting observed to predicted ratios greater than 1, may incur greater consequences than over prediction, the second aspect of the evaluation is termed conservatism and is expressed as the proportion of the observed to fitted or the observed to predicted ratios that are less than some upper limit such as 2 or 3. The major difference between these separate evaluations concerns those ratios < 0.5 and < 0.33 . Although these small ratios are inaccurate, the over prediction of the C_r is likely to be conservative of public health. Because of the inclusion of these small ratios, the estimates of conservatism will include a greater fraction of the ratios than the corresponding estimates of accuracy.

What constitutes acceptable standards of accuracy or conservatism is a decision for those employing the model, but for the purposes of discussion within this analysis, an acceptable level of accuracy will be defined as 80 % of the observed to predicted ratios within factors of 2 or 3 (*i.e.*, one prediction out of 5 outside of these factors). An acceptable level of conservatism will be defined as 90 % of the ratios < 2 or < 3 (*i.e.*, one prediction out of 10 being more than a factor of 2 or 3 lower than the observed value).

3. Results

The means, standard deviations and ranges of the predictor variables in the modeling data are summarized in Table 1. The number of observations per variable varies because there were 44 species distributed across 43 locations where K concentrations were measured but suspended sediment concentrations were only available for 14 locations.

3.1. The distribution of concentration ratios in the modeling data

Concentration ratios ranged over a factor of 100 (Fig. 1) in the modeling data from 99 and 122 L kg⁻¹ for *Blicca bjoerkna* (white bream) and *Abrami brama* (freshwater bream) to 15,156 and 15,250 L kg⁻¹ for *Esox lucius* and *Sander vitreus* (walleye). None of these species had the smallest or largest estimated trophic levels, and none of these extremes came from those locations with the smallest or largest K concentrations. The median C_r was 2000 L kg⁻¹, and 50 % of the C_r ranged from 875 to 3,800 L kg⁻¹. The frequency distribution of the concentration ratios was positively-skewed and approximated a lognormal distribution (Kolmogorov –Smirnov Test of Normality, D = 0.0581; P > 0.05; Conover, 1971).

Variance component analyses indicated that only 28 % of the variation in C_r was due to difference among species (Searle, 1959). This indicates that variation in C_r among species is smaller than the variation of C_r within a species and suggests that the direct application of a C_r observed for a species in one location to that species in another location without consideration of the possible differences in physical and chemical parameters between locations has questionable validity.

3.2. Distribution of trophic levels among species in the modeling data

The mean (\pm SE) trophic levels computed using resampling of food items (Fig. 2) ranged from 2.40 ± 0.21 for *Dorosoma cepedianum* (American gizzard shad) to 4.42 ± 1.06 for *M. salmoides* (largemouth bass) with most species having trophic levels between 3 and 4. The random resampling procedure results in relatively large standard errors for the trophic level estimates. The median ratio of standard error to estimate was 0.15 but increased with increasing estimates to 0.24.

There was not a 1:1 correspondence of the *fishbase.org* trophic levels with the piscivorous-nonpiscivorous classification of Rowan and Rasmussen (1994). Most of the piscivorous species from their study had trophic levels > 3.4, whereas most of the nonpiscivorous species had trophic levels < 3.4.

The variability in trophic levels was related, in part, to phylogenetic groupings. The trophic level data included species from 16 families and 33 genera. For the seven families represented by more than one genus, variance component analyses using nested-analysis of variance procedures (Searle, 1987) indicated that 48% of the variation in trophic levels was due to variation among families with 38 % due to variation among genera within families. The remaining 14 % was due to variation among species within genera.

3.3. Correlation among predictor and criterion variables in the modeling data

Spearman rank correlations (Conover, 1971) indicated statistically significant correlations of C_r with trophic levels, K concentrations and suspended sediment concentrations (Table 2). Spearman rank correlations were used rather than the more familiar Pearson product-moment correlations because 1) the latter procedure assumes bivariate normal distributions of the data and 2) may produce inflated estimates of the true correlation where the data, like those for the C_r , are positively skewed (Sokal and Rohlf, 1969).

Concentration ratios increased with increasing trophic level and decreased with increasing K and sediment concentrations. Correlations of trophic levels with either K or sediment concentrations were not statistically significant. The largest absolute magnitude correlation with C_r was the negative correlation of -0.567 for K, and K concentrations have previously been shown to have negative correlations with C_r for Cs (Rowan and Rasmussen, 1994; Smith et al., 2000; NCRP, 2007). An important experimental demonstration of the effect of K on C_r ratios for Cs in fish was the reduction of fish C_r following the addition of potassium chloride to a Chernobyl-contaminated lake (Smith et al., 2003). The largest correlation coefficient was a positive correlation between K and suspended sediment concentrations which suggests that these measures were providing similar information about the

location's water quality. The smallest absolute magnitude correlation coefficient between the predictor variables and fish C_r occurred for trophic level, and this smallest correlation was consistent with the variance component results that indicated differences among species were less important in affecting C_r than variation within species. Inspection of plots of C_r with the predictor variables indicated that the logarithm of C_r had a linear relationship with trophic level, but the logarithm of C_r had linear relationships with the logarithms of K and sediment concentrations.

3.4. The development of a predictive model

Initial model development involved using the predictor variables of trophic level and $\log_{10}(K)$ by themselves and all their possible combinations including interactions. These models were evaluated and compared using standard regression diagnostics including largest R^2 , largest adjusted R^2 (Draper and Smith, 1981), smallest error mean square, smallest AIC, and smallest BIC. All of these models were statistically significant (*i.e.*, $P \leq 0.05$), and two models, hereafter referred to as Model I and Model II (Table 3), were judged to have the best combinations of accuracy and simplicity expressed as the lack of redundant predictors as measured by AIC and BIC. Model I used the predictor variables 1) $\log_{10}(K)$ and 2) the interaction of trophic level and $\log_{10}(K)$. Model II used the predictor variables 1) trophic level and 2) $\log_{10}(K)$. The two models produced similar median observed to fitted ratios (Table 3) with a median ratio of the predicted C_r from Model I to the predicted C_r from Model II of 1.0 with 75 % of the predictions being within 10 % of one another.

Although Model I had marginally preferable statistical indices and fits to the data, Model II is more directly interpretable with clearly demonstrated increases in predicted C_r with increasing trophic level and decreases in predicted C_r with increasing K concentrations. Because the two models had

similar predicted C_r and Model II was more directly interpretable, Model II was used in all subsequent predictions of C_r .

