

A Field-Based Aquatic Life Benchmark for Conductivity in Central Appalachian Streams

[This page intentionally left blank.]

EPA/600/R-10/023F
March 2011

A Field-Based Aquatic Life Benchmark for Conductivity in Central Appalachian Streams

National Center for Environmental Assessment
Office of Research and Development
U.S. Environmental Protection Agency
Washington, DC 20460

DISCLAIMER

This document has been reviewed in accordance with U.S. Environmental Protection Agency policy and approved for publication. Mention of trade names or commercial products does not constitute endorsement or recommendation for use.

Cover Photo:

Used by permission, from Randall Sanger Photography. Photo of Ramsey Branch, West Virginia.

Back Cover Photos:

Used by permission, from Guenter Schuster. Photo of *Barbicambarus simmonsii*.

Used by permission, from North American Benthological Society. Photos of various invertebrates. © 2009 North American Benthological Society. <http://www.benthos.org/Education-and-Outreach/Media-Galleries/Invertebrates.aspx?Page=1>.

Preferred Citation:

U.S. EPA (Environmental Protection Agency). 2011. A Field-Based Aquatic Life Benchmark for Conductivity in Central Appalachian Streams. Office of Research and Development, National Center for Environmental Assessment, Washington, DC. EPA/600/R-10/023F.

CONTENTS

LIST OF TABLES	v
LIST OF FIGURES	vi
LIST OF ABBREVIATIONS AND ACRONYMS	vii
PREFACE	viii
AUTHORS, CONTRIBUTORS, AND REVIEWERS	ix
ACKNOWLEDGMENTS	xiii
EXECUTIVE SUMMARY	xiv
1. INTRODUCTION	1
1.1. CONDUCTIVITY	1
1.2. APPROACH	3
2. DATA SETS	6
2.1. DATA SET SELECTION	6
2.2. DATA SOURCES	6
2.3. DATA SET CHARACTERISTICS	9
3. METHODS	13
3.1. EXTIRPATION CONCENTRATION DERIVATION	13
3.2. TREATMENT OF POTENTIAL CONFOUNDERS.....	17
3.3. DEVELOPING THE SPECIES SENSITIVITY DISTRIBUTION	18
3.4. CONFIDENCE BOUNDS.....	19
3.5. EVALUATING ADEQUACY OF NUMBER OF SAMPLES.....	21
3.6. ESTIMATING BACKGROUND.....	22
4. RESULTS	23
4.1. EXTIRPATION CONCENTRATIONS.....	23
4.2. SPECIES SENSITIVITY DISTRIBUTIONS	23
4.3. HAZARDOUS CONCENTRATION VALUES AT THE 5 th CENTILE.....	23
4.4. UNCERTAINTY ANALYSIS	23
5. CONSIDERATIONS.....	25
5.1. CHOOSING TO USE FIELD VERSUS LABORATORY DATA.....	25
5.2. SELECTION OF THE EFFECTS ENDPOINT	26
5.3. TREATMENT OF MIXTURES.....	26
5.4. DEFINING THE REGION OF APPLICABILITY.....	27
5.5. BACKGROUND	33
5.6. SELECTION OF INVERTEBRATE GENERA.....	34
5.7. INCLUSION OF OTHER TAXA	34
5.8. TREATMENT OF LISTED SPECIES.....	34
5.9. INCLUSION OF REFERENCE SITES	35

CONTENTS (continued)

5.10. SEASONALITY, LIFE HISTORY, AND SAMPLING METHODS.....	35
5.11. FORMS OF EXPOSURE-RESPONSE RELATIONSHIPS.....	37
5.12. USE OF MODELED OR EMPIRICAL DISTRIBUTIONS.....	38
5.13. DUPLICATE SAMPLES.....	39
5.14. TREATMENT OF CAUSATION.....	39
5.15. TREATMENT OF POTENTIAL CONFOUNDERS.....	40
6. AQUATIC LIFE BENCHMARK.....	41
REFERENCES.....	42
APPENDIX A. CAUSAL ASSESSMENT.....	A-1
APPENDIX B. ANALYSIS OF POTENTIAL CONFOUNDERS.....	B-1
APPENDIX C. DATA SOURCES AND METHODS OF LAND USE/LAND COVER ANALYSIS USED TO DEVELOP EVIDENCE OF SOURCES OF HIGH CONDUCTIVITY WATER.....	C-1
APPENDIX D. EXTIRPATION CONCENTRATION VALUES FOR GENERA IN THE WEST VIRGINIA DATA SET.....	D-1
APPENDIX E. GRAPHS OF OBSERVATION PROBABILITIES FOR GENERA IN THE WEST VIRGINIA DATA SET.....	E-1
APPENDIX F. GRAPHS OF CUMULATIVE FREQUENCY DISTRIBUTIONS FOR GENERA IN THE WEST VIRGINIA DATA SET.....	F-1
APPENDIX G. VALIDATION OF METHOD USING FIELD DATA TO DERIVE AMBIENT WATER QUALITY BENCHMARK FOR CONDUCTIVITY USING A KENTUCKY DATA SET.....	G-1
APPENDIX H. EXTIRPATION CONCENTRATION VALUES FOR GENERA IN A KENTUCKY DATA SET.....	H-1
APPENDIX I. GRAPHS OF OBSERVATION PROBABILITIES FOR GENERA IN KENTUCKY DATA SET.....	I-1
APPENDIX J. GRAPHS OF CUMULATIVE FREQUENCY DISTRIBUTIONS FOR GENERA IN KENTUCKY DATA SET.....	J-1

LIST OF TABLES

1.	Differences between the method used to derive the conductivity benchmark and the method in Stephan et al. (1985) and the section of the report in which each is discussed	3
2.	Summary statistics of the measured water-quality parameters.....	10
3.	Number of samples with reported genera and conductivity meeting our acceptance criteria for calculating the benchmark value.....	11
4.	Samples excluded from the original data sets of 2,668 samples used to develop benchmark value	11
5.	Genera excluded from 95 th centile extirpation concentration calculation because they never occurred at reference sites	12
6.	Genera of threatened species included in the SSD	35

LIST OF FIGURES

1.	Points are sampling locations used to develop the benchmark from Level III Ecoregions 69 and 70 in West Virginia.	5
2.	Box plot showing seasonal variation of conductivity ($\mu\text{S}/\text{cm}$) in the reference streams of Ecoregions 69 and 70 in West Virginia from 1999 to 2006.....	7
3.	Box plot showing seasonal variation of conductivity ($\mu\text{S}/\text{cm}$) from a probability-based set of sample streams of Ecoregions 69 and 70 in West Virginia from 1997 to 2007.	7
4.	Box plot showing seasonal variation of conductivity ($\mu\text{S}/\text{cm}$) from the data set used to develop the benchmark.....	8
5.	Histograms of the frequencies of observed conductivity values in samples from Ecoregions 69 and 70 from West Virginia sampled between 1999 and 2006.....	14
6.	Examples of weighted CDFs and the associated 95 th centile extirpation concentration values.....	16
7.	Three typical distributions of observation probabilities	16
8.	The species sensitivity distribution.....	18
9.	Species sensitivity distribution (expanded).	19
10.	Diagram depicting the process for estimating the uncertainty of the HC_{05}	20
11.	The cumulative distribution of XC_{95} values for the 36 most sensitive genera and the bootstrap-derived means and two-tailed 95% confidence intervals.....	20
12.	Adequacy of the number of samples used to model the HC_{05}	21
13a.	Anions.	28
13b.	Cations	29
13c.	Dissolved metals	30
13d.	Total metals.....	31
13e.	Other water quality parameters.....	32
14.	Comparison of full data set and subsets of spring and summer collected samples	36
15.	Relationship of conductivity values sampled from the same site in spring and summer.....	37

LIST OF ABBREVIATIONS AND ACRONYMS

CDF	cumulative distribution function
CVs	chronic values
DC _x	depletion concentration
GAM	Generalized Additive Model
HC _x	hazardous concentration
LC _x	lethal concentration
LOWESS	locally weighted scatterplot smoothing
SSD	species sensitivity distribution
TMDL	total maximum daily load
U.S. EPA	United States Environmental Protection Agency
WABbase	Water Analysis Database (WABbase)
WVDEP	West Virginia Department of Environmental Protection
WVSCI	West Virginia Stream Condition Index
XC _x	extirpation concentration

PREFACE

At the request of U.S. Environmental Protection Agency's (EPA) Office of Water and Regions, the EPA Office of Research and Development has developed an aquatic life benchmark for conductivity for the Appalachian Region. The benchmark is applicable to mixtures of ions dominated by salts of Ca^{2+} , Mg^{2+} , SO_4^{2-} and HCO_3^- at a circum-neutral to alkaline pH. The impetus for the benchmark is the observation that high conductivities in streams below surface coal mining operations, especially mountaintop mining and valley fills, are associated with impairment of aquatic life. However, application of the benchmark is not limited to that source.

The benchmark was derived by a method modeled on the EPA's 1985 methodology for deriving ambient water-quality criteria for the protection of aquatic life. The methodology was adapted for use of field data, by substituting the extirpation of stream macroinvertebrates for laboratory toxicity data.

The methodology and derivation of the benchmark were reviewed by internal reviewers, external reviewers, and a review panel of the EPA's Science Advisory Board (SAB). The SAB panel's review was in turn reviewed by the Chartered SAB. The SAB review is available at <http://yosemite.epa.gov/sab/sabproduct.nsf/02ad90b136fc21ef85256eba00436459/984d6747508d92ad852576b700630f32!OpenDocument>.

The SAB concluded that the benchmark is applicable to the regions in which it was derived and the benchmark and the methodology may be applicable to other states and regions with appropriate validation. In addition, hundreds of public commenters provided their views. Comments from all of these sources were considered and used to improve the clarity and scientific rigor of the document.

AUTHORS, CONTRIBUTORS, AND REVIEWERS

AUTHORS

Susan M. Cormier, PhD
U.S. Environmental Protection Agency
Office of Research and Development, National Center for Environmental Assessment
Cincinnati, OH 45268

Glenn W. Suter II, PhD
U.S. Environmental Protection Agency
Office of Research and Development, National Center for Environmental Assessment
Cincinnati, OH 45268

Lester L. Yuan, PhD
U.S. Environmental Protection Agency
Office of Research and Development, National Center for Environmental Assessment
Washington, DC 20460

Lei Zheng, PhD
Tetra Tech, Inc.
Owings Mills, MD 21117

CONTRIBUTORS

R. Hunter Anderson, PhD
U.S. Environmental Protection Agency
Office of Research and Development, National Center for Environmental Assessment
Cincinnati, OH 45268

Jennifer Flippin, MS
Tetra Tech, Inc.
Owings Mills, MD 21117

Jeroen Gerritsen, PhD
Tetra Tech, Inc.
Owings Mills, MD 21117

Michael Griffith, PhD
U.S. Environmental Protection Agency
Office of Research and Development, National Center for Environmental Assessment
Cincinnati, OH 45268

AUTHORS, CONTRIBUTORS, AND REVIEWERS (continued)

CONTRIBUTORS (continued)

Michael McManus, PhD
U.S. Environmental Protection Agency
Office of Research and Development, National Center for Environmental Assessment
Cincinnati, OH 45268

John Paul, PhD
U.S. Environmental Protection Agency
National Health and Environmental Effects Research Laboratory
Research Triangle Park, NC 27711

Gregory J. Pond, MS
U.S. Environmental Protection Agency
Region III
Wheeling, WV 26003

Samuel P. Wilkes, MS
Tetra Tech, Inc.
Charleston, WV 25301

REVIEWERS

John Paul, PhD
U.S. Environmental Protection Agency
National Health and Environmental Effects Research Laboratory
Research Triangle Park, NC 27711

Samuel P. Wilkes, MS
Tetra Tech, Inc.
Charleston, WV 25301

Margaret Passmore, MS
U.S. Environmental Protection Agency
Office of Monitoring and Assessment, Freshwater Biology Team
Wheeling, WV 26003

Charles Delos, MS
U.S. Environmental Protection Agency
Office of Water, Health and Ecological Criteria Division
Washington, DC 20460

AUTHORS, CONTRIBUTORS, AND REVIEWERS (continued)

REVIEWERS (continued)

John Van Sickle, PhD
U.S. Environmental Protection Agency
National Health and Environmental Effects Research Laboratory, Western Ecology Division
Corvallis, OR 97333

Charles P. Hawkins, PhD
Western Center for Monitoring and Assessment of Freshwater Ecosystems
Department of Watershed Sciences
Utah State University
Logan, UT 84322

Christopher C. Ingersoll, PhD
U.S. Geological Survey
Columbia Environmental Research Center
4200 New Haven Road
Columbia, MO 65201

Charles A. Menzie, PhD
Exponent
2 West Lane
Severna Park, MD 21146

Science Advisory Board Panel on Ecological Impacts of Mountaintop Mining and Valley Fills

Duncan Patten, Chairman, PhD
Montana State University, Bozeman, MT

Elizabeth Boyer, PhD
Pennsylvania State University, University Park, PA

William Clements, PhD
Colorado State University, Fort Collins, CO

James Dinger, PhD
University of Kentucky, Lexington, KY

Gwendelyn Geidel, PhD
University of South Carolina, Columbia, SC

AUTHORS, CONTRIBUTORS, AND REVIEWERS (continued)

REVIEWERS (continued)

Kyle Hartman, PhD
West Virginia University, Morgantown, WV

Robert Hilderbrand, PhD
University of Maryland Center for Environmental Science, Frostburg, MD

Alexander Huryn, PhD
University of Alabama, Tuscaloosa, AL

Lucinda Johnson, PhD
University of Minnesota Duluth, Duluth, MN

Thomas W. La Point, PhD
University of North Texas, Denton, TX

Samuel N. Luoma, PhD
University of California – Davis, Sonoma, CA

Douglas McLaughlin, PhD
National Council for Air and Stream Improvement, Kalamazoo, MI

Michael C. Newman, PhD
College of William & Mary, Gloucester Point, VA

Todd Petty, PhD
West Virginia University, Morgantown, WV

Edward Rankin, MS
Ohio University, Athens, OH

David Soucek, PhD
University of Illinois at Urbana-Champaign, Champaign, IL

Bernard Sweeney, PhD
Stroud Water Research Center, Avondale, PA

Philip Townsend, PhD
University of Wisconsin–Madison, Madison, WI

AUTHORS, CONTRIBUTORS, AND REVIEWERS (continued)

REVIEWERS (continued)

Richard Warner, PhD
University of Kentucky, Lexington, KY

ACKNOWLEDGMENTS

Susan B. Norton, Teresa Norberg-King, Peter Husby, Peg Pelletier, Treda Grayson, Amy Bergdale, Candace Bauer, Brooke Todd, David Farar, Lana Wood, Heidi Glick, Cristopher Broyles, Linda Tackett, Stacey Herron, Bette Zwayer, Maureen Johnson, Debbie Kleiser, Crystal Lewis, Marie Nichols-Johnson, Sharon Boyde, Katherine Loizos, and Ruth Durham helped bring this document to completion by providing comments, essential fact checking, editing, formatting, and other key activities. Statistical review of the methodology was provided by Paul White, John Fox, and Leonid Kopylev.

EXECUTIVE SUMMARY

This report uses field data to derive an aquatic life benchmark for conductivity that can be applied to waters in the Appalachian Region that are dominated by salts of Ca^{2+} , Mg^{2+} , SO_4^{2-} and HCO_3^- at a circum-neutral to mildly alkaline pH. This benchmark is intended to protect the aquatic life in the region. It is derived by a method modeled on the EPA's standard methodology for deriving water-quality criteria (i.e., Stephan et al., 1985). In particular, the methodology was adapted for use of field data. Field data were used because sufficient and appropriate laboratory data were not available and because high-quality field data were available to relate conductivity to effects on aquatic life. This report provides scientific evidence for a conductivity benchmark in a specific region rather than for the entire United States.

The method used in this report is based on the standard methodology for deriving water-quality criteria, as explained in Stephan et al. (1985), in that it used the 5th centile of a species sensitivity distribution (SSD) as the benchmark value. SSDs represent the response of aquatic life as a distribution with respect to exposure. Data analysis followed the standard methodology in aggregating species to genera and using interpolation to estimate the centile. It differs primarily in that the points in the SSDs are extirpation concentrations (XCs) rather than median lethal concentrations ($\text{LC}_{50\text{s}}$) or chronic values. The XC is the level of exposure above which a genus is effectively absent from water bodies in a region. For this benchmark value, the 95th centile of the distribution of the probability of occurrence of a genus with respect to conductivity was used as a 95th centile extirpation concentration. Hence, this aquatic life benchmark for conductivity is expected to avoid the local extirpation of 95% of native species (based on the 5th centile of the SSD) due to neutral to alkaline effluents containing a mixture of dissolved ions dominated by salts of SO_4^{2-} and HCO_3^- . Because it is not protective of all genera and protects against extirpation rather than reduction in abundance, this level is not fully protective of sensitive species or higher quality, exceptional waters designated by state and federal agencies..

This field-based method has several advantages. Because it is based on biological surveys, it is inherently relevant to the streams where the benchmark may be applied and represents the actual aquatic life use in these streams. Another advantage is that the method assesses all life stages and ecological interactions of many species. Further, it represents the actual exposure conditions for elevated conductivity in the region, the actual temporal variation in exposure, and the actual mixture of ions that contribute to salinity as measured by conductivity.

The disadvantages of field data result from the fact that exposures are not controlled. As a result, the causal nature of the relationship between conductivity and the associated biological

impairments must be assessed. Also, any variables that are correlated with conductivity and the biotic response may confound the relationship of biota to conductivity. Assessments of causation and confounding were performed and are presented in the appendices. They demonstrate that conductivity can cause impairments and the relationship between conductivity and biological responses apparently is not appreciably confounded.

The chronic aquatic life benchmark value for conductivity derived from all-year data from West Virginia is 300 $\mu\text{S}/\text{cm}$. It is applicable to parts of West Virginia and Kentucky within Ecoregions 68, 69, and 70 (Omernick, 1987). It is expected to be applicable to the same ecoregions extending into Ohio, Pennsylvania, Tennessee, Virginia, Alabama, and Maryland, but data from those states have not been analyzed. This is because the salt matrix and background is expected to be similar throughout the ecoregions. The benchmark may also be appropriate for other nearby ecoregions, such as Ecoregion 67, but it has only been validated for use in Ecoregions 68, 69, and 70 at this time. This benchmark level might not apply when the relative concentrations of dissolved ions are not dominated by salts of Ca^{2+} , Mg^{2+} , SO_4^{2-} and HCO_3^- or the natural background exceeds the benchmark. However, the salt mixture dominated by salts of SO_4^{2-} and HCO_3^- is believed to be an insurmountable physiological challenge for some species.

1. INTRODUCTION

At the request of U.S. Environmental Protection Agency's (U.S. EPA) Office of Water and Regions, the Office of Research and Development has developed an aquatic life benchmark for conductivity that may be applied in the Appalachian Region associated with mixtures of ions dominated by salts of Ca^{2+} , Mg^{2+} , SO_4^{2-} and HCO_3^- at a circum-neutral to alkaline pH. The benchmark is intended to protect the aquatic life in streams and rivers in the region. It is derived by a method modeled on the EPA's standard methodology for deriving water-quality criteria (i.e., Stephan et al., 1985). In particular, the methodology was adapted for use of field data. Field data were used because sufficient and appropriate laboratory data were not available and because high quality field data were available to relate conductivity to effects on aquatic life in streams and rivers.

1.1. CONDUCTIVITY

Although the elements comprising the common mineral salts such as sodium chloride (NaCl) are essential nutrients, aquatic organisms are adapted to specific ranges of salinity and experience toxic effects from excess salinity. Salinity is the property of water that results from the combined influence of all disassociated mineral salts. The most common contributors to salinity in surface waters, referred to as matrix ions, are:

Cations: Ca^{2+} , Mg^{2+} , Na^+ , K^+

Anions: HCO_3^- , CO_3^{2-} , SO_4^{2-} , Cl^-

The salinity of water may be expressed in various ways, but the most common is *specific conductivity*. Specific conductivity (henceforth simply referred to as “conductivity”) is the ability of a material to conduct an electric current measured in microSiemens per centimeter ($\mu\text{S}/\text{cm}$) standardized to 25°C. (In this report, “conductivity” refers to the measurement, and resulting data and “salinity” refers to the environmental property that is measured.) Currents are carried by both cations and anions—but to different degrees depending on charge and mobility. Effectively, conductivity may be considered an estimate of the ionic strength of a salt solution. The ionic composition of mixtures of salts affects their toxicity (Mount et al., 1997). Therefore, a measure such as conductivity is necessary because the effects of the salts are a result of the magnitude of the exposure and the relative proportion of all of the ions in the mixture—not to any one individually. Hence, unless an individual ion occurs at a much higher concentration relative to its toxicity than other ions, the individual ion would not be the only potential cause, and a benchmark based on an individual ion could be under-protective. Therefore, this aquatic

life benchmark for conductivity is only appropriate for a mixture of salts dominated by the Ca^{2+} , Mg^{2+} , SO_4^{2-} , and HCO_3^- ions at a circum-neutral to mildly alkaline pH (6.0–10.0) in the Appalachian Region.

Salinity has numerous sources (Ziegler et al., 2007). Freshwater can become increasingly salty due to evaporation, which concentrates salts such as those in irrigation return waters (Rengasamy, 2002) or diversions that reduce inflow relative to evaporation (e.g., Pyramid Lake, Nevada). Intrusion of saltwater occurs when ground water withdrawal exceeds recharge especially near coastal areas (Bear et al., 1999; Werner, 2009). Freshwater can also become salty with the additions of brines and wastes (Clark et al., 2001), minerals dissolved from weathering rocks (Pond, 2004; U.S. EPA, 2011), and runoff from treating pavements for icy conditions (Environment Canada and Health Canada, 2001; Evans and Frick, 2000; Kelly et al., 2008).

Exposure of aquatic organisms to salinity is direct. Fish, amphibians, mussels, and aquatic macroinvertebrates are especially exposed on their gills or other respiratory surfaces that are in direct contact with dissolved ions in water. All animals have specific structures to transport nutrient ions and control their ionic and osmotic balance (Bradley, 2009; Evans, 2008a, b, 2009; Wood and Shuttleworth, 2008; Thorp and Covich, 2001; Komnick, 1977; Smith, 2001; Sutcliff, 1962; Hille, 2001). However, these cell membrane and tissue structures function only within a range of salinities. For example, some aquatic insects, such as most Ephemeroptera (mayflies), have evolved in a low-salt environment. Because they would normally lose salt in freshwater, their epithelium is selectively permeable to the uptake of certain ions and less permeable to larger ions and water. Many freshwater organisms depend heavily on specialized external mitochondria-rich chloride cells on the epithelium of their gills for the uptake of salts and export of metabolic waste (Komnick, 1977). Some life stages of animals may be particularly sensitive. For instance, ionic concentrations and transport processes are essential to regulate membrane permeability during external fertilization of eggs, including those of fish (Tarin et al., 2000).

Retention of ions is insufficient to maintain homeostasis and the actual uptake and export of ions occurs at semipermeable membranes. Anion, cation, and proton transport occurs by passive, active, uniport, and cotransport processes often in a coordinated fashion (Nelson and Cox, 2005; Hille, 2001). These numerous specific mechanisms are involved in the toxicity of solutions with relative ion concentrations different from what an organism typically encounters (see Appendix Section A.2.3 for more details on physiological mechanisms).

1.2. APPROACH

The approach used to derive the benchmark is based on the standard method for the EPA’s published Section 304(a) Ambient Water-Quality Criteria. Those criteria are the 5th centiles of species sensitivity distributions (SSDs) based upon laboratory toxicity tests, such that the goal is to protect at least 95% of the species in an exposed community (Stephan et al., 1985). SSDs are models of the distribution of exposure levels at which species respond to a stressor. That is, the most sensitive species respond at exposure level X_1 , the second most sensitive species respond at X_2 , etc. The species ranks are scaled from 0 to 1 so that they represent cumulative probabilities of responding, and the probabilities are plotted against the exposure levels (as seen in Posthuma et al., 2002). Centiles of the distribution can be derived using interpolation, parametric regression, or nonparametric regression. It should be noted that because SSDs are models of the distribution of sensitivity—and not just descriptions of the relative sensitivity of a particular set of species—they can be broadly applicable. In particular, SSDs derived using species from different continents are consistent for some chemicals (Hose and Van den Brink, 2004; Maltby et al., 2005).

For the conductivity benchmark, the SSDs are derived from field data. Some pollutants, such as suspended and bedded sediments (U.S. EPA, 2006; Cormier et al., 2008), and some assessment endpoints do not lend themselves to laboratory testing, and field data have some advantages for benchmark development (see Section 5.1). The differences between the method used here and the traditional method for deriving water-quality criteria are presented in Table 1, and the advantages are listed in Section 5.1.

Table 1. Differences between the method used to derive the conductivity benchmark and the method in Stephan et al. (1985) and the section of the report in which each is discussed

Difference	Section
Used field rather than laboratory data	5.1
Used extirpation as the response rather than a LC_{50} or CV	5.2
Used an integrative measure of a mixture rather than a single chemical	5.3
Used data from a particular region	5.4
Used the macroinvertebrate taxa from biological surveys rather than test species	5.6, 5.7, and 5.8

The choice to use field data to derive benchmarks of any kind poses some challenges. Because causal relationships in the field are uncontrolled, unreplicated, and unrandomized, they are more subject to a broader array of responses and to confounding. Confounding is the appearance of apparently causal relationships that are due to noncausal correlations. In addition, noncausal correlations and the inherent noisiness of environmental data can obscure true causal relationships. The potential for confounding is reduced, as far as possible, by identifying potential confounding variables, determining their contributions, if any, to the relationships of interest, and eliminating their influence when possible and as appropriate based on credible and objective scientific reasoning (see Appendix B). In addition, the evidence for and against salts as a cause of biological impairment is weighed using causal criteria adapted from epidemiology (see Appendix A).

Because relationships between conductivity and biological responses appear to vary among different mixtures of ions, this benchmark is limited to two contiguous regions with a particular dominant source of salinity. The regions are Level III 69 (Central Appalachian) and 70 (Western Allegheny Plateau) (see Figure 1) (U.S. EPA, 2007; Omernik, 1987; Woods et al., 1996). Low salinity rain water, sometimes so low as to not be accurately measured by conductivity, becomes salty as it interacts with the earth's surface. Along surface and ground water paths to the ocean, water contacts rocks. The rock demineralizes and contributes salts that accumulate. A large surface to volume ratio of unweathered rock increases dissolution of rock. For the most part, these salts are not degraded by natural processes but can be diluted by more rain or by less salty tributaries. Drought increases salt concentrations. Addition of wastes or waste waters also contributes salts. The prominent sources of salts in Ecoregions 69 and 70 are mine overburden and valley fills from large-scale surface mining, but they may also come from slurry impoundments, coal refuse fills, or deep mines. Other sources include effluent from waste water treatment facilities and brines from natural gas drilling and coalbed methane production. This benchmark for conductivity applies to waters influenced by current inputs from these sources in Ecoregions 69 and 70 with salts dominated by SO_4^{2-} and HCO_3^- anions at a circum-neutral to mildly alkaline pH.

