

Brief Communication

Step-by-Step Calculation and Spreadsheet Tools for Predicting Stressor Levels that Extirpate Genera and Species

Susan M Cormier,*† Lei Zheng,‡§ Erik W Leppo,‡ and Andrew Hamilton‡

†US Environmental Protection Agency, National Center for Environmental Assessment, Cincinnati, Ohio

‡Tetra Tech, Owings Mills, Maryland, USA

§Noblis, Inc, Windsor Mill, Maryland, USA

ABSTRACT

In 2011, the United States Environmental Protection Agency (USEPA) released a field-based method for estimating the extirpation of freshwater aquatic benthic invertebrates by ionic mixtures dominated by HCO_3^- , SO_4^{2-} , and Ca^{2+} measured as specific conductivity (SC). The estimate of extirpation was SC at the 95th centile (XC95) of a weighted cumulative frequency distribution (CFD) of a genus or species over a range of SC. A CFD of XC95 values was used to predict the SC at which 5% of genera were likely to be extirpated. Because there are many uses for XC95 values and many data sets that could be analyzed using this method, we laid out a step-by-step method for calculating XC95 values and the stressor level that predicts a 5% extirpation of genera (HC05). Although the calculations can be done with a handheld calculator, we developed 2 downloadable Microsoft Excel[®] spreadsheet calculation tools that are easy to use to calculate XC95 values, to plot a taxon's XC95 cumulative frequency distribution with increasing SC, and to plot probabilities of observing a taxon at a particular SC. They also plot cumulative frequency distributions of XC95 values and calculate HC05 values. In addition to the tools, we share an example and the output of XC95 values for 176 distinct aquatic benthic invertebrates in Appalachia, in West Virginia, USA. *Integr Environ Assess Manag* 2017;00:000–000. © 2017 SETAC

Keywords: Extirpation Specific conductivity Streams Appalachia Sensitivity distribution

INTRODUCTION

Recently, ecoepidemiological methods have been developed for estimating the extirpation of freshwater aquatic benthic invertebrates by ionic mixtures measured as specific conductivity (SC) (USEPA 2011, 2016b). The effect is extirpation, the depletion of a population of a species or genus to the point that it is no longer a viable resource or is unlikely to fulfill its function in the ecosystem (USEPA 2016a). This measure was selected because it was an unambiguously adverse effect. The threshold for extirpation was estimated as the SC at the 95th centile extirpation concentration (XC95) of a weighted cumulative frequency distribution (CFD) of a genus or species. A CFD of more than 100 different genus-level XC95 values was used to predict the SC at which 5% of benthic invertebrate genera were likely to be extirpated, that is, a hazardous concentration extirpating 5% of genera (HC05).

A distribution of XC95 values (XCD) is a form of species sensitivity distribution (SSD) (Posthuma et al. 2002). There are several available tools for developing an SSD using laboratory toxicity test data. The United States Environmental

Protection Agency (USEPA) SSD Generator produces SSDs by fitting the most commonly applied distribution, the log-probit, to toxicity data (Shaw-Allen 2010). An alternative is the Dutch program ETX, which provides 1 method for estimating the affected fraction and the output is a log-probit plot (van Vlaardingen et al. 2004). The Environment Canada SSD Master allows for 2 methods to estimate the fraction affected using 5 potential curve fits: log-normal, log-logistic, extreme value, Weibull, and Gumbell (CCME 2013).

The biological extirpation analysis tools (BEAT) presented in the present paper differ from the above tools in that they use the same equations used to develop field-based benchmarks (USEPA 2011) and criteria (USEPA 2016b). Those calculations involve several arithmetic manipulations that are tedious to perform by hand. Because large data sets and a variety of analyses are key to the development of the field-based method, the original work to develop the method used the open source statistical software R (R Core Team 2014). However, some researchers may not be familiar with R and not every application requires the flexibility of R. In fact, the calculations can be done with a handheld calculator. However, a spreadsheet tool makes the task easier.