Model II has the form:

$$C_r = 10^{(3.3348 + 0.2453*TI - 0.5713 * \log_{10}(K))} \quad (\text{Eq. 2})$$

where TI represents trophic level. The standard errors for the regression coefficients for the intercept, TI and $\log_{10}(K)$ were 0.218, 0.246, and 0.006, respectively.

The close correspondence of predicted values from the two models indicates that the interaction term of trophic level and K concentration in Model II provides little improvement over Model I, and this suggests that most fish were responding similarly to increasing K concentrations. A more pronounced difference between the models would have suggested greater differences among species in how K concentrations in the water affected their C_s concentrations. Similar evidence of fairly uniform responses to K concentrations among fish species included the similar declines in ^{137}Cs concentrations in seven species of fish in Lake Svyatoye in response to the addition of potassium chloride to the lake's waters (Smith et al.2003). The lack of pronounced interaction effects also suggests that Model II can be extended to predict C_r for species other than those in the modeling data.

The relatively large standard errors for the random resampling estimates of trophic level have the potential to produce a biased estimate of its regression coefficient and, consequently, the predicted C_r (Draper and Smith, 1981). This potential bias was estimated using Wald's (1940) test which involves smoothing the estimated regression coefficient by computing it from the means of the lower and upper thirds of the data. This test indicated the potential bias was approximately 2 %.

Attempts to develop models using suspended sediment concentrations did not produce results with greater accuracies than the above models due to: 1) the few locations (*i.e.*, 14) with suspended sediment concentrations; and 2) the strong correlation between K and suspended sediment concentrations (Table 2). Although suspended sediment based models were not developed, the accuracy of Model II in predicting C_r will be compared to that for the Rowan and Rasmussen (1994) model for those test locations with sediment concentrations (Smith et al., 2000).

3.5. Observed to fitted ratios for Model II

The median observed to fitted ratio for Model II was 0.965 with ratios ranging from 0.0632 to 5.676. The first and third quartiles for the ratios were 0.494 and 1.472, respectively. Accuracy, expressed as the percent of ratios within factors of 2 and 3 were 61 and 82 %, respectively. Conservatism, expressed as the percent of ratios < 2 and < 3, were 87 and 96 %, respectively. These levels of accuracy and conservatism were more consistent with factors of 3 than factors of 2.

3.6. The distribution of concentration ratios in the test data

For the 57 concentration ratios in the test data from 17 species in 25 locations, concentration ratios ranged over a factor of 100 from 82 and 90 L kg⁻¹ for *A. brama* and *Gymnocephalus cernuus* (grugeon) to 18,169 and 20,107 L kg⁻¹ for *P. fluviatilis*. The median C_r was 3,444 L kg⁻¹, and 50 % of the C_r were within the range 933 to 6008 L kg⁻¹. The concentration ratios were not lognormally distributed (Kolmogorov–Smirnov Test of Normality, $D = 0.1371$; $P < 0.01$; Conover, 1971). Although containing fewer C_r than the modeling data, the test data covered the full range of C_r in the modeling data. The larger median C_r for the test data indicates that a greater proportion of larger C_r occurred in the test data. A test of the model's accuracy involving a greater proportion of larger concentration ratios is

preferable to a test involving mostly smaller concentrations which might leave doubts about the model's accuracy for larger C_r .

3.7. Evaluation of Model II's approximations and conservatisms for predicting C_r

A comparison of the observed and the C_r predicted by Model II is shown in Figure 4, and the frequency distribution of the observed to predicted ratios for Model II are shown in Figure 5. The median observed to predicted ratio was 1.262 with 50 % of the ratios occurring between the first and third quartiles of 0.726 and 1.850, respectively. The median ratio of 1.262 was not significantly different from 1 (Zar, 1999). Accuracies, expressed as the percentage of ratios within factors of 2 or 3 of the measured value, were 67 and 82 %, respectively. Conservatisms expressed as the percentage of observed to predicted ratios being < 2 or < 3 were 79 and 88 %, respectively. These levels of accuracy and conservatism were more consistent with levels of 3 than levels of 2.

There were four observed to predicted ratios that were > 4 . These all occurred for *Perca fluviatilis*. They are indicated in Figure 4, and a possible cause for these large ratios is discussed in section 4.2 below.

3.8. Comparable accuracy of models for those test locations with suspended sediment concentrations

For tests of accuracies and conservatisms involving the locations with suspended sediment data, there were differences in accuracies between Model II and the Rowan and Rasmussen model (Table 4) but negligible differences in conservatism. A better accuracy result for the Rowan and Rasmussen model was reflected in 1) the median ratio being closer to 1, 2) a larger per cent of ratios being within a factor of 2, and 3) the maximum ratio being less. Although a model with a greater number of adequately measured predictive variables may, *a priori*, be expected to be more accurate than a model with fewer predictive variables, there are legitimate reasons for expecting the Rowan and Rasmussen

(1994) model to be a better predictor of C_r . These include: 1) that despite the close correlation of K and suspended sediment concentrations (Table 2), the sediment data may have been measuring unique and important aspects of the factors affecting C_s that are not explained by K alone (Blaylock, 1972), 2) the piscivorous or non-piscivorous classification may have more concisely expressed the variability among fish species than the trophic levels with their relatively large standard errors; and 3) the model was based on a greater number of locations with K and sediment concentrations in the Rowan and Rasmussen (1994) data which included data from both freshwater and marine systems. Each of the models met both the accuracy and conservatism criteria for factors of 3 and ratios < 3 , respectively.

4. Discussion

Although having observed to predicted ratios with levels of accuracy and conservatism within factors of 2 would be preferable to within factors of 3, Model II had levels of approximation and conservatism that were more consistent with factors of 3 than factors of 2. Whether these levels of accuracy and conservatism are sufficient is a user-driven decision, but obtaining levels of accuracy within factors of 2 and conservatisms < 2 using the predictive model approach may be problematic due to several potential sources of error.

It could be expected that improved predictive models would be obtained from more measures of C_r for more species from more locations with more predictor variables, however the accuracies of these improved models may still be limited by: 1) uncontrolled external effects such as pH and Ca concentrations (Smith et al., 2002); 2) biological effects such as seasonal differences in fish concentrations (*e.g.*, Peles et al., 2000), size or age dependent variations in fish concentrations (*e.g.*, Kryshev and Ryabov, 2000; Smith et al., 2002); 3) variations between benthic and pelagic feeding behaviors (Rowan et al., 1998); 4) methodological differences such whole-body versus muscle-only fish concentrations (*e.g.* Rowan and Rasmussen, 1994) and differences between filtered versus non-filtered

water samples (Rowan and Rasmussen, 1994); and 5) variations in fish trophic levels among sites differing compositions of the aquatic community (*e.g.*, Rasmussen et al., 1990) so that the actual trophic level of a species at a site may differ from the mean *fishbase.org* value. Model accuracies may also be affected by sampling errors in measuring the K and Cs concentrations in the water at a site. All of these sources of error and variations may affect both 1) the Cs concentration in the fish and 2) the model's predicted C_r , and, consequently, the resulting observed to predicted ratios.