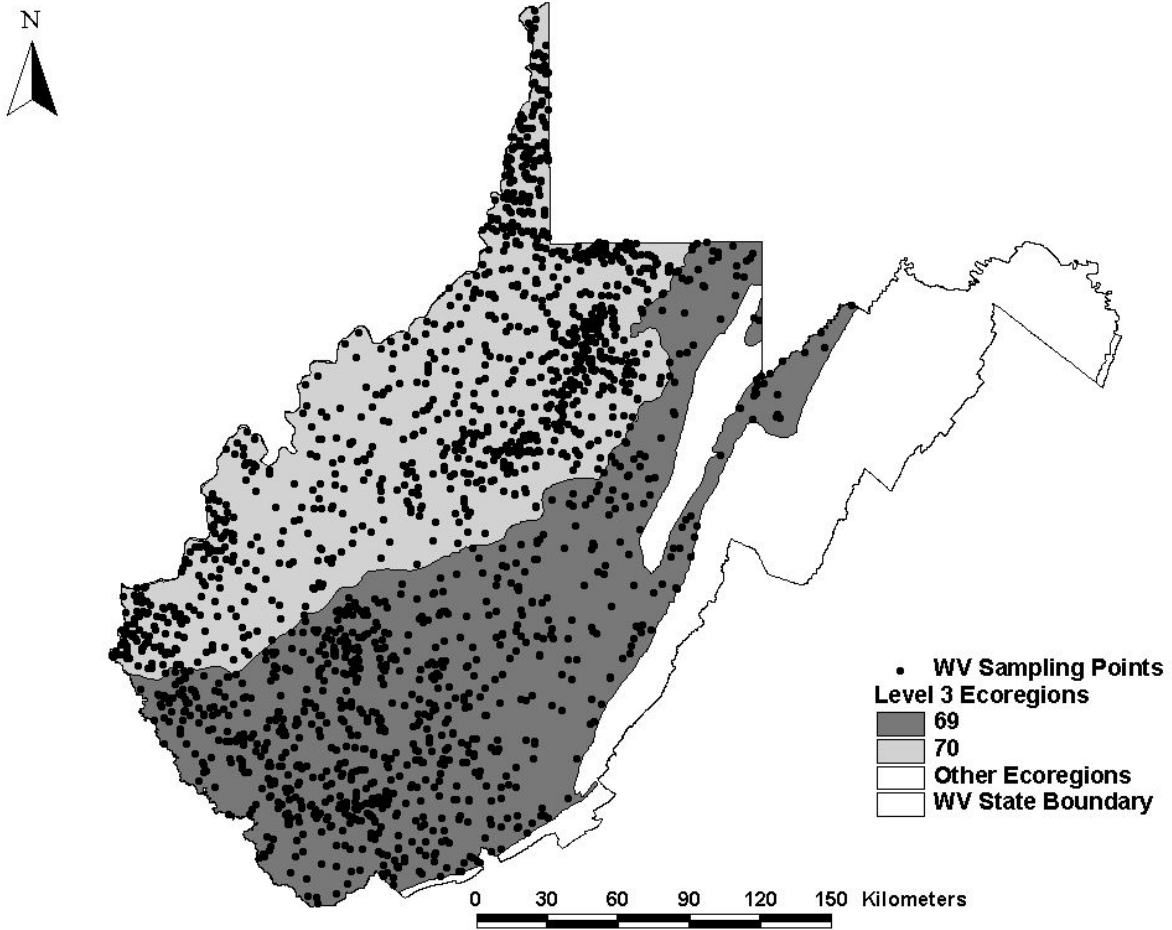


Figure 1. Points are sampling locations used to develop the benchmark from Level III Ecoregions 69 (light grey) and 70 in West Virginia.

2. DATA SETS

Data are required to develop and substantiate the benchmark. This section explains how the data were selected, describes the data that were used, and explains how the data set was refined to make it useful for analysis.

2.1. DATA SET SELECTION

The Central Appalachia (69) and Western Allegheny Plateau (70) Ecoregions were selected for development of a benchmark for conductivity because available data were of sufficient quantity and quality, and because conductivity has been implicated as a cause of biological impairment in these ecoregions (Pond et al., 2008; Pond, 2010; Gerritsen et al., 2010). These regions were judged to be similar in terms of water quality including resident biota and sources of conductivity. Confidence in the quality of reference sites in West Virginia was relatively high owing to the extensively forested areas of the region and well-documented process by which West Virginia Department of Environmental Protection (WVDEP) assigns reference status. They use a tiered approach. Only Tier 1 was used when analyses involved the use of reference sites, thus avoiding the use of conductivity as a characteristic of reference condition. Conductivity values from WVDEP's reference sites were low and similar in different months collected over several years (see Figure 2), providing evidence that the sites were reasonable reference sites. The 75th centiles were below 200 $\mu\text{S}/\text{cm}$ in most months. The 25th centiles from samples from a probability-based sample and from the full data set were below 200 $\mu\text{S}/\text{cm}$ in most months (see Figures 3 and 4). Also, a wide range of conductivity levels were sampled, which is useful for modeling the response of organisms to different levels of salinity.

2.2. DATA SOURCES

All data used for benchmark derivation were taken from the WVDEP's in-house Water Analysis Database (WABbase) 1999–2007. The WABbase contains data from Level III Ecoregions 66, 67, 69, and 70 in West Virginia (U.S. EPA, 2000; Omernik, 1987; Woods et al., 1996). In this assessment, only data from Ecoregions 69 and 70 were used (see Figure 1). Chemical, physical, and/or biological samples were collected from 2,542 distinct locations (2,668 samples) during the sampling years 1999–2007. WVDEP uses a tiered sampling design collecting measurements from long-term monitoring stations; targeted sites within watersheds on a rotating basin schedule; probability-based sites (Smithson, 2007); and sites chosen to further define impaired stream segments in support of total maximum daily load (TMDL) development (WVDEP, 2008b). Most sites have been sampled once during an annual sampling period, but

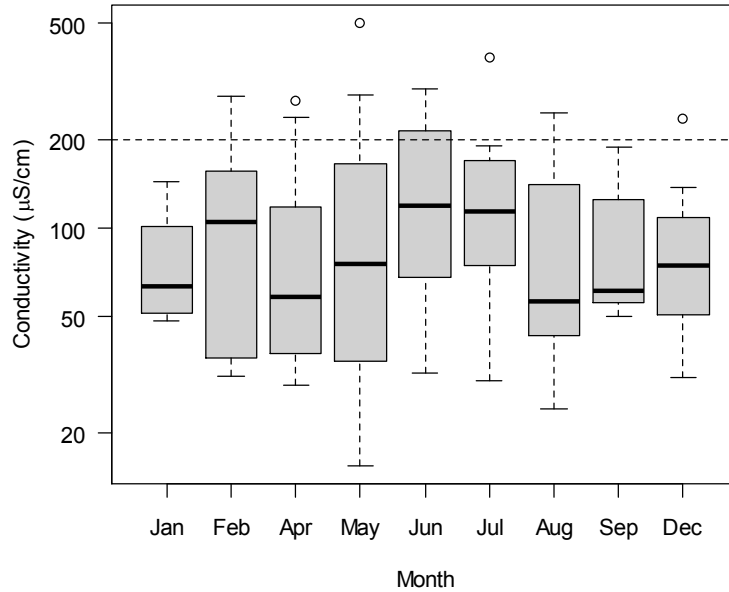


Figure 2. Box plot showing seasonal variation of conductivity ($\mu\text{S}/\text{cm}$) in the reference streams of Ecoregions 69 and 70 in West Virginia from 1999 to 2006. A total of 97 samples from 70 reference stations were used for this analysis. The 75th centiles were below 200 $\mu\text{S}/\text{cm}$ in all months except in June.

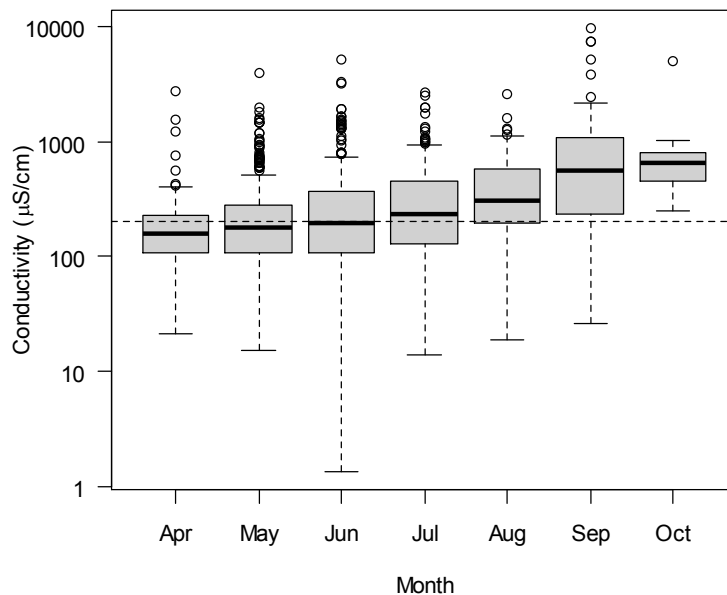


Figure 3. Box plot showing seasonal variation of conductivity ($\mu\text{S}/\text{cm}$) from a probability-based set of sample streams of Ecoregions 69 and 70 in West Virginia from 1997 to 2007. A total of 1,271 samples were used for this analysis. The 25th centiles were below 200 $\mu\text{S}/\text{cm}$ (horizontal dashed line) except in the September and October samples.

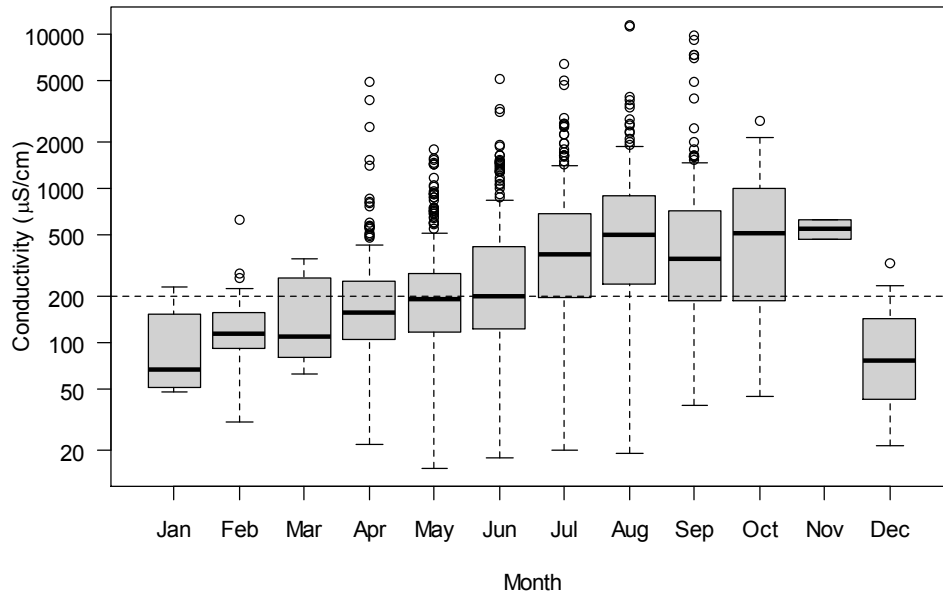


Figure 4. Box plot showing seasonal variation of conductivity ($\mu\text{S}/\text{cm}$) from the data set used to develop the benchmark. A total of 2,210 samples from 2000 to 2007 from Ecoregions 69 and 70 in West Virginia are represented. The 25th centiles were below 200 $\mu\text{S}/\text{cm}$ except in the August and November ($n = 2$) samples. The wide range of conductivities allows the XC_{95} to be well characterized.

TMDL sites have been sampled monthly for water-quality parameters. Some targeted sites represent least disturbed or reference sites that have been selected by a combination of screening values and best professional judgment (Bailey, 2009). Water quality, habitat, watershed characteristics, macroinvertebrate data (both raw data and calculated metrics), and supporting information are used by the State to develop 305(b) and 303(d) reports to the EPA (WVDEP, 2008b). All sites were in perennial reaches of streams.

Quality assurance and standard procedures are described by WVDEP (2006, 2008a). WVDEP collects macroinvertebrates from a 1- m^2 area of a 100-m reach at each site. When using 0.5-m-wide rectangular kicknet (595- μ mesh), four, 0.25- m^2 riffle areas are sampled. In narrow or shallow water, nine areas are sampled with a 0.33-m-wide D-frame dipnet of the same mesh size. Compositing samples are preserved in 95% denatured ethanol. A random subsample of 200 individuals $\pm 20\%$ are identified in the laboratory. All contracted analyses for chemistry and macroinvertebrate identification follow West Virginia's internal quality-control and quality-assurance protocols. This is a well-documented, regulatory database. EPA judged the quality assurance to be excellent based on the database itself, supporting documentation, and the experience of EPA Region 3 personnel.

Information was also obtained from the literature and other sources for the assessments of causality and confounding (see Appendices A and B):

- 1) Toxicity test results were obtained from peer-reviewed literature.
- 2) Information on the effects of dissolved salts on freshwater invertebrates was taken from standard texts and other physiological reviews.
- 3) An EPA Region 3 data set was obtained from Gregory Pond which includes the original data for Table 3 in Pond et al. (2008) and data collected for the Programmatic Environmental Impact Assessment (Bryant et al., 2002). It was used to evaluate the relative contribution of different ions in drainage from valley fills of large scale surface mining and for other analyses related to causation (see Appendix A) and confounding (see Appendix B). Some of these data were added since the 2010 public review draft.
- 4) The constituent ions for Marcellus Shale brine were provided by EPA Region 3 based on analyses by drilling operators (see Appendix A).
- 5) Data for Kentucky are from the Kentucky Department of Water database and are described in Appendix G, and results are presented in Appendices A, G, H, I, and J.
- 6) Geographic and related information are from WVDEP and various public sources and are described in Appendix C and also used in Appendix A.

2.3. DATA SET CHARACTERISTICS

Biological sampling usually occurred once per year with minimal repeat biological samples from the same location (5%). Multiple samples from the same location were not excluded from the data set (see Section 5.13). Summary statistics for ion concentrations and other parameters for the data set are provided in Table 2. The benchmark applies to waters with a similar composition to those in Table 2.

A total of 2,210 samples from Ecoregions 69 and 70 were used in the determination of the benchmark (see Figure 1 and Table 3). Data from a sampling event at a site were excluded from calculations if they lacked a conductivity measurement (see Table 4). Samples were excluded if the samples were identified as being from a large river ($>155 \text{ km}^2$) because the sampling methods differed (Flotermersch et al., 2001). They were excluded if the salt mixture was dominated by Cl^- rather than SO_4^{2-} (conductivity $> 1,000 \text{ }\mu\text{S/cm}$, $\text{SO}_4 < 125 \text{ mg/L}$, and $\text{Cl}^- > 250 \text{ mg/L}$). Four sites with elevated conductivity, high chloride, and low sulfate were removed in response to concerns that the benchmark might be biased by sites with salts dominated by Marcellus Shale brines.

Table 2. Summary statistics of the measured water-quality parameters

Parameter	Units	Min	25th centile	Median	75th centile	Max	Mean	Valid N
Conductivity	µS/cm	15.4	146	261	563	11,646	281.5	2,210
Hardness	mg/L	0.5	50.2	91.1	188	1,492	97.1	1,148
Alkalinity	mg/L	0.2	30.5	66.7	117	560	55	1,425
SO ₄ ²⁻	mg/L	1	17	37	159	6,000	51.6	1,428
Cl ⁻	mg/L	1	3	5.2	11.95	1,153	6.5	1,118
Ca, total	mg/L	0.002	13.6	25.1	49.2	430	25.5	1,154
Mg, total	mg/L	0.05	3.7	6.3	14	204	7.3	1,150
TSS	mg/L	1	3	4	6	190	4.3	1,442
Fe, total	mg/L	0.005	0.123	0.26	0.5	110	0.26	1,433
NO ₂ -NO ₃	mg/L	0.01	0.1	0.2	0.37	30	0.20	1,178
Al, total	mg/L	0.01	0.09	0.11	0.23	12	0.15	1,436
Al, dissolved	mg/L	0.01	0.02	0.05	0.06	0.93	0.04	1,287
Fe, dissolved	mg/L	0.001	0.02	0.042	0.06	11.8	0.05	1,259
Mn, total	mg/L	0.003	0.02	0.04	0.1	7.25	0.05	1,430
Mn, dissolved	mg/L	0.01	0.03	0.07	0.22	1.06	0.07	20
Total phosphate	mg/L	0.01	0.02	0.02	0.03	2.36	0.03	1,181
Se, dissolved	mg/L	0.001	0.001	0.001	0.001	1.26	0.001	313
Se, total	mg/L	0	0.001	0.001	0.005	1.26	0.002	496
Fecal coliform	counts/ 100 mL	0	36	170	600	250,000	151	2,035
DO	mg/L	1.02	8.2	9.2	10.3	18.35	9.3	2,182
pH	standard units	6.02	7.27	7.62	7.96	10.48	7.59	2,210
Catchment area	km ²	0.173	2.311	6.965	25.836	153.014	7.644	717
Temperature	°C	-0.28	15.1	18.4	21.3	31.9	17	2,210
Habitat	RBP score	49	115	130	145	192	127.8	2,186

Note: K⁺ and Na⁺ not measured; all means are geometric means except pH, DO, Temperature, and Habitat Score. DO = dissolved oxygen; RBP = rapid bioassessment protocol; TSS = total suspended solids.

Table 3. Number of samples with reported genera and conductivity meeting our acceptance criteria (see Table 4) for calculating the benchmark value.
Number of samples is presented for each month and ecoregion

Region	Month												Total
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
69	8	4	1	63	187	103	79	269	232	54	0	6	1,006
70	4	33	4	187	232	179	194	237	120	8	2	4	1,204
Total	12	37	5	250	419	282	273	506	352	62	2	10	2,210

Table 4. Samples excluded from the original data sets of 2,668 samples used to develop benchmark value

Parameter	Exclusion level	N of samples excluded
Catchment	>155 km ²	295
Conductivity	No measurement	10
pH	<6	147
High Cl ⁻	>1,000 μS/cm, SO ₄ <125 mg/L, and Cl ⁻ >250 mg/L	4
Taxonomic identification	Ambiguous taxa, family, or higher level	2

The effects of low pH were eliminated by excluding sites with a pH of <6. This prevented potential confounding of conductivity effects by the effects of acid mine drainage (see Appendix B). The freshwater chronic water-quality criterion requires waters to be maintained between a pH of 6.5 and 9 (U.S. EPA, 1986). The conductivity benchmark was derived from waters having a pH between 6.0 and 10. Thus the circum-neutral application brackets pH of 7 primarily in the range where pH is usually not toxic to freshwater organisms.

An organism was excluded from calculations if it was not identified to the genus level, and a genus was excluded if it was never observed at reference sites or it was observed in <25 samples. Invertebrate genera that did not occur at WVDEP Tier 1 reference sites represented 11.4% of all genera (see Table 5). They were excluded so that the data would be relevant to potentially unimpaired conditions—and so as to not include opportunistic salt-tolerant organisms. Genera were excluded that were observed at fewer than 25 sampling locations in the composited ecoregion thus ensuring reasonable confidence in the evaluation of the relationship

Table 5. Genera excluded from 95th centile extirpation concentration calculation because they never occurred at reference sites

<i>Argia</i>	<i>Ferrissia</i>	<i>Oecetis</i>	<i>Prostoma</i>	<i>Stictochironomus</i>
<i>Baetisca</i>	<i>Fossaria</i>	<i>Palpomyia</i>	<i>Saetheria</i>	<i>Tokunagaia</i>
<i>Calopteryx</i>	<i>Leucotrichia</i>	<i>Paracladopelma</i>	<i>Sphaerium</i>	<i>Tribelos</i>
<i>Corbicula</i>	<i>Nanocladius</i>	<i>Paratendipes</i>	<i>Stenochironomus</i>	<i>Tricorythodes</i>
<i>Dineutus</i>				

between conductivity and the presence and absence of a genus. This decision was made because an analysis showed that the benchmark varied within <5% when SSD models were constructed from ≥ 20 occurrences of each genus; whereas the benchmark steadily became lower when XC_{95} values were derived from <15 occurrences.

In the WABbase, 497 benthic invertebrate genera were identified in Ecoregions 69 and 70. Those ecoregions had 308 genera in common. Of these, 220 genera occurred at least once at 1 of the 70 reference sites in the two ecoregions. Genera that did not occur at reference sites were excluded from the SSD (see Table 5). Greater than 95% of genera observed at reference sites as defined by WVDEP occur in both Ecoregions 69 and 70. This indicates that the same sensitive genera exist in both ecoregions. Of the 220 genera, 163 occurred at ≥ 25 sampling locations in Ecoregions 69 and 70. Of the genera occurring at ≥ 25 sampling sites, 162 genera occurred in Ecoregion 69 and 163 in Ecoregion 70.

3. METHODS

The derivation of the benchmark for conductivity includes three steps: first, the extirpation values (XCs) for each invertebrate genus was derived. Second, the XC_{95} values for all genera were used to generate an SSD and the 5th centile of the distribution, the 5th centile hazardous concentration (HC_{05}). (The HC_X terminology for concentrations derived from SSDs is not in the EPA method [Stephan et al., 1985], but its usage has become common more recently [Posthuma et al., 2002]). Finally, background values were estimated for the regions to ensure that the benchmark is not in the background range. These steps are explained in this section.

We used the statistical package R, version Version: 2.12.1 (December 2010), for all statistical analyses (<http://www.r-project.org/>).

3.1. EXTIRPATION CONCENTRATION DERIVATION

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 (U.S. EPA, 2003). In this report, extirpation is operationally defined for a genus as “the conductivity value below which 95% of the observations of the genus occur and above which only 5% occur.” In other words, the probability is 0.05 that an observation of a genus occurs above its XC_{95} conductivity value. This is a chronic-duration endpoint because the field data set reflects exposure over the entire life cycle of the resident biota. The 95th centile was selected because it is more reliable than the maximum value, yet it still represents the extreme of a genus’s tolerance of conductivity. The maximum value is sensitive to occurrences due to drifting organisms, misidentifications, or other misleading occurrences.

The XC_{95} is estimated from the cumulative distribution of probabilities of observing a genus at a site with respect to the concurrently measured conductivity at that site. Observed conductivity values were nonuniformly distributed across a range of possible values (see Figure 5), and, therefore, we were more likely to observe a genus at certain conductivity values simply because more samples were collected at those values. To correct for the uneven sampling frequency, a weighted cumulative distribution function was used to estimate the XC_{95} values for each genus. The purpose of weighting is to avoid bias due to uneven distribution of observations with respect to conductivity by converting the sampling distribution to one that mimics an even distribution of sample across the gradient of conductivity. It creates a distribution more like the design of a toxicity test, which is appropriate when developing an exposure-response relationship. To compute weights for each sample, we first defined equally-sized bins, each

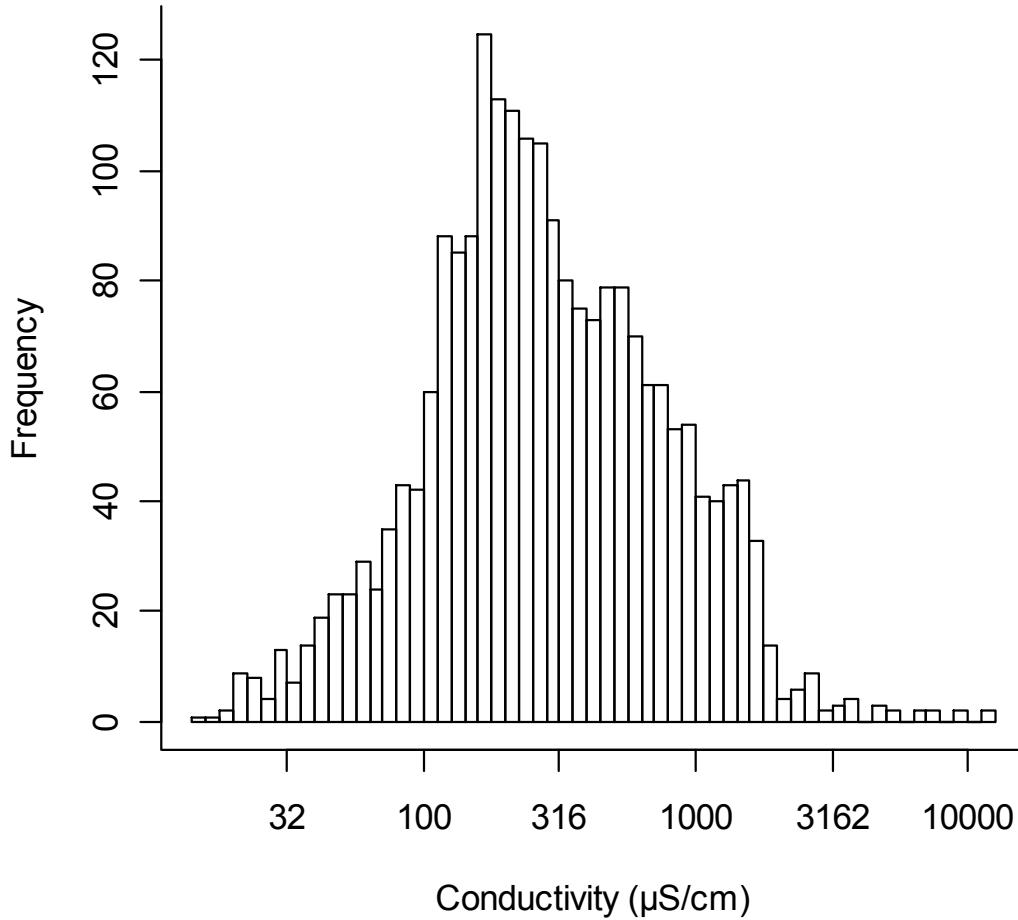


Figure 5. Histograms of the frequencies of observed conductivity values in samples from Ecoregions 69 and 70 from West Virginia sampled between 1999 and 2006. Bins are each $0.017 \log_{10}$ conductivity units wide.

$0.017 \log_{10}$ conductivity units wide, that spanned the range of observed conductivity values, a total of 60 bins. We then calculated the number of samples that occurred within each bin (see Figure 5). 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 cumulative distribution function, $F(x)$, of conductivity values associated with observations of a particular genus was computed for each unique observed value of conductivity, x , as follows:

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

where x_{ij} is the conductivity 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 genus of interest was observed in j^{th} sample of bin i ,

I is an indicator function that equals 1 if the indicated conditions are true, and 0 otherwise.

The XC_{95} value is defined as the conductivity value, x where $F(x) = 0.95$. Equation 1 is an empirical cumulative distribution function, and the output is the proportion of observations of the genus that occur at or below a given conductivity level. However, the individual observations are weighted to account for the uneven distribution of observations across the range of conductivities.