To enable a wide array of users and applications, we developed 2 downloadable Microsoft Excel[®] spreadsheet

* Address correspondence to cormier.susan@epa.gov

Published 14 August 2017 on wileyonlinelibrary.com/journal/ieam.

tools, XC95 BEAT and HC05 BEAT. The XC95 BEAT enables users to calculate the exposure that is predicted to extirpate a species or genera from a location or area. It generates an XC95 value, a plot of a taxon's proportional occurrence with respect to a stressor, and a plot of a taxon's probability of being observed within discrete stressor ranges or bins. The HC05 BEAT generates a CFD of XC95 values (XCD) and calculates the stressor level predicted to extirpate 5% of taxa. The HC05 BEAT enables analysis of a set of species or genera along a stressor gradient. By changing the inclusion criteria for data sets, one can estimate benchmarks (USEPA 2011, 2016b), assess local causes (Cormier et al. 2012; Coffey et al. 2014) and risks, and evaluate potential confounders (USEPA 2011, 2016b). This brief communication describes the tools and their uses. The tools and supporting information in Table 1 are available online as a set of appendices on GitHub, accessible through the link in the *Data Accessibility* section. We also discuss limitations of this method to help avoid misinterpretations and inferential errors.

METHODS AND EXAMPLES

Loading the XC95 BEAT and HC05 BEAT

There are 2 separate workbooks: one for calculation of an XC95 (XC95 BEAT, Appendix S2) and one for calculation of an HC05 (HC05 BEAT, Appendix S3). The XC95 and HC05 BEAT are Excel files (*.xlsm), designed to work in 2007 or later versions, that depend on macros for operation, so you must select "enable macros" when you open the file. Directions to enable macros are included on the first worksheet ("Notes") of the workbook. Instructions for using the workbooks are

also included on the "Notes" worksheet of each workbook but are summarized in this section.

Entering new data into the BEAT

To use a new data set with the XC95 BEAT, prepare a sample x taxon matrix worksheet using any nonzero number to indicate presence (e.g., 1) and zero (0) to indicate absence; the corresponding stressor value (e.g., "Conductivity") also needs to be included. To use a new data set to calculate an HC05, prepare a data set of taxa names and XC95 values and upload to the HC05 BEAT.

Running the BEAT

To run the XC95 BEAT, select the taxon of interest from the pull-down list, then click on the "Calculate XC95" button. The results will appear in the "Graphs" tab. The XC95 values can be copied and saved to another sheet. After results are viewed, click on the "Clear Results" button to return to the Input Sheet.

To run the HC05 BEAT with the preloaded XC95 values in the worksheet, select a set of taxa from the list by making "Present?" equal to "TRUE" in the column titled "XC95_Select." Or, if you add your own XC95 values and taxa to the same list, these additional taxa will be automatically included in the calculation if "Present?" is equal to "TRUE." After you click the "Calculate XCD" button, the results will appear in the "XCD_Chart" worksheet.

Operations performed by the XC95 BEAT

Calculation of the probability plot (Figure 1), the empirical weighted CFD (Figure 2), and the resulting XC95 involves

Table 1. Tools, step-by-step calculation, and example outputs^a

Appendix	Contents	Description
S1	Step-by-step	Describes the underlying calculation of the XC95 and HC05 values with a detailed example using specific conductivity as the stressor and benthic invertebrates as the affected taxa
S2	XC95 BEAT	Calculates exposure that is predicted to extirpate a species or genus from a location or area (XC95 value)
		Plots a taxon's cumulative proportional occurrence with respect to a stressor
		Plots a taxon's probability of being observed within discrete stressor ranges or bin
S3	HC05 BEAT	Generates a cumulative frequency distribution of XC95 values (XCD)
		Calculates the stressor level predicted to extirpate 5% of taxa
S4	Metadata	Description of data set from Ecoregions 69 and 70, West Virginia, USA
S5	Data set	Example data set from Ecoregions 69 and 70
S6	XC95 examples	Example XC95 values for 176 distinct aquatic benthic invertebrate genera
S7	Scatter plots	Example plots of probability of observing each of 176 genera
S8	CFD plots	Example plots of cumulative proportion of occurrence of each of 176 genera
S9	Taxa maps	Geographic distributions of 462 genera from an example data set

CFD = cumulative frequency distribution; HC05 = stressor level that predicts a 5% extirpation of genera; XC95 = 95th centile extirpation concentration; XCD = CFD of XC95 values.