In addition, more improved predictive models will not improve the resolution of the measurement of C_r in modeling and test data sets. For the purposes of this analysis, models have been constructed and tested assuming that concentration ratios have been measured without error, which is clearly not the case. Where the C_r is computed as the ratio of two means, its sample variance is affected by both the sample variance of the concentrations in fish and the water (van Kempen and van Vliet, 2000). Where the C_r is computed as the ratio of k_1 to k_2 , its sample variance is affected by the sample variances of k_1 and k_2 and the covariance of the sample estimates of k_1 and k_2 (van Kempen and van Vliet, 2000). Although procedures have been developed to compute confidence intervals for ratios that are analogous to concentration ratios (*e.g.* Gardiner et al., 2001), confidence intervals for concentration ratios have rarely been reported. Without measures of the sampling errors for C_r , the extent to which these sampling errors contribute to the variability of observed to predicted ratios and the assessments of model accuracy remains to be determined.

Although each of these sources of error may be unlikely to cause factor of 2 errors on their own, their individual effects may combine to result in ratios outside the range of 0.5 to 2.

4.2. Possible model applications

If Model II's levels of accuracy and conservatism are deemed acceptable for practical uses, the model may be used to predict concentration ratios or to simply identify those species in those locations that may be most in need of monitoring and possible limitations of consumption. Whatever uses this model may or may not have, it does suggest the desirability of compiling measures of K concentrations in freshwater environments at risk from accidental Cs contamination. These compiled K concentrations would then be available to be combined with the more readily available trophic level data to estimate potential exposures following an accidental release.

Although the model might be used in its present form, its usefulness in specific situations could be enhanced by 1) a consideration of the alternative estimators of trophic levels presented in *fishbase.org* and 2) a consideration of pelagic versus benthic feeding with a possible adjustment of the predicted C_r .

4.3. *The use of alternative estimates of mean trophic level*

In the previous analyses, the mean trophic level for species was estimated using the method of random resampling of food items, but it was noted that alternative diet based estimates were also available for some species such as *P. fluviatilis*. The mean trophic level from the resampling procedure for this species was 3.66 but an estimate of 4.35 is also available for adult fish from diet studies (*fishbase.org*). Nineteen of the 57 fish in the test data sets were *P. fluviatilis*, and these 19 fish accounted for 9 of the 12 observed to predicted ratios > 2 , which included all four of the ratios that were > 4 . If the trophic level estimate of 4.35 is used, *P. fluviatilis* accounts for only 5 ratios > 2 with none of these being > 4 . Using the 4.35 estimate rather than the 3.66 estimate also 1) reduces the median ratio from 1.262 to 1.033, 2) increases the percentage of ratios within a factor of 3 from 82 % to 86 %, and 3) increases the percentage of ratios ≤ 3 from 88 % to 95 %.

This case for *P. fluviatilis* serves as only an example of the use of an alternative trophic level estimate. Selecting an alternative trophic level for a species should be based on knowledge that the alternative is more appropriate for the species in the location being assessed. However, the example does illustrate the improvement in predictions that might occur from being able to select a more appropriate alternative.

4.4. Potential differences between benthic and pelagic food chains

Food items and diet sources in *fishbase.org* are separated into pelagic and benthic components (e.g., zooplankton and zoobenthos), but no distinction is made between these components in the computation of mean trophic levels. This is appropriate for the objectives of the FishBase organization, but it is a limitation for predicting C_r for C_s . The greater accumulation of C_s from feeding on benthic sources is well-documented (Rowan et al., 1998; NCRP, 2007). Thus, a primarily benthic feeder may be expected to have a greater C_r for C_s than a pelagic feeder from the same system due to using the same C_s concentration in the water as a base for computing the C_r .

Examples of the effect that this difference in feeding locations has on predicted C_r are the observed to fitted C_r for *Ambloplites rupestris*, the rock bass. Studies have shown that this species accumulates greater C_s than co-occurring pelagic feeders (Rowan et al., 1998), and its benthic feeding behavior is reflected in *fishbase.org* with 75% of its food items and 94 % of its feeding habits being benthic. The species has a mean (\pm SE) trophic level of 3.33 ± 0.44 and therefore, predicted C_r of 2,407 and 1,645 L kg⁻¹, respectively, from Model II for sites with 22 and 43 $\mu\text{M K L}^{-1}$. However, its two observed C_r in the modeling data for these sites are 6,000 and 7,583 L kg⁻¹ (Rowan and Rasmussen, 1994) with corresponding observed to fitted ratios of 2.46 and 4.55.

These observations for *A. rupestris* imply that a more accurate predictive model could result from 1) the development of a benthic feeding index to accompany the trophic level estimate or 2) increasing the predicted C_r by a factor of 1.7X to 2X for benthic species as suggested from the results of Rowan et al. (1998). Such a 1.7X increase for *A. rupestris* would result in observed to fitted ratios of 1.45 and 2.68, but this 1.7X increase may not be appropriate for all primarily benthic feeders. For example, the primarily benthic feeding species *C. commersoni* and *Coregonus clupeaformis* (lake whitefish) in the modeling data have median observed to fitted ratios of 1.47 and 0.71, respectively, with all but one of the 16 ratios being < 2 .

4.5. Extrapolating C_s concentration ratios between sampled and unsampled fish species.

Beresford et al. (2013) have developed procedures based on phylogenetic similarities among fish species to allow the C_r determined for a species at a site to be extrapolated to either 1) species that were not sampled at the site or 2) species that were not present but are characteristic of that site. This extrapolation allows a more complete development of species information in databases to be used for predicting the environmental transfer of radionuclides into fish and other wildlife (Yankovich et al., 2012).