An example of a weighted cumulative distribution function (CDF) is shown in Figure 6 for the mayfly, *Drunella*. The horizontal dashed line indicates the point of extirpation where $F(x) = 0.95$ intersects the CDF. The vertical dashed line indicates the XC_{95} conductivity value on the x-axis.

This method for calculating the XC_{95} will generate a value even if the genus is not extirpated. For example, the occurrence of *Nigronia* changes little with increasing conductivity (see Figure 6). In order to examine the trend of taxa occurrence along the conductivity gradient, we used a nonparametric function (Generalized Additive Model [GAM] with 3 degrees of freedom) to model the likelihood of a taxon being observed with increasing conductivity (see Figure 7). The solid line is the mean smoothing spline fit. The dots are the mean observed probabilities of occurrence, estimated as the proportion of samples within each conductivity bin. The conductivity at the red, vertical, dashed line is the estimated XC_{95} from the weighted cumulative distribution (see Appendix E).

Because of the data distributions, not all 95th centiles correspond to extirpation, and some imprecisely estimate the extirpation threshold. The following rules were applied to the XC_{95} values. If the GAM mean curve at maximum conductivity is approximately equal to 0 (defined as less than 1% of the maximum modeled probability), then the XC_{95} is listed without qualification. If the GAM mean curve at maximum conductivity is >0 , but the lower confidence

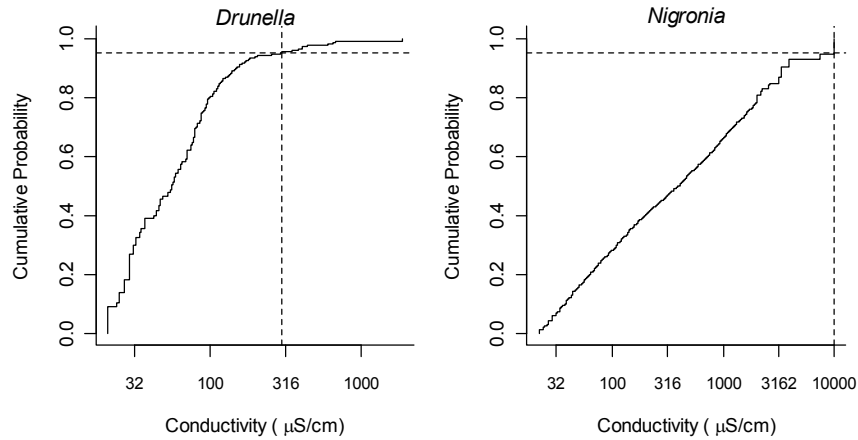


Figure 6. Examples of weighted CDFs and the associated 95th centile extirpation concentration values. The step function shows weighted proportion of samples with *Drunella* or *Nigronia* present at or below the indicated conductivity value ($\mu\text{S/cm}$). The XC_{95} is the conductivity at the 95th centile of the cumulative distribution function (CDF) (vertical dashed line). In a CDF, genera that are affected by increasing conductivity (e.g., *Drunella*) show a steep slope and asymptote well below the maximum exposures; whereas, genera unaffected by increasing conductivity (e.g., *Nigronia*) have a steady increase over the entire range of measured exposure and do not reach a clear asymptote.

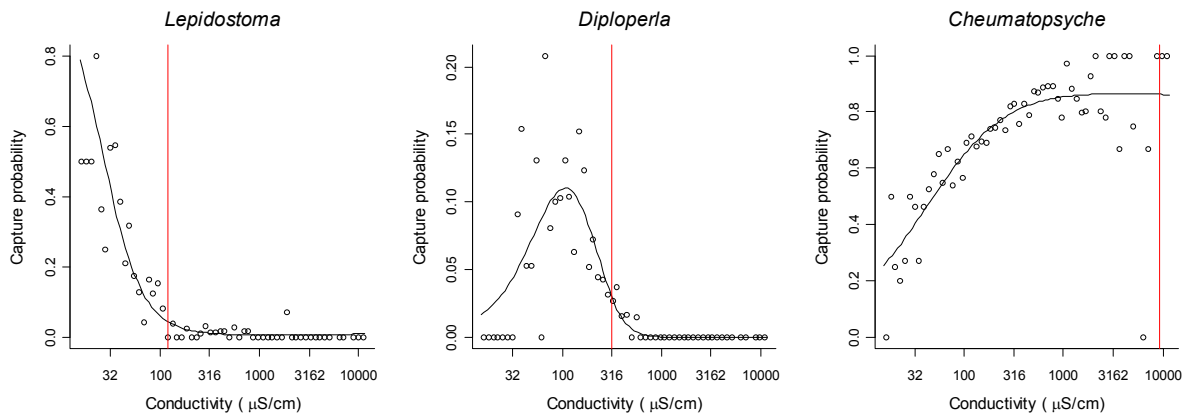


Figure 7. Three typical distributions of observation probabilities. Open circles are the probabilities of observing the genus within a range of conductivities. Circles at zero probability indicate no individuals at any sites were found at these conductivities. The lines fitted to the probabilities are for visualization. The vertical red line indicates the XC_{95} . Note that different genera respond differently to increasing salinity. *Lepidostoma* declines, *Diploperla* has an optimum, and *Cheumatopsyche* increases. The XC_{95} for genera like *Cheumatopsyche* are reported as “greater than” because extirpation did not occur in the measured range.

limit is approximating to 0 (<1% of the maximum mean modeled probability) the value is listed as approximate (~). If the GAM lower confidence limit is >0, then the XC₉₅ is listed as greater than (>) the 95th centile. All model fits and the scatter of points were also visually inspected for anomalies and if the model poorly fit the data, the uncertainty level was increased to either (~) or (>). This procedure was applied to the distributions in Appendices E and I, and the results appear in Appendices D and H. Also, these models were used to evaluate when genera began to decline as evidence of alteration and sufficiency in Appendix A.

For example, the XC₉₅ for *Cheumatopsyche* (an extremely salt tolerant genus) is >9,180 µS/cm (see Appendices D and E). Whereas, although *Pteronarcys* is declining, the upper confidence bound is >0; therefore, its XC₉₅ is ~634 µS/cm. The assignment of (>) and (~) does not affect the HC₀₅, but are provided to alert users of the uncertainty of the XC₉₅ values for other uses such as comparison with toxicity test results or with results from other geographic regions.

3.2. TREATMENT OF POTENTIAL CONFOUNDERS

Potentially confounding variables for the relationship of conductivity with the extirpation of stream invertebrates were evaluated in several ways, which are described in Appendix B. Based on the weight of evidence, only low pH was a likely confounder. As mentioned previously in Section 2.3, because low pH waters violate existing water-quality criteria and because the data set was large, sites were excluded with pH <6 before identifying the XC₉₅ values.

We evaluated the effects of spring benthic invertebrate emergence, temperature, and different conductivities associated with season by partitioning the data set into spring (March–June) and summer (July–October) subsets. However, we found that the SSDs for spring and all year were similar. Because high and low exposures occurred in all seasons, we chose to include the occurrence of a genus whenever it was observed. Therefore, although we explored season, we could not justify excluding an observation of a genus just because it was seen outside an imposed time frame.

Other potential confounders were evaluated by weighing the available evidence. Because confounders are by definition correlated with the cause of concern and the effect, we determined the degree of correlation of the confounder with conductivity and with the number of ephemeropteran genera. We also evaluated contingency tables of the occurrence of any Ephemeroptera at a site with respect to high and low levels of conductivity and the potential confounder. Ephemeroptera were selected as an effect endpoint that allowed us to evaluate a greater range of exposures and confounding factors than occurs for individual genera. The confounding analysis focused on Ephemeroptera because they are among the most sensitive genera. Other evidence of confounding was included when appropriate data were available.

3.3. DEVELOPING THE SPECIES SENSITIVITY DISTRIBUTION

The SSDs are cumulative distribution plots of XC_{95} values for each genus relative to conductivity (see Figures 8 and 9). The cumulative proportion for each genus P is calculated as $P = R / (N + 1)$ where R is the rank of the genus and N is the number of genera. Some salinity-tolerant genera are not extirpated within the observed range of conductivity. So, like laboratory test endpoints reported as “greater than” values, we retained field data that do not show the field endpoint effect (extirpation) in the database. In this way, they can be included in N when calculating the proportions responding because they fall in the upper portion of the SSD. The HC_{05} was derived by using a 2-point interpolation to estimate the centile between the XC_{95} values bracketing $P = 0.05$ (i.e., the 5th centile of modeled genera). The benchmark is obtained by rounding the HC_{05} to two significant figures as directed by Stephan et al. (1985).

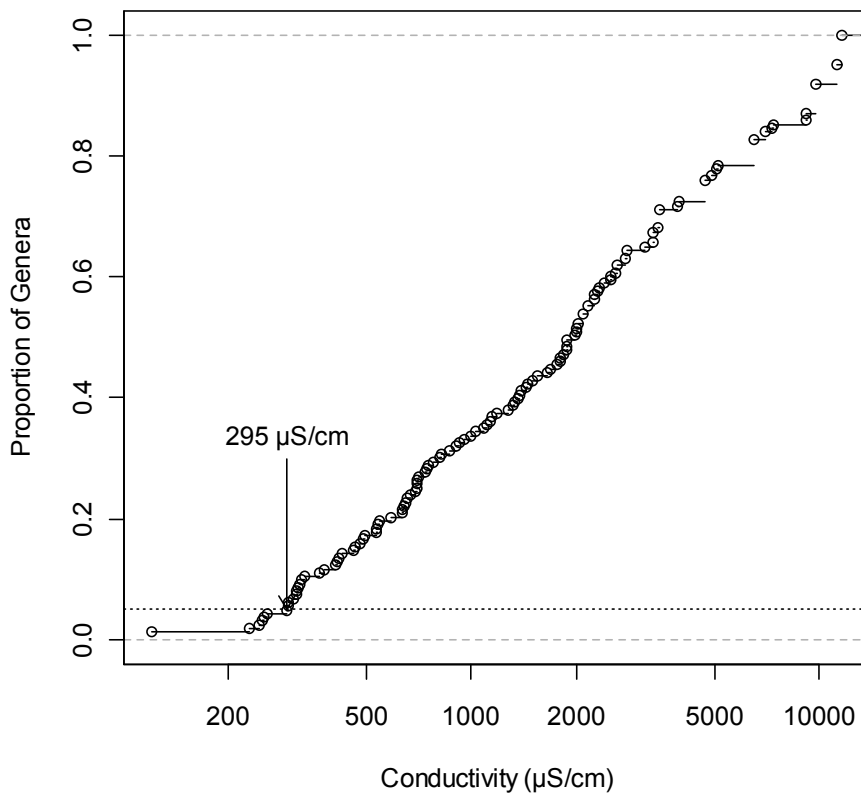


Figure 8. The species sensitivity distribution. Each point is an XC_{95} value for a genus. There are 163 genera. The HC_{05} (295 $\mu\text{s}/\text{cm}$) is the conductivity at the intercept of the SSD with the horizontal line at the 5th centile.

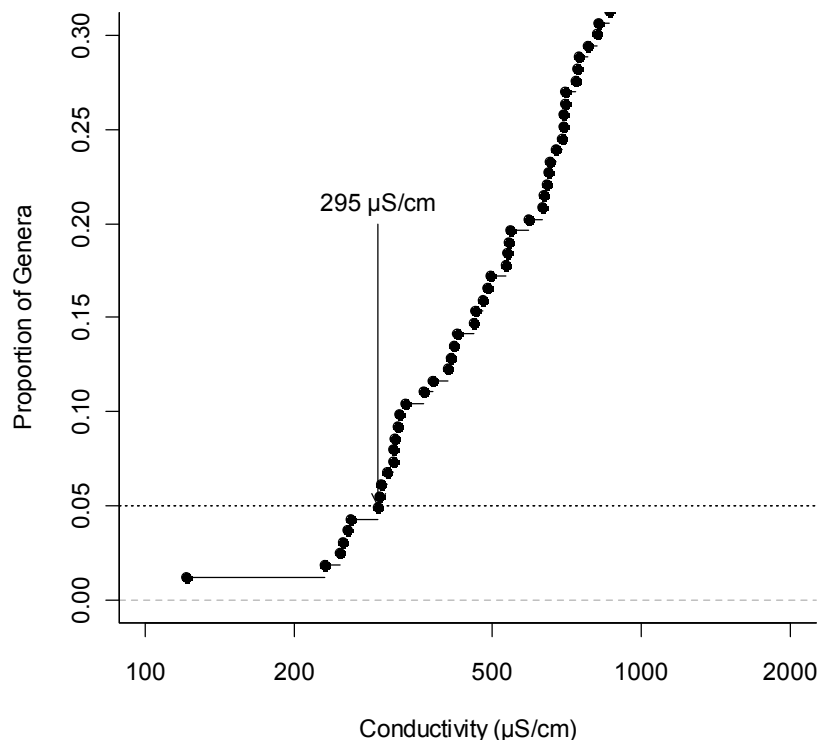


Figure 9. Species sensitivity distribution (expanded). The dotted horizontal line is the 5th centile. The vertical arrow indicates the HC₀₅ of 295 µS/cm. Only the lower 50 genera are shown to better discriminate the points in the left side of the full distribution.

3.4. CONFIDENCE BOUNDS

The purpose of this analysis is to characterize the uncertainty in the benchmark value by calculating confidence bounds on the HC₀₅ values. Because the XC₉₅ values were estimated from field data and then the HC₀₅ values were derived from those XC₉₅ values, we used a method that generated distributions and confidence bounds in the first step and propagated the statistical uncertainty of the first step through the second step (see Figure 10). Bootstrapping is commonly used in environmental studies to estimate confidence limits of a parameter, and the method has been used in the estimation of HC₀₅ values (Newman et al., 2000, 2002).

Bootstrap estimates of the XC₉₅ were derived for each genus used in the derivation of the benchmark by resampling 2,210 times (the number of observations in the data set) with replacement (see Figure 10) (Efron and Tibshirani, 1993). From each bootstrap sample, the XC₉₅ was calculated for each genus by the same method applied to the original data (see Section 3.1). That process was repeated 1,000 times to create a distribution of XC₉₅ values for each genus. These distributions were used to calculate a two-tailed 95% confidence interval on

the XC_{95} for each genus. The XC_{95S} from the original data set, the mean XC_{95S} of the bootstrap distributions, and the confidence intervals are shown for the 36 most sensitive genera (see Figure 11).

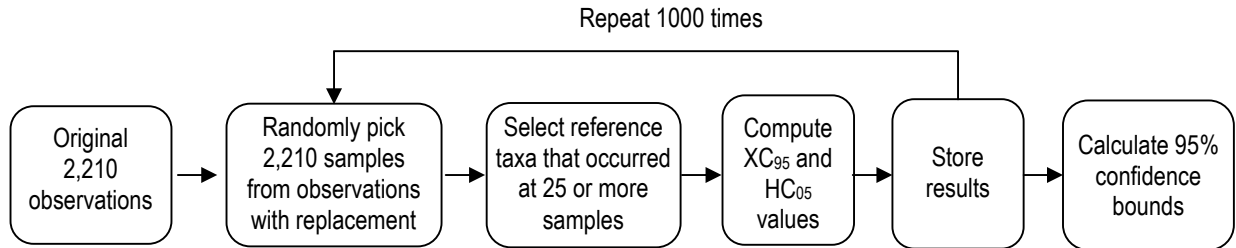


Figure 10. Diagram depicting the process for estimating the uncertainty of the HC_{05} .

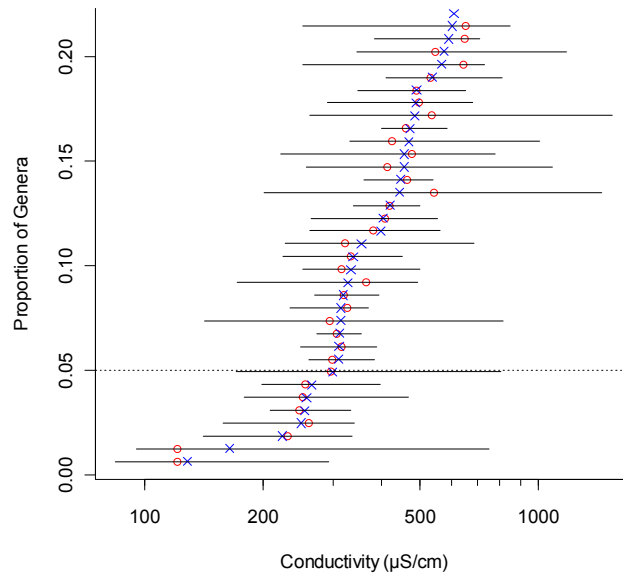


Figure 11. The cumulative distribution of XC_{95} values for the 36 most sensitive genera (red circles) and the bootstrap-derived means (blue x symbol) and two-tailed 95% confidence intervals (whiskers). The 5th centile is shown by the dashed line.

Uncertainty in the HC_{05} value was evaluated by generating an HC_{05} from each of the 1,000 sets of bootstrapped XC_{95} estimates. The distribution of 1,000 HC_{05} values was used to generate two-tailed 95% confidence bounds on these bootstrap-derived values.

3.5. EVALUATING ADEQUACY OF NUMBER OF SAMPLES

Bootstrapping was performed to evaluate the effect of sample size on the HC_{05} and their confidence bounds. This process is similar to the method used to calculate confidence bounds on the HC_{05} values (see Figure 10). A data set with a selected sample size was randomly picked with replacement from the original 2,210 samples. From the bootstrap data set, the XC_{95} was calculated for each genus by the same method applied to the original data and the HC_{05} was also calculated. The uncertainty in the HC_{05} value was evaluated by repeating the sampling and HC_{05} calculation 1,000 for each sample size. The distribution of 1,000 HC_{05} values was used to generate two tailed 95% confidence bounds on these bootstrap-derived values. The whole process was repeated for a selected sample size range from 100 to 2,210 samples. The mean HC_{05} values, the numbers of genera used for HC_{05} calculation, and their 95% confidence bounds, were plotted to show the effects of sample sizes. The HC_{05} values stabilize at approximately 800 samples in this data set, which suggests that 800 is a minimum sample size for this method (see Figure 12). Note that, the mean HC_{05} value is lower than the actual HC_{05} value at a similar sample size, because the Monte Carlo results are asymmetrical (i.e., there are more ways that the sample variance can result in lower values than higher values).

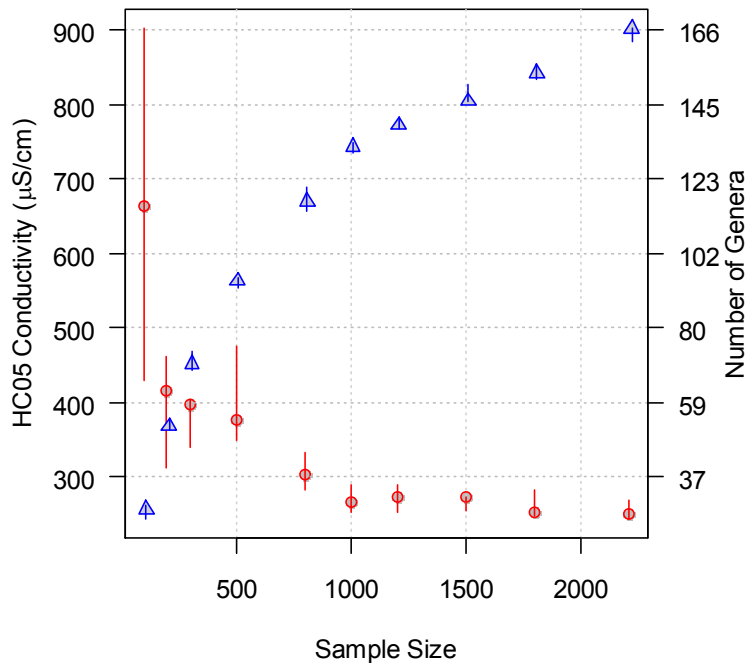


Figure 12. Adequacy of the number of samples used to model the HC_{05} . As sample size increases the number of genera included in the SSD increases (triangles). As sample size increases, the confidence bounds on the HC_{05} decreases and the mean HC_{05} is asymptotic at $<300 \mu s/cm$ (circles).

3.6. ESTIMATING BACKGROUND

In general, a benchmark should be greater than natural background. The background conductivities of streams were estimated using that portion of the WABbase that consists of probability-based samples. Those are samples from locations that were selected to represent streams within a stream order with equal probability. The 25th centile of the probability-based samples was selected as the estimate of the upper limit of background because disturbed and even impaired sites are included in the sample (U.S. EPA, 2000). A total of 1,271 probability-based samples were collected from Ecoregions 69 and 70. The background values on the 25th centile were 72 $\mu\text{S}/\text{cm}$ for Ecoregion 69, 153 $\mu\text{S}/\text{cm}$ for Ecoregion 70, and 116 $\mu\text{S}/\text{cm}$ when samples from Ecoregions 69 and 70 are combined (see Figure 3). We also estimated the background conductivity using reference sites in WABbase (see Figure 2). The 75th centiles from 43 sites in Ecoregion 69 and 27 sites in Ecoregion 70 are 66 $\mu\text{S}/\text{cm}$ for Ecoregion 69, 214 $\mu\text{S}/\text{cm}$ for Ecoregion 70. When samples from Ecoregions 69 and 70 are combined, the 75th centile is 150 $\mu\text{S}/\text{cm}$. Sampling locations were among the least disturbed based on WVDEP's best professional judgment (WVDEP, 2008a, b); therefore, the 75th centile was selected (U.S. EPA, 2000). The bases for selecting centiles are explained in Section 5.5.

4. RESULTS

4.1. EXTIRPATION CONCENTRATIONS

The XC_{95} values are presented in Appendix D. Values are calculated for all macroinvertebrate genera that were observed at a reference site and at a minimum of 25 sampling sites in the two ecoregions. Distributions of occurrence with respect to conductivity are presented for each genus of macroinvertebrate in Appendix E and the CDFs used to derive the XC_{95} values are presented in Appendix F.

4.2. SPECIES SENSITIVITY DISTRIBUTIONS

A SSD for invertebrates is derived from XC_{95} values of 163 genera (see Figure 8). The SSDs do not reach a horizontal asymptote at 100% of genera because salt-tolerant genera are included in the SSD that are not extirpated within the observed range of conductivity values. The lower third of the SSD is shown in Figure 10 for better viewing of the points near the 5th centile of genera.

4.3. HAZARDOUS CONCENTRATION VALUES AT THE 5TH CENTILE

The hazardous concentration value at the 5th centile of the SSDs is 295 $\mu\text{S}/\text{cm}$. Rounding the HC_{05} to two significant figures yields a benchmark value of 300 $\mu\text{S}/\text{cm}$.

4.4. UNCERTAINTY ANALYSIS

The bootstrap statistics yield 95% confidence bounds of 228 and 303 $\mu\text{S}/\text{cm}$ (see Figure 11). The asymmetry of the confidence bounds with respect to the point estimate of 295 $\mu\text{S}/\text{cm}$ is not unusual. In bootstrap-generated estimates, such as those used here, asymmetry occurs because statistical resampling from the distribution of data generates more realizations that produce values lower than the point estimate than realizations that produce higher values.

Confidence bounds represent the potential range of HC_{05} values using the SSD approach, given the data and the model. Conceptually, these confidence bounds may be thought of representing the potential range of HC_{05} values that one might obtain by returning to West Virginia and resampling the streams. The contributors to this uncertainty include measurement variance in determining conductivity and sampling variance in the locations for monitoring and in collecting and enumerating organisms. It also includes variance due to differences in stream reaches, weather, and other random factors.

The confidence bounds do not address potential systematic sources of variance such as differences between geographic areas or between different organizations performing the

sampling using different protocols. The contributions of those sources of uncertainty—in addition to the sampling uncertainty—can best be evaluated by comparing the results of independent studies. One estimate of that uncertainty is provided by comparing the all-year HC₀₅ values derived from West Virginia and Kentucky data. Even though the data were obtained in different areas by different agencies using different laboratory processing protocols, the values (West Virginia: 295 μS/cm, Kentucky: 282 μS/cm) differ by <5% (see Appendix G for details). In addition, the 95% confidence bounds on the HC₀₅ values for the two states overlap, suggesting that the sampling variance (i.e., the uncertainty captured by the confidence intervals) may be the largest component of total uncertainty. While this result is from only one comparison of two states, it does provide a reassuring validation of the West Virginia results.

5. CONSIDERATIONS

Because of the complexity of field observations, decisions must be made when deriving field-based benchmark values that are not required when using laboratory data. In the case of conductivity, additional decisions must be made to address a pollutant that is a mixture and a naturally occurring constituent of water.

5.1. CHOOSING TO USE FIELD VERSUS LABORATORY DATA

The standard methodology for deriving water-quality criteria uses results from laboratory toxicity studies (Stephan et al., 1985); however, we have adapted the method to use field data because suitable laboratory data are not available. Furthermore, SSDs based on laboratory studies cannot replicate the range of conditions, effects, or interactions that occur in the field (Suter et al., 2002). Although field data require additional assurance of attributable causation due to potential confounders (Section 5.15, Appendices A and B), field data have many advantages over laboratory data.

- 1) Field exposures include realistic levels, proportions, and variability of pollutant mixtures.
- 2) Field exposures occur in inherently realistic physical and chemical conditions.
- 3) Field exposures include regionally appropriate taxa and relative abundances of taxa.
- 4) Field studies can include more taxa than are available in laboratory data sets.
- 5) Field data include appropriately sensitive taxa and life stages.
- 6) Field data include pollutant interactions with migration, predation, competition, and other behaviors.
- 7) Organisms in the field have realistic nutrition and levels of stress.
- 8) Organisms in the field realistically integrate effects of pollutants and other conditions into a population response.
- 9) The field chronic endpoint (extirpation of a population) is inherently relevant, but the chronic laboratory test endpoints correspond to no particular effect (chronic values—CVs).

This study can benefit from these inherent advantages of field data because of the availability of large, high quality data sets with clear effects of the pollutant and little evidence of confounding.

5.2. SELECTION OF THE EFFECTS ENDPOINT

We have used the extirpation concentration as the effects endpoint because it is easy to understand that an adverse effect has occurred when a genus is lost from an ecosystem. However, for the same reason, it may not be considered protective. An alternative is to use a depletion concentration (DC_x) based on a percent reduction in abundance or capture probability. Another option is to use only those taxa sensitive to the stressor of concern, thus developing an SSD for the most relevant taxa. DC_x values or other more sensitive endpoints may be considered when managing exceptional resources.

In this study, an invertebrate genus may represent several species, and this approach identifies the pollutant level that extirpates all sampled species within that genus, that is, the level at which the least sensitive among them is rarely observed. In a review of extrapolation methods, Suter (2007) indicated that although species within a genus respond similarly to toxicants, different species within a genus could have evolved to partition niches afforded by naturally occurring causal agents such as conductivity (Remane, 1971). Hence, an apparently salt tolerant genus may contain both sensitive species and tolerant species. A potential solution would be to use distinct species. However, this may not be practical because some taxa are very difficult to identify except as late instars. We chose to follow Stephan et al. (1985) by using genera until such time that the advantages and disadvantages of using species can be more fully studied.