^aAvailable at <https://github.com/smcomier/Biological-Extirpation-Analysis-Tools-BEAT/releases/tag/v.1.0.2>

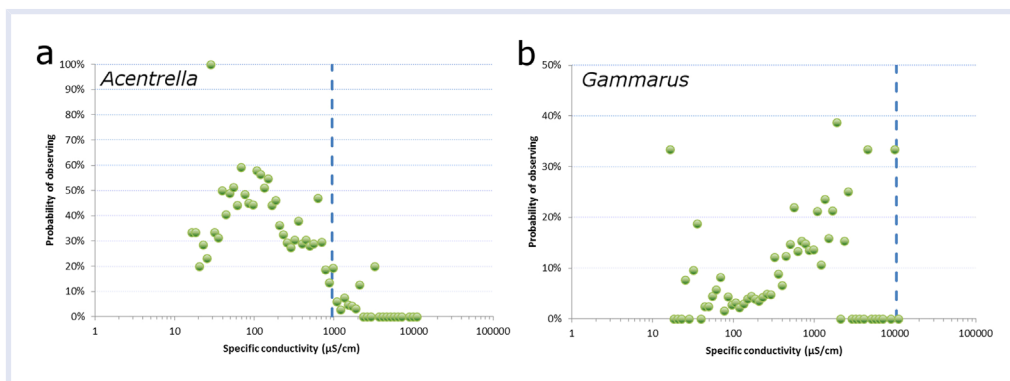


Figure 1. Probability plots of *Acentrella* (a) and *Gammarus* (b) relative to specific conductivity (SC). Probability of observing a taxon is the percentage of sites in a discrete range of SC (bin) where the individuals of a taxon were observed. For *Acentrella* (a), the rare occurrence of 100% probability represents 7 sites sampled within a bin from 26.8 to 29.9 $\mu\text{S}/\text{cm}$. The extirpation concentration (XC95) shown by the vertical dashed line is confident because the plot clearly declines to zero probability. The XC95 for *Gammarus* (b) is not defined because the probability increases within the measured range of SC in this data set. *Gammarus*' XC95 is therefore undefined and should be characterized as greater than the listed XC95 value.

several steps, all of which are automated in the XC95 BEAT. For example, for each taxon (genus or species) that meets your data-selection conditions, an empirical CFD is constructed that is weighted to correct for any potential bias from the unequal distribution of sampling of sites across the range of stressor exposures such as SC (Figure 3). Corrections for potential confounding require independent analyses and are not performed during this process (USEPA 2011, 2016b). This weighted CFD represents the proportion of observations of a taxon that occurs below given exposure levels. The extirpation effect threshold for a taxon is defined as the 95th centile point on the weighted CFD of the total observations of the taxon. Above this threshold, occurrences of the genus are spurious or so low that the presence of the taxon is unlikely. The two log₁₀ exposure levels bracketing the 95th centile are linearly interpolated to give an XC95 for a taxon. Users wishing to display underlying worksheets may do so by selecting the following Excel sequence: Home>Format>Visibility>Hide & Unhide>Unhide Sheet.

An example probability plot (Figure 1a) generated by the XC95 BEAT shows the distribution of the genus *Acentrella*

(presence at $N = 1291$), a mayfly nymph, from 3734 sampled sites collected along a SC gradient. The y-axis shows the probability of observing a taxon: the fraction of sampling sites within a given bin of SC values in which this genus was observed. This probability plot is used to assess whether an XC95 is defined or greater than the XC95 calculated by the CFD. The decreasing scatter plot for *Acentrella* with a clear y-intercept of 0 indicates that the XC95 is well defined. In contrast, Figure 1b shows an example of a scatter plot for *Gammarus* amphipods (presence at $N = 317$). The trend of the scatter plot is increasing and therefore the probability of observing the genus is not unequivocally near 0; therefore, the XC95 for *Gammarus* is not defined within the sampling range of SC in the data set. This is likely to occur because different species within the genus *Gammarus* have widely different optima ranging from obligate freshwater species to marine species (Beadle and Cragg 1940); therefore,

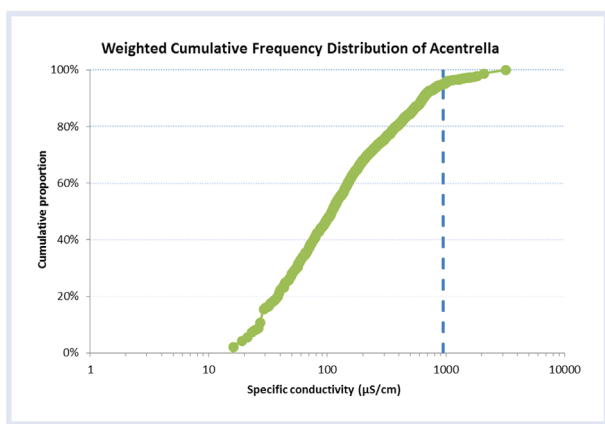


Figure 2. Weighted cumulative frequency distribution (CFD) for empirical observations of *Acentrella* versus specific conductivity (SC). Vertical dashed blue line marks 95% of observations of *Acentrella* and the SC below which 95% of the genus occurred in the data set. In this example, extirpation (defined as an XC95) is estimated at 947 $\mu\text{S}/\text{cm}$.