Phylogenetic similarities used by Beresford et al. (2013) include feeding behaviors (*e.g.*, piscivorous), feeding locations (*i.e.*, benthic versus pelagic), and habitat (*i.e.*, lentic versus lotic), and these similarities are evaluated to produce expected relative ratios of C_r between species. For example, the relative ratios for *D. cepedianum* and *M. salmoides* are 4.3 and 8.5, respectively (Beresford et al. 2013, Table 2). The procedures of Beresford et al. (2013) do not use trophic level estimates from *fishbase.org*, but the expected relative ratios for these species are similar to the relative trophic levels for these species of 2.4 and 4.42.

Model II may also be used to perform similar extrapolations of Cs concentration ratios among sampled and unsampled species at a site. To compute the extrapolated C_r for the unsampled species (C_{r_unsamp}), the procedure requires 1) the C_r measured for the sampled species (C_{r_samp}), 2) the trophic level for the sampled species (TI_samp), and 3) the trophic level for the unsampled species (TI_unsamp). The extrapolated $\log_{10}(C_{r_unsamp})$ is estimated using equation (3).

$$\text{Log}_{10}(C_{r_unsamp}) = \log_{10}(C_{r_samp}) + 0.2453 * (TI_unsamp - TI_samp) \quad (\text{Eq. 3})$$

Neither the procedures of Beresford et al. (2013) nor equation (3) require data on chemical or physical factors of the site's water because 1) these procedures are extrapolating values between species within the same system and 2) are assuming that these parameters are affecting the different species in the same manner. This assumption of similar effects on different species is supported for the effects of K on fish Cs concentrations and fish Cs C_r by 1) the minor contribution of the $\log_{10}(K)$ -trophic level interaction term to the predictive models, 2) the subsequent accuracy of the no interaction Model II in predicting C_r in the test data, and 3) the similar fish responses to potassium chloride additions in Lake Svyatoye (Smith et al., 2003).

4.6. Potential additional predictor variables

More accurate models are possible using a greater number of predictor variables, and there are, besides the obvious inclusion of suspended sediment concentrations, several other reported correlates with fish Cs concentrations that might be included in the models. These correlates may be divided into those related to 1) additional measures of water quality, 2) fish properties, and 3) food web structure. The measures of water quality include pH (Morgan et al., 1993; Smith et al., 2002), large Ca concentrations in the water especially in conjunction with low K concentrations (Smith et al., 2002), and mean annual water temperatures (Rowan and Rasmussen, 1994). Those related to fish properties

include fish size (Rowan et al., 1998; Smith et al., 2002), fish age (Rowan et al., 1998), and whether the species inhabits the epilimnion or hypolimnion (Rowan and Rasmussen, 1994). More complex food web structures with greater food chain lengths have also been related to increased contaminant concentrations in fish (Rasmussen et al., 1990; Rowan and Rasmussen, 1994).

Some of these possible predictors are correlated among themselves or with the existing predictors. Fish size and fish age are possible correlates with trophic level estimates and might be represented by the selection or modification of trophic level estimates. Estimates of trophic levels for adult fishes are available for numerous species in *fishbase.org*. Calcium and pH are also likely correlates (Smith et al., 2002) whose effects might be jointly represented by pH. Correlations among K, suspended sediments, pH and Ca may occur due to the gradient from hard-water to soft-water environments (Rowan and Rasmussen, 1994). The epilimnion versus hypolimnion distinction appears to be related to water temperatures (Rowan and Rasmussen, 1994). Determining food web complexity may be too difficult and too costly to qualify as a readily-obtainable predictor variable, but 1) the diversity of potential prey species and 2) the numbers of fish species, and therefore, presumably greater food web diversity, can be correlated with lake area (Barbour and Brown, 1974; Jackson and Harvey, 1989; Karatayev et al., 2005). These variables, with their potential intercorrelations that can compromise model accuracy, suggest that future modeling activities may benefit primarily from the use of pH, temperature and area.

Although there are these and possibly other useful predictors that could be added to expanded models, there can be a tradeoff between model accuracy and model utility. As the number of predictor variables increases, there is the possibility that fewer and fewer sites will have the requisite data on all or most of the predictor variables required by the model. The model developed in this study is hopefully one with ease of utility if perhaps less than ideal accuracy.

5. Advantages using the *fishbase.org* trophic levels compared to using discrete diet categories

There are two major advantages of using the *fishbase.org* trophic levels in place of the simpler classification of fish species into discrete categories such as nonpiscivorous and piscivorous. The first advantage is the avoidance of potentially arbitrary decisions as to which diet category a particular species should be assigned. The second is an increased resolution in predicted concentration ratios.

The increased resolution results from the use of separate trophic levels for each species with a resulting separate predicted concentration ratio for each species. Where fishes are classified into discrete categories, separate predicted concentration ratios are available only for each category and not for each species. For the median observed K concentration of $36 \mu\text{M L}^{-1}$ and the range in trophic levels from 2 to 4.5, the predicted species-specific concentration ratios from Model II range from 864 L kg^{-1} to 3544 L kg^{-1} with a separate prediction available for each 0.01 increment in trophic level. This greater than a factor of 4 range of predicted concentration ratios also occurs at all other K concentrations. For models where discrete classifications of diet are employed, the number of predicted concentration ratios would be the same as the number of classifications. For the classification of species as nonpiscivorous and piscivorous, only two predicted C_r would be available within the range of 864 L kg^{-1} to 3544 L kg^{-1} . Although using a greater number of categories would increase the number of predictions, it would also increase the complexity in deciding which species belongs in which category.

If discrete categories are preferred or required, the *fishbase.org* trophic levels may serve as 1) a basis for their construction and 2) the computation of an appropriate predicted concentration ratio. For example, the discrete classifications may be formed using ranges of trophic levels such as 2.0 to 3.5 for nonpiscivorous, or primarily nonpiscivorous, species and > 3.5 to 4.5 for piscivorous, or primarily piscivorous, species with a predicted C_r being computed using the midpoint trophic levels of 2.75 and 4 for the two intervals.

6. Summary

A model to predict species-specific concentration ratios for Cs radionuclides in fish is discussed that achieves reasonable levels of accuracy while requiring only two input variables. One variable is the K concentration in the water. The other variable is a trophic level estimate that is readily available from the internet source fishbase.org.

Acknowledgements

Resources and funding were provided by the Radiation Protection and Measurements Section of the Department of Environmental and Radiological Health Sciences of Colorado State University. Statistical analyses were supported by the Rocky Mountain Center for Nuclear Computations.

