EPA Public Access

Author manuscript

Environ Toxicol Chem. Author manuscript; available in PMC 2019 March 01.

About author manuscripts

Submit a manuscript

Published in final edited form as:

Environ Toxicol Chem. 2018 March ; 37(3): 871–883. doi:10.1002/etc.4022.

USING EXTIRPATION TO EVALUATE IONIC TOLERANCE OF FRESHWATER FISH

Michael B. Griffith^{a,*}, Lei Zheng^b, and Susan M. Cormier^a

^aU.S. Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment, Cincinnati, Ohio, USA

^bTetra Tech, Owings Mills, Maryland, USA

Abstract

Field data of fish occurrences and specific conductivity were used to estimate the tolerance of freshwater fish to elevated ion concentrations and to compare the differences between species- and genus-level analyses for individual effects. We derived extirpation concentrations at the 95th percentile (XC95) of a weighted cumulative frequency distribution for fish species inhabiting streams of the central and southern Appalachians by customizing methods used previously with macroinvertebrate genera. Weighting factors were calculated based on the number of sites in basins where each species occurred, reducing overweighting observations of species restricted to fewer basins. Comparing the species- and genus-level fish XC95 values, XC95s for fish genera were near the XC95s for the most salt-tolerant species in the genus. Therefore, a genus-level effect threshold is not reliably predictive of species-level extirpation, unless the genus is monospecific in the assessed assemblage. Of the 101 fish species XC95 values, 5% were less than 509 $\mu\text{S}/\text{cm}$, 10% were less than 565 $\mu\text{S}/\text{cm}$. The lowest XC95 for a species was 322 $\mu\text{S}/\text{cm}$, which is greater than 300 $\mu\text{S}/\text{cm}$, the exposure estimated to extirpate 5% of macroinvertebrate genera in the central Appalachians. Above 509 $\mu\text{S}/\text{cm}$, 41 of the 101 species are expected to decline in occurrence.

Keywords

Fish; Salinity; Specific conductivity; Species sensitivity distributions; Freshwater toxicology; Aquatic toxicology

INTRODUCTION

One cause of adverse effects in freshwater organisms is increased concentrations of dissolved ions associated with human activity [1–4]. Freshwater fish, as with freshwater

* Address correspondence to Michael B. Griffith, griffith.michael@epa.gov.

Supplemental Data—The Supplemental Data (Tables S1–S2, Fig. S1–S3) are available on the Wiley Online Library at DOI:

Disclaimer—This manuscript was prepared at the U.S. Environmental Protection Agency (U.S. EPA), National Center for Environmental Assessment, Cincinnati Division. It has been subjected to the agency's peer and administrative review and approved for publication. However, the views expressed are those of the authors and do not necessarily represent the views or policies of the U.S. EPA.

The author declares that there are no conflicts of interest.

Data availability—The compiled data set used in these analyses can be viewed or downloaded at <https://doi.org/10.23719/1376690>.

invertebrates, are hyper-regulators that maintain greater ion concentrations in their blood and cytoplasm than in the naturally hypo-osmotic freshwaters [5–7]. Because the water is more dilute than their blood or cytoplasm, fish must deal with absorption of water across their epithelial membranes and diffusional loss of salts through semi-permeable membranes and by excretion. Several anatomical and physiological mechanisms are involved. Water is excreted as dilute urine by their renal system. Salts are absorbed by various transporter proteins or permeases on gill and renal epithelia that transport ions against concentration gradients [6, 8, 9]. All ions contribute to the osmolality of blood and cytoplasm, but the concentrations of specific ions are maintained at differing concentrations both in the blood and cytoplasm. These ion concentration gradients facilitate the movement of solutes among these two compartments and the external water [5–7]. Increased concentrations of some ions in water can alter osmotic stress [8, 10], but increased concentrations of specific ions, such as H^+ , K^+ , and HCO_3^- , alter electrochemical gradients, effecting ion transport and associated energy expenditures that can cause ionoregulatory stress and acid-base imbalances [7, 10, 11]. Differences among species in the capacity or affinity of the ion transporter proteins for specific ions can affect the tolerance of fish species to increased ionic concentrations in water [7, 12]. In our analyses of mid-Atlantic Highlands ecoregions, anthropogenic sources of ions particularly increase the cations, Ca^{2+} and Mg^{2+} , and the anions, HCO_3^- and SO_4^{2-} [13, 14].

To characterize adverse exposure levels for aquatic life, the U.S. Environmental Protection Agency (EPA) developed a field-based method to estimate benchmarks or criteria for exposure to ionic mixtures, measured as specific conductivity (SC) [15]. A benchmark is a non-regulatory reference point that enables comparison and goal setting. The method was illustrated with a field data set from the central Appalachians that estimated the SC likely to extirpate 5% of aquatic benthic invertebrates [15–17]. Recently, the EPA has released a draft method adapting the field-based method for developing criteria, which are recommendations for identifying harmful levels [18]. To emphasize the broader application of these methods to estimate fish species extirpation, we use the more general term of benchmark or threshold in this manuscript.

Extirpation is defined as “the depletion of a population to the point that it is no longer a viable resource or is unlikely to fulfill its function in the ecosystem” [19]. Extirpation in this field-based method was operationally defined for a taxon as the SC value below which 95% of the observations of a species or genus occur (XC95) [15]. The 95th percentile characterizes the extreme of an organism’s tolerance but is more stable than the maximum value at which a taxon is observed, as outlier observations may be due to transient occurrences or sink habitats [17]. At the EPA’s benchmark for central Appalachian streams, 5% of benthic invertebrate genera in a stream are likely to be extirpated at 300 $\mu S/cm$ [5].

Because the original field-based SC benchmark was derived from benthic macroinvertebrate data [16], it was not clear whether freshwater fish might also be extirpated. To resolve this, we estimated both XC95 values for fish species and the SC at which the probability of observing a fish species begins to decline. These fish values were compared to the 300 $\mu S/cm$ benthic invertebrate benchmark.

We also investigated whether genus XC95 values and species XC95 values systematically differ. Fish are generally identified to species and are easily identified to genus. Aquatic insects are typically identified only to genus. To address this difference in taxonomic identification, we derived and compared XC95 values for 101 fish species and their genera.

Lastly, because many aquatic insects can fly between watersheds, biogeographical distributions differ between fish and macroinvertebrates [20, 21]. We asked whether changing the weighting of observations of fish to adjust for biogeography affected the XCD and XCD05. To answer, we compared an analysis using the U.S. EPA [15] method that includes weighting by the number of all samples within an SC range to an analysis where weighting considered only those samples within an SC range that were collected from the biogeographical range of each fish species.

Answering these questions and performing tests to evaluate the effects of sample sizes on the calculations allowed us to provide methods for estimating the effects of pollutants on fish species using a field-based approach. We applied that method to assess the effects of increased concentrations of dissolved ions on freshwater fish. It allowed us to consider the consequences of selecting genus or species as an endpoint and to provide SC levels likely to extirpate fish species in the central and southern Appalachians.

METHODS AND MATERIALS

Data sources

Because no single data set was available for fish that was similar to the West Virginia data set used for macroinvertebrates, data from several sources were compiled into a fish data set. The fish data set was constructed to include 4 contiguous mid-Atlantic Highlands Level III ecoregions: 67 (Ridge and Valley), 68 (Southwestern Appalachians), 69 (Central Appalachians), and 70 (Western Allegheny Plateau) (Fig. 1) [22, 23]. These ecoregions are characterized by mountain ridges and valleys underlain by sedimentary bedrock and having extensive areas of forest and agriculture with a few large metropolitan areas (Pittsburgh, PA and Charleston-Huntington, WV being the largest). These 4 ecoregions are placed in the Ozark, Ouachita-Appalachian Forests Level II ecoregion [24] and are physiographically part of the Ridge and Valley and the Appalachian Plateau provinces of the Appalachian Highlands [25]. Large-scale land disturbances in the region are the result of forestry, some agriculture, and resource extraction, primarily coal mining. Eight data sets were combined prior to calculating XC95 values for fish in this region (Table 1).