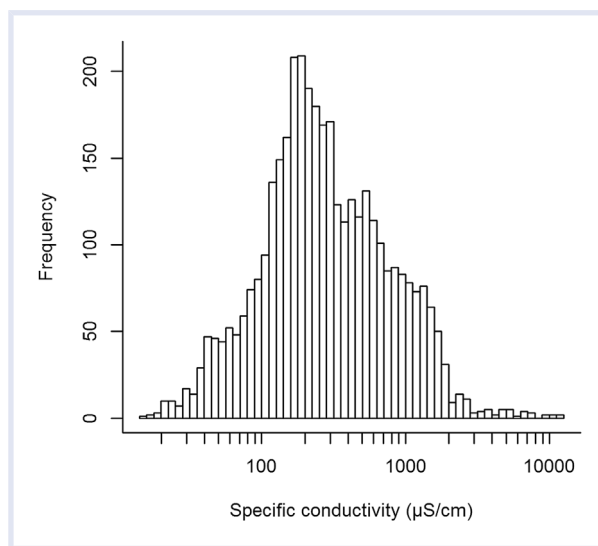


Figure 3. Histogram of the frequencies of observed specific conductivity (SC) values in samples from Ecoregions 69 and 70 in West Virginia, USA, sampled between 1997 and 2010. Bins are each 0.017 (1/60) of the range of log₁₀ SC units wide.

genus-level identification may not be informative for identifying a relevant SC tolerance range for *Gammarus* and the XC95 is greater than the estimated value, that is, it remains undefined. For a quantitative method, statistical software such as R can be used to fit a model to the scatter plot with defined confidence bounds (USEPA 2011, 2016b). The XC95 BEAT does not perform this analysis but example generalized additive models of the scatter plots are provided in Appendix S7, and the confidence in the XC95 values was defined using those plots and the methods described in USEPA (2011, 2016b) and provided in Appendix S1. The alternative is to assess the shape of the scatter by inspection and determine the occurrences of a genus at SC values greater than the calculated XC95. There is greater confidence in XC95 values followed by more bins with zero occurrences.

Each point in Figure 2 represents the cumulative fraction of observations of *Acentrella* relative to the SC gradient. It depicts the same data set of the weighted observations of *Acentrella* in Figure 1a but without calculating probabilities. The SC value for the interpolated 95th centile point (947 $\mu\text{S}/\text{cm}$) is the estimated XC95 and is shown at the 95th centile of the CFD as a vertical dashed line (Figure 2). For comparison, the XC95 is also shown on the scatter plot but it is not calculated from that plot (Figures 1a and 1b).

The XC95 BEAT automatically calculates the probability plot, the empirical weighted CFD, and the resulting XC95, which if done manually, involves several steps. First, equally sized bins are defined to compute weights for each sample. The selection of bin size depends on the size of the data set and requires balancing the requirements of sufficient observations in a bin to define the proportion and sufficient bins to define the form of the response. The default number of bins in the XC95 BEAT is 60. For the example, each of the 60 bins is assigned a width equal to 0.017 (1/60) multiplied by the range of the log₁₀-transformed SC values within the data set (Figure 3). The tool provides the user with an option for changing the number of bins on the data input worksheet.

Next, the bins are weighted to correct for any potential bias from unequal distribution of sampling of sites across the range of exposure concentrations. The assigned weight for each sample within a given bin is $w_i = 1/n_i$, where n_i is the number of samples in the i^{th} bin. The weighting ensures that sites in bins with many observations are not overly influential. The value of the weighted CFD, $F(x)$, is computed using the following equation for each unique observed value (x_{ij}) of the stressor associated with observations of a particular taxon:

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

where x_{ij} is the stressor 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 taxon of interest is observed in j^{th} sample of bin i , and I is an indicator function that equals 1 if the indicated conditions are true, and 0 otherwise.

It should be noted that within the observed ranges of SC, the weighted CFDs of some genera demonstrate a response to SC (e.g., *Acentrella*) as shown by a steep slope and asymptote well below the maximum exposures (Figure 2). Conversely, genera unaffected within this SC range have a steady increase or remain flat over the entire range of measured exposure and do not reach a clear asymptote. It is somewhat easier to evaluate genera with undefined XC95 values from the plots of the probability of observing a taxon. For example, an XC95 is undefined in the plots of the probabilities of observing a taxon that increase (e.g., Figure 2b, *Gammarus*) or shows no directional response as SC increases (e.g., *Nigronia*, Appendix S7). In such cases, the XC95 should be considered as greater than the value calculated.