References

- Barbour, C. D., Brown, J. H., 1974. Fish species diversity in lakes. *Am. Nat.* 108, 473-489.
- Beauchamp, J. J., Olson, J. S., 1973. Corrections for bias in regression estimates after logarithmic transformation. *Ecology* 54, 1403-1407.
- Beresford, N. A., Yankovich, T. L., Wood, M. D., Fesenko, S., Andersson, P., Muikku, M., Willey, N. J. 2013. A new approach to predicting environmental transfer of radionuclides to wildlife: A demonstration for freshwater fish and caesium. *Sci. Total Environ.*, in press.
- Blaylock, B. G., 1982. Radionuclide databases available for bioaccumulation factors for freshwater biota. *Nuclear Safety* 23, 427-438.
- Conover, W. J. 1971. *Practical non parametric statistics*. John Wiley and Sons, Inc. New York.
- Doi, H., Takahara, T., Tanaka, K. 2012. Trophic position and metabolic rate predict the long-term decay process of radioactive cesium in fish: A Meta-Analysis. *PLoS ONE* 7, e29295.
- Draper, N., Smith, H., 1981. *Applied Regression Analysis, Second Ed.* John Wiley and Sons, Inc. New York.
- Fesenko, S., Fesenko, J., Sanzharova, N., Karpenko, E., Titov, I., 2011. Radionuclide transfer to freshwater biota species: Review of Russian language. *J. Environ. Radioact.* 102, 8-25.
- Froese, R., Pauly, D. Eds., 2011. *FishBase*. World Wide Web Electronic Publishing. www.fishbase.org.
- Gardiner, J. C., Huebner, M., Jetto, J., Bradley, C. J., 2001. On parameter confidence intervals for the cost-effectiveness ratio. *Biomet J.* 43, 283-296.

- Hosseini, A., Thorrying, H. brown, J. E., Saxen, R. llus, E., 2008. Transfer of radionuclides in aquatic ecosystems – default concentration ratios for aquatic biota in the Erica tool. J. Environ. Radioact. 99, 1409-1429.
- Jackson, D. A., Harvey, H. H., 1989. Biogeographic associations in fish assemblages: Local vs. regional processes. Ecology 70, 1472-1484.
- Karatayev, A. Y., Burlakova, L. E., Dodson, S. I., 2005. Community analysis of Belarusian lakes; relationships of species diversity to morphology, hydrology and land use. J. Plank. Res. 27, 1045-1053.
- Kline, T., Pauly, D., 1988. Cross-validation of trophic level estimates from a mass-balance model of Prince William Sound using ^{15}N : ^{14}N data. In, Fishery stock assessment models. Alaska Sea Grant College Program. Report AK-SG-98-01.
- Kryshev, A. I., Ryabov, I. N., 2000. A dynamic model of ^{137}Cs accumulation by fish of different age classes. J. Environ. Radioact., 50, 221-233.
- Morgan, I. J., Tyler, P., Bell, M. V., 1993. The accumulation of ^{137}Cs by alevins and fry of Atlantic salmon and brown trout. J. Fish Biol. 43, 877-888.
- NCRP. 2007. Cesium-137 in the Environment: Radioecology and Approaches to Assessment and Management. National Council on Radiation Protection and Measurements. NCRP Report No. 154. Bethesda, Maryland, USA.
- Newman, M. E., Brisbin, I. L. Jr., 1990. Variation of ^{137}Cs levels between sexes, body sizes and collection localities of mosquitofish, *Gambusia holbrooki* (Girard 1859), inhabiting a reactor cooling reservoir. J. Environ. Radioact. 12, 131-141.

- Pinder, J. E. III., Hinton, T. G., Taylor, B. E., Whicker, F. W., 2011. Cesium accumulation by aquatic organisms at different trophic levels following an experimental release into a small reservoir. *J. Environ. Radioact.* 102, 283-293.
- Pinder, J. E. III., Hinton, T. G., Whicker, F. W., Smith, J. T., 2009. Cesium accumulation by fish following acute input to lakes: a comparison of experimental and Chernobyl-impacted systems. *J. Environ. Radioact.* 100, 456-467.
- Ramsey, F. L., Schafer, D. W., 2002. *The statistical sleuth: A course in methods of data analysis.* Second Edition. Duxbury, Pacific Grove CA.
- Rasmussen, J. B., Rowan, D. J., Lean, D. R. S., Cary, J. H., 1990. Food chain structure in Ontario lakes determines PCB levels in lake trout (*Salvelinus namaycush*) and other pelagic fish. *C. J. Fish. Aquat. Sci.* 47, 2030-2038.
- Rowan, D. J., Chant, L. A., Rasmussen, J. B., 1998. The fate of radiocesium in freshwater communities – Why is biomagnification variable both within and between species? *J. Environ. Radioact.* 40, 15-36.
- Rowan, D. J., Rasmussen, J. B., 1994. Bioaccumulation of radiocesium by fish: The influence of physiochemical factors and trophic structure. *Can. J. Fish. Aquat. Sci.* 51, 2388 – 2410.
- Searle, S. R., 1987. *Linear models for unbalanced data.* John Wiley and Sons. New York.
- Smith, J. T., Kudelsky, A. V., Ryabov, I. N., Hadderingh, R. H., 2000. Radiocesium concentration factors of Chernobyl-contaminated fish: a study of the influence of potassium and a “blind” testing of a previously developed model. *J. Environ. Radioact.* 48, 359-369.

- Smith, J. T., Kudelsky, A. V., Ryabov, I. N., Daire, S. E., Boyer, L., Blust, R. J., Fernandez, J. A., Hadderingh, R. H., Voitsekhovitch, O. V., 2002., Uptake and elimination of radiocesium in fish and the size effect. *J. Environ. Radioact.* 62, 145-164.
- Smith, J. T., Kudelsky, A. V., Ryabov, I. N., Hadderingh, R. H., Bugakov, A. A., 2003. Application of potassium chloride to a Chernobyl-contaminated lake: modeling the dynamics of radiocesium in an aquatic ecosystem and decontamination of fish. *Sci. Total Environ.* 305, 217-227.
- Sokal, R. R., Rohlf, F. J. 1969. *Biometry*. W. H. Freeman , San Francisco, CA.
- Vander Zander, M. J., Rasmussen, J. B., 1996. A trophic position model of pelagic food webs: Impact on contaminant bioaccumulation in lake trout. *Ecol. Monogr.* 66, 451-477.
- van Kempen, G. M. P., van Vliet, L. J. 2000. Mean and variance of ratio estimators used in fluorescence ratio imaging. *Cytom.* 39, 300-305.
- Vanderploeg, H. A., Parzyck, D. C., Wilcox, W. H., Kercher, J. R., Kaye, S. V., 1975. Bioaccumulation factors for radionuclides in freshwater biota. Oak Ridge National Laboratory, Environmental Sciences Division Publication N. 783. ORNL-5002. Oak Ridge, TN.
- Wald, A., 1940. The fitting of straight lines if both variables are subject to error. *Ann. Math. Stat.* 11, 284-300.
- Whicker, F. W., Pinder, J. E., Bowling, J. W., Alberts, J. J., Brisbin, I. L., 1990. Distribution of long-lived radionuclides in an abandoned reactor cooling reservoir. *Ecol. Monogr.* 60, 471-496.
- Yankovich, T., Beresford, N. A., Fesenko, S., Fesenko, J. Phaneuf, M., Dagher, E., Outola, I., Andersson, P., Thiessen, K., Ryan, J. Wood, M. D., Bollhofer, A., Barnett, C. L., Copplestone, D., 2012.