Fish survey data along with chemical and physical data were collected from a total of 3,465 distinct sites during the sampling years 1990–2014. The Environmental Monitoring and Assessment Program (data sets 1 and 2), Regional Applied Research Effort (data set 5), and West Virginia Department of Environmental Protection (WVDEP) (data set 6) sites were probability samples selected as part of regional surveys [26–28]. Sites sampled by Kentucky Department of Environmental Protection, Division of Water (KDEP-OW); Stauffer and Ferreri [29]; Ohio Environmental Protection Agency, Division of Surface Water (OEPA-DSW); and Pennsylvania Fish and Boat Commission (PFBC) (data sets 3, 4, 7, and 8,

respectively) included targeted-sampling sites that were part of bioassessment studies. All sites were in perennial reaches of streams.

Most sites in the combined data set were sampled once, but 293 sites were revisited and resampled one or more times, usually within the same year or in consecutive years. Sites were not identified as “least disturbed” or reference sites [30]. However, at least 134 sites occurred in catchments described as >90% forested. Water quality, habitat, and fish data (both raw data and calculated metrics) were collected as part of these regional bioassessment surveys. Quality assurance and standard procedures are described by Lazorchak et al. [31], U.S. Environmental Protection Agency [32], KDEP-DOW [33–36], Stauffer and Ferreri [29], Bryant et al. [37], West Virginia Department of Environmental Protection [38], OEPA-DSW [39–41], and Pennsylvania Department of Environmental Protection [42]. Biological sampling usually occurred once from March through November (Table 2) with fish sampling protocols, including electrofishing or electrofishing and seining, that are designed to collect all except very rare species [31].

Data set characteristics

The combined data set used in the analyses included 3,277 sites out of 3,456 total sites from Ecoregions 67, 68, 69, and 70 (Fig. 1) and were used in the calculation of the XC95 values for fish (Table 2). Data from a sampling event at a site were excluded from the analysis if they lacked an SC measurement ($n = 53$). Observations from 26 sites where no fish were collected were excluded to minimize bias from sites that were too small to support fish. To prevent potential confounding by the effects of acid mine drainage or acid deposition, 100 sites with a pH <6 were excluded from the analysis and 2.7% of sites lacking pH measurements were retained. Analyses, therefore, represent waters having a pH between 6.0 and 9.5, a range that is not notably toxic to most fish [43]. The ionic matrix at most sites in the data set was dominated by SO_4^{2-} and HCO_3^- anions. We did not exclude sites that may have a different toxicity (i.e., $[\text{Cl}^-] - [\text{SO}_4^{2-}] + [\text{HCO}_3^-]$), as was done in the benchmark derivation for macroinvertebrates [5], because such sites appeared to be rare in the data set. Furthermore, of the 6 sites with higher $[\text{Cl}^-]$ than $[\text{SO}_4^{2-}] + [\text{HCO}_3^-]$ in mg/L, the total ionic concentrations and SC were very low and well below levels likely to cause effects.

Observations of fish were excluded from calculations if the fish were not identified to species. Such fish were generally immature specimens, and identifiable, mature specimens of the species were generally present in the same sample. All fish were considered to be freshwater species. Species observed at fewer than 25 sampling locations in the aggregated ecoregions were excluded to ensure reasonable confidence in the evaluation of the relationship between SC and the occurrence of a species. Although stocking could artificially raise the XC95 estimates and influence the identification of naturally reproducing populations of species, the highly valued, native salmonid species, brook trout (*Salvelinus fontinalis*), was included. Three non-native fish species, rainbow trout (*Oncorhynchus mykiss*), brown trout (*Salmo trutta*), and common carp (*Cyprinus carpio*), were included in the analyses. The two non-native salmonids were included because some trout populations are established in the region and they are recreationally important. However, it is uncertain how stocking may have affected the estimation of their XC95 values. Common carp was

included because it has become irreversibly established in the region. All other fish species analyzed were native to at least one of the river basins.

Fish extirpation analysis

The derivation of the XC95 values for SC and fish follows the field-based method for macroinvertebrates [15] with some variations to answer particular research questions. For all analyses, the XC95 values are estimated as the 95th percentile of a cumulative distribution of weighted observations of a taxon over a range of SC that was concurrently measured at each site [24]. Weighted observations of a taxon were used to adjust for uneven sampling along the SC exposure gradient. Because the distribution, and thus, the occurrence of fish species are affected by biogeography [20, 21] and stream size [44], two different sample weighting procedures were compared. Using the original method [15], the observations of a taxon were weighted by the number of samples in defined SC bins (Fig. 2). For the biogeographically adjusted weighting procedure, observations of a taxon were weighted by the number of samples in defined SC bins in river systems with catchment areas in which a species occurred (Table S1). For example, the calculation of sample weights for species that only occur in the Cumberland drainage were calculated using only the number of samples from the Cumberland drainage and not based on samples from the entire combined data set. Except when evaluating the effect of sample size on the analysis, analyses were performed to calculate XC95 values for each fish species and each genus with a minimum of 25 observations. The selection of a minimum of 25 occurrences for inclusion of a species or genus was chosen to provide sufficient occurrences to estimate the species or genus XC95 values while including as many species or genera as possible. The fewer the number of samples, the less likely the tails of the distribution of XC95 values are sampled. The sample size was judged as sufficient when the addition of species did not disproportionately extend the lower 5th percentile of the overall distribution of species XC95 values (XCD05). The effect of varying sample size was assessed based on an analysis that is described in the Results (see *Effect of Sample Size*).

Analysis of biogeographical weighting

For the biogeographically weighted method, the range of stream sizes (based on the log₁₀ transformed catchment area [in km²]) and river basins (based on 4-digit hydrological unit codes [HUCs]) were identified where fish species were collected. Before calculating weights and XC95 values for each fish species or genus, data subsets were prepared for each species to limit the total of sampled sites to those in the 4-digit HUCs where the species occurred in the entire combined data set. Also for each fish species or genus, sites were excluded within catchments greater than the maximum and less than the minimum size where a fish species or genus occurred in the data set (Supplementary Data Table S1). Thus, weighting was adjusted for drainage size and biogeography.

To compare the effect of weighting, we computed weights for each sample in two ways: using an analysis for the entire area and an analysis for sub-areas defined for species. In both the original and biogeographical weighting procedure, we first defined equally sized bins, each 0.048 log₁₀ SC units wide, that spanned the range of observed SC values. The bin size was selected to balance the requirements of sufficient occurrences in a bin to define the

proportion and sufficient bins to define the form of the response [15]. We then calculated the number of samples that occurred within each bin (Fig. 2). Each sample was then assigned a weight $w_i = 1/n_i$, where n_i is the number of samples in the i^{th} bin. The value of the weighted CFD, $F(x)$, of SC values associated with observations of a particular species was computed for each unique observed value of SC, x , as:

$$F(x) = \frac{\sum_{i=1}^{N_b} \sum_{j=1}^{M_i} w_i I(x_{ij} < x \text{ and } G_{ij})}{\sum_{i=1}^{N_b} \sum_{j=1}^{M_i} w_i (G_{ij})} \quad (1)$$

where x_{ij} is the SC value in the j^{th} sample of bin i , N_b is the total number of bins, M_i is the number of samples in the i^{th} bin, G_{ij} is true if the species of interest was observed in the j^{th} sample of bin i , and I is an indicator function that equals 1 if the indicated conditions are true, and 0 otherwise. The XC95 value is defined as the SC value, x where $F(x) = 0.95$. It is graphed as the SC value at the intersection of this line and the CFD.

Analysis of threshold of decline

To estimate the SC above which observations of a species begin to decline and to evaluate whether each XC95 value is well characterized by the available data and exposure range, we examined plots of a taxon's probability of being observed within discrete SC ranges or bins with increasing SC. A generalized additive model (GAM), which is the mean smoothing spline fit with 3 degrees of freedom, was added to plots of probabilities of observing a taxon, estimated as the proportion of samples within each bin in which the species or genus was observed. The fitted line characterizes the trend of observations for each species or genus along the SC gradient. The maximum of the GAM fitted model was used to identify the SC at which a species began to decline. For convenient comparison, the estimated XC95 derived from the weighted CFD was added to the GAM plots.

Qualifications of the XC95 values were assigned based on the intercept of the mean fitted line and the lower confidence limit of the GAM. If the GAM mean curve at the XC95 value was less than 1% of the maximum modeled probability (i.e., approximately 0), then the XC95 is listed without qualification (i.e., assessed as a reliable estimate of an XC95 value). If the GAM mean curve at the XC95 value is >0 but the lower confidence limit is less than 1% of the maximum modeled probability, then the XC95 was listed as approximate ($\approx$). If the GAM lower confidence limit at the XC95 value was >0 , then the XC95 is listed as greater than ($>$). The assignment of $>$ or $\approx$ qualifications are provided to alert users of the uncertainty of the XC95 values for this and other applications, such as causal assessment [2, 45].