In summary, the XC95 value is defined as the stressor value corresponding to $F(x) = 0.95$. Equation 1 is an empirical weighted CFD, and the output is the proportion of observations of the taxon that occur at or below a given exposure level where the individual observations are weighted to account for the uneven distribution of observations across the range of stressor values.

Operation for calculating an HC05

Given a data set with various taxa and their related XC95 values for a particular stressor, the HC05 BEAT calculates an XCD and an HC05 for the stressor. The XCD is a cumulative distribution plot showing the proportion of taxa extirpated as a function of the stressor value. The HC05 is the estimated value of the stressor at which 5% of species are predicted to become extirpated.

For example, the inset (Figure 4a) shows the XCD for SC ($\mu\text{S}/\text{cm}$) based upon macroinvertebrate data from Ecoregions 69 and 70 from West Virginia, USA. This XCD was calculated and plotted automatically by HC05 BEAT, given a sample data set of 176 genera and their associated XC95 values for SC. For this example data set, the HC05 indicates

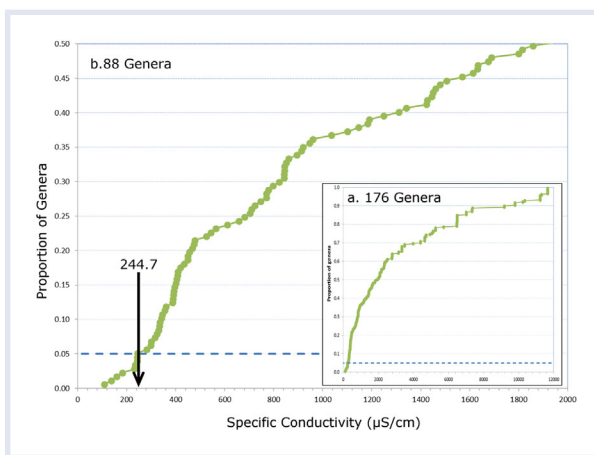


Figure 4. Cumulative frequency distribution of XC95 values (XCD) for specific conductivity (SC) ($\mu\text{S}/\text{cm}$). Each dot is an XC95 value for 1 unique benthic invertebrate genus. The SC at the intersection of the XCD and the 5th centile (horizontal line) is the HC05 (244.7 $\mu\text{S}/\text{cm}$). $N = 176$ XC95 (a); $N = 88$ (b).

that approximately 5% of species will be extirpated at an SC of 245 $\mu\text{S}/\text{cm}$.

In data sets from freshwater systems, it is not uncommon that the tolerance range of half of the taxa will be undefined, such that the species is not extirpated at the highest measured SC. But that does not influence the HC05 because the XC95 values for these taxa are not in the SC range near the 5th centile of the distribution of XC95 values. For other applications such as specific site evaluations, the list of XC95 values should be consulted to determine where the plotted points represent XC95 values greater than the calculated value. To avoid misinterpretation of an XCD constructed from other data sets, a reasonable heuristic is to restrict predictions of the proportion of affected genera to the lower third of the curve. For example, because the XC95 values at SC greater than 1500 $\mu\text{S}/\text{cm}$ have been characterized as greater than the assigned XC95 value in the example data set, the upper portion of the XCD in Figure 4a does not predict the proportion of genera likely to be extirpated at that concentration. About half of the XC95 values are undefined in Figure 4a, and therefore we also show Figure 4b, which depicts the range of XC95 values that are defined in the sampled range of the data set.

To estimate an HC05 without the HC05 BEAT, one simply counts the number of taxa and ranks their XC95 values from least to greatest value. These can be plotted as a CFD with the XC95 value on the x-axis and the "Proportion of Taxa" extirpated on the y-axis. The HC05 is the SC at the 5th centile of the distribution of XC95 values. To identify the rank at the 5th centile, one identifies the XC95 bracketing the 5th centile rank and interpolates the log₁₀ SC at the 5th centile of the ranked XC95 values.

The XC95 values for the 176 genera are provided in Appendix S6. For genera with XC95 values >1500 $\mu\text{S}/\text{cm}$, the probability of observing those genera often increases as SC increases without an apparent optimum in the measured SC range of the data set.