Establishing a database of radionuclide transfer parameters for freshwater wildlife. *J. Environ. Radioact.* In press.

Zar, J. H., 1999. *Biostatistical analysis, Fourth Edition.* Prentice Hall, Upper Saddle River, New Jersey.

Table 1.

The mean, n = number, standard deviation, median, minimum and maximum for the predictor variables trophic level, concentrations of K ($\mu\text{M L}^{-1}$), and suspended samples (mg L^{-1}) in the water. Trophic levels are computed for 44 species. Concentrations of K and suspended sediments are computed for 43 and 14 locations, respectively.

Variable	n	Mean	Standard Deviation	Median	Minimum	Maximum
Trophic Level	44	3.47	0.48	3.53	2.40	4.42
K	43	67.9	90.6	36	8	512
Suspended Sediment	14	10.4	18.6	3.4	0.21	70.7

Table 2.

Spearman rank correlations among trophic levels, K concentrations ($\mu\text{M L}^{-1}$), suspended sediment concentrations (mg L^{-1}) and Cs concentration ratios (kg L^{-1}). n = number of samples. P = probability.

	Variables		
	K concentrations	Sediment Concentrations	Concentration Ratios
Trophic Levels	-0.114 n = 207 P > 0.05	0.083 n = 90 P > 0.05	0.337 n = 207 P < 0.001
K Concentrations		0.760 n = 14 P < 0.01	-0.567 n = 207 P < 0.001
Sediment Concentrations			-0.483 n = 90 P < 0.001

Table 3.

Comparison of regression statistics for two predictive models. Model I uses the predictive variables $\log_{10}(k)$ and the interaction of trophic level and $\log_{10}(K)$. Model II uses the predictive variables trophic level and $\log_{10}(K)$. AIC and BIC refer, respectively, to Akaike's and Bayes Information Criteria. * = $P \leq 0.01$.

Model	F-test	Error Mean Square	R^2	Adjusted R^2	AIC	BIC	Median Observed to Fitted Ratio	Median Predicted Concentration Ratio
I	55.82*	0.1306	0.353	0.347	-177.01	-177.01	0.984	2227
II	54.47*	0.1317	0.348	0.342	-175.29	-176.24	0.965	2191

Table 4.

Comparison of accuracy and conservatism between Model II using trophic level and $\log_{10}(K)$ concentration and the equation 5 model of Rowan and Rasmussen (1994) for 28 observations of C_r for 6 species from 10 eastern European lakes and reservoirs with measures of suspended sediment concentrations from Smith et al. (2002).

Model	No. of Ratios	Distribution of Observed to Predicted Ratios					% Ratios within factors of		% Ratios <	
		Minimum	First Quartile	Median	Third Quartile	Maximum	2	3	2	3
II	28	0.21	0.55	0.83	1.47	5.19	64	82	86	93
5	28	0.37	0.67	0.97	1.53	4.17	71	93	86	93

Figure Legends

Fig. 1. The frequency distribution of the 207 concentration ratios in the modeling data.

Fig. 2. The frequency distribution of trophic levels for the 44 fish species used to build the predictive model.

Fig. 3. The frequency distribution of the 57 concentration ratios in the test data.

Fig. 4. A comparison of the observed and predicted concentration ratios for the 57 observations in the test data. Lines denote the observed to predicted ratios of 3:1, 1:1 and 1:3. The four cases where the observed to predicted ratios for *Perca fluviatilis* were > 4 are indicated.

Fig. 5. The cumulative frequency distribution of the 57 observed to predicted ratios from the test data.