Estimation of confidence bounds by bootstrapping

Derivation of the SC at the 5th percentile of the cumulative distribution of XC95 values (XCD05) for fish follows the field-based method for macroinvertebrates for SC [15]. The XC95 values are used to generate a species or genus XCD by ordering the taxa from lowest to highest XC95 value, and the XCD05 value is derived from the XCD as the 5th percentile

of the distribution of the species or genera. The confidence intervals (CI) on the XC95 values and XCD05 for fish were calculated using bootstrapping as described by the field-based method for deriving SC criteria [15], again using the sample size and weights defined for each fish species.

Specifically, bootstrap estimates of the XC95 values were derived for each species used in the derivation of the XCD05 by resampling 3,277 times (the number of sites in the combined data set) with replacement [46]. For each bootstrap sample, the XC95 for each species and the XCD05 were calculated by the same method applied to the original data. That process was repeated 1,000 times to create distributions of XC95 and XCD05 values, and these distributions were used to calculate a 2-tailed 95% CI for the XCD and XCD05. The statistical package R version 3.1.2 (October 2014) was used for all statistical analyses [47].

Analysis of effect of sample size

The shape of the XCD and XCD05 and their associated uncertainties are influenced by the number of species and by the number of sites sampled. The number of sampled sites affects the potential to observe species often enough to reliably estimate an XC95. Inclusion of more species increases the representativeness of the low SC tail of the XC95 distribution and hence the XCD05. However, a low number of occurrences of a species may affect the reliability of such an XC95. We evaluated the effect of the number of sites that were sampled and its effect on the number of species and consequently its effect on the XCD05. We used a bootstrapping protocol to estimate 1,000 XCDs, the number of species in each XCD, and their median from 1,000 XCD05 values for a range of sample sizes (100 to 3,000 sites). A bootstrapping protocol was then used to calculate CIs on the XCD05 values. To evaluate the effect of data set sample size, subsamples of the combined data sets with 100 to 3,000 sites (1,000 samples each) were randomly picked with replacement from the original 3,277 samples. From each subsample of the combined data set, the XC95 was calculated for each species by the same procedure applied to the original data, and the XCD05 was calculated. The uncertainty in the XCD05 value was evaluated by repeating the sampling and XCD05 calculation 1,000 times for each data set sample size (Supplementary Data Fig. S1). The distribution of 1,000 XCD05 values was used to generate a median XCD05 and 2-tailed 95% CI bounds on these bootstrap-derived values.

RESULTS

The fish XC95 values and proportion of species extirpated at different SC levels were calculated for water with a relatively consistent mixture of ions (i.e., $[\text{Cl}^-] + [\text{SO}_4^{2-}] + [\text{HCO}_3^-]$ in mg/L) at circumneutral to mildly alkaline pH (6–9.5). Where HCO_3^- , SO_4^{2-} , and Cl^- were measured in mg/L ($n = 979$), Cl^- was the dominant anion at 6 sites, but 4 of these sites had $\text{SC} < 300 \mu\text{S}/\text{cm}$ and so no sites were eliminated based on Cl^- dominance. For the circumneutral to alkaline streams, Ca^{2+} ($r = 0.93$), Mg^{2+} ($r = 0.91$), and $\text{SO}_4^{2-} + \text{HCO}_3^-$ ($r = 0.95$) are highly correlated with SC, and Na^+ and Cl^- are moderately correlated ($r = 0.77$ and 0.67 , respectively). Table 3 summarizes the descriptive statistics for ion concentrations and other variables for the 3,277 observations in the combined data set used in the analyses, and Supplementary Data Table S2 shows the Spearman correlations among

these variables. Based on these characteristics, the following results are relevant to streams with a similar ionic composition as those represented in the combined data set (i.e., $[\text{Cl}^-]$ $[\text{SO}_4^{2-}] + [\text{HCO}_3^-]$ in mg/L).

Fish species exposure response

In the combined database, 210 fish species were identified. Of those, 101 species were observed in at least 25 sites and XC95 values were calculated for them (Table 4). The 4 ecoregions had 36 of these 101 species in common, with 76 species in Ecoregion 67, 47 species in Ecoregion 68, 97 species in Ecoregion 69, and 86 species in Ecoregion 70. Some of the 109 species that were observed in less than 25 sites were rare and had small ranges, but most of these infrequently observed species are found in river basins, such as the Tennessee, James, and Roanoke, that were primarily limited to upper tributaries within the 4 ecoregions across 7 states included in the combined database. Genus XC95 values were also calculated for those fish genera with more than one species if there were ≥ 25 occurrences of the genus (Table 4). Supplementary Data Fig. S2 illustrates the CFD for each species used to derive the XC95 values. Supplementary Data Fig. S3 provides the GAM plots used to assign qualifying designations of “approximately” or “greater than” to the calculated values.

Biogeographical weighting procedure

The species XCD05 values resulting from the biogeographical weighting procedure and from the original weighting procedure using the entire combined data set were 509 and 518 $\mu\text{S}/\text{cm}$, respectively. This represents less than a 2% difference between the two weighting schemes. The biogeographical weighting method may be slightly more precise for estimating XC95 values, because the weights are tailored to the individual species geographical distribution. Therefore, the following analyses use the XC95 values calculated with the more complex biogeographical weighting procedure.

Effect of sample size

To evaluate the effect on the XCD05 of requiring a minimum of 25 observations of a species, XC95 values were calculated using different minima of species observations and then calculating XCD05 values for each set of XC95 values. As the minimum number of observations of a species required for inclusion in the data set increases, fewer species are included in the XCD and the XCD05 increases toward a temporary asymptote (Fig. 3). When species that do not meet the minimum number of samples are removed, the XCD05 increases when a species is in the lower 5th percentile, while the XCD05 decreases when a species has an XC95 value greater than the 5th percentile. The number of samples in the data set affects the number of observations of a species and their inclusion in the XCD, thereby also affecting the XCD05. To ensure inclusion of as many fish species as possible while also having sufficient numbers of occurrences to estimate species XC95 values, a minimum of 25 occurrences of a species was selected for inclusion. Of the 210 species, 109 species occurred < 25 sites. An analysis of the effect of varying the minimum number of occurrences before calculating an XC95 showed that the XCD05 decreased by about 200 $\mu\text{S}/\text{cm}$ (about 60%) with a sample size of 15, and for sample sizes of 25 to 60 observations, the XCD05 varied by less than 10% (Fig. 3).

The effect on the XCD05 of the number of sampled sites was evaluated by estimating the median XCD05 and 95% CI from the 1,000 subsamples of the combined data set with resampling for subsampled data sets ranging from 100 to 3,000 sites. As the data set sample sizes increase, the number of species increases 9-fold, and the XCD05 95% CI decreases by more than 2-fold (Fig. 4). The median XCD05 varies little for sample sizes between 100 and 3,000 sites, but the 95% CI around the median decreases from approximately 440 to about 150 $\mu\text{S}/\text{cm}$. The combined data set was considered adequate for characterizing the distribution of 101 species represented in the XCD.

Distribution of XC95 values

If one assumes that the XCD represents a model of the range of SC tolerance of fish in Appalachian streams, then 5th percentile of the XCD, an XCD05, is a benchmark that estimates extirpation of 5% of fish species at 509 $\mu\text{S}/\text{cm}$ with a 95% CI of 355–534 $\mu\text{S}/\text{cm}$ (Fig. 5). The uncertainty around the XCD05 value was estimated using bootstrapping (Fig. 6). The CI captures uncertainty associated with measurement of SC and the collection and enumeration of fish. Variance due to differences in stream reaches, weather, and other random factors is also included. The confidence bounds in the analyses for fish may also characterize some additional potential systematic sources of variance, such as geographic area and different sampling protocols and organizations performing the sampling. The distribution of XC95 values is more reliable in the lower half of the distribution (Figure 5). The XC95 values in the upper half are uncertain and these values are designated as greater than the assigned XC95 value.