Example calculation using the HC05 BEAT

Details of the manual calculation of the HC05 are shown in an example (Appendix S1). The detailed steps are performed automatically by the spreadsheet tool. A simple quantile-based method is performed. Users may load XC95 values that they have calculated using the XC95 BEAT. Users who choose to use the preloaded values in the XC95 BEAT (Appendix S6) from Ecoregions 69 and 70 (Omernik 1987) should keep in mind that the species comprising the genera in their region may differ. As a result, regional XC95 values at the genus level may be higher or lower because different species with different tolerances and different proportions of species may be represented in a genus XC95 value (Griffith et al. 2017). The BEAT HC05 reports the HC05 and provides a plot of the cumulative proportion of genera versus XC95.

DISCUSSION

The step-by-step description of the derivation of an XC95 value, an XCD, and an HC05 is intended to make these effect

endpoints accessible to a broad variety of interested entities. The spreadsheet calculation tool will save time for water quality practitioners and researchers who would like to calculate XC95 values for their own data sets and applications but prefer not to develop a calculation tool.

An example using the method was provided for a combined area in the Central Appalachians using freshwater aquatic benthic invertebrate genera (USEPA 2011; Cormier and Suter 2013; Cormier et al. 2013). However, this general method can be applied to either a species or genus level of taxonomic identification and is not limited to invertebrates. For example, the method was applied using data for freshwater fish species in the same region (Griffith et al. 2017; USEPA 2016b). Although not tested, it may be applicable to other taxonomic or functional groups and other ecosystems, including terrestrial ones.

The USEPA (2011) field-based method can be used for other applications in addition to the development of effect levels or benchmarks. The XC95 values of macroinvertebrates in streams have been shown to be useful for determining the causes of altered biological communities. For example, Cormier et al. (2012) showed that increased ionic concentration was a likely cause of altered benthic invertebrate communities in the Ten Mile Creek watershed in southwestern Pennsylvania, USA. Conversely, Coffey et al. (2014) used XC95 values to show that ionic concentration was less likely than organic enrichment as the cause of biological impairments identified by the Tennessee Department of Environmental Conservation in Pigeon Roost Creek in central Tennessee, USA.

Because the XC95 represents the limit of a taxon's tolerance and not a central tendency such as a lethal concentration that kills 50% of test animals (LC50), there are risks that the BEAT could be misapplied. Thus, the following considerations should be carefully evaluated when using the tools. Some recommended heuristics are listed below:

- Measurements of the agents are paired in space and time with biological sampling;
- high-quality or reference sites are included in the data set;
- background SC levels are similar throughout the region;
- characteristics of the agent (i.e., ionic composition) are similar across the region for the paired data (i.e., other mixtures or agents may occur but they are analyzed separately);
- some biological sampling occurs when intolerant genera are likely to be collected (e.g., March through June in Appalachia) and where they are likely to occur (e.g., leaf packs, riffles);
- the exposure gradient is broad enough to include no effects, weak effects, and strong effects;
- sensitivity analyses are performed to evaluate effect of bin size, sample size, and other methodological parameters;
- data are available and sufficient to evaluate potential confounding factors; and
- an independent data set or statistical models validate the results.

It is also important to keep in mind that a formula will yield a value whether the data are appropriate or not. One mistake is to assume that differences between XCD plots are due to biological or ecological causes when there is some other explanation. For example, divergences of XCD curves greater than the 50th centile are often due to differences in the exposure range of particular data sets. The XC95 values greater than the 50th centile of the XCD are often not precisely estimated because the range of conductivity values of most data sets does not include the full range of SC tolerated by every organism. That is, the SC tolerated by a particular organism is greater than the highest measured SC in the data set. Therefore, the SC range is not wide enough to measure XC95 values for some taxa, there are too few samples at the higher SC ranges, or there are other confounding factors affecting those values. For example, a comparison of XCDs with similar background SC from Hun-Tai River Basin in Northeast China and Central Appalachia areas in Kentucky and West Virginia of the United States are statistically different; however, the narrower range of SC in the Hun-Tai River Basin data set does not measure any XC95 values greater than 1160 $\mu\text{S}/\text{cm}$ (Zhao et al. 2016), whereas XC95 values that can potentially be measured at the maxima SC for data sets are 2390 $\mu\text{S}/\text{cm}$ in Kentucky and 11 646 $\mu\text{S}/\text{cm}$ in West Virginia (USEPA 2011). Consequently, the maximum XC95 and maximum of the XCD are limited to the highest measured SC in each data set, and so XCDs diverge when the measured ranges of SC are substantially different.