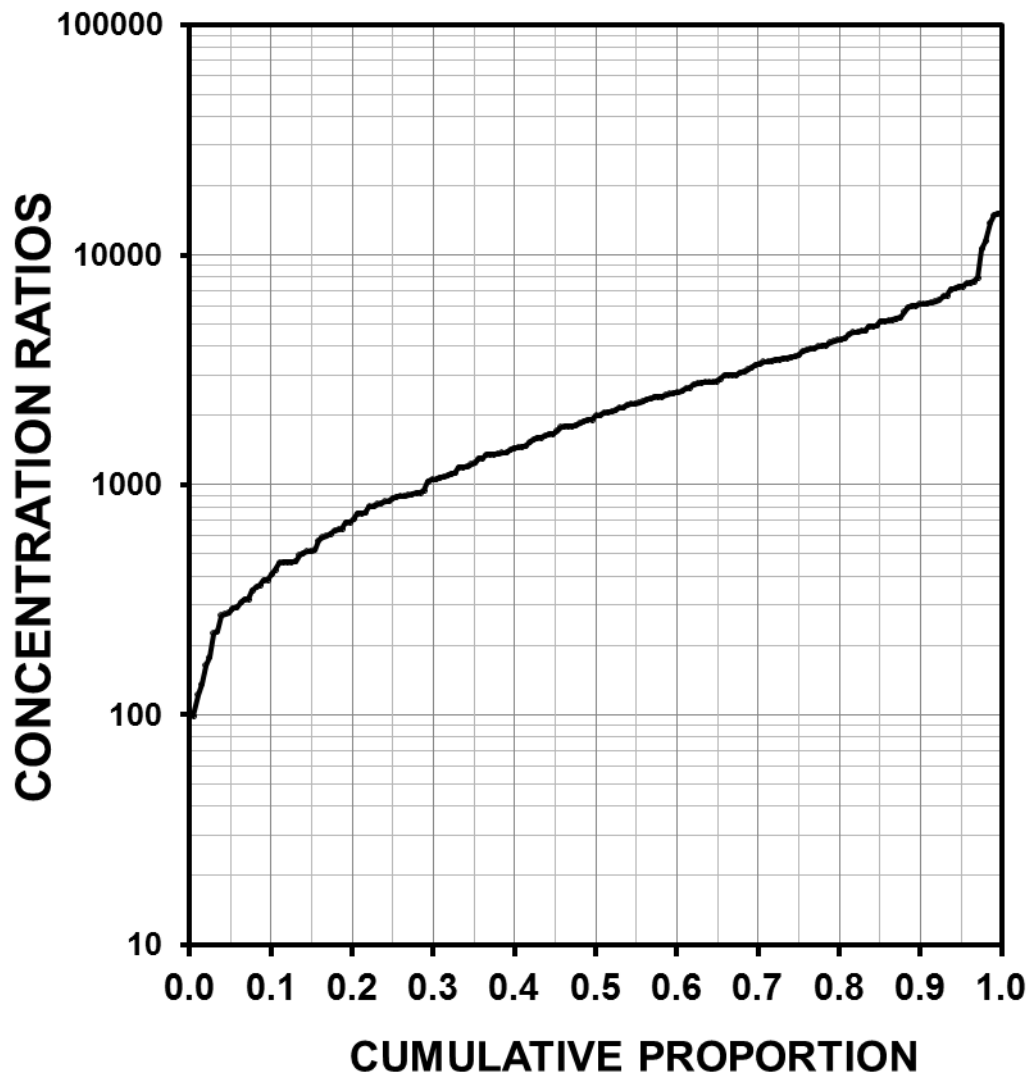


Figure 1

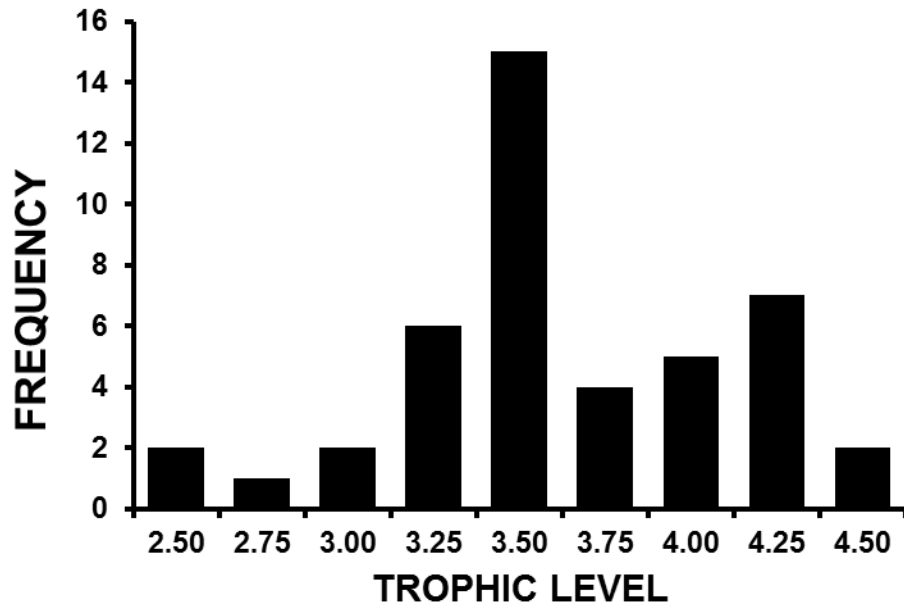


Figure 2

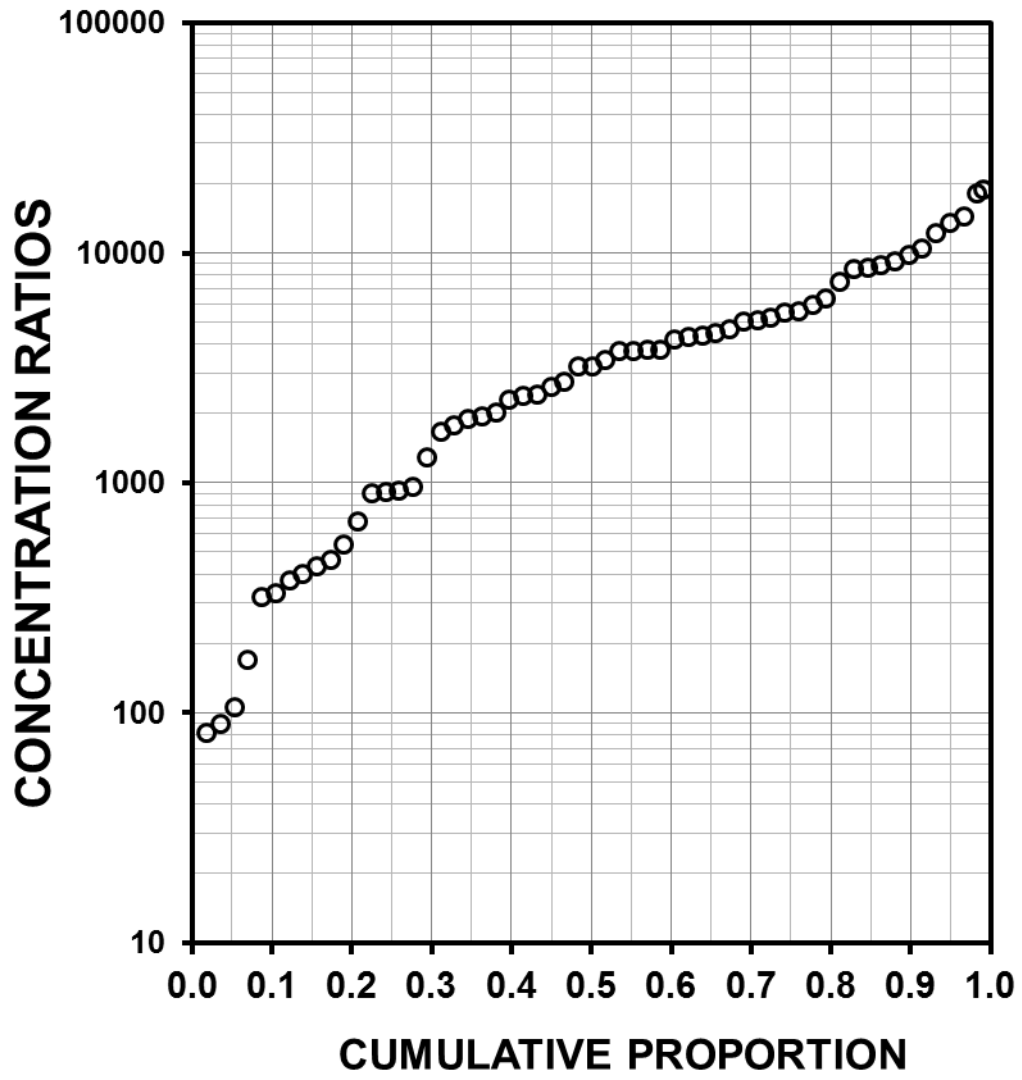