Comparison of species and genus extirpation concentrations

We also calculated the genus XC95 values and compared those values to the species XC95 values (Table 4). Each genus XC95 tends toward the high end of the range of XC95 values of the species in that genus (Fig. 7). For example, the XC95 for the genera *Etheosoma*, *Noturus*, *Esox*, and *Cottus* occurred in the XCD at an SC XC95 of the most salt-tolerant species in each genus (*E. baileyi*, *N. insignis*, *E. niger*, *C. girardi*, respectively). However, the genus XCD05 for fish was only slightly higher (545 $\mu\text{S}/\text{cm}$) than the species XCD05 (509 $\mu\text{S}/\text{cm}$), probably because many of the more salt-intolerant genera in the genus XCD were represented by only one species.

Threshold of decline

In the GAM plots, the scatter plots depict the probability of occurrence of a fish species given a stream's SC. The maximum probability of occurrence of a species is in streams with an SC at the apex of the GAM plot (Supplementary Data Fig. S3). Occurrences of a species become less probable as SC increases and, at times, at much lower SC values than the XC95 value. In streams measuring 509 $\mu\text{S}/\text{cm}$, the fish XCD05, 41 of 101 of the fish species decrease in probability of occurrence. Therefore, these plots can also demonstrate effects of increased exposure to ion mixtures prior to extirpation.

DISCUSSION

Natural and anthropogenic agents can locally extirpate organisms [48, 49]. Increasing ionic concentrations are known to affect physiological processes required for sustainable populations [2, 7]. A method for identifying the intensity of a stressor that can lead to extirpation was developed using aquatic benthic invertebrate data [15]. With minor modification, we determined that the method could be used to calculate potential extirpation values of fish from streams. However, interpretation of the results needs to take into account differences between fish and invertebrates with respect to sampling and natural history.

The combined data used in these analyses were collected from a study region in the mid-Atlantic Highlands from eastern Kentucky north to northwestern New Jersey. While the 4 ecoregions have similar sedimentary geology and sources of dissolved ions in freshwaters, the study region includes 5 river basins that each drain separately to the Chesapeake Bay or Atlantic Ocean (i.e., Delaware, Susquehanna, Potomac, James, and Roanoke rivers) along with 2 major tributaries of the Mississippi River (i.e., Tennessee and Ohio rivers). Fish that are primary freshwater species are well known to currently have a limited ability to naturally disperse among major river systems, particularly those that drain separately to the ocean, and the distributions of some fish species collected are limited to one or more, but not all, of these river basins [20, 21]. Moreover, the distribution of individual fish species is often dependent on stream size [44]. Therefore, given our definition of the XC95, we considered it appropriate and logical to limit the analysis of each species to those sites where the species was expected to occur based on biogeography and stream size [3]. Although this modification removes potential concern of effects from the weighting process, in practice, the choice of weighting procedure had little influence on the XCD05s. The XCD05 for the customized fish weighting procedure and from the entire combined data set weighting procedure were 509 and 518 $\mu\text{S}/\text{cm}$, respectively, because species near the 5th percentile occurred in most basins and so weighting was the same. This similarity may not hold for species in other areas or data sets.

The lowest XC95 for a fish species, *Etheostoma bailey*, in our data was 322 $\mu\text{S}/\text{cm}$. Thus, the benchmark from the genus SC derived from macroinvertebrates in the central Appalachians of 300 $\mu\text{S}/\text{cm}$ [16] would be protective of fish. At 509 $\mu\text{S}/\text{cm}$, SC, one would predict extirpation of 5% and reduced probability of observing 40% of fish species in a stream at that SC. Although no stream supports all of the species included in the XCD; proportionally, one might expect a similar distribution of XC95 values. For small streams with few species, the XCD may be less representative and individual species XC95 values may be better predictors of potential extirpation.

Fish are more easily identified to species than benthic macroinvertebrates [1] and this allowed us to compare fish species and genus XC95 values. We found significant variation in salt tolerance among species within genera. For example, the XC95 values among the 3 species of the madtom genus *Noturus* range from 349 $\mu\text{S}/\text{cm}$ to >2,578 $\mu\text{S}/\text{cm}$ (Fig. 7). The 10 species of the darter genus *Etheostoma* range from 322 $\mu\text{S}/\text{cm}$ to >4,000 $\mu\text{S}/\text{cm}$. Moreover, genus XC95 values tend toward the high end of the range of species XC95 values within each genus. Thus, the genus XC95 is not necessarily predictive of species-level

effects, and inferences based on genus are biased toward higher SC. This bias is expected because the XC95 is not a central tendency of a distribution but near the upper range of the distribution (i.e., the 95th percentile). For example, there are 4 species in the genus *Cottus*, 3 of which have XC95 values near 500 $\mu\text{S}/\text{cm}$ and 1 near 2,000 $\mu\text{S}/\text{cm}$; as a result, the 95th percentile of the combined observations yields the genus XC95 near 2,000 $\mu\text{S}/\text{cm}$. Therefore, the XC95 is dependent on the occurrence of the most tolerant species in the genus, including those that were not numerous enough to confidently calculate a distinct species XC95.

These results have implications for the derivation of benchmarks or criteria for chemicals that allow for some niche partitioning across the gradient of naturally occurring concentrations. For example, U.S. EPA guidance for developing water quality criteria from laboratory data uses the geometric mean of species within a genus and develops a distribution from those genus mean effect levels [50]. For xenobiotic chemicals often tested with laboratory bioassays, there is usually less variation between congeneric species and using a geometric mean may be appropriate for these types of toxicants [50, 51]. However, for ionic mixtures using the genus level, SC effect values would be less protective than the species level, at least in the case of fish, because the variation in tolerance to increased SC among species is large and each XC95 represents an upper threshold rather than a central tendency.

Laboratory and field studies are complementary but not directly comparable. A search of the ECOTOX database [52] of toxicity data for major ions or their salts found data for 47 fish species, but almost 40% (18) were not native to North America (not including common carp or brown trout as non-native) and may not be representative of streams with naturally low background SC as occurs in Appalachia. Also, the toxicity data in the ECOTOX database are not easily comparable to our data, because the ECOTOX data are for single ions or salts and not for mixtures of ions and are measured in $\mu\text{g}/\text{L}$ rather than $\mu\text{S}/\text{cm}$. Exposure times and effect endpoints are very different for laboratory and field studies. Therefore, both are informative, but not equivalent.

Physiological studies have described epithelial cellular processes that affect the tolerance of freshwater fish to elevated ionic concentrations [7]. However, the effect gradients of decreasing species occurrence (Supplementary Material Fig. S3) and the extirpation of species (Fig. 5) may not be entirely due to direct physiological effects. These observed effects in streams may also be caused by changes in food and habitat resources because of changes in algal [53] and invertebrate [2, 54] assemblages that are also affected by increased exposure to dissolved ions or by indirect physiological or behavioral changes [55, 56].

Although fish appear to be generally more tolerant of SC than macroinvertebrates, this may be related to our ability to observe fish throughout the year rather than to actual differences in salt-tolerance. Most salt-intolerant benthic macroinvertebrates are observed for a shorter period of the year, because they are univoltine, reproducing and surviving within a single year. Salt-intolerant benthic macroinvertebrates were primarily observed during the spring [15], because many aquatic insects complete their life cycle by early summer and pass the summer, often in diapause, as eggs or early instar larvae that are too small to sample

effectively [31, 57]. Fish were more often sampled in June through September (Table 2) when SC levels tend to be higher [13, 16]; therefore, effect endpoints may represent a maximum exposure compared to a level tolerated for an annual average as estimated for benthic invertebrate XC95 values. This is also one of the reasons that fish and invertebrate XC95 values were not combined into a single XCD. For these reasons, direct comparison of the salt-intolerant fish and macroinvertebrate XC95 values and XCDs are intended to be illustrative and should be interpreted cautiously because they possibly represent different exposure conditions.

CONCLUSIONS

This research confirmed that field-based data could be used to characterize the tolerance and extirpation thresholds for fish species. Interestingly, it showed that fish and invertebrate assemblages have a very similar range of specialization for natural niches influenced by ionic concentration. Of the estimated aquatic insect XC95 values [15], 8% were lower than the lowest estimated XC95 for fish (322 $\mu\text{S}/\text{cm}$). This result may be related to the ability to observe fish during the SC annual maximum in this area during late July to September; whereas, many salt-intolerant insects are too small to be collected at that time of year. Therefore, the XC95 for fish may be an annual maximum whereas for salt intolerant invertebrates the XC95 may be an annual exposure value. Also, the number of invertebrates and fish included in the XCD differed. The HCO5 for the invertebrates was based on 163 genera. At 163 fish, the XCD05 could be less than 350 $\mu\text{S}/\text{cm}$ (Figure 3). Therefore, the fish and invertebrate benchmarks are not directly comparable, but are informative how different taxonomic groups respond to ionic concentration.