To illustrate the effect of a limited stressor range on the XCD, we prepared 2 data sets from the example data set. One set is the full exposure range between 15.4 and 11 646 $\mu\text{S}/\text{cm}$, and the other is the same data set restricted to samples $\leq 1000 \mu\text{S}/\text{cm}$. The HC05 values remain fairly similar; however, the XCDs diverge at higher proportions of affected genera (Figure 5). This statistical divergence is a data set limitation and not a measure of ecological similarity or tolerance of the biological assemblage. However, for the taxa near the 5th centile, the range is sufficient to estimate XC95 values and thus estimate the HC05. It is necessary to examine the scatter plots of the probability of occurrence of taxa to ascertain if the estimated XC95 values of the lower part of the XCD have been measured or if they are greater than the measured range of the SC in the data set.

The original method (USEPA 2011; Cormier and Suter 2013) used confidence bounds of a generalized additive model fitted to the probability histogram and was calculated using R statistical software. This capability is not provided with the XC95 BEAT due to the complexity of the coding. However, confidence in an XC95 value may be assessed by examining the shape and asymptote of the plots of the probability of observing a taxon, for example, Figures 1a and 1b. In lieu of a more sophisticated analysis, users should assume that most XC95 values in the upper portion of the XCD are greater than the calculated value.

Potential confounding can be another concern. Because the method was developed to generate an HC05, methods

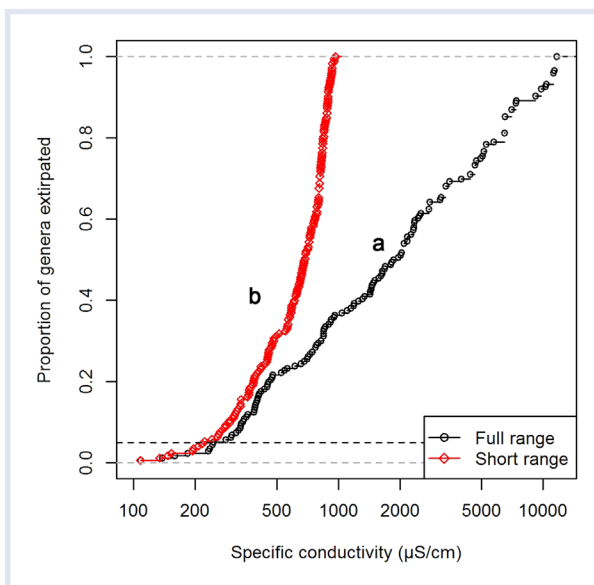


Figure 5. Cumulative frequency distribution of XC95 values (XCD) for specific conductivity (SC) ($\mu\text{S}/\text{cm}$). Permutation of the same data set and same taxa occurrence within the full SC range of specific conductivity from 15 to 11 646 $\mu\text{S}/\text{cm}$ (a) and 15 to 1000 $\mu\text{S}/\text{cm}$ (b). Divergence of the XCD is due to SC range, not different biological responses. In plot b, the maximum XC95 is restricted to 1000 $\mu\text{S}/\text{cm}$, causing the entire plot to shift left because higher exposures are not possible.

for assessing the potential confounding of the XCD model were developed and performed (Suter and Cormier 2013; USEPA 2016b). In this data set, the HC05 is not appreciably confounded by other variables (USEPA 2016b). Similar analytical approaches might be applied to individual genera or species but have not been explored by us.

SUMMARY

In summary, the extirpation of an entire genus or species is a serious consequence of exposure to the measured stressor, and an extirpation of genera or species over a wide area may lead to extinction. The tools described here need to be used with care and the outputs interpreted conservatively with consideration of the full suite of scientific issues more fully explained in USEPA (2016b). We hope that scientists and practitioners will use these calculation tools to identify areas where organisms are at a high risk of extirpation, as well as help to set protective and restoration goals for stressors.

Acknowledgment—The authors declare no conflicts of interest. This work was supported by and prepared at the USEPA, National Center for Environmental Assessment (NCEA), Cincinnati Division. The authors are indebted to the work of field and laboratory personnel who collected and analyzed the primary data. The manuscript 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 USEPA. Kathleen Bland, Linda Tackett, and Michael Gallagher edited and formatted the document. Constructive comments from Glenn Suter, Dan

Petersen, and from anonymous reviewers helped to substantially improve earlier versions of this manuscript.

Disclaimer—The peer-review process for this article was handled by the Editorial Board without the involvement of S Cormier.