Because this endpoint is based on full life-cycle exposures and responses of populations to multigenerational exposures, it is considered a chronic-duration endpoint.

5.3. TREATMENT OF MIXTURES

In natural waters, salinity is a result of mixtures of ions. A metric is required to express the strength of that mixture. We use conductivity because it is a measure of the ionic strength of the solution, because it is related to biological effects, and because it is readily measured accurately. However, conductivity per se is not the cause of toxic effects, and waters with different mixtures of salts but the same conductivity may have different toxicities. In this case, the benchmark value was calculated for a relatively uniform mixture of ions in those streams that exhibit elevated conductivity in the Appalachian Region associated with salts dominated by Ca^+ , Mg^+ , SO_4^{2-} , and HCO_3^- ions at circum-neutral to mildly alkaline pH (pH 6–10). Recent increases in drilling for natural gas may change the toxicity of salinity in this region, and monitoring should be designed to evaluate differences. The relative contributions of individual salts from large-scale surface coal mining are described by Pond et al. (2008). Whereas Ca^{2+} , Mg^{2+} , SO_4^{2-} , and HCO_3^- are the four most common ions to drain from surface coal mines (Bryant et al., 2002), ions of Na^+ and Cl^- are the two most common in seawater and brines from

Marcellus Shale drilling operations (see Appendix A, Table A-16). Because the few sites with very elevated Cl^- were found to be outliers in the distributions of occurrence, they were deleted from the data set used to derive the XC_{95} values. Hence, the use of the benchmark value in other regions or in waters that are contaminated by other sources, such as road salt or irrigation return waters, may not be appropriate. However, for the circum-neutral to alkaline drainage from surface mines and valley fills, these four primary ions are highly correlated with conductivity (see Figures 13a–e).

5.4. DEFINING THE REGION OF APPLICABILITY

If the method for developing a benchmark as described here is applied to a large region, the increased range of environmental conditions and a greater diversity of anthropogenic disturbances may obscure the causal relationship. However, if the region is too small, the available data set may be inadequate, and the resulting benchmark value will have a small range of applicability. In this case, we chose two adjoining regions that have abundant data, >95% of genera in common, and a common dominant source of the stressor of concern.

Although Ecoregions 69 (Central Appalachia) and 70 (Western Allegheny Plateau) are very similar, including similar bedrock types, the relative abundances differ. The coal-bearing subregions of the Central Appalachians are 69a (Forested Hills and Mountains), and 69d (Cumberland Mountains). According to Woods et al. (1996), “Ecoregion 69 ... is a high, dissected, and rugged plateau made up of sandstone, shale, conglomerate, and coal of Pennsylvanian and Mississippian age. The plateau is locally punctuated by a limestone valley (the Greenbrier Karst; subregion 69c) and a few anticlinal ridges” (p.30). Ecoregion 70 has more heterogeneous bedrock formations than subregions 69a and 69d. It is underlain by shale, siltstone, limestone, sandstone, and coal, including the interbedded limestone, shale, sandstone, and coal of the Monongahela Group and the Pennsylvanian sandstone, shale, and coal of the Conemaugh and Allegheny Groups (Woods et al., 1996).

Individual analyses of Ecoregions 69 and 70 result in a somewhat lower HC_{05} value for Ecoregion 69 and a somewhat higher value for 70 (254 $\mu\text{S}/\text{cm}$ in Ecoregion 69 and 345 $\mu\text{S}/\text{cm}$ in Ecoregion 70). This difference might be attributed to the background water chemistry (see the following section). However, if the genera were adapted to high conductivity in Ecoregion 70 and low conductivity in 69, or if they were represented by more resistant species in 70 and more sensitive species in 69, it would be expected that the XC_{95} values would consistently go up in Ecoregion 70 and down in Ecoregion 69 relative to the values in the combined data set. However, XC_{95} values go up and down in both ecoregions when they are analyzed individually.

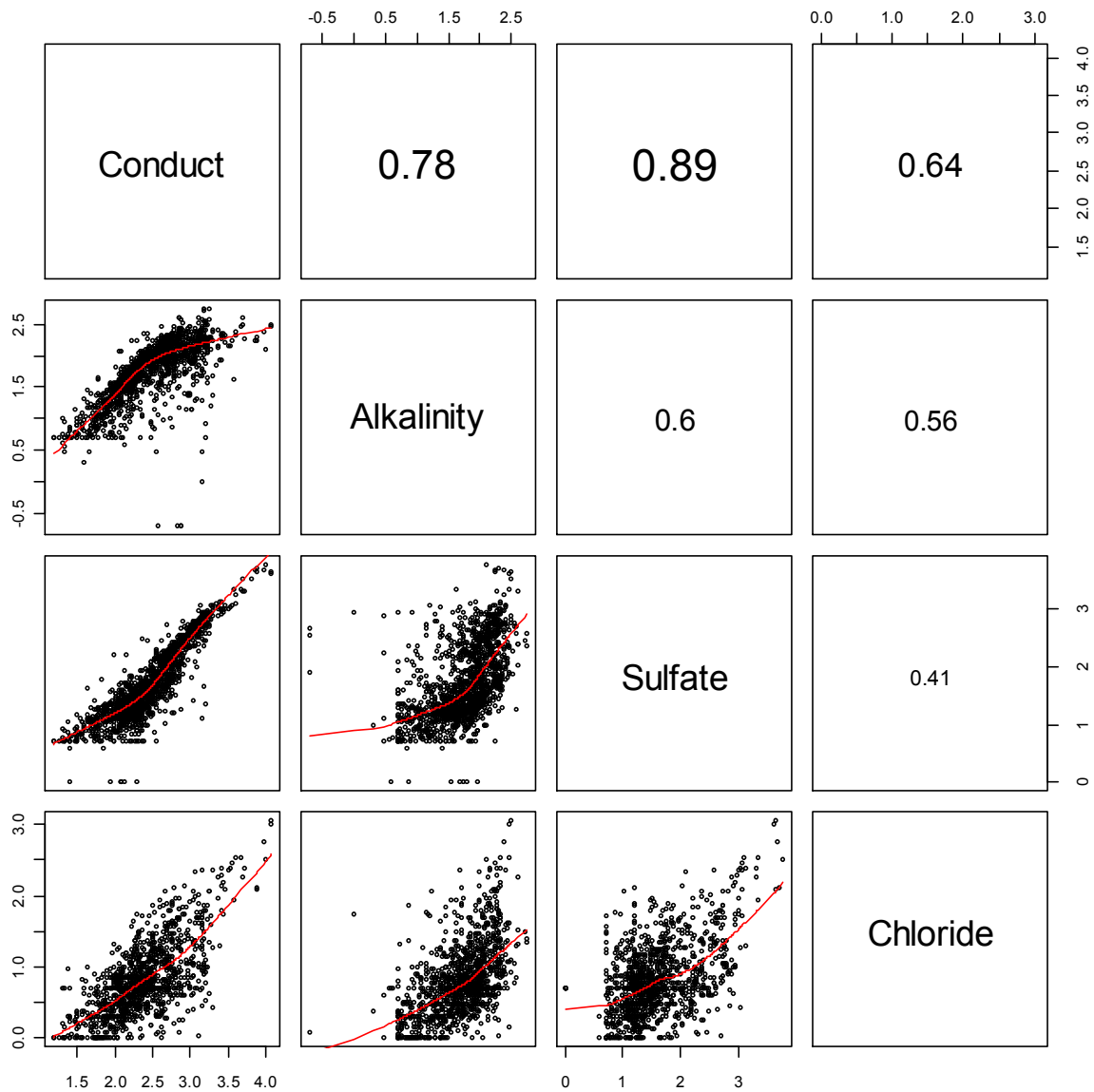


Figure 13a. Anions. Matrix of scatter plots and absolute Spearman correlation coefficients between conductivity ($\mu\text{S}/\text{cm}$), alkalinity (mg/L), sulfate (mg/L), and chloride (mg/L) concentrations in streams of Ecoregions 69 and 70 in West Virginia. All variables are logarithm transformed. The smooth lines are the locally weighted scatter plot smoothing (LOWESS) lines (span = $2/3$).

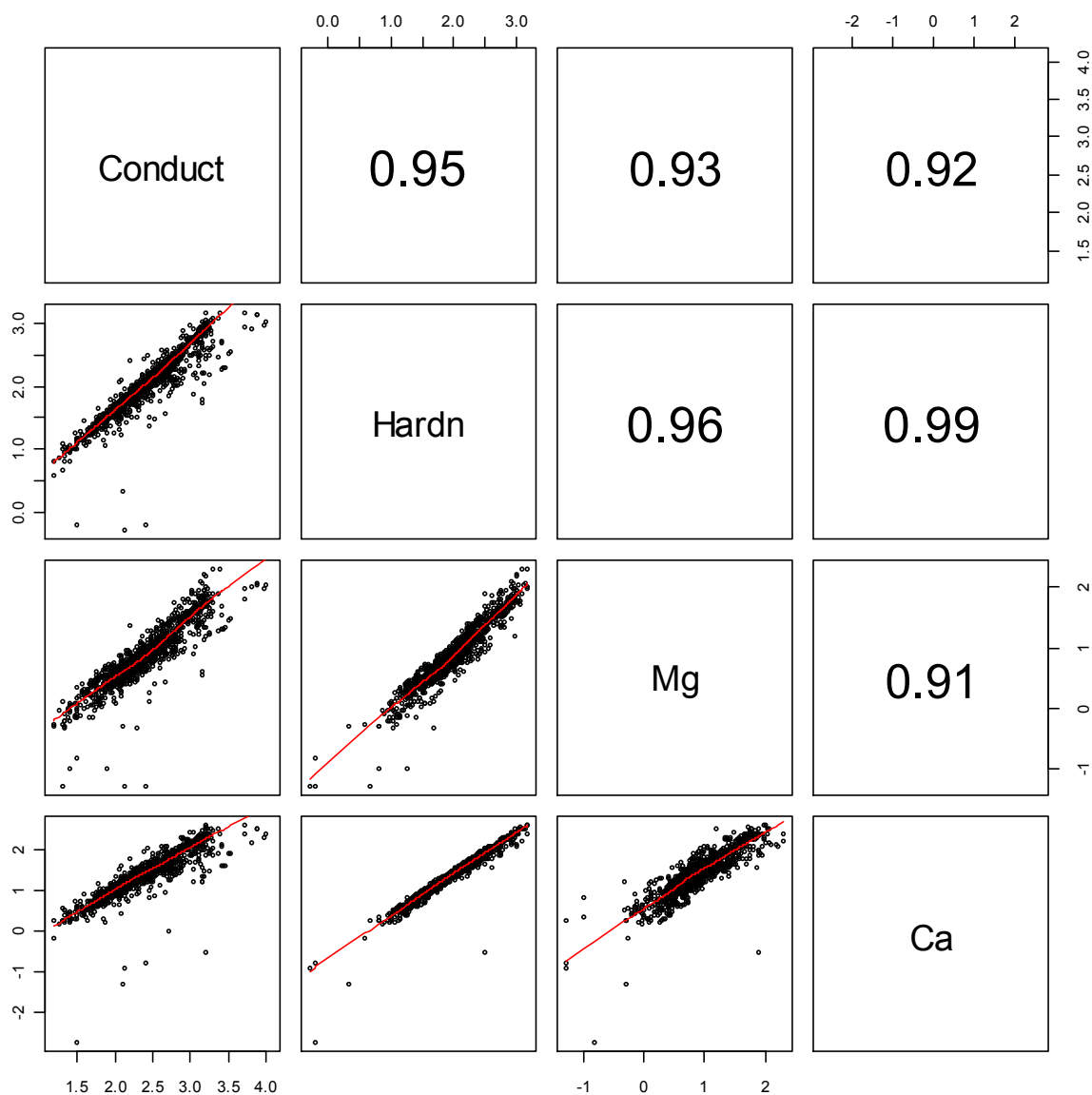


Figure 13b. Cations. Matrix of scatter plots and absolute Spearman correlation coefficients between conductivity ($\mu\text{S}/\text{cm}$), hardness (mg/L), Mg (mg/L), and Ca (mg/L), in the streams of Ecoregions 69 and 70 in West Virginia. All variables are logarithm transformed. The smooth lines are the locally weighted scatter plot smoothing (LOWESS) lines (span = 2/3).

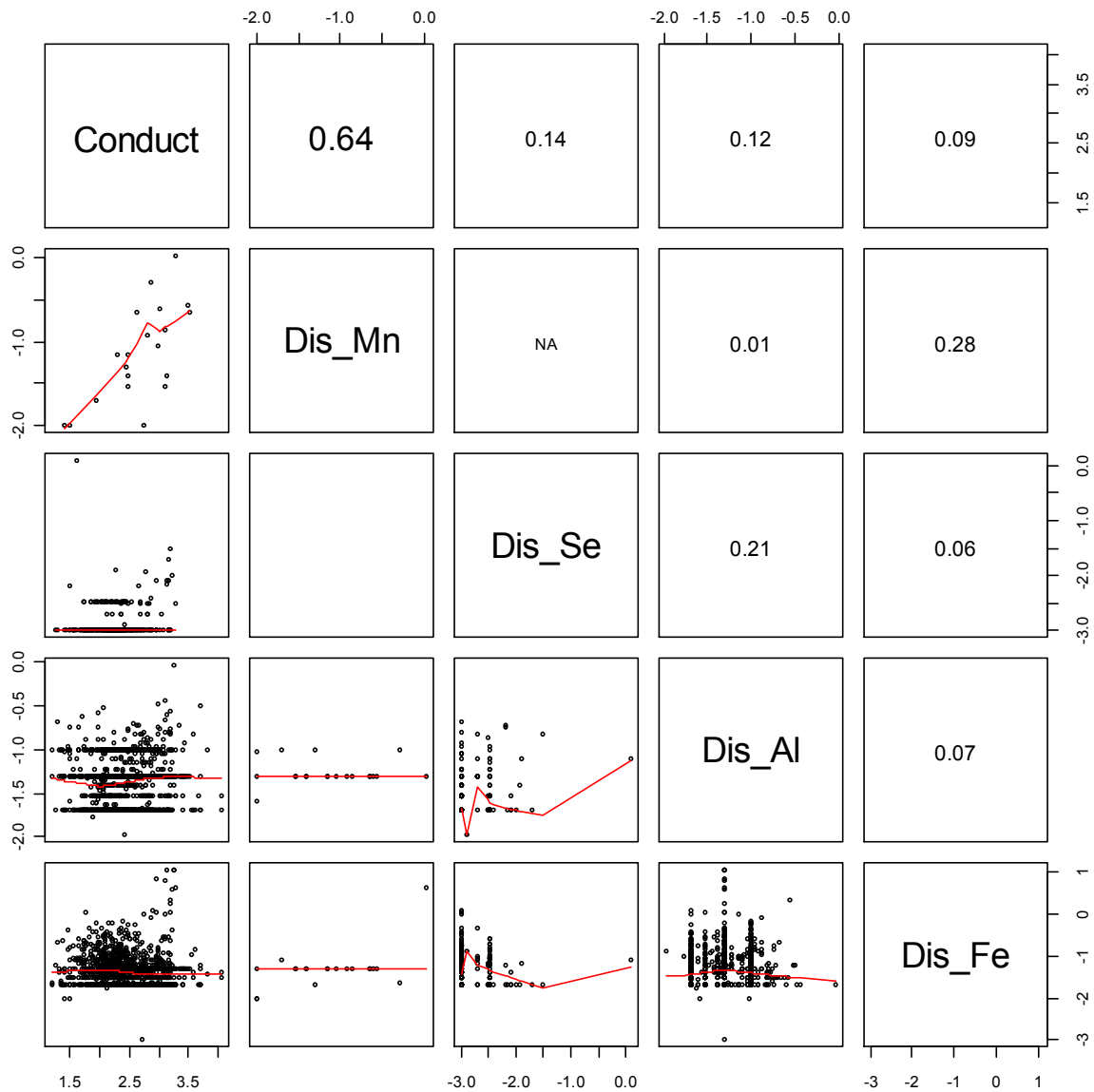


Figure 13c. Dissolved metals. Matrix of scatter plots and absolute Spearman correlation coefficients among conductivity ($\mu\text{S}/\text{cm}$) and dissolved metal concentrations (mg/L) in the streams of Ecoregions 69 and 70 in West Virginia. All variables are logarithm transformed. The smooth lines represent the locally weighted scatter plot smoothing (LOWESS) lines (span = 2/3).

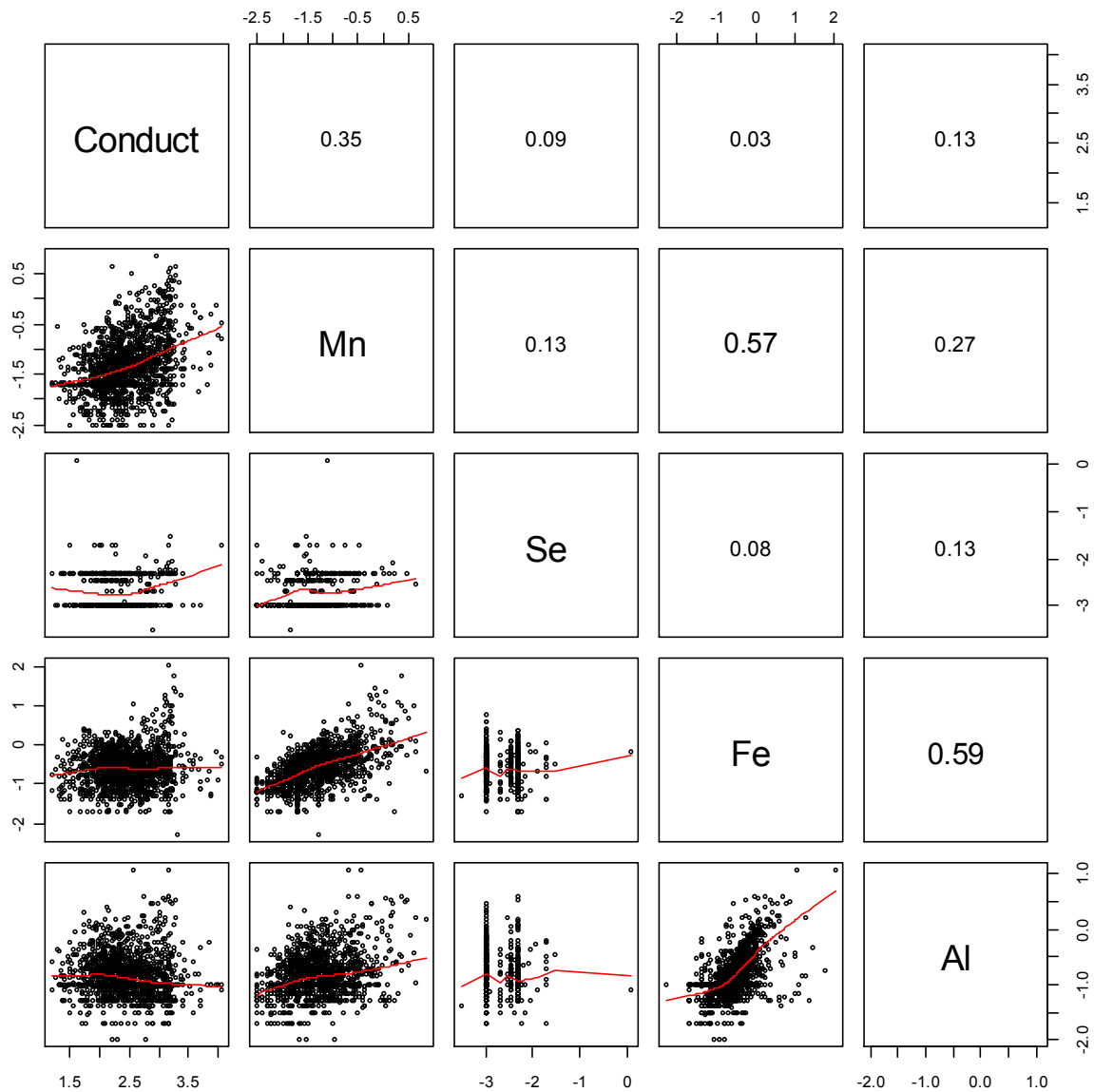


Figure 13d. Total metals. Matrix of scatter plots and absolute Spearman correlation coefficients between conductivity ($\mu\text{S}/\text{cm}$) and total metal concentrations (mg/L) in the streams of Ecoregions 69 and 70 in West Virginia. All variables are logarithm transformed. The smooth lines represent the locally weighted scatter plot smoothing (LOWESS) lines (span = 2/3).

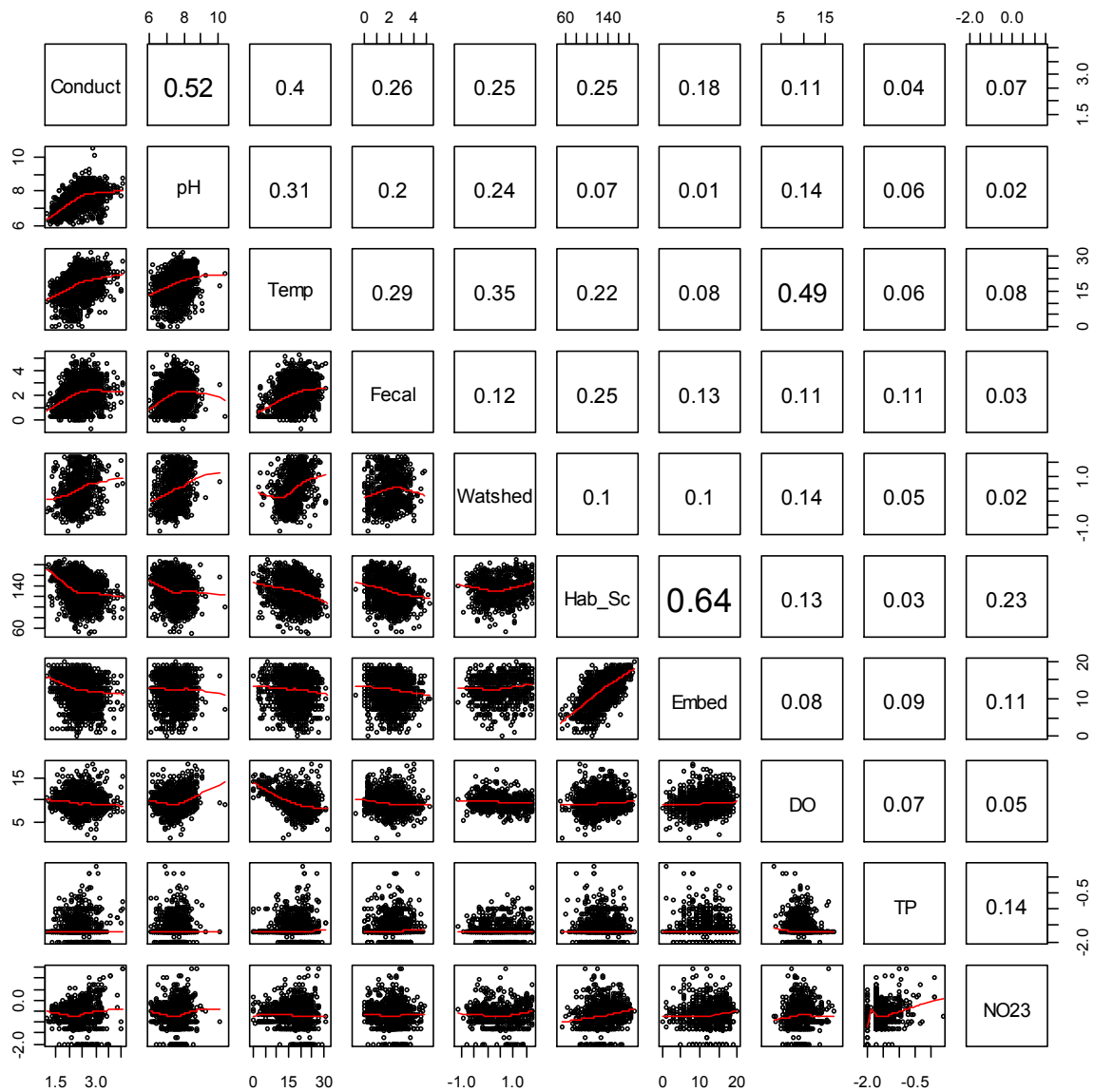


Figure 13e. Other water-quality parameters. Matrix of scatter plots and absolute Spearman correlation coefficients between environmental variables in the streams of Ecoregions 69 and 70 in West Virginia. The smooth lines are locally weighted scatter plot smoothing (LOWESS) lines (span = 2/3). Conductivity is logarithm transformed specific conductance ($\mu\text{S}/\text{cm}$); Temp is water temperature ($^{\circ}\text{C}$); RBP is Rapid Bioassessment (Habitat) Protocol score (possible range from 0 to 200); Fecal is logarithm transformed fecal coliform bacteria count (per 100 mL water); Watershed is logarithm transformed watershed area (km^2); embeddedness is a parameter score from the Rapid Bioassessment Protocol (possible range from 0 to 20); DO is dissolved oxygen (mg/L); TP is logarithm transformed total phosphorus (mg/L); NO23 is logarithm-transformed nitrate and nitrite (mg/L).

The differences in HC₀₅ values appear to be due primarily to random differences in which rarer genera do not meet the minimum sample size of 25 occurrences in a region. When the data set is split by ecoregion, the SSD model is reduced by 31 genera for Ecoregion 69 and 35 genera for Ecoregion 70. Furthermore, the two Ecoregions had similar genera, and, although Ecoregion 70 had a slightly higher estimated background, there were sites that had conductivity below 100 suggesting that the truly undisturbed background would be low. Overall we could not justify the increase in uncertainty associated with the reduced sample size and number of genera. Therefore, EPA did not derive benchmarks for individual ecoregions.

5.5. BACKGROUND

For naturally occurring stressors, it would not, in general, be appropriate to derive a benchmark value that is within the background range. Background levels may be estimated from reference sites, which are sites that are judged to be among the best within a category. However, because disturbance is pervasive, reference sites are not necessarily pristine or representative of natural background. Many reference sites have unrecognized disturbances in their watersheds or have recognized disturbances that are less than most others in their category. Some may have extreme values of a stressor because of measurement error or unusual conditions at the time the sample was taken. For those reasons, when estimating background concentrations, it is conventional to use only the best 75% of reference values. The cutoff centile is based on precedent and on the collective experience of EPA field ecologists (U.S. EPA, 2000). Estimated background conductivities for Ecoregions 69, 70, and both combined are 66, 214, and 150 $\mu\text{S}/\text{cm}$, respectively, using 75th centiles of reference sites in West Virginia.

Alternatively, background values may be estimated using samples from a probability-based design. Such samples include all waters within the sampling frame, including impaired sites, with defined probability. In some regions, there are no undisturbed streams. To characterize the best streams, the 25th centile is commonly used by EPA field ecologists (U.S. EPA, 2000). Based on the 25th centiles, estimated background conductivities for Ecoregions 69, 70, and both combined are 72, 153, and 116 $\mu\text{S}/\text{cm}$ for probability-based samples in West Virginia.