Methodological assessment to fine-tune the weighting associated with uneven sampling along SC exposure range allayed concern about working with biogeographically constrained species of fish. Fish are easier to identify to species than benthic macroinvertebrates. Thus, it was easier to identify species unlikely to occur at particular sites because of biogeography [21] or stream size [44], rather than SC. Therefore, we used subsets of the data set to analyze single fish species to exclude sites where a species did not occur because of biogeography or stream size. Although this more complicated calculation does offer more confidence in the analysis, biogeographical weighting and the original weighting procedures produce XCD05s that are within the 95% CI. Therefore, if computational resources are lacking, the simpler weighted analysis is still informative.

The comparison of species and genus XC95 values showed that genus XC95 values are dependent on the occurrence of the most tolerant species in the genus. Therefore, the genus-level XC95 values tend toward the upper range of XC95 values of the species in that genus. However, for the data set analyzed herein, 5% of salt-intolerant genera were represented by a single species, and therefore, there was little difference between the species- (509 $\mu\text{S}/\text{cm}$) and the genus-level (545 $\mu\text{S}/\text{cm}$) XCD05.

The XCD05 of 509 $\mu\text{S}/\text{cm}$ for fish species from these analyses and even the lowest XC95 (i.e., 322 $\mu\text{S}/\text{cm}$) for the darter *Etheostoma baileyi* were greater than the benchmark based on macroinvertebrates for Ecoregions 69 and 70 (300 $\mu\text{S}/\text{cm}$) [16] or example criteria for

Ecoregions 69 (310 $\mu\text{S}/\text{cm}$), although Ecoregion 70 was slightly higher (340 $\mu\text{S}/\text{cm}$) [18]. Therefore, the thresholds based on macroinvertebrates would also be as protective against extirpation of most fish due to increases in this salt mixture in this region of the central and southern Appalachians. However, additional studies might be able to elucidate whether individual species are limited directly by ion concentration or by an indirect pathway such as changes in food resources, predation, or competition.

At 509 $\mu\text{S}/\text{cm}$, 40% of fish species are less likely to occur because this value exceeds the species' optimal SC range. A 40% decline is a serious effect and may be a harbinger of extirpation of fish species in a stream or larger drainage. For example, at 300 $\mu\text{S}/\text{cm}$, there is about a 20% probability of brook trout occurring in a stream compared to a >90% probability at the species' optimum of <20 $\mu\text{S}/\text{cm}$. For some species like brook trout, which are particularly valued, an XC95 value may not be socially acceptable and a different assessment endpoint may be preferred.

Having distinct estimates of the SC affecting many species of fish and macroinvertebrate genera from field estimates has many advantages. Most importantly, it makes assessing each group possible and does not limit the assessment to the least distinct level of taxonomic identification (i.e., combining fish and invertebrates at the genus level, which we have shown can be biased toward the more tolerant species in a genus). Rather, it allows assessment at the lowest taxonomic level. This suggests that it would also be advantageous to identify invertebrates to species [58]. Other potentially informative areas of research include investigating physiological traits of the most salt-intolerant invertebrate and fish species, and examining the effect of key functions or groups that would be affected by changes in SC.

We believe that these species-level exposure-response curves and XC95 values makes it possible to assess the potential effects of increased ionic concentration on distinct fish species, including reductions in their occurrence. We expect that these exposure-response relationships will be useful for causal assessments for fish, just as they have been useful for causal assessments of benthic invertebrates [45].

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We would like to thank G.J. Pond and L. Reynolds for their assistance obtaining some of the data sets and the staff of the Kentucky Department of Environmental Protection, Division of Water; West Virginia Department of Natural Resources and Department of Environmental Protection, Watershed Branch; Ohio Environmental Protection Agency, Division of Surface Water, Ecological Assessment Branch; and Pennsylvania Fish and Boat Commission, who conducted the fish surveys and produced some of the data sets used in these analyses. The other data sets were produced by U.S. EPA's Office of Research and Development and Region 3 with cooperating state agencies and universities as part of EMAP and the Programmatic Environmental Impact Statement for mountaintop mining and valley fills. Earlier drafts of the manuscript were reviewed and improved by comments from C. Flaherty (U.S. EPA, OW, Washington, DC), G. Suter (U.S. EPA, ORD, Cincinnati, OH), and L. Reynolds (U.S. EPA, Region 3, Wheeling, WV). L. Tackett and K. Bland (Highlight Technologies, LLC) assisted in the preparation of the manuscript for publication. The present study is based on work supported by the U.S. EPA.

References

1. U.S. Environmental Protection Agency. The effects of mountaintop mines and valley fills on aquatic ecosystems of the central Appalachian coalfields. National Center for Environmental Assessment, Office of Research and Development, U.S. Environmental Protection Agency; Washington, D.C.: 2011. p. i-xc.p. 1-141.
2. Cormier SM, Suter GW, Zheng L, Pond GJ. Assessing causation of the extirpation of stream macroinvertebrates by a mixture of ions. *Environ Toxicol Chem.* 2013; 32:277–287. [PubMed: 23147750]
3. Pond GJ, Passmore ME, Borsuk FA, Reynolds L, Rose CJ. Downstream effects of mountaintop coal mining: comparing biological conditions using family- and genus-level macroinvertebrate bioassessment tools. *J N Am Benthol Soc.* 2008; 27:717–737.
4. Palmer MA, Bernhardt ES, Schlesinger WH, Eshleman KN, Foufoula-Georgiou E, Hendryx MS, Lemly AD, Likens GE, Loucks OL, Power ME, White PS, R WP. Mountaintop mining consequences. *Science.* 2010; 327:148–149. [PubMed: 20056876]
5. Evans DH. Teleost fish osmoregulation: what have we learned since August Krogh, Homer Smith, and Ancel Keys. *Am J Physiol Regul Integr Comp Physiol.* 2008; 295:R704–R713. [PubMed: 18525009]
6. Perry SF, Shahsavarani A, Georgalis T, Bayaa M, Furimsky M, Thomas SLY. Channels, pumps, and exchangers in the gill and kidney of freshwater fishes: Their role in ionic and acid-base regulation. *J Exp Zool Part A.* 2003; 300A:53–62.
7. Griffith MB. Toxicological perspective on the osmoregulation and ionoregulation physiology of major ions by freshwater animals: teleost fish, Crustacea, aquatic insects, and Mollusca. *Environ Toxicol Chem.* 2017; 36:576–600. [PubMed: 27808448]
8. Evans, DH. Osmotic and ionic regulation by freshwater and marine fishes. In: Ali, MA., editor. *Environmental Physiology of Fishes.* Plenum Press; New York, NY: 1980. p. 93-122.
9. Gouaux E, MacKinnon R. Principles of selective ion transport in channels and pumps. *Science.* 2005; 310:1461–1465. [PubMed: 16322449]
10. Tseng Y-C, Hwang P-P. Some insights into energy metabolism for osmoregulation in fish. *Comp Biochem Physiol C Toxicol Pharmacol.* 2008; 148:419–429. [PubMed: 18539088]
11. Bøeuf G, Payan P. How should salinity influence fish growth? *Comp Biochem Physiol C Toxicol Pharmacol.* 2001; 130:411–423. [PubMed: 11738629]
12. Evans DH. Kinetic studies of ion transport by fish gill epithelium. *Am J Physiol Regul Integr Comp Physiol.* 1980; 238:R224–R230.
13. Griffith MB. Natural variation and current reference for specific conductivity and major ions in wadeable streams of the conterminous USA. *Freshw Sci.* 2014; 33:1–17.
14. Griffith MB, Norton SB, Alexander LC, Pollard AI, LeDuc SD. The effects of mountaintop mines and valley fills on the physicochemical quality of stream ecosystems in the central Appalachians: A review. *Sci Total Environ.* 2012; 417:1–12. [PubMed: 22264919]
15. U.S. Environmental Protection Agency. EPA/600/R-10/023F. U.S. Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment; Washington, DC: 2011. A field-based aquatic life benchmark for conductivity in central Appalachian streams.
16. Cormier SM, Suter GW, Zheng L. Derivation of a benchmark for freshwater ionic strength. *Environ Toxicol Chem.* 2013; 32:263–271. [PubMed: 23161648]
17. Cormier SM, Suter GW II. A method for deriving water-quality benchmarks using field data. *Environ Toxicol Chem.* 2013; 32:255–262. [PubMed: 23147651]
18. U.S. Environmental Protection Agency. EPA/822/R-07/010. U.S. Environmental Protection Agency, Office of Water; Washington, DC: 2017. Draft Field-based methods for developing aquatic life criteria for specific conductivity.
19. U.S. Environmental Protection Agency. EPA/630/P-02/004B. U.S. Environmental Protection Agency, Risk Assessment Forum; Washington, DC: 2003. Generic ecological assesment endpoints (GEAEs) for ecological risk assessment.