Data Accessibility—The tools, data sets, example calculations, and example outputs are available online at <https://wiley.figshare.com/ieam> and <https://github.com/smcormier/Biological-Extirpation-Analysis-Tools-BEAT/releases/tag/v.1.0.2>.

REFERENCES

- Beadle LC, Cragg JB. 1940. Studies on adaptation to salinity in *Gammarus* spp. *J Exp Biol* 17(2):153–163. [cited 2017 October 13]. <http://jeb.biologists.org/content/jexbio/17/2/153.full.pdf>
- [CCME] Canadian Council of Ministers of the Environment. 2013. Determination of hazardous concentrations with species sensitivity distributions: SSD master, version 3.0 (unpublished). Gatineau (QC): National Guidelines and Standards Office, Environment Canada. Available from email: ceqg-rcqe@ec.gc.ca.
- Coffey DB, Cormier SM, Harwood J. 2014. Using field-based species sensitivity distributions to infer multiple causes. *HERA* 20(2):402–432.
- Cormier SM, Coffey DB, Griffith M. 2012. Letter to the Editor in Chief concerning the article “Status of fish and macroinvertebrate communities in a watershed experiencing high rates of fossil fuel extraction: Tenmile Creek, a major Monongahela River tributary” by Kimmel and Argent, 2012. *Water Air Soil Poll* 223:4659–4662.
- Cormier SM, Suter II GW. 2013. A method for deriving water-quality benchmarks using field data. *Environ Toxicol Chem* 32:255–262.
- Cormier SM, Suter GW, Zheng L. 2013. Derivation of a benchmark for freshwater ionic strength. *Environ Toxicol Chem* 32(2):263–271.
- Griffith MB, Zheng L, Cormier SM. 2017. Using extirpation to evaluate ionic tolerance of freshwater fish. *Environ Toxicol Chem* DOI: 10.1002/etc.4022.
- Omerik JM. 1987. Ecoregions of the conterminous United States. *Ann Assoc Am Geogr* 77:118–125.
- Posthuma L, Suter GW II, Traas TP. 2002. General introduction to species sensitivity distributions. In: Posthuma L, Suter GW II, Traas TP, editors. *Species sensitivity distributions in ecotoxicology*. Boca Raton (FL): Lewis. p 3–10.
- R Core Team. 2014. R: A language and environment for statistical computing. Vienna (AT): R Foundation for Statistical Computing. [cited 2015 October 10]. <http://www.R-project.org/>
- Shaw-Allen P. 2010. Species sensitivity distribution generator. SSD_Generator_V1.xlt. CADDIS Volume 4: Data analysis. Washington (DC): US Environmental Protection Agency. [cited 2017 October 13]. <https://www.epa.gov/caddis-vol4/caddis-volume-4-data-analysis-download-software>
- Suter GW, Cormier SM. 2013. A method for assessing the potential for confounding applied to ionic strength in Central Appalachian streams. *Environ Toxicol Chem* 32(2):288–295.
- [USEPA] US Environmental Protection Agency. 2011. A field-based aquatic life benchmark for conductivity in central Appalachian streams. Washington (DC): USEPA Office of Research and Development, National Center for Environmental Assessment. EPA/600/R-10/023F.
- [USEPA] US Environmental Protection Agency. 2016a. Generic ecological assessment endpoints (GEAEs) for ecological risk assessment (2nd Edition) and Technical Background Paper. Washington (DC): Risk Assessment Forum. EPA/100/F15/005. [cited 2017 October 13]. <https://www.epa.gov/osa/ecosystem-services-ecological-risk-assessment-end-points-guidelines>
- [USEPA] US Environmental Protection Agency. 2016b. Public review draft: Field-based methods for developing aquatic life criteria for specific conductivity. Washington (DC): USEPA Office of Water. EPA-822-R-07-010.
- van Vlaardingen PLA, Traas TP, Wintersen AM, Aldenberg T. 2004. ETX 2.0. A program to calculate hazardous concentrations and fraction affected, based on normally distributed toxicity data. Bilthoven (NL): Rijksinstituut voor Volksgezondheid en Milieu (RIVM). [cited 2017 October 13]. <http://www.rivm.nl/bibliotheek/rapporten/601501028.html>
- Zhao Q, Jia X, Xia R, Lin J, Zhang Y. 2016. A field-based method to derive macroinvertebrate benchmark for specific conductivity adapted for small data sets and demonstrated in the Hun-Tai River Basin, Northeast China. *Environ Pollut* 216:902–910. DOI: 10.1016/j.envpol.2016.06.065