Background between Ecoregions 69 and 70 appear to be different; however, none of these values exceed the benchmark value of 300 $\mu\text{S}/\text{cm}$. The higher estimates of background conductivity in Ecoregion 70 relative to Ecoregion 69 may be attributed to the variable occurrence of limestone and limestone-derived soils. The higher level of development and population density in Ecoregion 70 may also contribute, but it was not evaluated.

5.6. SELECTION OF INVERTEBRATE GENERA

Selection of genera to model can affect the results. Using the data set of all taxa includes species that may occur due to a competitive advantage in polluted water. Some taxa, such as *Corbicula*, are not native to streams in North America. Using only genera found at sites with minimal disturbance as defined by reference sites somewhat alleviates this problem. The reference site genera are often linked to state narrative water-quality standards; thus, they represent the aquatic life use that state water-quality criteria should be designed to protect. Furthermore, the importance of losing species that inhabit minimally disturbed sites may be clearer to decision makers and stakeholders. In this particular case, using all genera, including invasive species, would increase the HC₀₅ by less than 2%.

Genera are also selected for statistical reasons. We restricted genera used in the analyses to those recorded at a minimum of 25 sampling sites to reduce the chance that an apparent extirpation is due to sampling variance and to increase the likelihood that the models and quantitative analyses for potential confounding are reasonably strong.

5.7. INCLUSION OF OTHER TAXA

Inclusion of other taxa are recommended under the EPA's 1985 criteria derivation methodology (Stephan et al., 1985) solely to ensure that other taxonomic groups are not more sensitive than those already evaluated. Fish were not included because their occurrence is strongly affected by stream size making it difficult to determine XC₉₅ values. Indeed, some of the affected streams naturally have no fish. In addition, the WABbase data set used to derive the benchmark does not contain data for fish. Other data sets that do contain fish are not as large and do not contain as great a range of conductivity values. A separate SSD might be developed for fish, once these technical issues are resolved. Data for plants and amphibians are not available. Additional findings regarding mussels could change this analysis if they are found to be more sensitive to conductivity than the invertebrates used here. Mussels were not represented because genera did not occur in a minimum of 25 samples probably owing to the WVDEP sampling methods. Additional analyses may be necessary to ensure protection of federally or state listed rare, threatened, or endangered species of fish, amphibians, and mussels.

5.8. TREATMENT OF LISTED SPECIES

Species listed by West Virginia Department of Natural Resources (WVDNR, 2007) as threatened were among the genera observed. Because taxa were identified to genus, we are not certain if the species are included. Therefore, we recommend that the invertebrate taxa in Table 6, that were included in the SSD, be identified to species in subsequent monitoring to evaluate the risk to these threatened taxa. Also, some genera of listed species were not included

Table 6. Genera of threatened species included in the SSD (WVDNR, 2007)

Genus	Common Family Name	Genus	Common Family Name
<i>Allocapnia</i>	stonefly	<i>Diploperla</i>	stonefly
<i>Alloperla</i>	stonefly	<i>Ephemera</i>	mayfly
<i>Caecidotia</i>	isopod	<i>Orconectes</i>	stonefly
<i>Calopteryx</i>	jewelwing	<i>Pteronarcys</i>	stonefly
<i>Cambarus</i>	crayfish	<i>Stenonema</i>	mayfly
<i>Cordulegaster</i>	spiketail	<i>Sweltsa</i>	stonefly
<i>Crangonyx</i>	amphipod	<i>Utaperla</i>	stonefly

in the SSD because the genus was not collected in sufficient numbers, such as from the genera *Gomphus*, *Hansonoperla*, *Macromia*, and *Ostrocerca*. Furthermore, freshwater mussels were not well represented in the samples perhaps due to the sampling methods. Stephan et al. (1985) recommend lowering the concentration below the 5th centile when necessary to protect threatened, endangered, or otherwise important species. Rare species may be ecologically important.

5.9. INCLUSION OF REFERENCE SITES

If high-quality (i.e., reference) sites are not included in the data set, effects on sensitive species will not be incorporated into the benchmark. That is, the lower end of the SSD will be missing. For example, in a region where all watersheds include tilled agricultural land uses, all sites are affected by sediment, so a legitimate SSD for sediment could not be derived by this method in that region. In this case, WVDEP's reference sites were included as well as many probability-based sites with >90% forest cover, which are believed to be representative of good-to high-quality systems.

5.10. SEASONALITY, LIFE HISTORY, AND SAMPLING METHODS

The seasonality of life history events such as emergence of aquatic insects can affect the probability of detecting a species because eggs and early instars are not captured by the sampling methods used. As a result, annual insects that emerge in the spring are present but unlikely to be detected in the summer, when conductivities increase in some streams.

The effects of seasonality and life history were evaluated by comparing HC₀₅ values partitioned for season. The data set was partitioned into spring and summer based on seasonal

patterns of conductivity in the full data set (see Figure 4) and the HC₀₅ was calculated. The spring season is March through June. The summer season is July through October. The HC₀₅ values in the truncated data sets are 317 $\mu\text{S}/\text{cm}$ for spring that included 132 genera and 415 $\mu\text{S}/\text{cm}$ for summer that included 120 genera. The greater summer HC₀₅ is due to the loss of sensitive taxa from the SSD. The lower end of the SSD for the full data set and spring samples are fairly similar (see Figure 14). Lower effects levels in the spring were not due to an insufficient test range of conductivities because exposures as high as 5,200 $\mu\text{S}/\text{cm}$ occurred in the spring samples. Because the spring data set included both sensitive genera and a full range of exposures, it was judged more reliable than the summer model.

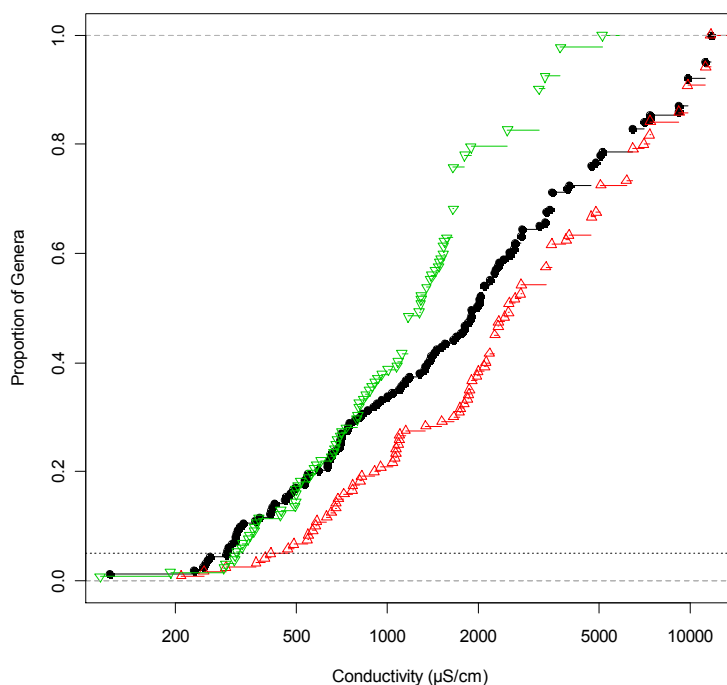


Figure 14. Comparison of full data set (circles) and subsets of spring (inverted triangles) and summer (triangles) collected samples. Spring consists of 132 genera, summer of 120 genera. The SSD for the full data set and summer are similar until XC_{95S} of 1,000 $\mu\text{S}/\text{cm}$. The summer SSD lacks sensitive genera.

Because we cannot be sure whether the greatest exposures in summer are tolerated by the spring-emergent genera, we estimated the likelihood that conductivity would increase in the summer. Sampling locations with at least one spring and summer conductivity measurement were identified. The spring season is March through June. The summer season is July through October. High and low conductivity streams are represented in both spring and summer samples. The conductivity in certain streams was three times greater in the summer than the spring.

However, streams with conductivity $<300 \mu\text{S}/\text{cm}$ in summer are below the benchmark in spring 98% of the time (see Figure 15). So, if a stream meets the benchmark in summer, it is likely to meet it year-round. Therefore, seasonal variation should be considered when planning monitoring of conductivity and should include ample samples in the spring to ensure inclusion of sensitive genera.

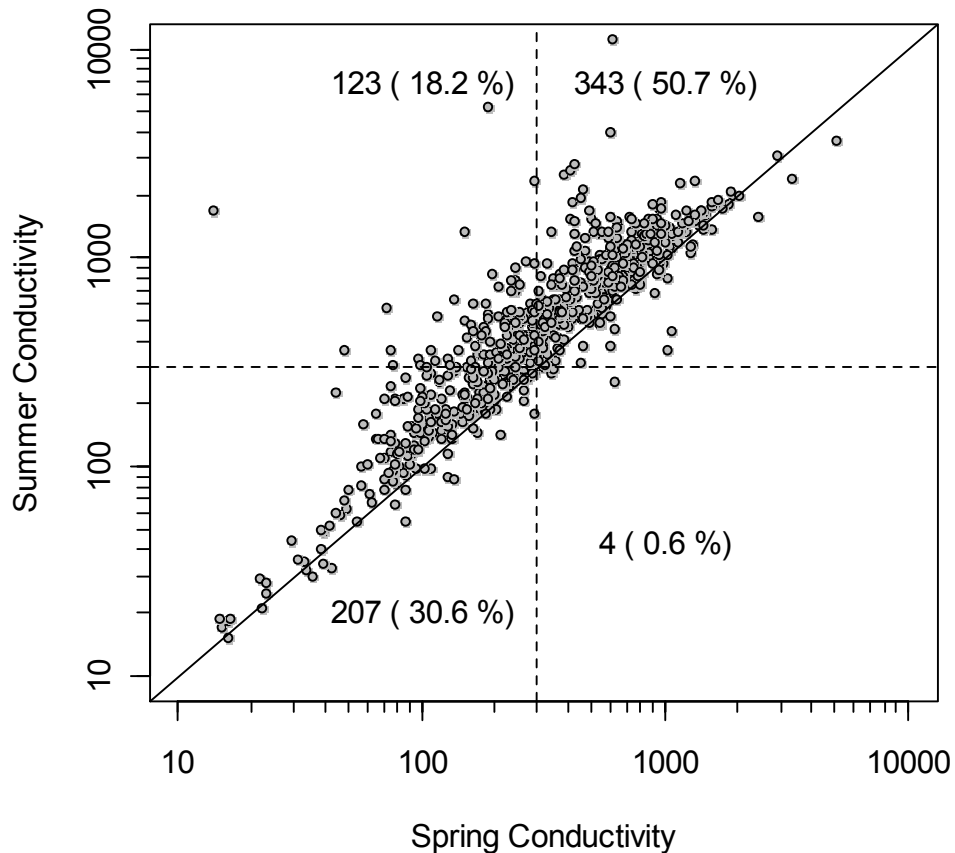


Figure 15. Relationship of conductivity values sampled from the same site in spring and summer. When conductivity is $<300 \mu\text{S}/\text{cm}$ (broken lines) in March thru June, the conductivity is $<300 \mu\text{S}/\text{cm}$ in the same stream 63% of the time July through October. When the conductivity is $<300 \mu\text{S}/\text{cm}$ in July through October, the conductivity in the same stream March through June is $<300 \mu\text{S}/\text{cm}$ 98% of the time.

5.11. FORMS OF EXPOSURE-RESPONSE RELATIONSHIPS

The diversity of the forms of the exposure-response relationships (i.e., decreasing, unimodal, decreasing, and no relationship) (see Figure 7 and Appendix E) has required some

methodological decisions. The forms are expected given the nature of the salts and the variance in sensitivity. The salt mixture includes nutrient elements, and, like other pollutants that are nutrients at low exposure levels (e.g., copper and selenium), the response to this mixture is expected to have a unimodal distribution (see Figure 7, *Diploperla*). In the ascending (left) limb, nutrient needs are increasingly being met. In the descending (right) limb, toxicity is increasing. However, many of the empirical exposure-response relationships do not display both limbs. They may show: (a) the descending portion of the curve because none of the observed conductivity levels are sufficiently low to show deficiency for the taxon (see Figure 7, *Lepidostoma*); (b) the ascending portion because none of the observed conductivity levels are sufficiently high to show toxicity for the taxon (see Figure 7, *Cheumatopsyche*); (c) the entire unimodal curve because their optimum is near the center of observed conductivities and the range from deficiency to toxicity is relatively narrow (see Figure 7, *Diploperla*); or (d) no trend because the optimum is more of a plateau than a peak so it extends across the range of observed conductivities (see Appendix E, *Nigronia*).

In order to estimate effects to sensitive taxa, it may be necessary to exclude genera favored by the pollutant if the region is highly modified. This was not done with the Appalachian data set. All genera regardless of the exposure-response form were included in the SSD. However, the XC values for those such as *Cheumatopsyche* that do not descend to zero in the observed range are treated as “greater than values.” Because the 5th centile of the SSD is derived by interpolation, it is not necessary to provide point estimates of the XC values for resistant taxa. The setting of the benchmark in a conductivity range in which the occurrence of some genera is increasing suggests that the benchmark could result in the extirpation of some genera. However, that is not the case. All but one of the 163 genera occur in sites with low conductivity (<100 $\mu\text{S}/\text{cm}$). Even if that were not the case, the concern for resistant taxa is unwarranted. The EPA sets water-quality criteria to protect the taxa that occur prior to pollution—not taxa that require pollution.

5.12. USE OF MODELED OR EMPIRICAL DISTRIBUTIONS

When deriving XC and HC values, one might use a centile of an empirical distribution or fit a function to the data and calculate the value from the resulting model. Models use all of the data and, therefore, are resistant to biases associated with any peculiar data at the centiles of interest or to uneven distributions of data. However, there is no a priori reason to believe that these distributions have a prescribed mathematical form, and fitted models may fit the data poorly at the centiles of interest. In particular, standard models are symmetrical but many SSDs are not, so the data are poorly fit at the extremes. The use of a nonparametric regression method to alleviate the problem of assuming a particular functional form can result in biologically

unlikely forms, may reduce the potential generality of the model, and is not readily understood. The use of empirical distribution functions without fitted models eliminates the problems of model selection and makes the method easier to understand and implement. With respect to SSDs, this issue is unresolved, and assessors are encouraged to consider the properties of their distributions when deciding whether to fit or not (Newman et al., 2002; Suter et al., 2002). In the interest of conceptual and operational simplicity, we identify the XC_{95} as the conductivity value at which the empirical cumulative probability is 0.95. The HC_{05} is determined by 2-point interpolation of points on the empirical distributions of XC_{95} values as described in Stephan et al. (1985).

5.13. DUPLICATE SAMPLES

Although most sites in the WABbase were sampled only once, 4% were sampled more than once and 5% of samples were from sites with duplicates. This situation may be confused with pseudoreplication, but that statistical error is not an issue in this analysis because we are estimating a value rather than testing a hypothesis. Duplicates provide more information especially when samples originate from different seasons when different genera may be present, but they could be problematical if they introduce a bias (e.g., if low conductivity sites were more likely to be sampled repeatedly). However, the duplicated sites do not appear to be biased in this case. In fact, if a simple inverse weighting scheme is applied (e.g., if a site is sampled twice; each observation is weighted 0.5) it does not materially change the result ($HC_{05} = 293$). Therefore, for the sake of simplicity and to avoid the possibility of inadvertently introducing bias by inappropriately weighting, we have not deleted or differentially weighted the duplicated samples. However, if there is a potential for bias due to duplication of some samples in future applications of this method, an appropriate weighting scheme could be applied. It was not necessary in this case.

5.14. TREATMENT OF CAUSATION

Causation should not be an issue in laboratory toxicity tests, but, even with rigorous treatment of confounders, scientists will question whether observed field relationships are truly causal (Kriebel, 2009). Like many epidemiologists, we believe that statistical analysis of relationships should be supplemented by the consideration of qualitative criteria for causation. In this case, we used evidence of causal characteristics derived from Hill's considerations (Cormier et al., 2010) to evaluate the causal relationship of conductivity and extirpation of organisms (see Appendix A).

5.15. TREATMENT OF POTENTIAL CONFOUNDERS

The use of field data to understand and manipulate causal relationships is limited by the possibility that the apparent relationship used to estimate the benchmark is confounded. Confounding is a bias in the analysis of causal relationships due to the influence of extraneous factors (confounders). Confounding occurs when a variable is correlated with both the cause and its effect. The correlations are usually due to a common source of multiple, potentially causal agents. However, they may be observed for other reasons (e.g., when one variable is a by-product of another) or due to chance associations. Confounding can bias a causal model resulting in uncertainty concerning the actual magnitude of the effects. Therefore, a variety of types of evidence are used to determine whether confounders significantly affect the results (see Appendix B). This is done because statistics alone cannot determine the causal nature of relationships (Pearl, 2009; Stewart-Oaten, 1996).

Potential confounders include the following: habitat, organic enrichment, nutrients, deposited sediments, pH, selenium, temperature, lack of headwaters, catchment area, settling ponds, dissolved oxygen, and metals. One potential confounder, low pH, was known to cause effects and was controlled by removing sites with pH <6 (see also Section 2.3). The influence of selenium is unclear due to poor data and should be investigated. The signal from conductivity was strong so that other potential confounders that were not strongly influential could be ignored with reasonable or greater confidence. These variables do affect species in the region, but their effects do not alter the signal from conductivity or the aquatic life benchmark.

6. AQUATIC LIFE BENCHMARK

The aquatic life benchmark of 300 $\mu\text{S}/\text{cm}$ was developed for year-round application. This level is intended to prevent the extirpation of 95% of invertebrate genera in this region. The estimated two-tailed 95% lower confidence bound of the HC_{05} point estimate is 228 $\mu\text{S}/\text{cm}$ and the upper bound is 303 $\mu\text{S}/\text{cm}$.

The aquatic life benchmark has been validated by an independent data set. Application of the same methodology to data from the State of Kentucky gave a very similar result, 282 $\mu\text{S}/\text{cm}$ with a lower confidence bound of 169 $\mu\text{S}/\text{cm}$ and an upper bound of 380 $\mu\text{S}/\text{cm}$ (see Appendix G).

The method used to develop the benchmark is an adaptation of the standard method for deriving water-quality criteria for aquatic life (i.e., Stephan et al., 1985), so it is supported by precedent. Because the organisms are exposed throughout their life cycle, this is a chronic value. Acute exposures were not evaluated.

The aquatic life benchmark for conductivity is provided as scientific advice for reducing the increasing loss of aquatic life in the Appalachian Region associated with a mixture of salts dominated by Ca^+ , Mg^+ , SO_4^{2-} , and HCO_3^- at circum-neutral pH. The aquatic life benchmark for conductivity is applicable to the parts of West Virginia, that provided the data for its derivation, and to Kentucky, which gave essentially the same result. It may be applicable to Ohio, Tennessee, Pennsylvania, Virginia, Alabama, and Maryland in Ecoregions 68, 69, and 70. This is because the salt matrix and background is expected to be similar throughout the ecoregions. (Region 68 [Southwestern Appalachia] does not occur in WV and is not included in the derivation of the benchmark value, but it is included in the validation data set from Kentucky [see Appendix G]). The aquatic life benchmark may also be appropriate for other nearby regions. However, this benchmark level may not apply when the relative concentrations of dissolved ions are different (see Table 2 for the ranges of concentrations in the data set used to derive the benchmark value).

REFERENCES

- Bailey, J. (2009) Watershed assessment section's 2009 standard operating procedures. West Virginia Department of Environmental Protection, Division of Water and Waste Management, Watershed Branch, Charleston, VA. Available online at <http://www.dep.wv.gov/WWE/getinvolved/sos/Documents/SOPs/WABSOP.pdf>
- Bear, JA; Cheng, HD; Sorek, S; et al. (1999) Seawater intrusion in coastal aquifers: concepts, methods and practices. Norwell, MA: Kluwer Academic Publishers.
- Bradley, TJ. (2009) Animal osmoregulation. New York:Oxford University Press.
- Bryant, G; McPhilliamy, S; Childers, H. (2002) A survey of the water quality of streams in the primary region of mountaintop / valley fill coal mining, October 1999 to January 2001. In: Draft programmatic environmental impact statement on mountaintop mining / valley fills in Appalachia - 2003. Appendix D. U.S. Environmental Protection Agency, Region 3, Philadelphia, PA. Available online at <http://www.epa.gov/Region3/mtntop/pdf/appendices/d/stream-chemistry/MTMVFCchemistryPart1.pdf>.
- Clark, ML; Miller, KA; Brooks, MH. (2001) US Geological Survey monitoring of Powder River Basin stream-water quantity and quality. U.S. Geological Survey, Cheyenne, WY. Available online at <http://pubs.usgs.gov/wri/wri014279/pdf/wri014279.pdf>.
- Cormier SM; Paul, JF; Spehar, RL; et al. (2008) Using field data and weight of evidence to develop water quality criteria. *Integr Environ Assess Manag* 4(4):490–504.
- Cormier, SM; Suter, GW, II; Norton, SB. (2010) Causal characteristics for ecoepidemiology. *Hum Ecol Risk Assess* 16(1):53–73.
- Efron, B; Tibshirani, R. (1993) An introduction to the bootstrap. Monographs on statistics and applied probability, 57. Boca Raton, FL: Chapman & Hall/CRC.
- Environment Canada and Health Canada. (2001) Priority substances list assessment report: road salts. Available online at <http://www.ec.gc.ca/Substances/ese/eng/psap/final/roadsalts.cfm>.
- Evans, DH. (2008a) Teleost Fish Osmoregulation: What have we learned since August Krogh, Homer Smith, and Ancel Keys? *Am J Physiol Regul Integr Comp Physiol* 295: R704-R713.
- Evans, DH. (2008b) Osmotic and ionic regulation: cells and animals. Boca Raton: CRC Press.
- Evans, M; Frick, C. (2002) The effects of road salts on aquatic ecosystems. National Water Research Institute (NWRI), Saskatoon, Saskatchewan NWRI Contribution No 02-308.
- Flotemersch, JE; Cormier, SM; Autrey, BC. (2001) Comparisons of boating and wading methods used to assess the status of flowing waters. U.S. Environmental Protection Agency, Office of Research and Development, Cincinnati, OH. EPA/600/R-00/108. Available online at http://www.epa.gov/nerleerd/MCD_nocover.pdf.
- Gerritsen, J; Zheng, L; Burton, J; et al. (2010) Inferring causes of biological impairment in the Clear Fork Watershed, West Virginia. U.S. Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment, Cincinnati, OH. EPA/600/R-08/146. Available online at http://oaspub.epa.gov/eims/eimscomm.getfile?p_download_id=496962
- GNU (2010) "R" version 2.12.1. GNU operating system. Available online at <http://www.r-project.org>
- Hille, B. (2001) Ion channels of excitable membranes, 3rd edition. Sunderland, MA: Sinauer Associates, Inc.
- Hose, GC; Van den Brink, P. J. (2004) Confirming the species-sensitivity distribution concept for endosulfan using laboratory, mesocosm, and field data. *Arch Environ Contam Toxicol* 47:511-520.

- Kelly, VR; Lovett, GM; Weathers, KC; et al. (2008) Long-term sodium chloride retention in a rural watershed: legacy effects of road salt on streamwater concentration. *Environ Sci Technol* 42(2):410–415.
- Komnick, H. (1977) Chloride cells and chloride epithelia of aquatic insects. *Int Rev Cytol* 49:285–328.
- Kriebel, D. (2009) How much evidence is enough? Conventions on causal inference. *Law Contemp Probl* 72:121–136.
- Maltby, L; Blake, N; Brock, TCM; et al. (2005) Insecticide species sensitivity distributions: Importance of test species selection and relevance to aquatic ecosystems. *Environ Toxicol Chem* 24(2):379–388.
- Mount, DR; Gulley, DD; Hockett, R; et al. (1997) Statistical models to predict the toxicity of major ions to *Ceriodaphnia dubia*, *Daphnia magna*, and *Pimephales promelas* (fathead minnows). *Environ Toxicol Chem* 16(10):2009–2019.
- Nelson, D; Cox, M. (2005) *Lehninger principles of biochemistry*. 4th edition. New York: W. H. Freeman & Co.; pp. 395–397.
- Newman, MC; Ownby, DR; Mezin, LCA; et al. (2000) Applying species sensitivity distributions in ecological risk 34 assessment: Assumptions of distribution type and sufficient number of species. *Environ Toxicol Chem* 19(2):508–515.
- Newman, MC; Ownby, DR; Mezin, LCA; et al. (2002) Species sensitivity distributions in ecological risk assessment: distributional assumptions, alternate bootstrap techniques, and estimation of adequate number of species. In: Posthuma, L; Suter, GW, II; Traas, TP; eds. *Species sensitivity distributions in ecotoxicology*. Boca Raton, FL: Lewis Publishers; pp 119–132.
- Omernik, JM. (1987) Ecoregions of the conterminous United States. *Ann Assoc Am Geograph* 77:118–125.
- Pearl, J. (2009) *Causality: models, reasoning, and inference*, 2nd edition. Cambridge, UK: Cambridge University Press.
- Pond, GJ. (2004) Effects of surface mining and residential land use on headwater stream biotic integrity in the eastern Kentucky coalfield region. Kentucky Department of Environmental Protection, Division of Water, Frankfort, KY.
- Pond, GJ. (2010) Patterns of Ephemeroptera taxa loss in Appalachian headwater streams (Kentucky, USA). *Hydrobiologia*. 641(1):185–201.
- Pond, GJ; Passmore, ME; Borsuk, FA; et al. (2008) Downstream effects of mountaintop coal mining: comparing biological conditions using family- and genus-level macroinvertebrate bioassessment tools. *J N Am Benthol Soc* 27(3):717–737.
- Posthuma, L; Suter, GW, II; Traas, TP; eds. (2002) *Species sensitivity distributions in ecotoxicology*. Boca Raton, FL: Lewis Publishers.
- Remane, A. (1971) Ecology of brackish water. In: Remane A; Schlieper, C, eds. *Biology of brackish water*. New York, NY: John Wiley and Sons.
- Rengasamy, P. (2002) Transient salinity and subsoil constraints to dryland farming in Australian sodic soil: an overview. *Aust J Exp Agri* 42(3):351–361.
- Smith, DG. (2001) *Pennak's freshwater Invertebrates of the United States: Porifera to Crustacea*, 4th edition. New York, NY: John Wiley & Sons Inc.
- Smithson, J. (2007) West Virginia Stream/River Survey Design 2007—2111. WV Department of Environmental Protection, Division of Water and Waste Management, Charleston, WV 25304.