20. Matamoros WA, Hoagstrom CW, Schaefer JF, Kreiser BR. Fish faunal provinces of the conterminous United States of America reflect historical geography and familial composition. *Biol Rev.* 2016; 91:813–832. [PubMed: 26031190]
21. Stauffer JR, Boltz JM, White LR. The fishes of West Virginia. *Proc Acad Nat Sci Phila.* 1995; 146:1–389.
22. U.S. Environmental Protection Agency. Level III ecoregions of the continental United States (revision of Omernik, 1987). U.S. Environmental Protection Agency, Office of Research and Development, National Health and Environmental Effects Research Laboratory; Corvallis, OR: 2013. April 2013 ed.pp map scale 1:7,500,000
23. Omernik JM. Ecoregions of the conterminous United States. *Ann Assoc Am Geogr.* 1987; 77:118–125.
24. Wilken, E., Nava, FJ., Griffith, G. North American Terrestrial Ecoregions - Level III. Commission for Environemental Cooperation; Montreal, QC, Canada: 2011.
25. Fenneman, NM. Physiography of Eastern United States. 1st. McGraw-Hill Book Company, Inc.; New York, NY: 1938.
26. Herlihy AT, Larsen DP, Paulsen SG, Urquhart NS, Rosenbaum BJ. Designing a spatially balanced, randomized site selection process for regional stream surveys: The EMAP Mid-Atlantic pilot study. *Environ Monit Assess.* 2000; 63:95–113.
27. Detenbeck NE, Cincotta D, Denver JM, Greenlee SK, Olsen AR, Pitchford AM. Watershed-based survey designs. *Environ Monit Assess.* 2005; 103:59–81. [PubMed: 15861987]
28. Smithson, J. West Virginia stream/river survey design 2007–2111. West Virginia Department of Environmental Protection, Division of Water and Waste Management; Charleston, WV: 2007.
29. Stauffer, JR., Ferreri, CP. Characterization of stream fish assemblages in selected regions of mountain top removal/valley fill coal mining. - Draft programmatic environmental impact statement on mountaintop mining/valley fills in Appalachia (2003). Appendix D. U.S. Environmental Protection Agency, Region 3; Philadelphia, PA: 2002.
30. Stoddard JL, Larsen DP, Hawkins CP, Johnson RK, Norris RH. Setting expectations for the ecological condition of streams: The concept of reference condition. *Ecol Applic.* 2006; 16:1267–1276.
31. Lazorchak, JM.Klemm, DJ., Peck, DV., editors. Environmental monitoring and assessment program - Surface waters: Field operations and methods for measuring the ecological condition of wadeable streams. U.S. Environmental Protection Agency, Office of Research and Development; Washington, DC: 1998.
32. U.S. Environmental Protection Agency. EPA-600/4-87-026. U.S. Environmental Protection Agency, Office of Research and Development; Washington, DC: 1987. Handbook of methods for acid deposition studies: Laboratory analyses for surface water chemistry.
33. Kentucky Department of Environmental Protection. In situ water quality measurements and meter calibration standard operating procedures. Kentucky Department of Environmental Protection, Division of Water; Frankfort, KY: 2009.
34. Kentucky Department of Environmental Protection. Sampling the surface water quality of lotic systems. Kentucky Department of Environmental Protection, Division of Water; Frankfort, KY: 2009.
35. Kentucky Department of Environmental Protection. Standard operation procedure: Laboratory procedures for fish processing and taxonomic identification. Kentucky Department of Environmental Protection, Division of Water; Frankfort, KY: 2009.
36. Kentucky Department of Environmental Protection. Collection methods for fish in wadeable streams, Version 2.1. Kentucky Department of Environmental Protection, Division of Water; Frankfort, KY: 2010.
37. Bryant, G., McPhilliamy, S., Childers, H. A survey of the water quality of streams in the primary region of mountaintop/valley fill coal mining, October 1999 to January 2001.-Draft programmatic environmental impact statement on mountaintop mining/valley fills in Appalachia (2003). Appendix D. U.S. Environmental Protection Agency, Region 3; Philadelphia, PA: 2002.
38. West Virginia Department of Environmental Protection. 2009 Standard operating procedures. West Virginia Department of Environmental Protection; Watershed Branch, Charleston, WV: 2009.

39. Ohio Environmental Protection Agency. Biological criteria for the protection of aquatic life: Volume III. Standardized biological field sampling and laboratory methods for assessing fish and macroinvertebrate communities. Ohio Environmental Protection Agency, Division of Surface Water, Ecological Assessment Section; Columbus, OH: 1989.
40. Ohio Environmental Protection Agency. Surface water field sampling manual for water column chemistry, bacteria and flows. Ohio Environmental Protection Agency, Division of Surface Water, Ecological Assessment Section; Columbus, OH: 2013.
41. Ohio Environmental Protection Agency. 2013 updates to Biological criteria for the protection of aquatic life: Volume III. Standardized biological field sampling and laboratory methods for assessing fish and macroinvertebrate communities. Ohio Environmental Protection Agency, Division of Surface Water, Ecological Assessment Section; Columbus, OH: 2013.
42. Pennsylvania Department of Environmental Protection. Wadeable semi-quantitative fish sampling protocol for streams. Pennsylvania Department of Environmental Protection, Bureau of Point and Non-point Source Management; Harrisburg, PA: 2013.
43. U.S. Environmental Protection Agency. 440/5-86/001 E. Office of Water, U.S. Environmental Protection Agency; Washington, D.C: 1986. Quality criteria for water 1986; p. 1-396.
44. McCormick FH, Hughes RM, Kaufmann PR, Peck DV, Stoddard JL, Herlihy AT. Development of an index of biotic integrity for the Mid-Atlantic Highlands region. *Trans Amer Fish Soc.* 2001; 130:857–877.
45. Coffey DB, Cormier S, Harwood J. Using field-based species sensitivity distributions to infer multiple causes. *Hum Ecol Risk Assess.* 2014; 20:402–432.
46. Efron, B., Tibshirani, RJ. *An Introduction to the Bootstrap.* Capman & Hall/CRC; Boca Raton, FL: 1993.
47. R Development Core Team. *R: A language and environment for statistical computing.* R Foundation for Statistical Computing; Vienna, Austria: 2011.
48. Norton SB, Cormier SM, Suter GW, Subramanian B, Lin E, Altfater D, Counts B. Determining probable causes of ecological impairment in the Little Scioto River, Ohio, USA: Part 1. Listing candidate causes and analyzing evidence. *Environ Toxicol Chem.* 2002; 21:1112–1124. [PubMed: 12069294]
49. Haake DM, Wilton T, Krier K, Stewart AJ, Cormier SM. Causal assessment of biological impairment in the Little Floyd River, Iowa, USA. *Hum Ecol Risk Assess.* 2010; 16:116–148.
50. Stephan, CE., Mount, DL., Hansen, DJ., Gentile, JH., Chapman, GA., Brungs, WA. Guidelines for deriving numerical national water quality criteria for the protection of aquatic organisms and their uses. U.S. Environmental Protection Agency, Office of Research and Development; Washington, DC: 1985. p. 106
51. Nyman A-M, Schirmer K, Ashauer R. Importance of toxicokinetics for interspecies variation in sensitivity to chemicals. *Environ Sci Technol.* 2014; 48:5946–5954. [PubMed: 24758734]
52. U. S. Environmental Protection Agency. ECOTOX User Guide: ECOTOXicology Knowledgebase System. 4.0. Washington, DC: 2016.
53. Potapova M, Charles DF. Distribution of benthic diatoms in U.S. rivers in relation to conductivity and ionic composition. *Freshwater Biol.* 2003; 48:1311–1328.
54. Hitt NP, Floyd M, Compton M, McDonald K. Threshold Responses of Blackside Dace (*Chrosomus cumberlandensis*) and Kentucky Arrow Darter (*Etheostoma spilotum*) to Stream Conductivity. *Southeast Nat.* 2016; 15:41–60.
55. Killen SS, Marras S, Metcalfe NB, McKenzie DJ, Domenici P. Environmental stressors alter relationships between physiology and behaviour. *Trends Ecol Evol.* 2013; 28:651–658. [PubMed: 23756106]
56. Scott GR, Sloman KA. The effects of environmental pollutants on complex fish behaviour: integrating behavioural and physiological indicators of toxicity. *Aquat Toxicol.* 2004; 68:369–392. [PubMed: 15177953]
57. Barbour, MT., Gerritsen, J., Snyder, BD., Stribling, JB. Rapid bioassessment protocols for use in streams and wadeable rivers: Periphyton, benthic macroinvertebrates and fish. 2nd. U.S. Environmental Protection Agency, Office of Water; Washington, DC: 1999.