Figure 3

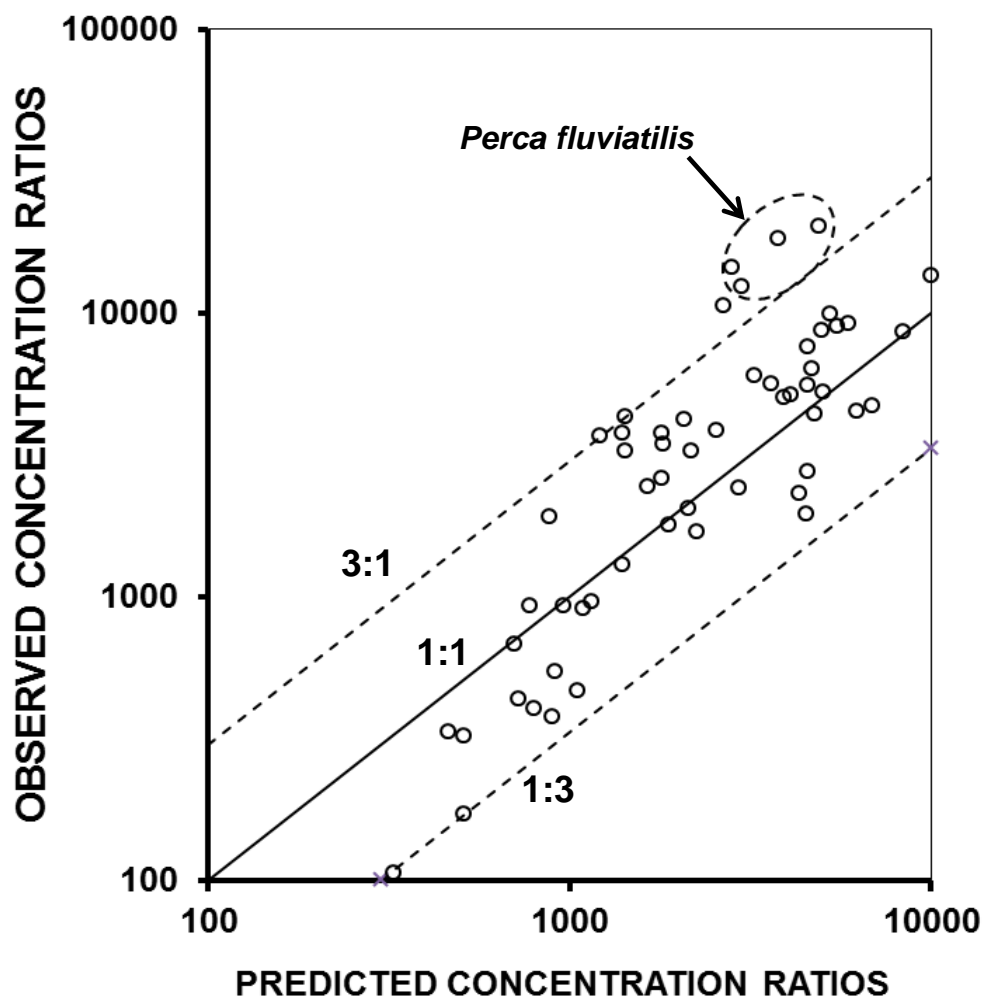


Figure 4

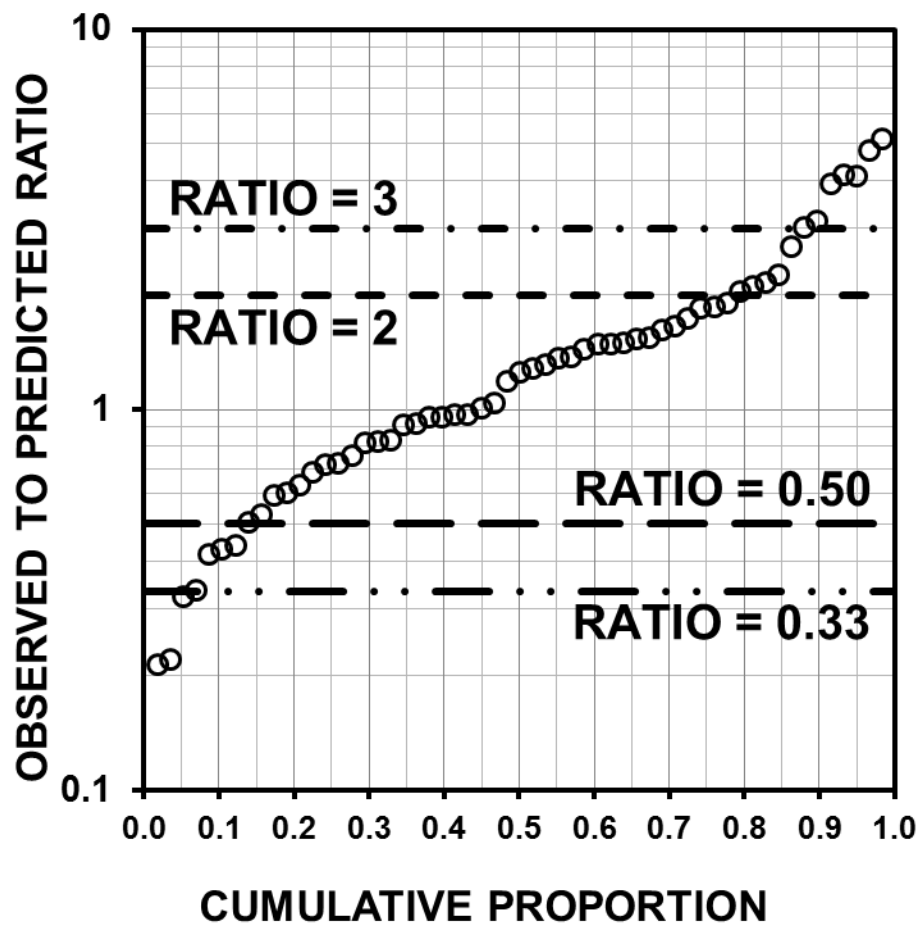


Figure 5