- Stephan, CE; Mount, DI; Hanson, DJ; et al. (1985) Guidelines for deriving numeric National Water Quality Criteria for the protection of aquatic organisms and their uses. U.S. Environmental Protection Agency, Office of Research and Development, Washington, D.C. PB85-227049. Available online at <http://www.epa.gov/waterscience/criteria/library/85guidelines.pdf>.
- Stewart-Oaten, A. (1996) Problems in the analysis of environmental monitoring data. In: Detecting environmental impacts, Schmitt, RJ; Osenberg, CW; ed. New York, NY: Academic Press. p. 109–131.
- Sutcliffe, DW. (1962) The Composition of Haemolymph in Aquatic Insects. *Exp. Biol* 39:325–343.
- Suter, GW, II. (2007) Ecological risk assessment, 2nd edition. Boca Raton, FL: CRC Press.
- Suter, GW, II; Traas, T; Posthuma, L. (2002) Issues and practices in the derivation and use of species sensitivity distributions. In: Posthuma, L; Suter, II, GW; Traas, T; eds. Species sensitivity distributions in ecotoxicology. Boca Raton: Lewis Publishers; pp 437–474.
- Tarin, JJ; Cano, A. (2000) Fertilization in protozoa and metazoan animals: Cellular and molecular aspects. Berlin: Springer-Verlag.
- Thorp, JH; Covich, AP. (2001) Ecology and classification of North American freshwater invertebrates. New York, NY: Academic Press.
- U.S. EPA (Environmental Protection Agency). (1986) Quality criteria for water 1986. Office of Water, Washington, DC; EPA-440/5-86/001. Available online at http://water.epa.gov/scitech/swguidance/waterquality/standards/criteria/aqlife/upload/2009_01_13_criteria_goldbook.pdf
- U.S. EPA (Environmental Protection Agency). (2000) Nutrient criteria technical guidance manual: rivers and streams. Office of Water, Washington, DC; EPA/822/B-00/002. Available online at <http://www.epa.gov/waterscience/criteria/nutrient/guidance/rivers/rivers-streams-full.pdf>.
- U.S. EPA (Environmental Protection Agency). (2003) Generic ecological assessment endpoints (GEAEs) for ecological risk assessment. Risk Assessment Forum Washington, DC; EPA/630/P-02/004B. Available online at http://oaspub.epa.gov/eims/eimscomm.getfile?p_download_id=429201.
- U.S. EPA (Environmental Protection Agency). (2006) Framework for developing suspended and bedded sediments (SABS) water quality criteria. Office of Water, Washington, DC; EPA/822/R-06/001. Available online at http://www.epa.gov/med/Prods_Pubs/framework_sabs_20060601.pdf.
- U.S. EPA (Environmental Protection Agency). (2007) Level III ecoregions of the Continental United States (revision of Omernik 1987). National Health and Environmental Effects Research Laboratory, Corvallis, OR. Available online at <http://kgs.uky.edu/kgsweb/download/geology/useco.pdf>.
- U.S. EPA (Environmental Protection Agency). 2011. The Effects of Mountaintop Mines and Valley Fills on Aquatic Ecosystems of the Central Appalachian Coalfields. Office of Research and Development, National Center for Environmental Assessment, Washington, DC. EPA/600/R-09/138A.
- Werner, AD. (2009) A review of seawater intrusion and its management in Australia. *Hydrogeol J* 18:281–285. Available online at <http://www.springerlink.com/content/q75147m6j6766470/fulltext.pdf> (accessed 10/14/2009).
- WVDEP (West Virginia Department of Environmental Protection). (2006) Quality assurance project plan for watershed branch monitoring activities. Division of Water and Waste Management., WVDEP, Charleston, WV; pp. 279.

WVDEP (West Virginia Department of Environmental Protection). (2008a) West Virginia integrated water quality monitoring and assessment report. West Virginia Department of Environmental Protection, Charleston, WV. Available online at http://www.dep.wv.gov/WWE/watershed/IR/Documents/IR_2008_Documents/WV_IR_2008_Supplements_Complete_Version_EPA_Approved.pdf.

WVDEP (West Virginia Department of Environmental Protection). (2008b) 2008 Standard operating procedures. Vol. 1. Watershed Assessment Branch, WVDEP, Charleston, WV.

WVDNR (West Virginia Department of Natural Resources). (2007) Rare, threatened and endangered animals. West Virginia Natural Heritage Program. February 2007. Available online at <http://www.wvdnr.gov/Wildlife/documents/Animals2007.pdf> (accessed 10/12/2009).

Wood, CM; Shuttleworth, TJ. (2008) Cellular and molecular approaches to fish ionic regulation. Vol 14: Fish Physiology. San Diego, CA: Academic Press, Inc.

Woods, AJ; Omernik, JM; Brown, DD; et al. (1996) Level III and IV ecoregions of Pennsylvania and the Blue Ridge Mountains, the Ridge and Valley, and the Central Appalachians of Virginia, West Virginia, and Maryland. U.S. Environmental Protection Agency, Office of Research and Development, Corvallis, OR; EPA/600/R-96/077.

Ziegler, CR; Suter, GW, II; Kefford, BJ. (2007) Candidate cause: ionic strength. CADDIS (The Causal analysis/diagnosis decision information system). Available online at www.epa.gov/caddis (accessed 10/13/09).

APPENDIX B ANALYSIS OF POTENTIAL CONFOUNDERS

ABSTRACT

The purpose of Appendix B is to evaluate the ability of factors that may co-occur with conductivity (i.e., potential confounders) to weaken our ability to model the relationship between conductivity and occurrence of genera. The analyses in this appendix do not determine whether those factors cause effects in the region. Rather, they evaluate how the potential confounders may affect our ability to model the relationship between conductivity and the loss of macroinvertebrate genera.

The appendix addresses its purpose in two ways. First, it supports Appendix A by demonstrating that none of the potential confounders is responsible for the association between conductivity and biological effects. Second, it supports the development of the benchmark value by determining whether the confounders have substantive influence on the causal relationship between salts and macroinvertebrate assemblages. Twelve potential confounders were evaluated: habitat, organic enrichment, nutrients, deposited sediments, pH, selenium, temperature, lack of headwaters, catchment area, settling ponds, dissolved oxygen, and metals. The inference was performed by identifying potential confounders and then determining the occurrence and strength of 10 types of evidence of confounding for each of them. The term “confounding” refers to a bias in the analysis of causal relationships due to the influence of extraneous factors (confounders), in this case, the stressors listed above.

The effect of confounders was found to be minimal and manageable. Potential confounding by low pH was minimized by removing sites with pH <6 from the data set when calculating the aquatic life benchmark. The signal from conductivity was strong, so that potential confounders that were not strongly influential could be ignored with reasonable or greater confidence. No analysis can demonstrate that these variables have no influence at any place or time, but, this analysis does demonstrate that their influence on the relationship of conductivity and extirpation of genera is minimal given the streams that would be affected by the aquatic life benchmark.

B.1. INTRODUCTION

Having established that salt mixtures dominated by bicarbonate and sulfate cause biological impairments in the region (see Appendix A), this appendix addresses other potential causes of impairment in the region that might confound that relationship. The goal of this analysis is not to eliminate confounding variables. They are natural variables such as temperature and habitat structure that cannot be literally eliminated like eliminating smokers in

an epidemiological study. Nor is the goal to equate the levels of confounders to an ideal or pristine level. High conductivity effluents do not enter wilderness streams. Rather, the streams are subject to some level of current or historic disturbance. The overall goal of the Report is to estimate conductivity levels that would protect against the unacceptable effects of salts in those streams (i.e., typical streams receiving salty effluents in the region of concern). The goal of the assessment in Appendix B is to determine if the model developed for that purpose is a reliable predictor of harmful effects and protective levels. We do this by trying to discover if there are factors that bias that model.

Confounding is a bias in the analysis of causal relationships due to the influence of extraneous factors (confounders). Confounding occurs when a variable is correlated with both the cause and its effect. The correlations are usually due to a common source of multiple, potentially causal agents. However, they may be observed for other reasons (e.g., when one variable is a by-product of another) or due to chance associations.

Confounding may have two consequences. First, it can result in identification of a cause that is in fact a noncausal correlate. That possibility is commonly addressed by applying Hill's (1965) considerations or some equivalent set of criteria for causation as in Appendix A. This is done because statistics alone cannot determine the causal nature of relationships (Pearl, 2009; Stewart-Oaten, 1996). Second, confounding can bias a causal model resulting in uncertainty concerning the actual magnitude of the effects. That can be addressed by considering the magnitudes of correlations with and without the potential confounder or by considering the change in the results when the potential confounder is removed.

A variety of types of evidence may be used to determine whether confounders significantly affect the results; we have identified 10 types of evidence. They are related to three of the characteristics of causation used to determine that elevated conductivity is a cause of impairment of stream communities in Appendix A: co-occurrence, sufficiency, and alteration.

1. **Co-occurrence of confounder and cause:** Confounders are correlated with the cause of interest. A low correlation coefficient is evidence against the potential confounder.
2. **Co-occurrence of confounder and effect:** Confounders are correlated with the effect of interest. A low correlation coefficient is evidence against the potential confounder.
3. **Co-occurrence of confounder and cause:** Even when the confounder is not correlated with the cause of interest, it may be influential at extreme levels. A lack of influence at extreme levels of the cause and the potential confounder is evidence against the potential confounder.

4. **Co-occurrence of confounder and effect:** If the frequency of the effect does not diminish when the potential confounder is never present or is present in all cases, the confounder can be discounted in that subset.
5. **Sufficient confounder:** The magnitude of the potential confounder (e.g., concentration of a co-contaminant) may be compared to exposure-response relationships from elsewhere (e.g., laboratory toxicity tests) to determine if the exposure to the potential confounder is sufficient. If it is not sufficient, that is evidence that it is not acting as a confounder.
6. **Sufficient confounder:** If the confounder is estimated to be sufficient in a subset of cases, those cases may be removed from the data set, and the remaining set reanalyzed to determine the influence of their removal on the results.
7. **Sufficient confounder:** Multivariate statistical techniques may be used to estimate the magnitude of confounding or to adjust the causal model for confounding—if their assumptions hold.
8. **Sufficient confounder:** If the potential confounder occurs in a sufficiently small proportion of cases, it can be ignored.
9. **Alteration:** If a potential confounder has characteristic effects that are distinct from those of the cause of concern, then the absence of those effects can eliminate the potential confounder as a concern in either individual cases or the entire data set.
10. **Alteration:** If the effects are characteristic of the cause of concern and not of the potential confounder, then the potential confounder can be eliminated as a concern in either individual cases or the entire data set.

Weighing evidence for confounding differs from weighing evidence for causation. The causal assessment in Appendix A determines whether dissolved salts are an important cause of biological impairment in the region. This assessment of confounding accepts the result of the causal assessment and attempts to determine whether any of the known potential confounders interfere with estimating the effects of conductivity to a significant degree. If there is significant interference, the confidence in the model predictions would be weakened unless the model is modified. That requires a different weighting and weighing method from the one in Appendix A, which would be used if the goal were to determine whether the potential confounder is itself a cause.

As in Appendix A, the number of ephemeropteran (mayfly) genera is used as a standard metric for the effects of conductivity, which may or may not be confounded. Because the endpoint effect is extirpation of 5% of genera and the sensitive genera are primarily

Ephemeroptera, this is an appropriate metric. However, because of a resistant mayfly genus (see Figure A-1), it is not expected that all Ephemeroptera will be missing at high conductivities.

Some commenters recommended using multivariate statistics in place of weight-of-evidence analysis as the sole means to address potential confounders. However, because of the goals of the analysis and the nature of the data, it is not appropriate to use multivariate statistics alone to try to model the relationship between conductivity and extirpation or to eliminate the effects of confounders or estimate the magnitude of their effects. First, no statistical test can demonstrate that an association is causal. Second, violation of assumptions prevents reliable estimation of the influence of one potentially causal variable on another. Multiple regression depends on assumptions of independence, additivity, and normality that are not met. In sum, multivariate statistical associations are just associations, and association is not causation. However, they can be used as evidence in the weight-of-evidence analysis along with other incomplete or imperfect pieces of evidence to help reach the best-supported conclusion.

B.2. WEIGHTING

The evidence is weighted using a system of plus (+) for supporting the potential confounder (i.e., the evidence suggests that the potential confounder is actually causing the effect to a significant degree), minus (–) for weakening the potential confounder (i.e., the evidence suggests that the potential confounder does not contribute to the effect to a significant degree), and zero (0) for no effect. One to three plus or minus symbols are used to indicate the weight of a piece of evidence.

+++ or ---	Convincingly supports or weakens
++ or --	Strongly supports or weakens
+ or -	Somewhat supports or weakens
0	No effect

Any relevant evidence receives a single plus, minus, or zero to register the evidence as relevant and to indicate a decreased or increased potential for confounding (see Table B-1). The strength of evidence is considered next. Criteria for scoring the strength of evidence are presented below for the common types. They were developed for transparency and consistency and are based on the best professional judgments. After strength, the other possible unit of weight is assigned depending on the type of evidence.

Table B-1. Relationships between qualities of evidence and scores for weighing evidence

Qualities of the evidence	Score, not to exceed three minus or three plus
Logical implications and relevance	+, 0, -
Strength	Increase score
Other qualities	Increase score

For co-occurrence (Evidence Types 1–4), strength or consistency of the association is the primary consideration. The primary measure of association is Spearman’s correlation coefficients. For comparison to the potential confounders, the correlation coefficient for conductivity and number of ephemeropteran genera are –0.61 for the West Virginia (WV) data set and –0.72 for the EPA Region 3 data set, values in the upper end of the moderate range. Correlations, as measures of co-occurrence, can be scored as in Table B-2.

The scores in this appendix are based on conventional expectations for a confounder that is itself a cause. That is, a potential confounder—such as a metal by itself—might cause extirpation of invertebrate genera (independent combined action) or might act in combination with conductivity to extirpate invertebrate genera (additive or more than additive combined action). However, sometimes correlations are anomalous. For example, a potential confounder may actually decrease effects as when calcium reduces effects of metals. Such anomalous results require case-specific interpretation based on knowledge of mechanisms and characteristics of the ecosystems being analyzed.

Table B-2. Weighting co-occurrence using correlations for Evidence Types 1–2

Assessment	Strength	Score
Absent	$r \leq 0.1 $	--
Weak	$ 0.1 < r < 0.25 $	-
Moderate	$ 0.75 \geq r \geq 0.25 $	+
High	$r > 0.75 $	++

Anomalous results may also result from violation of the expectation that a confounder should be correlated with both conductivity and the effect. If only one of the correlations is

observed, that result requires additional interpretation. If the potential confounder is correlated with the effect, but not with conductivity, the result may be due to chance or to a partitioning of causation in space. That is, they are independent because the confounder impairs communities at different locations than conductivity. This could occur if the potential confounder and conductivity have different sources. In any case, it is not a confounder of conductivity.

In the contingency tables (Evidence Type 3), the frequency of occurrence of any Ephemeroptera (i.e., of the failure to extirpate all ephemeropteran genera) is presented for combinations of high and low levels of conductivity and of the potential confounder. If the frequency of occurrence is much lower when the potential confounder is present at high levels, this is supporting evidence for confounding. Note that the goal here is not to determine the effects of exceeding a criterion or other benchmark. Rather the goal is to clarify the co-occurrence of conductivity, confounders, and effects by determining the frequency of effects at each possible combination of extremely high and low levels of conductivity and the potential confounder. It is expected that, if a variable is indeed a confounder, its influence on the occurrence of effects would be seen at an extreme level. This use of contingency tables could reveal influences of confounders that are obscured when the entire ranges of data are correlated by, for example, a step function or other discontinuity in the relationship. Therefore, clearly high and low levels of conductivity and the potential confounder are used in contingency tables.

When scoring evidence from contingency tables, a potential confounder gets a plus score if its presence at a high level reduces the probability of occurrence by more than 25% and a minus score if it does not (see Table B-3). It gets a double plus score if its presence at a high level reduces the probability of occurrence by more than 75% and a double minus score if it raises it by less than 10%. Any decrease in effects at high levels of a potential confounder is anomalous and is treated as strong negative evidence.

Table B-3. Weighting co-occurrence for Evidence Type 3 using contingency tables

Assessment	Strength	Score
High levels of a confounder should increase the probability that a site lacks Ephemeroptera at low conductivity, and low levels of the confounder should decrease the effect at high conductivities	Increased effect >25%	+ for co-occurrence
	Increased effect >75%	+ + for co-occurrence and strength
	Increased effect <25%	- for co-occurrence
	Increased effect <10% or decreased effect	- - for co-occurrence and strength

The evidence concerning sufficiency of the confounder (Evidence Types 5–8) is diverse. Only Evidence Type 6 was sufficiently common and consistent to develop scoring criteria. For Evidence Type 6, the primary consideration is the degree of departure of the correlation in the truncated data set from the correlation of conductivity and Ephemeroptera in the full data set (see Table B-4). However, no more than one negative score was given if less than 10% of the data were removed.

Table B-4. Weighting sufficiency for Evidence Type 6: alteration of the correlation of conductivity with the number of ephemeropteran genera after removal of elevated levels of a confounder

Assessment	Strength	Score
Removal of elevated levels of a confounder should change the correlation coefficient	Coefficients decrease by <10% ($0.55 < r$ for WV data)	-- for a lack of change in effect with removal of confounder
	Coefficients decrease by <20% ($0.49 \leq r$ for WV data)	- for a small change in effect with removal of confounder
	Coefficients decrease by >20% ($0.49 > r$ for WV data)	+ for a strong change in effect with removal of confounder
	Coefficients increase	-- because removal of a true confounder should decrease the effect of conductivity

For alteration, the primary consideration is the degree of specificity of the effects of the confounder relative to those of the salts. This type of evidence is rare and is scored ad hoc when it occurs. Additional considerations that may result in a higher score are presented in Table B-5.

The primary data source for evidence of confounding is the Watershed Analysis Data Base (WABbase), which was used to derive the benchmark. Except where indicated, reported results are derived from those data, which are referred to as the West Virginia data. However, where possible and appropriate, the EPA Region 3 data set from West Virginia samples (referred to as the EPA data set) is used for independent corroboration. The EPA data set is much smaller and often does not have enough extreme values of the potential confounder to calculate reliable contingency tables or regressions of censored data.

Table B-5. Considerations used to weight the evidence concerning the influence of potentially confounding variables

Quality of evidence	Descriptor
Logical implication	Negative or positive
Directness of cause	Proximate cause, sources, or intermediate causal connections
Specificity	Effect attributable to only one cause or to multiple causes
Relevance to effect	From the case or from other similar situations
Nature of the association	Quantitative or qualitative
Strength of association	Strong relationships and large range or weak relationships and small range
Consistency of information	All consistent or some inconsistencies
Quantity of information	Many data or few data
Quality of information	Good study or poor study

Source: Cormier et al. (2010).

B.3. WEIGHING

After the individual pieces of evidence have been weighted, the body of evidence for a potential confounder is weighed based primarily on the consistency of the evidence and secondarily on the strength of the pieces of evidence (see Table B-6). The body of evidence—rather than any one piece of evidence—determines how strongly these potential confounders might affect the model.

B.4. POTENTIAL CONFOUNDERS

Potential confounders were chosen because they were believed to be associated with mountaintop mining, valley fills, or other sources of salts or because of suggestions from reviewer or public comments. Each of the discussions in this section begins with a statement of the reason that the potential confounder was chosen for evaluation.

Table B-6. Weighing confidence in the body of evidence for a potential confounder

Assessment	Score	Body of evidence	Action
Very confident	---	All minus, some strongly negative evidence	No treatment for confounding
Moderately confident	--	All minus, no strongly negative evidence	No treatment for confounding
Reasonably confident	-	Majority minus	No treatment for confounding
Undetermined	0	Approximately equal positive and negative, ambiguous evidence, or low quality evidence	Additional study advised
Potential confounding	+	Majority plus	Correction for confounding may be advised

B.4.1. Habitat Quality

Stream habitat may be modified by physical disturbance, changes in flow or increased sediment loads in reaches that receive high conductivity effluents. Habitat quality was represented by a qualitative index, the Rapid Bioassessment Protocol Habitat Evaluation (RBP) derived by the WVDEP, which increases as habitat quality increases. Component metrics were not used because they were less correlated with Ephemeroptera than the index.

Habitat quality was analyzed as part of groups of variables that were judged a priori to be more likely than others to have combined effects. Therefore, sites at which RBP and pH were low and fecal coliform count was high were removed to determine whether the 5th centile hazardous concentration (HC₀₅) was affected (see Figure B-1). Similarly, RBP was used with fecal coliform count and temperature in a multiple linear regression with conductivity (see Table B-7).

The body of evidence was mixed. Habitat scores were moderately correlated with both conductivity and biological response, which indicates a potential for confounding. However, removal of poor habitat had little effect on the correlation of conductivity with Ephemeroptera or on the derivation of the HC₀₅ for conductivity (see Table B-7 and Figure B-1). Habitat score had a very slight effect on the intercept and the slope for conductivity in a multiple regression (see Table B-7). In addition, Ephemeroptera occur even when habitat is poor (see Table B-8). The weight of the scored body of evidence indicated habitat was not a confounder (see Table B-9).

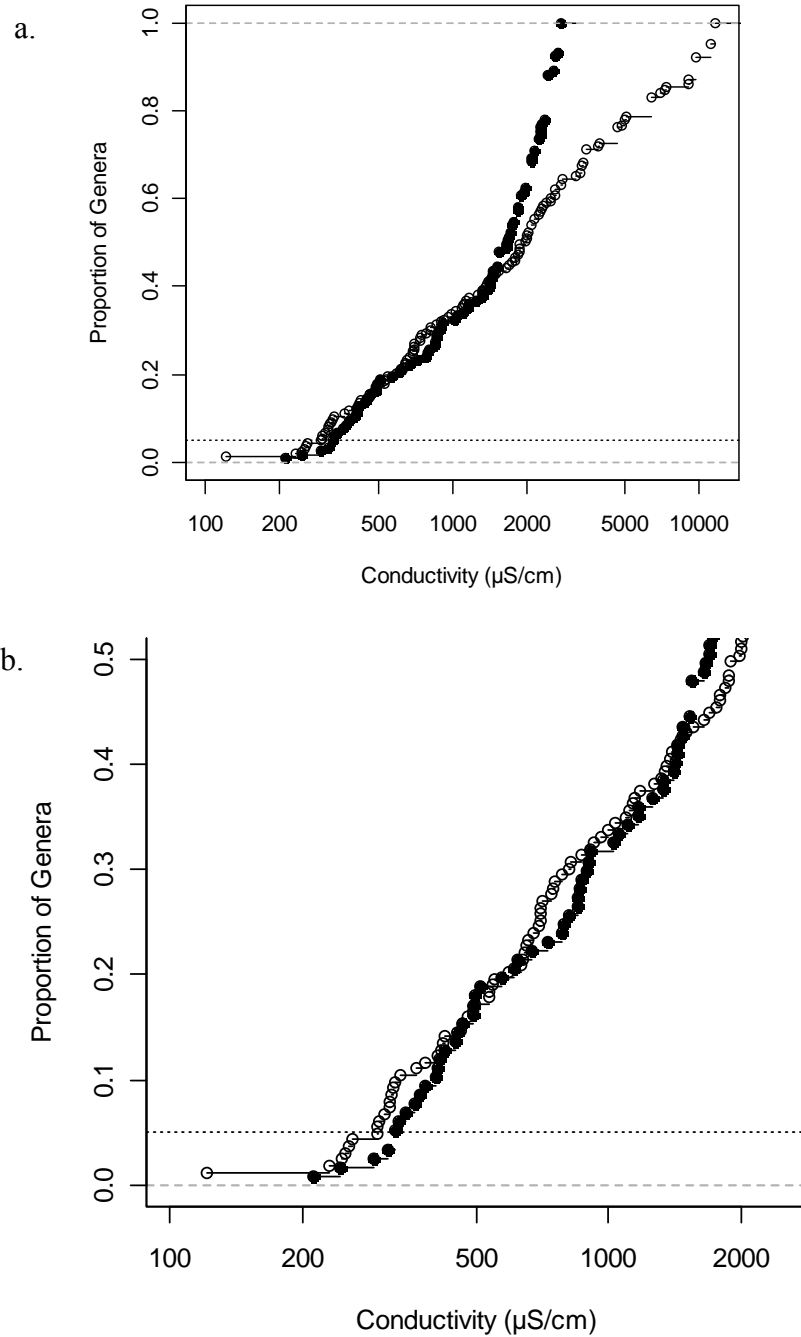


Figure B-1. Species sensitivity distribution for all year, pH >6 and all sites (open circles) and for sites with pH ≥6, Rapid Bioassessment Protocol score ≥135 and fecal coliform ≤400 colonies/100 mL (closed circles). (a) Uncensored with 163 genera and censored dataset with 117 genera. (b) Only the lower half of the SSD is shown to better discriminate the points in the left side of the full distribution. Habitat disturbance and organic enrichment have little influence; the HC₀₅ for the constrained data set is 326 μS/cm based on 117 genera. The upper and lower confidence bounds on that value are 229 μS/cm and 343 μS/cm, respectively.

Table B-7. An output table for two linear regression models. The first is the simple model predicting ephemeropteran genera from conductivity. The second is a multivariate model with the additional covariates RBP score, temperature, and fecal coliform count. These variables were chosen a priori as likely confounders that could co-occur and have combined effects.

Parameter	Estimate	Standard error
Univariate model		
Intercept	3.65	0.055
Conductivity slope	-0.93	0.024
Multivariate model		
Intercept	3.39	0.11
Conductivity slope	-0.92	0.029
RBP slope	0.0014	0.0005
Temperature slope	0.0068	0.0026
Fecal coliform slope	0.037	0.012

Table B-8. Number and percent of sites with high and low quality habitat and high and low conductivity with Ephemeroptera in streams (pH >6)

	Conductivity <200 μ S/cm	Conductivity >1,500 μ S/cm
Habitat score <115	140/142 (99%)	12/31 (39%)
Habitat score \geq 140	373/375 (99%)	13/22 (59%)

Table B-9. Evidence and weight for confounding by habitat quality

Type	Score	Evidence
1. Correlation of cause and confounder	+	RBP score was barely moderately correlated with conductivity, ($r = -0.25, n = 2,192$).
2. Correlation of effect and confounder	+	RBP score was barely moderately correlated with the number of ephemeropteran genera ($r = 0.26, n = 2,192$).
3. Contingency of high level of cause and confounder	-	In a contingency table (see Table B-8), Ephemeroptera are present at 99% of sites with low conductivity (<200 μ S/cm) even when habitat is poor (<115). However, with high conductivity, Ephemeroptera are present at only about half of sites regardless of habitat.
6. Removal of confounder	-	When sites with moderate to poor habitat (an RBP score <140) were removed from the analysis, conductivity is a little less negatively correlated with the number of Ephemeroptera ($r = -0.55, n = 747$).
	--	The SSD and HC ₀₅ are very similar when the XC ₉₅ values were calculated with a subset of the data set with sites removed with pH of <6, RBP score <135, and fecal coliform >400 colonies/100 mL (see Figure B-1).
7. Multivariate statistics	--	Habitat quality, temperature and fecal coliform together had essentially no effect on the slope in multiple regression and the slope for RBP score is particularly small (see Table B-7).
Weight of evidence	-	Reasonably confident. The correlations are marginal; RBP explains only 6.7% of the variance in ephemeropteran occurrence, based on r^2 . However, the contingency table gives relatively strong negative evidence (Ephemeroptera occur even when habitat is poor), and elimination of poor habitat (along with high coliform counts) has almost no effect on the SSD or HC ₀₅ (see Figure B-1). Habitat has very little effect in the multiple regression. Therefore, we did not correct for habitat, but more detailed habitat studies could be worthwhile.