58. Yang J, Zhang X, Xie Y, Song C, Sun J, Zhang Y, Giesy JP, Yu H. Ecogenomics of zooplankton community reveals ecological threshold of ammonia nitrogen. *Environ Sci Technol*. 2017; 51:3057–3064. [PubMed: 28191935]

Fig. 1.

The fish sampling locations ($N = 3,465$) are from Level III Ecoregions 67, 68, 69, and 70 spanning the states of Kentucky, West Virginia, Virginia, Ohio, Maryland, Pennsylvania, and New Jersey. State outlines from the U.S. Environmental Protection Agency (U.S. EPA) Base Map Shapefile. Omernik Level III Ecoregions are from National Atlas (NationalAtlas.gov) Projection NAD1982UTM17N.

Fig. 2. Histogram of the overall sampling frequencies of observed SC values in samples from the combined data set from Ecoregions 67, 68, 69, and 70 from March through November. Histograms are customized for each fish species before assigning weights. More of the sampled sites are near the median than at the extremes.

Fig. 3. Relationship between the minimum number of observations for inclusion of a species on the number of species (red squares, right y-axis) included in the extirpation concentration distribution (XCD) and on the 5th percentile of the cumulative distribution of XC95 (XCD05) (blue circles, left y-axis).

Fig. 4.

Adequacy of the number of samples used to model the 5th percentile of the cumulative distribution of XC95 (XCD05). As sample size increases, the number of species (red squares, red vertical lines are the 95% confidence intervals (CI), right y-axis) included in the extirpation concentration distribution (XCD) increase and the XCD05 values (blue circles, blue vertical lines are the 95% CI, left y-axis) decrease. The mean XCD05 95% CI becomes fairly constant when 1,000 sites and 75–90 species are evaluated.

Fig. 5.

The species extirpation concentration distribution (XCD) for fish. The 5th percentile of the cumulative distribution of XC95 (XCD05) is the specific conductivity (SC) (509 $\mu\text{S}/\text{cm}$) at the intercept of the XCD with the horizontal, gray dashed line at the 5th percentile. The black dots represent species for which the calculated extirpation concentrations at the 95th percentile (XC95) values are not qualified, while gray dots are species for which the calculated XC95 values are qualified as “approximately,” and open dots are species for which the calculated XC95 values are qualified as “greater than.” The XCD is truncated to allow the names to be listed for the first 41 species.

Fig. 6.

Cumulative distribution of the first 25% of the extirpation concentrations at the 95th percentile (XC95) values (blue circles) and the 95% confidence interval (CI) (dotted lines) based on 1,000 species extirpation concentration distribution (XCD) bootstrapping results. The small gray dots represent XC95 values for a bootstrapping iteration (note that the species in each percentile may vary with each species XCD iteration). The larger dark dots represent the calculated XC95 values in the species XCD. The median bootstrapped 5th percentile of the cumulative distribution of XC95 (XCD05) is 456 μS/cm with 95% CI of 355–534 μS/cm.

Fig. 7.

The genus extirpation concentration distribution (XCD) for fish. The genus extirpation concentrations at the 95th percentile (XC95) values (open circles) for fish genera observed 25 times are plotted along with the species XC95 value (small solid circles) for fish species observed 25 times, although some species are obscured by plotting at the same location. For visualization, the horizontal lines connect fish species in the same genus. The 5th percentile of the cumulative distribution of XC95 (XCD05) is the specific conductivity (SC) (545 µS/cm) at the intercept of the genus XCD with the horizontal gray dashed line at the 5th percentile. The black circles represent species or genera where the calculated XC95 values are either not qualified or qualified as “approximately,” and the gray circles are species or genera for which the calculated XC95 values are qualified as “greater than.” The genus XCD is truncated to allow the names to be listed for the first 35 genera.

Table 1

Data sets that were combined to calculate the XC95 values for fish in Level III Ecoregions 67, 68, 69, and 70.

Data set	Conducted or collected by	Sampling period	No. of sites
1. The Mid Atlantic Highlands Assessment	U.S. EPA's Environmental Monitoring and Assessment Program	1993–1996	172
2. The Mid Atlantic Integrated Assessment	U.S. EPA's Environmental Monitoring and Assessment Program	1997–1998	119
3. Fish bioassessment data	Kentucky Department for Environmental Protection, Division of Water stream bioassessment program	1991–2004	285
4. Fish and chemistry data	Stauffer and Ferreri [29] and Bryant, McPhilliamy and Childers [37] for the Programmatic Environmental Impact Statement for mountaintop mining and valley fills	1999–2001	34
5. Fish and chemistry data	U.S. EPA's Regional Applied Research Effort program in cooperation with the West Virginia Department of Natural Resources [27]	2001–2002	118
6. Fish bioassessment data	West Virginia Department of Environmental Protection pilot program	2007–2009	43
7. Fish bioassessment data	Ohio Environmental Protection Agency, Division of Surface Water	1999–2013	593
8. Fish assessment data	Pennsylvania Fish and Boat Commission (PFBC)	1990–2014	2,101

Table 2

Number of samples with reported fish species and specific conductivity (SC) meeting the acceptance criteria for calculating the hazardous concentration (HC05). The number of sites is presented for each month and ecoregion.

Ecoregion	Month												Total
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
67	0	2	89	75	85	383	325	325	97	23	12	3	1,419
68	0	0	0	3	0	4	9	15	6	2	0	0	39
69	1	0	12	51	33	175	231	170	70	60	5	0	808
70	1	0	7	9	29	237	332	250	93	52	0	1	1,011
Total	2	2	108	138	147	799	897	760	266	137	17	4	3,277

Table 3

Summary statistics of the water quality variables from the eight data sets.

Variable	Min	25 th percentile	Median	75 th percentile	Max	Mean	Valid <i>n</i>
Sp cond (µS/cm)	9.4	84.0	217	430	4,000	328	3,277
Hardness (mg/L)	0.00	20.0	42.0	118	772	83.4	1,488
Alkalinity (µeq/L)	6.28	983	1,960	3,160	7,670	2,120	995
HCO ₃ ⁻ (µeq/L)	0.00	887	1,910	3,120	7,680	2,060	1,014
SO ₄ ²⁻ (µeq/L)	44.4	365	1,000	3,160	52,900	3,240	1,014
Ca ²⁺ (µeq/L)	29.9	1,100	2,150	3,660	18,300	2,900	1,029
Mg ²⁺ (µeq/L)	28.8	637	1,150	1,970	21,600	1,810	917
Na ⁺ (µeq/L)	4.35	223	478	1,070	27,900	1,160	877
K ⁺ (µeq/L)	6.39	51.2	76.7	102	1,240	87.0	872
Cl ⁻ (µeq/L)	0.726	139	310	673	8,610	587	1,035
Total Fe (µg/L)	1.00	10.0	36.3	110	2,690	143	369
NO ₃ ⁻ (µg/L)	6.00	125	298	794	875,000	2,270	1,099
Total N (µg/L)	45.0	210	436	860	875,000	2,400	956
Total Al (µg/L)	1.00	6.00	16.0	31.0	1,060	52.8	360
Total Mn (µg/L)	1.10	10.0	20.0	82.0	2,090	82.6	367
Total P (µg/L)	1.0	6.0	13.0	24.0	971	28.1	532
Total Se (µg/L)	1.0	1.0	2.0	3.0	1,300	98.9	85
Dissolved O ₂ (mg/L)	1.2	7.3	8.6	9.6	18.6	8.5	822
pH	6.00	6.90	7.31	7.80	9.50	7.36	3,190
Water temperature (°C)	0.4	14.0	17.0	19.7	31.0	16.7	2,601
RBP habitat score	38	75	114	139	191	111	801
Catchment area (km ²)	0.111	11.47	28.79	88.70	18,640	272	1,280

Sp cond = specific conductivity, RBP = Rapid Bioassessment Protocol [31]

Table 4

Extirpation concentration (XC95) values for fish that were observed at a minimum of 25 sites. Rank is the order of the fish species from smallest to greatest species XC95 in the species sensitivity distribution. The XC95 is listed as approximate ($\approx$) if the generalized additive model (GAM) mean curve at maximum specific conductivity is greater than 0 but the lower confidence limit is approximately 0 (<1% of the maximum mean modeled probability). The XC95 is listed as greater than (>), if the GAM lower confidence limit is greater than 0.