SSD = species sensitivity distribution.

B.4.2. Organic Enrichment

Sources of organic enrichment such as domestic sewage and animal wastes are also sources of salts that contribute to conductivity. Fecal coliform counts are an indicator of organic enrichment and the presence of sources that may contain other toxicants such as household waste. The evidence is mixed, but, overall, the evidence against significant confounding associated with fecal coliform counts was much stronger than the supporting evidence (see Tables B-7, B-10, and B-11).

	Conductivity <200 μ S/cm	Conductivity >1,500 μ S/cm
Coliform <400 colonies/100 mL	610/613 (99%)	30/69 (43%)
Coliform >400 colonies/100 mL	184/187 (98%)	14/34 (41%)

Table B-11. Evidence and weights for confounding by organic enrichment

Type	Score	Evidence
1. Correlation of cause and confounder	+	Fecal coliform counts were barely moderately correlated with conductivity ($r = 0.26$, $n = 2,040$).
2. Correlation of effect and confounder	+	Coliform counts were barely moderately correlated with the number of ephemeropteran genera ($r = -0.25$, $n = 2,040$).
3. Contingency of high level of cause and confounder	--	In a contingency table (see Table B-10), the presence of high coliform counts did not change the probability of finding Ephemeroptera at either high or low conductivity.
6. Removal of confounder	--	When samples >400 colonies/100 mL were removed from the analysis, the correlation of conductivity with Ephemeroptera barely changed ($r = -0.61$, $n = 1,364$).
	--	The species sensitivity distribution (SSD) and HC ₀₅ are very similar to those used in the benchmark, when calculated from subset of the data with sites removed with pH of <6, RBP score <135, and fecal coliform >400 colonies/100 mL (see Figure B-1).
7. Multivariate statistics	--	Habitat quality, temperature and fecal coliform together had essentially no effect on the slope for conductivity in multiple regression (see Table B-7).
Weight of evidence	-	Reasonably confident: the correlations producing the two positive scores were exactly on the margin, and negative evidence was strong. No treatment for confounding.

B.4.3. Nutrients

Nitrogen and phosphorus may come from sewage and animal wastes or from fertilizers used in agriculture or mine reclamation. Because neither nutrient was correlated with conductivity or Ephemeroptera, effects could not be confounded by nutrients when conductivity increased (see Table B-12).

Table B-12. Evidence and weights for confounding by nutrients

Type	Score	Evidence
1. Correlation of cause and confounder	--	Conductivity was uncorrelated with nitrate and nitrite in the WV data set ($r = 0.07$, $n = 1,182$) and moderately correlated in the EPA data set ($r = 0.33$, $n = 39$).
	--	Conductivity was uncorrelated with total phosphorus in the WV data set ($r = 0.04$, $n = 1,185$) and the EPA data set ($r = 0.03$, $n = 45$).
2. Correlation of effect and confounder	-	Ephemeroptera was uncorrelated with nitrate and nitrite in the WV data set ($r = -0.04$, $n = 1,182$) and barely moderately correlated in the EPA data set ($r = -0.26$, $n = 39$).
	--	Ephemeroptera was uncorrelated with total phosphorous ($r = 0.001$, $n = 1,185$) and the EPA data set ($r = 0.06$, $n = 45$).
3. Contingency of high level of cause and confounder	NA	Contingency table analyses were not used because extreme nutrient levels were rare at high conductivities.
6. Removal of confounder	-	When samples with nitrate plus nitrite >0.6 mg/L were removed from the analysis, the correlation of conductivity with the number of Ephemeroptera was little changed ($r = -0.54$, $n = 999$).
	-	When samples with total phosphorus >0.04 mg/L were removed from the analysis, the correlation of conductivity with the number of Ephemeroptera was little changed ($r = -0.56$, $n = 998$).
Weight of evidence	---	Very confident: all negative, some strongly negative. No treatment for confounding.

NA = not applicable.

B.4.4. Deposited Sediment

Mining and other activities that result in crushing and exposing rocks are sources of salts and potentially of silt that may affect stream organisms. A qualitative measure of embeddedness (WABase embeddedness score) was evaluated by contingency table and by correlation (see Table B-13 and B-14). No evidence supported embeddedness as a confounder (see Table B-14).

Table B-13. Number of sites with high and low embeddedness scores and high and low conductivity with Ephemeroptera present in streams (pH >6)

	Conductivity <200 μ S/cm	Conductivity >1,500 μ S/cm
Embeddedness score <7	42/44 (95%)	7/16 (44%)
Embeddedness score >15	210/211 (99%)	6/15 (40%)

Table B-14. Evidence and weights for confounding by deposited sediment

Type	Score	Evidence
1. Correlation of cause and confounder	–	The WABbase embeddedness score is weakly correlated with conductivity ($r = -0.18$, $n = 2,197$).
2. Correlation of effect and confounder	–	The WABbase embeddedness score is weakly correlated with Ephemeroptera ($r = 0.22$, $n = 2,197$).
3. Contingency of high level of cause and confounder	– –	In a contingency table (see Table B-13), high embeddedness (score >15) has little effect at either high or low conductivity.
6. Removal of confounder	– –	When samples with an embeddedness score <13 are removed from the analysis, the correlation of conductivity with the number of Ephemeroptera was virtually unchanged ($r = -0.62$, $n = 1,088$).
Weight of evidence	– – –	Very confident: all negative, some strongly. No treatment for confounding.

B.4.5. High pH

The dissolution of limestone and dolomite increases as unweathered surface area of rock increases. Waters draining crushed limestone and dolomite contain HCO_3^- , which contributes to higher pH and alkalinity. The HCO_3^- that raises the pH is also a major anion moiety that contributes to conductivity. Hence, pH directly reflects a major constituent of conductivity (HCO_3^-), so it could not be a conventional confounder. In addition, salts influence hydrogen ion activity—which is measured as pH. In any case, available evidence indicates that the variance in pH has little effect on the derivation of the HC_{05} for conductivity in waters above pH 7 (see Tables B-15 and B-16).

B.4.6. Low pH

Because low pH from acid mine drainage is known to be an important cause of impairment where it occurs and was judged a priori to be a potentially important environmental variable. That preconception was supported by the evidence summarized here (see Table B-17).

Therefore, sites with pH <6 were not used to calculate the XC values. However, Table B-15 suggests that even below pH 4.5, conductivity is more important than acidity to the occurrence of Ephemeroptera (see Tables B-15 and B-17). In sum, although the benchmark applies to waters with neutral or basic pH, high conductivity appears to also cause effects at low pH.

Table B-15. Number of sites with high and low conductivity with high and low levels of pH with Ephemeroptera present

	Conductivity <200 $\mu\text{S}/\text{cm}$	Conductivity >1,500 $\mu\text{S}/\text{cm}$
pH <4.5	16/19 (84%)	0/14 (0%)
pH >8.5	3/3 (100%)	4/8 (50%)

Table B-16. Evidence and weights for confounding by high pH

Type	Score	Evidence
1. Correlation of cause and confounder	0	Conductivity was moderately correlated with pH between 7 and 9 in the WV data set ($r = 0.45$, $n = 1,900$) and weakly correlated in the EPA data set ($r = 0.14$, $n = 45$).
2. Correlation of effect and confounder	–	High pH was weakly correlated with Ephemeroptera in the WV data set ($r = -0.19$, $n = 1,906$) and in the EPA data set ($r = -0.10$, $n = 45$).
3. Contingency of high level of cause and confounder	–	In a contingency table (see Table B-15), high pH at high conductivities has the same frequency of Ephemeroptera as high conductivity without elevated levels of another variable in other contingency tables (approximately 50%).
5. Levels of confounder known to cause effects	–	EPA (1976) Water Quality Standards indicate that water with pH 6.5–9 is protective of freshwater fish and nearly all data were within that range.
	–	Tests of the mayfly <i>Isonychia bicolor</i> found sublethal effects at pH 10 and lethality at pH 11 (Peters et al., 1985).
6. Removal of confounder shows it is important	–	When samples with pH >8.5 are removed from the analysis, the correlation of conductivity with the number of Ephemeroptera was unchanged ($r = -0.62$, $n = 2,151$). However, this evidence is weak because relatively few sites were removed.
8. Potential confounding evaluated by frequency	–	The number of sites with a pH >8.5 is a very small proportion of the sample (<2.5%), so high pH is unlikely to influence the conductivity relationship.
Weight of evidence	–	Reasonably confident: majority negative. No treatment for confounding.

Table B-17. Evidence and weights for confounding by low pH

Type	Score	Evidence
1. Correlation of cause and confounder	+	Conductivity was moderately correlated with pH <6 ($r = -0.48, n = 145$).
2. Correlation of effect and confounder	+	Low pH was moderately correlated with Ephemeroptera ($r = 0.46, n = 145$).
3. Contingency of high level of cause and confounder	-	Even at low pH some low conductivity streams support some Ephemeroptera but not at high conductivities (see Table B-15).
5. Levels of confounder known to cause effects	+	Hatching success of the mayfly <i>Habrophlebia vibrans</i> was reduced a pH of 5.0 and lower (Rowe et al., 1988).
	-	WVSCI was not reduced at pH 4–6 unless aluminum was elevated in the Clear Fork, WV, study (Gerritson et al., 2010).
Weight of evidence	+	Potential confounding: majority positive. Correction for confounding was preformed.

B.4.7. Selenium

Selenium (Se) is a potential confounder because it is commonly associated with coal, and elevated levels have been reported in the region, but the evidence does not support confounding (see Table B-18). No correlations were found between selenium and Ephemeroptera or between selenium and conductivity in the West Virginia data set or in the EPA Region 3 data set. This result is unreliable because most of the selenium values were detection limits, and many of the detection limits were relatively high, even equaling the water quality criterion of 5.0 µg/L. In addition, there were too few high selenium concentrations in the West Virginia data to perform a contingency table analysis. For these reasons, correlational evidence of confounding was ambiguous.

Evidence of the sufficiency of observed selenium levels to cause extirpation of stream macroinvertebrates is weakly negative. The National Ambient Water Quality Criterion (5 µg/L) is irrelevant because it is based on more sensitive vertebrates (U.S. EPA, 2004). Field and laboratory studies have found invertebrates to be relatively insensitive and unaffected at levels observed in WV streams (Lemly, 1993; Chapman et al., 2010). In outdoor artificial streams dosed with selenium, insects were less sensitive than fish, crustaceans, and oligochaetes; baetid mayfly nymphs (*Baetis*, *Callibaetis*), damselfly nymphs (*Enallagma*), and chironomid larvae were not statistically significantly reduced—even at 30 µg/L (Swift, 2002). Relatively few invertebrate species have been tested and highly sensitive species may be identified in the future

(DeBruyn and Chapman, 2007), but the available toxicological evidence does not indicate that selenium confounds the relationship between conductivity and invertebrate extirpation.

The effects of removing high selenium on the conductivity relationship (Evidence Type 6) were addressed using the West Virginia data set. When data from streams with selenium concentrations above the water quality criterion ($5 \mu\text{g/L}$) were removed, the linear correlation coefficient for number of ephemeropteran genera and log conductivity is barely changed ($r = -0.56$, $n = 339$) relative to the full data set. When the same analysis was performed with the EPA data set, the correlation was actually greater than that for the full data set ($r = -0.84$, $n = 32$) (see Figure B-2), which is contrary to expectations for a confounder. This result indicates that the conductivity relationship is not confounded by toxic effects of selenium.

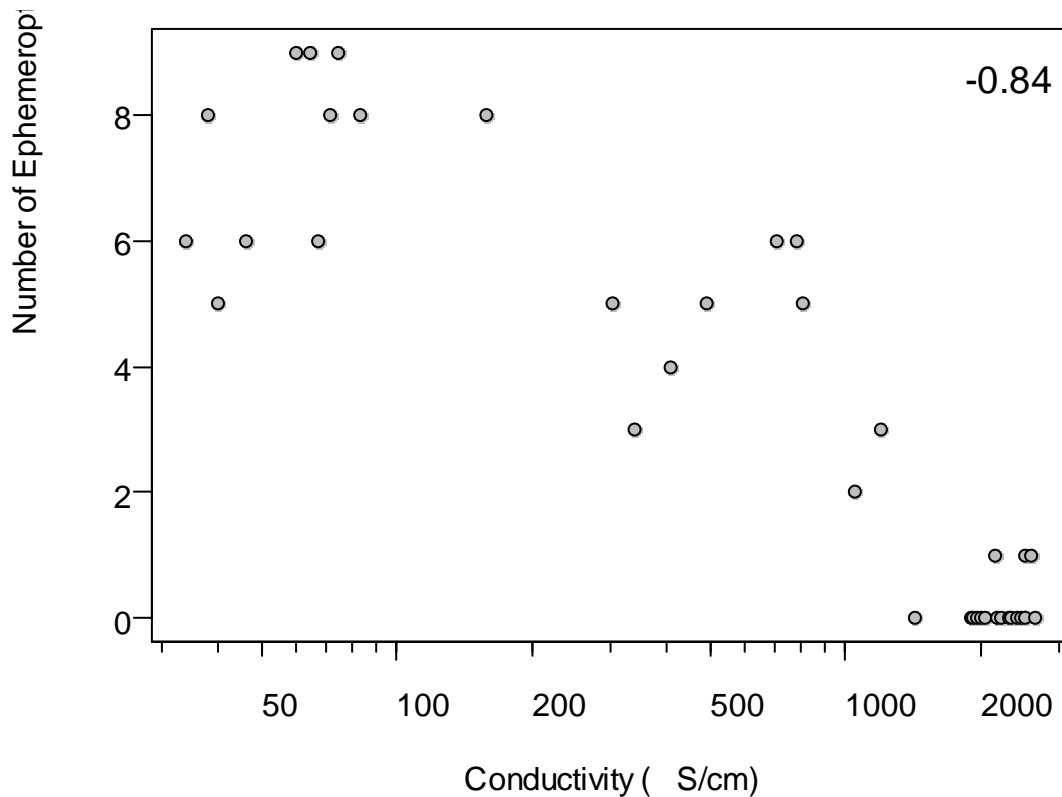


Figure B-2. Spearman's correlation coefficient and scatterplot between the number of ephemeropteran genera and conductivity for 32 sites with low selenium concentrations ($<5 \mu\text{g/L}$). Data from the EPA Region 3 data set.

Table B-18. Evidence and weights for confounding by selenium

Type	Score	Evidence
1. Correlation of cause and confounder	0	Conductivity was not correlated with total selenium in the WV data set ($r = 0.09$, $n = 501$) and in the EPA data set ($r = -0.07$, $n = 46$), but the evidence is ambiguous due to poor selenium data.
2. Correlation of effect and confounder	0	Ephemeroptera were not correlated with total selenium in the WV data set ($r = -0.04$, $n = 501$) and in the EPA data set ($r = -0.07$, $n = 46$), but the evidence is ambiguous due to poor selenium data.
5. Levels of confounder known to cause effects	–	In the most relevant toxicity test, effects on insects in an artificial stream over an exposure of >2 years, occurred at >0.030 mg/L (Swift, 2002). The 90 th centiles for dissolved and total selenium were (0.003 and 0.005 mg/L).
6. Removal of confounder shows it is important	–	After removing high selenium sites ($\geq 5 \mu\text{g/L}$), the correlation of Ephemeroptera with conductivity is barely changed ($r = -0.56$, $n = 339$) relative to the full data set, but the evidence is not strong because few sites have high selenium. The same analysis performed with the EPA data set also found no reduction in correlation (Figure B-2).
9. Specific effects of the confounder	– –	Selenium affects fish more than invertebrates and, in Swift (2002), crustaceans and oligochaetes more than insects, which is not the pattern seen in the streams.
	– –	Selenium causes characteristic deformities in fish, which have not been seen in the streams.
	– –	Selenium effects occur primarily in top predators, not herbivores and detritivores such as the Ephemeroptera.
	– –	Selenium at ambient concentrations causes effects in lentic systems but not lotic systems such as the streams sampled in WV. Deformities typical of selenium have been found in a reservoir in the region but not in streams (WVDEP, 2010).
Weight of evidence	–	Reasonably confident: majority negative. No treatment for confounding.

Consideration of the specific effects of selenium (Evidence Type 9) suggests that it is not an important contributor to the impairment. First, the most sensitive organisms to aqueous selenium are fish and other oviparous vertebrates (Chapman et al., 2010) but, in this case, relatively selenium-insensitive insects are most affected. Second, selenium causes characteristic deformities in fish, which have not been reported in WV streams. Third, the effects of selenium

at low concentrations are seen in lentic ecosystems (lakes, reservoirs, ponds, wetlands)—not in streams like those from which the conductivity relationship and benchmark were derived (Chapman et al., 2010). Finally, because selenium is biomagnified, it primarily affects top predators not the herbivores and detritivores that are affected in this case. This specificity is supported by the fact that, in the region, the only reported effects of selenium are greatly elevated body burdens and associated deformities in a top predator fish (largemouth bass) in a lentic system (Upper Mud River Reservoir) (WVDEP, 2009, 2010).

The weight of evidence does not support confounding by selenium, so no action was taken to adjust the dataset or analysis. However, because existing selenium data are poor, the occurrence of selenium in central Appalachian streams should be investigated further.

B.4.8. Temperature

Elevated temperature may occur with elevated conductivity if the sources of salts are associated with reduced stream shading or if saline effluents are warmed. In an evaluation using contingency tables, Ephemeroptera were present at 99–100% of sites at low conductivity at both high and low temperature (see Table B-19). However, the differences between low and high temperature are not large and that in itself suggests that temperature would not be a confounder (a variant of Evidence Type 5). Correlations of temperature with conductivity are inconsistent (see Table B-20). More importantly, elevated temperature does not appear to be associated with the loss of Ephemeroptera and the relationship of conductivity to Ephemeroptera is not influenced by elevated temperatures.

Table B-19. Number of sites with high and low temperatures and high and low conductivity with Ephemeroptera present in streams (pH >6)

	Conductivity <200 μS/cm	Conductivity >1,500 μS/cm
Temperature <17°C	468/474 (99%)	9/27 (33%)
Temperature >22°C	78/78 (100%)	24/43 (56%)

Table B-20. Evidence and weights for confounding by temperature

Type	Score	Evidence
1. Correlation of cause and confounder	0	Temperature was moderately correlated with conductivity year-round in the WV data set ($r = 0.39$, $n = 2,216$) but weakly correlated in the EPA data set ($r = 0.17$, $n = 46$).
2. Correlation of effect and confounder	–	Temperature was weakly correlated with Ephemeroptera year round in the WV data set ($r = -0.22$, $n = 2,216$) and uncorrelated in the EPA data set ($r = -0.06$, $n = 46$)
3. Contingency of high level of cause and confounder	– – –	Ephemeroptera were present at 99–100% of sites at low conductivity at both high and low temperature. In the high conductivity categories, Ephemeroptera occurred in more sites with elevated temperatures (see Table B-19), which is contrary to expectations, if temperature were contributing to the impairment.
5. Levels of confounder known to cause effects	–	Temperature limits are highly taxon specific but temperatures rarely exceeded the WV limits for reference sites ($<30.6^{\circ}\text{C}$ May–November and $<22.8^{\circ}\text{C}$ December–April) and, therefore, are not likely to cause extirpation.
6. Removal of confounder shows it is important	– –	When high temperatures ($>22^{\circ}\text{C}$) were deleted, the correlation of conductivity and Ephemeroptera was unchanged ($r = -0.61$, $n = 1,787$).
7. Multivariate statistics	– –	Habitat quality, temperature and fecal coliform together had essentially no effect on the slope in multiple regression (see Table B-7).
Weight of evidence	– –	Moderately confident: none positive, some strongly negative. No treatment for confounding.

B.4.9. Lack of Headwaters

The loss of headwaters due to mining and valley fill eliminates a source of recolonization for downstream reaches. Hypothetically, this could result in extirpation of invertebrates if the sampled sites are sink habitats that must be recolonized by headwater source habitats. This is plausible in stream reaches immediately below valley fills. However, where there are other headwaters on tributaries above the sampling site, they serve as alternative sources for recolonization. No regional data are available to address this issue. However, examination of individual watersheds shows that many if not most of the sampled sites have at least one upstream intact headwater. Two examples are presented here.

Ballard Fork, a tributary to the Mud River in West Virginia, is downstream of several valley fills but has unmined tributaries upstream such as Spring Branch (see Figures B-3, B-4, and B-5). Conductivity in Spring Branch measured $<44\text{--}66\ \mu\text{S}/\text{cm}$. Conductivity in Ballard Fork was $464\text{--}2,300\ \mu\text{S}/\text{cm}$. In Spring Branch, the benthic invertebrate assemblages in the springs of 1999, 2000, and 2006 had 6–8 genera of Ephemeroptera representing 29–45% of the sample. In contrast, on the same dates Ballard Fork had 1–3 genera of Ephemeroptera representing only 2–4% of the sample and those may be immigrant specimens. Hence, even when a source of recolonization was available from Spring Branch, ephemeropteran genera were extirpated in Ballard Fork where conductivity was elevated. Also, habitat quality (total RBP habitat score), embeddedness, and pH are not related to biological quality, so they are not confounders in these streams (see Table B-21).

In the Twentymile Creek watershed, the most upstream catchment above river kilometer (Rkm) 44 is a small headwater that is 99% forested. Between Rkm 44 and 13, the tributary catchments are heavily mined with valley fills. Below Rkm 25 to the mouth, benthic invertebrate assemblages are depauperate. Two catchments that enter Twentymile Creek near Rkm 17 and 14 are 100% forested with diverse benthic invertebrate assemblages. Nevertheless, at Rkm 12, the benthic assemblage in Twentymile Creek remains depressed.

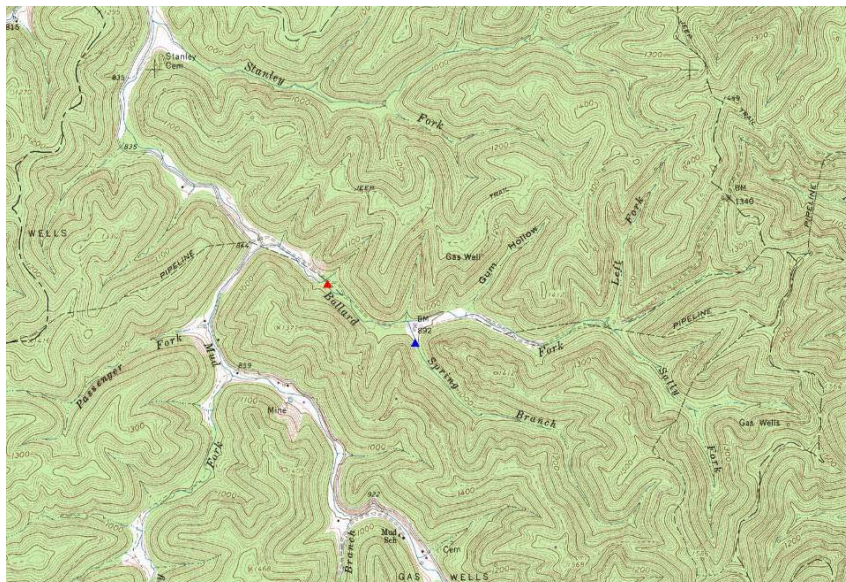


Figure B-3. Topographical map of Spring Branch (blue triangle) and Ballard Fork (red triangle) sampling stations.



Figure B-4. Aerial imagery (June 13, 2007) with superimposed sampling locations of Spring Branch (turquoise square) and Ballard Fork (yellow square). Mined land drains into Ballard Fork (upper section of image) and forested land drains into Spring Branch (lower right quadrant). Two valley fills indicated by white arrows as examples.



Figure B-5. Aerial imagery (April 10, 1996) with superimposed sampling locations of spring branch (turquoise square) and Ballard Fork (yellow square). Same area as Figure 3. The many upstream valley fills in Ballard Fork are easily seen.

Table B-21. Comparison of low conductivity Spring Branch with high conductivity Ballard Fork

Stream name	Date	Embed.	Total RBP score	pH	µS/cm	# E	% E	Total count
Spring Branch	5/9/2006	16	149	7.7	66	8	29.27	205
Spring Branch	4/18/2000	16	163	7.5	44	6	44.76	143
Spring Branch	4/20/1999			7.7	51	8	34.72	337
Ballard Fork	5/9/2006	14	149	8.1	1,195	3	2.96	203
Ballard Fork	4/18/2000	12	148	7.1	464	1	2.08	48
Ballard Fork	1/25/2000			7.9	1,050	0	0	52
Ballard Fork	7/26/1999			8.2	2,300	0	0	88
Ballard Fork	4/20/1999			8.1	1,201	3	4.12	291

Embed. = embeddedness score from RBP; RBP = Rapid Bioassessment Protocol Habitat Evaluation; # E = Number of ephemeropteran genera; % E = percent of ephemeropteran individuals in the sample; Total count = count of all individuals of all taxa.

Source: data from U.S. EPA mountaintop mining studies (Green et al., 2000; Pond et al., 2008).

Downstream from Rkm 12, there are mixed mining and forest land uses. Near Rkm 2 there are legacy mining and urban land uses (see Table B-22). WVSCI scores, number of ephemeropteran families and number of ephemeropteran, plecopteran, and trichopteran (EPT) families were low when conductivity was high regardless of the condition of catchments that provided sources of benthic macroinvertebrates including salt-sensitive genera.

In these two examples, the evidence indicates that the reduction in ephemeropteran genera or EPT is not caused by a lack of sources of recolonization from headwaters. This is not to say that recolonization is never an issue. The sources of salts in this region are primarily chronic and localized, so lack of recolonization is unlikely to confound their effects. However, if an episodic agent caused the loss of aquatic organisms (e.g., drought or forest treatment with insecticides), sources of recolonization could be important.

Table B-22. Twentymile Creek sampling locations, conductivity, habitat score, number of EPT taxa, and WVSCI scores

Year	River kilometer	Tributary catchment land use ^a	Max reported conductivity (µS/cm)	RBP habitat score	# E	# EPT	WVSCI
2003	44.6	Forested	44	148	4	15	90.72
2004	44.6	Forested	37				–
1998	25.1	Mined	805	155	3	7	67.62
2003	25.1	Mined	2,087	153	1	5	58.45
2003	11.9	Mixed Forest and Mine	1,702	157	2	7	64.74
2004	11.9	Mixed Forest and Mine	1,282	–	–	–	–
2003	1.8	Mixed Forest, Mine, & Urban	987	–	–	–	–
2004	1.8	Mixed Forest, Mine, & Urban	1,138	–	–	–	–
2003	0.5	Mixed Forest, Mine, & Urban	845	146	2	6	66.73
2004	0.5	Mixed Forest, Mine, & Urban	836	–	–	–	–
1998	0	Mixed Forest, Mine, & Urban	590	131	3	8	65.94

^aLand use refers to catchment land use of tributaries upstream from the sampled sites in Twentymile Creek.

E families = Number of ephemeroptera families; #EPT = ephemeropteran, plecopteran, and trichopteran families; WVSCI = West Virginia Stream Condition Index.