Rank	Species	Species XC95	Genus XC95	Rank	Species	Species XC95	Genus XC95
1	<i>Etheostoma baileyi</i>	322	>3,226	51	<i>Lepomis gulosus</i>	>1,958	-
2	<i>Noturus insignis</i>	349	>2,578	52	<i>Phenacobius mirabilis</i>	>2,000	>2,000
3	<i>Erimyzon oblongus</i>	376	376	53	<i>Clinostomus elongatus</i>	>2,009	-
4	<i>Esox niger</i>	$\approx$ 467	>1,572	54	<i>Cottus bairdii</i>	>2,046	-
5	<i>Salvelinus fontinalis</i>	508	508	55	<i>Aplodinotus grunniens</i>	>2,099	>2,099
6	<i>Cottus girardi</i>	$\approx$ 518	>1,961	56	<i>Etheostoma camurum</i>	>2,122	-
7	<i>Clinostomus funduloides</i>	535	>1,790	57	<i>Lepomis gibbosus</i>	>2,157	-
8	<i>Cottus caroliniae</i>	542	_a	58	<i>Notropis volucellus</i>	>2,122	-
9	<i>Cottus cognatus</i>	$\approx$ 557	-	59	<i>Pylodictis olivaris</i>	>2,122	>2,122
10	<i>Nocomis leptocephalus</i>	$\approx$ 565	>2,303	60	<i>Notropis atherinoides</i>	>2,157	-
11	<i>Etheostoma kennicotti</i>	$\approx$ 586	-	61	<i>Pomoxis annularis</i>	>2,278	-
12	<i>Chrosomus oreas</i>	592	$\approx$ 3,094	62	<i>Nocomis micropogon</i>	>2,303	-
13	<i>Notropis telescopus</i>	675	>4,000	63	<i>Lampetra aepyptera</i>	2,323	2,323
14	<i>Cyprinella analostana</i>	682	>4,000	64	<i>Notropis buccatus</i>	>2,323	-
15	<i>Margariscus margarita</i>	>685	>685	65	<i>Percina caprodes</i>	>2,359	-
16	<i>Lythrurus fasciolaris</i>	707	1,081	66	<i>Lepomis megalotis</i>	2,578	-
17	<i>Luxilus cornutus</i>	$\approx$ 724	>4,000	67	<i>Etheostoma nigrum</i>	>2,578	-
-	<i>Erimystax</i> spp.	_b	744	68	<i>Etheostoma variatum</i>	>2,578	-
18	<i>Fundulus diaphanus</i>	759	1,090	69	<i>Moxostoma duquesnei</i>	>2,578	>2,578
19	<i>Salmo trutta</i>	$\approx$ 759	$\approx$ 759	70	<i>Moxostoma erythrurum</i>	>2,578	-
20	<i>Exoglossum maxillingua</i>	>760	576	71	<i>Noturus flavus</i>	>2,578	-
21	<i>Percina peltata</i>	$\approx$ 822	>2,578	72	<i>Pimephales vigilax</i>	>2,578	-
22	<i>Lepomis auritus</i>	851	>2,750	73	<i>Micropterus salmoides</i>	>2,630	>3,066
23	<i>Cyprinella whipplei</i>	854	-	74	<i>Notropis rubellus</i>	>2,630	-
24	<i>Etheostoma olmstedii</i>	>898	-	75	<i>Micropterus dolomieu</i>	>2,641	-

Rank	Species	Species XC95	Genus XC95	Rank	Species	Species XC95	Genus XC95
25	<i>Anguilla rostrata</i>	>898	>898	76	<i>Dorosoma cepedianum</i>	>2,750	>2,750
26	<i>Hybopsis amblops</i>	≈982	≈982	77	<i>Ictalurus punctatus</i>	>2,750	>2,750
27	<i>Semotilus corporalis</i>	>1,000	>3,066	78	<i>Labidesthes sicculus</i>	>2,750	>2,750
28	<i>Moxostoma carinatum</i>	1,040	>2,578	79	<i>Lepomis macrochirus</i>	>2,750	–
29	<i>Oncorhynchus mykiss</i>	>1,075	>1,075	80	<i>Catostomus commersoni</i>	>2,755	>2,755
30	<i>Esox lucius</i>	>1,103	–	81	<i>Etheostoma zonale</i>	>3,066	–
31	<i>Percopsis omiscomaycus</i>	≈1,105	≈1,105	82	<i>Semotilus atromaculatus</i>	>3,066	–
32	<i>Noturus miurus</i>	1,150	–	83	<i>Notropis photogenis</i>	>3,066	–
33	<i>Lepisosteus osseus</i>	≈1,170	≈1,170	84	<i>Micropterus punctulatus</i>	≈3,094	–
34	<i>Lythrurus umbratilis</i>	≈1,193	–	85	<i>Chrosomus erythrogaster</i>	>3,094	–
35	<i>Rhinichthys cataractae</i>	≈1,343	>3,535	86	<i>Pimephales notatus</i>	>3,094	–
36	<i>Percina macrocephala</i>	1,351	–	87	<i>Rhinichthys obtusus</i>	>3,094	–
37	<i>Ameiurus nebulosus</i>	≈1,358	>4,000	88	<i>Ambloplites rupestris</i>	>3,266	>3,266
38	<i>Myxetremia melanops</i>	≈1,372	≈1,372	89	<i>Etheostoma flabellare</i>	>3,266	–
39	<i>Notemigonus crysoleucas</i>	≈1,400	≈1,400	90	<i>Lepomis cyanellus</i>	>3,266	–
40	<i>Notropis hudsonius</i>	≈1,400	≈1,400	91	<i>Rhinichthys atratulus</i>	≈3,590	–
41	<i>Pomoxis nigromaculatus</i>	>1,413	>2,278	92	<i>Campostoma anomalum</i>	>3,590	>3,590
42	<i>Perca flavescens</i>	>1,580	>1,580	93	<i>Cyprinus carpio</i>	>3,590	>3,590
43	<i>Esox americanus</i>	>1,625	–	94	<i>Hypentelium nigricans</i>	>3,590	>3,590
44	<i>Percina maculata</i>	≈1,643	–	95	<i>Ameiurus natalis</i>	>4,000	–
45	<i>Carpiodes cyprinus</i>	1,672	1,672	96	<i>Cyprinella spiloptera</i>	>4,000	–
46	<i>Ictiobus bubalus</i>	1,672	1,672	97	<i>Etheostoma blennioides</i>	>4,000	–
47	<i>Moxostoma anisurum</i>	>1,693	–	98	<i>Etheostoma caeruleum</i>	>4,000	–
48	<i>Pimephales promelas</i>	>1,732	>3,094	99	<i>Luxilus chrysocephalus</i>	>4,000	–
49	<i>Etheostoma spectabile</i>	1,824	–	100	<i>Notropis stramineus</i>	>4,000	–
50	<i>Lepomis microlophus</i>	>1,858	–	101	<i>Sander canadensis</i>	>4,000	>4,000

^aHyphens indicates fish species where the genus XC95 is provided for a congeneric species above it in the table.

^bOnly a genus XC95 was calculated for *Erimystax* spp. because none of the 4 species collected in the combined data set (*E. cahni*, *E. insignis*, *E. x-punctatus*, or *E. dissimilis*) were observed in 25 samples, but together, they were observed in 38 samples. All the other information is for the 4 species combined.