Source: data from WABbase.

B.4.10. Catchment Area

Larger streams tend to have more moderate chemical properties than small streams because they receive waters from more sources, both natural and anthropogenic. Consequently, extreme values, in this case both low and high conductivity, tend to occur less frequently in large streams. One of the initial data filters for this analysis was to exclude streams larger than 155 km² (or 60 mi²). Small streams are numerically more abundant than large streams and the inclusion of large streams might introduce extraneous variance. This raises the issue whether

stream size is a potential confounder and whether the results from small streams might be extrapolated to larger streams. That is, do the same effects of conductivity occur in larger streams as were found in the detailed analysis of smaller streams? We examined these issues by analyzing the influence of stream size (as catchment area) on the effects of conductivity and on the occurrence of Ephemeroptera.

We categorized streams by catchment area into three groups: small catchments less than 6 mi² (15.5 km²), medium catchments of 6 to 60 mi² (15.5 km² to 155 km²), and large catchments greater than 60 mi² (155 km²). In all three stream size categories, if conductivity was <200 µS/cm, 99% or more of all streams had Ephemeroptera, but if conductivity was above 1,500 µS/cm, fewer streams had Ephemeroptera (see Table B-23). The number of Ephemeroptera taxa declines with increasing conductivity in all streams with measured catchment areas, independent of classification of catchment area ($r = -0.59$). Correlation of log conductivity with log catchment area is weak (see Table B-24).

The weight of evidence for confounding by catchment area (see Table B-24) is uniformly negative, so we conclude that catchment area has little or no effect on invertebrate response to conductivity.

Table B-23. Number and percent of streams with Ephemeroptera present: small, medium and large streams and low and high conductivity (pH >6)

	Conductivity <200 µS/cm	Conductivity >1,500 µS/cm
Small streams (<15.5 km ²)	302/303 (100%)	6/15 40%
Medium streams (≥15.5 km ² and ≤155 km ²)	118/119 (99%)	10/14 (71%)
Large streams (>155 km ²)	37/37 (100%)	1/2 (50%)

Table B-24. Evidence and weights for confounding by catchment area

Type	Score	Evidence
1. Correlation of cause and confounder	–	Log catchment area was very weakly correlated with log conductivity ($r = 0.18$, $n = 926$).
2. Correlation of effect and confounder	– –	Log catchment area was not correlated with the number of ephemeropteran genera ($r = -0.009$, $n = 926$).
3. Contingency of high level of cause and confounder	–	In a contingency table (see Table B-23), catchment area did not affect the probability of finding Ephemeroptera at low conductivity. Medium size somewhat increased the probability of occurrence at high conductivity.
6. Removal of confounder	– –	When large streams were removed, the correlation of conductivity and number of ephemeropteran genera was barely changed ($r = -0.60$, $n = 837$).
Weight of evidence	– – –	Very confident: all negative, some strongly negative. No treatment for confounding.

B.4.11. Ponds

The effluents from most valley fills flow into settling ponds, and it has been suggested that those ponds are the actual cause of downstream community impairments. This issue was addressed using the EPA Region 3 data set because it identifies the presence of ponds. When data from only streams with ponds are used (i.e., the occurrence of ponds is removed as a variable—Evidence Type 4), the correlation coefficient for number of ephemeropteran genera and log conductivity is $r = -0.84$ (see Figure B-6). This result is somewhat higher than those for the uncensored EPA Region 3 data set ($r = -0.73$), which is contrary to the expectation if ponds were the cause. This result clearly shows that the conductivity relationship is not a result of co-occurrence with ponds. In addition, when ponds are removed and the streams are reclaimed, conductivity remains high and the effects continue. For example, Venter’s Branch and Jones Branch in Martin County, KY, were mined in the mid 1990s, and the ponds were removed. When the streams were sampled in 2009, conductivity was $>2,000 \mu\text{S}/\text{cm}$ and no Ephemeroptera were found in either stream (Greg Pond, U.S. EPA, personal communication).

The weight of evidence for confounding from ponds is uniformly negative, so we conclude that the presences of ponds have little or no effect on invertebrate response to conductivity.

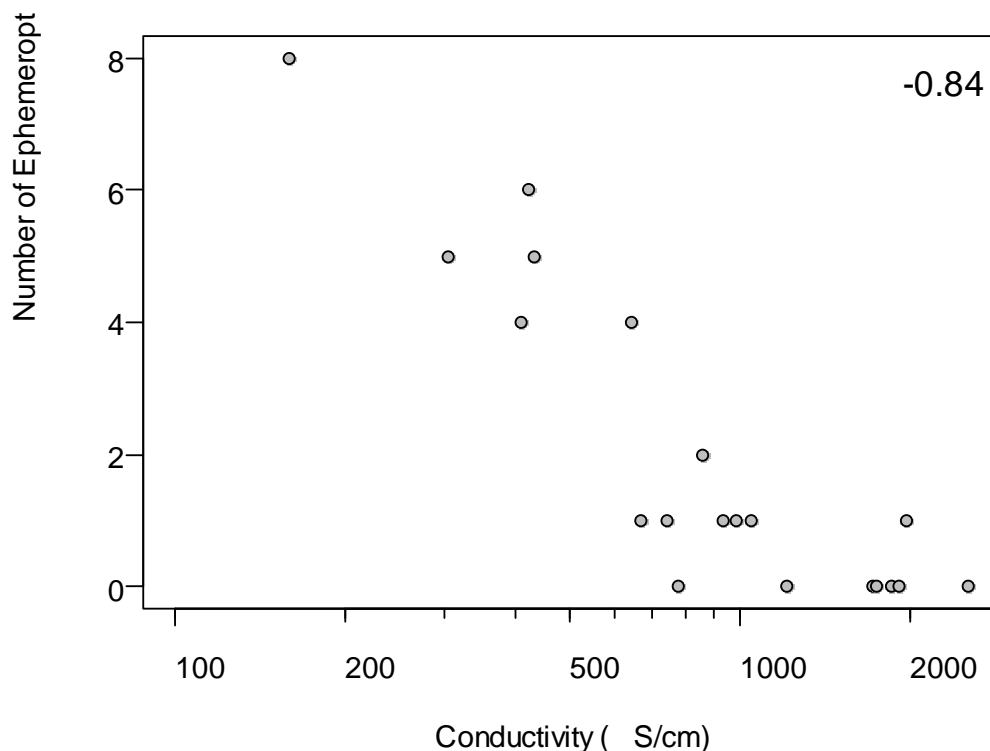


Figure B-6. Spearman’s correlation coefficient and scatterplot between the number of ephemeropteran genera and conductivity for 20 sites below settling ponds for valley fills. Data from the EPA Region 3 data set.

B.4.12. Dissolved Oxygen

Dissolved oxygen (DO) is not expected to be a confounder because these relatively shallow and high gradient streams are generally well oxygenated, but reviewer comments suggested that DO might be a confounder.

The 30-day mean water quality criteria for DO are 6.5 mg/L for coldwater and 5.5 mg/L for warm water (U.S. EPA, 1986). Ephemeropterans showed slightly reduced body condition and survivorship at concentrations below 7 mg/L DO in laboratory studies (Love et al., 2005; Pucket and Cook, 2004). A recent assessment of the Clear Fork watershed, WV, derived a plausible threshold for DO of 5 mg/L and a substantial threshold of 4 mg/L (Gerritson et al., 2010). DO is rarely low (see Table B-25), but because the hour of sampling was not consistent in the data set, some uncertainty remains. Nevertheless, Ephemeroptera are present at 99% of sites with low conductivity even when DO is low for these streams and at high conductivity the presence of Ephemeroptera is unaffected by DO. Correlations of DO with conductivity were weak and with Ephemeroptera were very weak (see Table B-26). The available evidence shows no signs of confounding by low DO (see Tables B-25 and B-26).

Table B-25. Number of sites with high and low dissolved oxygen and high and low conductivity with Ephemeroptera present in streams (pH >6)

	Conductivity <200 μ S/cm	Conductivity >1,500 μ S/cm
DO >10.3 mg/L	244/246 (99%)	11/28 (39%)
DO <8.2 mg/L	172/174 (99%)	12/30 (40%)

Table B-26. Evidence and weight for confounding by dissolved oxygen (DO)

Type	Score	Evidence
1. Correlation of cause and confounder	–	DO was weakly correlated with conductivity ($r = -0.11$, $n = 2,188$).
2. Correlation of effect and confounder	– –	DO was uncorrelated with the number of ephemeropteran genera ($r = 0.09$, $n = 2,188$).
3. Contingency of high level of cause and confounder	– –	In a contingency table (see Table B-25), Ephemeroptera are present at 99% of sites with low conductivity (<200 μ S/cm) even when DO is low (<8 mg/L) and at high conductivity the presence of Ephemeroptera is unaffected by DO.
5. Level of confounder known to cause effects	–	The 30 day mean water quality criteria for DO of 6.5 mg/L for coldwater and 5.5 mg/L for warm water (U.S. EPA, 1986) are below the lower 10 th centile of WV sites (7.3 mg/L).
	–	Reduced body condition and survivorship in ephemeropterans occur below 7 mg/L DO in laboratory studies (Love et al., 2005; Pucket and Cook, 2004).
	–	In the Clear Fork watershed, WV, a plausible threshold for DO of 5 mg/L and a substantial threshold of 4 mg/L were derived (Gerritsen et al., 2010).
6. Removal of confounder	– –	When sites with moderate to low DO (<8.2 mg/L) were removed from the analysis, the correlation of conductivity with the number of Ephemeroptera is slightly increased ($r = -0.63$, $n = 1,642$).
Weight of evidence	– – –	Very confident: all negative, some strongly negative. No treatment for confounding.

B.4.13. Metals

Iron (Fe), aluminum (Al), and manganese (Mn) are the metals most associated with acid mine drainage and commenters have suggested that they may cause the impairment associated with conductivity. However, for the following reasons, the circum-neutral to moderately alkaline streams are unlikely to experience toxicity from these metals (Luoma and Rainbow, 2008). The most toxic form of iron (free Fe^{2+}) does not occur in oxygenated waters above pH 4. Under those conditions, iron occurs as hydroxide particles or, if significant dissolved organic matter is present, as iron colloids. In these forms, iron is thought to serve primarily to reduce the toxicity of co-occurring metals by adsorption and co-precipitation. Toxic divalent aluminum precipitates similarly above pH 5 as hydroxide flocs or polymeric aluminum. Divalent manganese is converted to insoluble Mn^{4+} in mildly alkaline waters. The precipitates of these metals may adversely modify habitats and directly affect organisms. However, the valley fill effluents that are primarily responsible for the relationship between conductivity and extirpation of invertebrates are not equivalent to the acid drainage into neutralizing streams that results in heavy accumulations of precipitates. Finally, the toxicity of these divalent anions is mitigated by divalent calcium, which is the dominant cation in the saline mixtures. Hence, it is expected that, as conductivity increases, the toxicity of these metals will decrease per unit concentration.

Because of concern for combined effects of metals, multiple linear regression of conductivity, iron, aluminum and manganese was performed. The metals reduced the coefficient for conductivity by only 8.6% (see Table B-27).

Iron and aluminum are clearly not confounders, based on contingency table analyses (see Tables B-28 and B-29), weak correlations (see Tables B-31 and B-32) and other evidence (see Tables B-31 and B-32).

However, manganese is more ambiguous since it is moderately correlated with both conductivity and ephemeropteran genera (see Tables B-30 and B-33). Manganese has been relatively poorly studied because it has seldom been found at toxic levels. Like other divalent cationic metals, Mn^{2+} is less toxic in hard (i.e., high Ca) waters and the high conductivity waters in this region are inherently hard. Based on a linear relationship of hardness to conductivity in the WV data, 300 $\mu\text{S}/\text{cm}$ conductivity is equivalent to a hardness of approximately 200 mg/L CaCO_3 . The equivalent hardness-adjusted British Columbia Chronic Water Quality Guideline for manganese is 1.5 mg/L (BC, 2001). Dittman and Buchwalter (2010) provide the laboratory study with the most directly relevant taxa: aquatic insects from Appalachia. They quantified bioaccumulation and performed biomarker studies that found reduced levels of cysteine and glutathione at 0.10 and 0.50 mg/L but they saw no overt toxic effects. The most relevant conventional toxicity tests of aquatic invertebrates were 21 day reproduction tests of *Daphnia magna* which yielded IC_{25} values of 5.4 and 9.4 mg/L for hardness levels of 100 and 250 mg/L,

respectively (Reimer, 1999). A recent assessment of the Clear Fork watershed, WV, concluded that total manganese at 0.002–0.50 mg/L was a minor contributor to biotic impairment, because manganese was weakly correlated ($r = -0.16$) with the WVSCI index when corrected for stronger causes (Gerritsen et al., 2010).

In sum, iron and aluminum are clearly not confounders. Equivocal evidence suggests that manganese is potentially a weak confounder.

Table B-27. An output table for two linear regression models. The first is the simple model predicting ephemeropteran genera from conductivity. The second is a multivariate model with the additional covariates iron, aluminum, and manganese.

Parameter	Estimate	Standard error
Univariate model		
Intercept	3.65	0.056
Conductivity slope	-0.93	0.024
Multivariate model		
Intercept	3.05	0.092
Conductivity slope	-0.85	0.031
Iron	-0.028	0.042
Aluminum	-0.066	0.044
Manganese	-0.30	0.033

Table B-28. Number of sites with high and low total iron and high and low conductivity with Ephemeroptera present in streams (pH >6)

	Conductivity <200 μ S/cm	Conductivity >1,500 μ S/cm
Iron >0.5 mg/L	122/124 (98%)	8/28 (29%)
Iron <0.12 mg/L	139/140 (99%)	13/24 (54%)

Table B-29. Number of sites with high and low total aluminum and high and low conductivity with Ephemeroptera present in streams (pH >6)

	Conductivity <200 μS/cm	Conductivity >1,500 μS/cm
Aluminum >0.23 mg/L	177/178 (99%)	5/22 (23%)
Aluminum <0.09 mg/L	103/103 (100%)	14/31 (45%)

Table B-30. Number of sites with high and low total manganese and high and low conductivity with Ephemeroptera present in streams (pH >6)

	Conductivity <200 μS/cm	Conductivity >1,500 μS/cm
Mn >0.1 mg/L	69/72 (96%)	13/50 (26%)
Mn <0.02 mg/L	158/158 (100%)	3/5 (60%)

Table B-31. Evidence and weight for confounding by iron

Type	Score	Evidence
1. Correlation of cause and confounder	---	Dissolved iron was uncorrelated with conductivity in the WV data ($r = -0.08$, $n = 1,265$) and weakly correlated in the EPA data ($r = -0.17$, $n = 12$). Both signs are incorrect for confounding.
	--	Total iron was uncorrelated with conductivity in the WV data ($r = 0.03$, $n = 1,439$) and weakly correlated with the wrong sign in the EPA data ($r = -0.14$, $n = 46$).
2. Correlation of effect and confounder	--	Dissolved iron was uncorrelated with the number of ephemeropteran genera in the WV data ($r = -0.08$, $n = 1,265$) and in the EPA data ($r = -0.04$, $n = 12$).
	-	Total iron was weakly correlated with the number of ephemeropteran genera in the WV data ($r = -0.14$, $n = 1,436$) and in the EPA data with the wrong sign ($r = 0.12$, $n = 46$).
3. Contingency of high level of cause and confounder	--	In a contingency table (see Table B-28), Ephemeroptera are present at $\geq 98\%$ of sites with low conductivity ($< 200 \mu\text{S/cm}$) even when total iron is high ($> 0.1 \text{ mg/L}$). There are too few observations at extreme conductivities to derive a contingency table for dissolved iron.
5. Level of confounder known to cause effects	-	The most relevant criteria are the British Columbia Chronic Water Quality Guidelines of 1 mg/L for total iron and 0.35 mg/L for dissolved iron (BC, 2008), which are above the 90 th centiles in WV (0.93 and 0.14 mg/L , respectively).
	-	The most relevant conventional toxicity tests were 120 h tests of the mayfly <i>Leptophlebia marginata</i> with an LC_{50} of 106.3 mg/L and reduced predator avoidance at 70 mg/L at pH 7 and low conductivity ($7.0 \mu\text{S/cm}$) (Gerhardt, 1994) which are well above the maximum dissolved iron in WV.
	-	Two highly relevant field studies use data from the same source. Total iron caused no or minimal change at 0.21 mg/L and slight to moderate changes at 1.74 mg/L using benthic macroinvertebrate abundances in the WVDEP data set (Linton et al., 2007). Acid drainage sites were not excluded. Gerritson et al. (2010) found no effects of iron in the WVDEP data set.
6. Removal of confounder	--	When sites with moderate to high dissolved iron ($> 0.06 \text{ mg/L}$) were removed from the analysis, conductivity is more negatively correlated with the number of Ephemeroptera ($r = -0.72$, $n = 949$), which is contrary to expectations for a confounder. This result is corroborated by the EPA data set ($r = -0.77$, $n = 9$).
	--	When sites with moderate to high total iron ($> 0.5 \text{ mg/L}$) were removed from the analysis, conductivity is slightly more negatively correlated with the number of Ephemeroptera ($r = -0.66$, $n = 1,076$), which is contrary to expectations for a confounder. This result is corroborated by the EPA data set ($r = -0.64$, $n = 34$).
7. Multivariate statistics	--	In the multiple linear regression, the slope for iron is less than a tenth that of conductivity (see Table B-27).
Weight of evidence	---	Very confident: all negative, some strongly negative. No treatment for confounding.

Table B-32. Evidence and weight for confounding by aluminum

Type	Score	Evidence
1. Correlation of cause and confounder	–	Dissolved aluminum was weakly correlated with conductivity in the WV data ($r = 0.12$, $n = 1,293$) and in the EPA data ($r = 0.18$, $n = 12$).
	–	Total aluminum was weakly correlated with conductivity and in the wrong direction in the WV data ($r = -0.12$, $n = 1,442$) and uncorrelated in the EPA data ($r = 0.03$, $n = 46$).
2. Correlation of effect and confounder	–	Dissolved aluminum was weakly correlated with the number of ephemeropteran genera in the WV data ($r = -0.16$, $n = 1,293$) and uncorrelated in the EPA data ($r = -0.02$, $n = 12$).
	--	Total aluminum was uncorrelated with the number of ephemeropteran genera in the WV data ($r = 0.03$, $n = 1,442$) and weakly correlated in the EPA data ($r = 0.15$, $n = 46$); both have the wrong sign
3. Contingency of high level of cause and confounder	–	In a contingency table (see Table B-29), Ephemeroptera are present at >99% of sites with low conductivity (<200 $\mu\text{S}/\text{cm}$) even when total aluminum is high (>0.1 mg/L). However, there are fewer Ephemeroptera at high conductivity with high total aluminum so some confounding is possible but only at levels far above the benchmark. There are too few observations at extreme conductivities to derive a contingency table for dissolved aluminum.
5. Level of confounder known to cause effects	0	The most relevant criteria are the British Columbia Acute and Chronic Water Quality Criteria of 0.1 and 0.05 mg/L, respectively, for dissolved aluminum above pH 6.5 (BC, 2001). The chronic value equals the median value in WV; the acute value equals the 90 th centile. However, the criteria are based on effects on sensitive fish.
	0	The most relevant conventional toxicity tests were 48 h tests of <i>Ceriodaphnia dubia</i> in neutralized acid mine drainage which gave a mean LC_{50} for total aluminum of 2.9 mg/L (Soucek et al., 2001). This value is well above the 90 th centile (0.1 mg/L) but its relevance to stream insects is unclear.
	–	In the most relevant field study, the plausible and substantial effects thresholds were >0.2 mg/L and >0.4 mg/L dissolved aluminum in WV Ecoregion 69 (Gerritson et al., 2010). These are above the 90 th centile in WV (0.1 mg/L).
6. Removal of confounder	–	When sites with moderate to high dissolved aluminum (>0.06 mg/L) were removed from the analysis, conductivity is slightly more negatively correlated with the number of Ephemeroptera ($r = -0.68$, $n = 973$) which is contrary to expectations for a confounder.
	–	When sites with moderate to high total aluminum (>0.23 mg/L) were removed from the analysis, conductivity is slightly more negatively correlated with the number of Ephemeroptera ($r = -0.66$, $n = 1,063$) which is contrary to expectations for a confounder. This result is corroborated by the EPA data ($r = -0.79$, $n = 15$).
7. Multivariate statistics	--	In the multiple regression, the slope for aluminum is less than a tenth that of conductivity (see Table B-27).
Weight of evidence	--	Moderately confident: none positive, some strongly negative. No treatment for confounding.

Table B-33. Evidence and weight for confounding by manganese

Type	Score	Evidence
1. Correlation of cause and confounder	0	Dissolved Mn was moderately correlated with conductivity in the WV data ($r = 0.64$, $n = 20$) but weakly correlated in the EPA data ($r = 0.22$, $n = 12$).
	+	Total Mn was moderately correlated with conductivity in the WV data ($r = 0.35$, $n = 1,436$) and in the EPA data ($r = 0.55$, $n = 46$).
2. Correlation of effect and confounder	+	Dissolved Mn was moderately correlated with the number of ephemeropteran genera in the WV data ($r = -0.73$, $n = 20$) and in the EPA data ($r = -0.37$, $n = 12$).
	+	Total Mn was moderately correlated with the number of ephemeropteran genera in the WV data ($r = -0.41$, $n = 1,436$) and in the EPA data ($r = -0.49$, $n = 46$).
3. Contingency of high level of cause and confounder	-	In a contingency table (see Table B-30), Ephemeroptera are present at $\geq 96\%$ of sites with low conductivity ($< 200 \mu\text{S/cm}$) even when total Mn is high ($> 0.1 \text{ mg/L}$). However, there are fewer Ephemeroptera at high conductivity with high total Mn suggesting that some confounding is possible at levels far above the benchmark. There are too few dissolved Mn observations at extreme conductivities to derive a contingency table.
5. Level of confounder is known to cause effects	-	The most relevant criterion is the British Columbia Chronic Water Quality Guideline for Mn of 1.5 mg/L (BC, 2001). This is above the maximum dissolved Mn.
	0	The most relevant conventional toxicity tests were 21 day reproduction tests of <i>Daphnia magna</i> which yielded IC_{25} values of 5.4 and 9.4 mg/L for hardness levels of 100 and 250 mg/L , respectively (Reimer, 1999). This is far above the 90 th centile and maximum dissolved Mn (0.29 and 1.06 mg/L), but its relevance to stream insects is uncertain.
	-	In the most relevant field study, total Mn in the Clear Fork watershed, WV, was weakly correlated ($r = -0.16$) with the WVSCI index when corrected for stronger causes and there were no substantial effects (Gerriston et al., 2010).
6. Removal of confounder	-	When sites with moderate to high dissolved Mn ($> 0.05 \text{ mg/L}$) were removed from the analysis, the correlation of conductivity with the number of Ephemeroptera is little changed ($r = -0.58$, $n = 16$). There were too few high dissolved Mn sites in the EPA data to corroborate.
	--	When sites with moderate to high total Mn ($> 0.1 \text{ mg/L}$) were removed from the analysis, the correlation of conductivity with the number of Ephemeroptera is slightly increased which is contrary to expectations for a confounder ($r = -0.63$, $n = 1,067$). This result is corroborated by the EPA data ($r = -0.74$, $n = 34$, compared to $r = -0.72$).
7. Multivariate statistics	-	In the multiple regression, the slope for Mn is only 35% that of conductivity and the conductivity slope is reduced by only 8.6% relative to the univariate slope (see Table B-27).
Weight of evidence	-	Reasonably confident. Majority negative. No treatment for confounding.

B.5. SUMMARY OF ACTIONS TAKEN TO ADDRESS POTENTIAL CONFOUNDING

Low pH is an apparent confounder, but sites with pH <6 were removed from the data set when calculating the benchmark value. Other potential confounders were eliminated from consideration with some confidence. We do not argue that these variables do not cause impairment at some locations in the region. Neither do we argue that they have no influence at all on salt-impaired sites. Rather, given the inevitable variability in sites to which the benchmark would be applied and the relatively strong relationship of conductivity and loss of sensitive genera, the evaluated confounders do not substantially affect the model that is used to develop and apply the conductivity benchmark.

REFERENCES

- BC (British Columbia) (2001) Ambient water quality guidelines for manganese. Ministry of Environment, Province of British Columbia, Canada. Available online at <http://www.env.gov.bc.ca/wat/wq/BCguidelines/manganese/manganese.html>.
- BC (British Columbia) (2008) Ambient water quality guidelines for iron. Ministry of Environment, Province of British Columbia, Canada. Available online at http://www.env.gov.bc.ca/wat/wq/BCguidelines/iron/iron_overview.pdf
- Chapman, PM; Adams, WJ; Brooks, ML; et al. (2010) Ecological assessment of selenium in the aquatic environment. Boca Raton: SETAC/CRC Press.
- DeBruyn, AMH; Chapman, PM. (2007) Selenium toxicity to invertebrates: will proposed thresholds for toxicity to fish and birds also protect their prey? *Environ Sci Technol* 41(5):1766–1770.
- Dittman, E; Buchwalter, DB. (2010) Manganese bioaccumulation in aquatic insects: Mn oxide coatings, molting loss, and Mn(II) thiol scavenging. *Environ Sci Technol* 44(23):9182–9188.
- Gerhardt, A. (1994) Short term toxicity of iron (Fe) and lead (Pb) to the mayfly *Leptophlebia marginata* (L.) (Insecta) in relation to freshwater acidification. *Hydrobiologia* 284:157–168.
- Gerritsen, J; Zheng, L; Burton, J; et al. (2010) Inferring causes of biological impairment in the Clear Fork Watershed, West Virginia. U.S. Environmental Protection Agency, Office of Research and Development, National Center for Environmental Assessment, Cincinnati, OH. EPA/600/R-08/146. Available online at http://cfpub.epa.gov/si/si_public_record_report.cfm?dirEntryId=201963
- Green, J; Passmore, M; Childers, H. (2000) A survey of the condition of streams in the primary region of mountaintop mining/valley fill coal mining. Mountaintop mining/valley fill programmatic environmental impact statement. Region 3, U.S. Environmental Protection Agency, Philadelphia, Pennsylvania. Available online at <http://www.epa.gov/region03/mtntop/pdf/appendices/d/streams-invertebrate-study/FINAL.pdf>
- Hill, AB. (1965) The environment and disease: Association or causation. *Proc Royal Soc Med* 58(5):295–300.
- Lemly AD. (1993) Guidelines for evaluating selenium data from aquatic monitoring and assessment. *Environ Monitor Assess* 28:83–100.
- Linton, TK; Pacheco, MAW; McIntyre, DO; et al. (2007) Development of bioassessment-based benchmarks for iron. *Environ Toxicol Chem* 26(6):1291–1298.
- Love, JW; Taylor, CM; Warren, ML. (2005) Predator density and dissolved oxygen affect body condition of *Stenonema tripunctatum* (Ephemeroptera, Heptageniidae) from intermittent streams. *Hydrobiologia* 543:113–118.

- Luoma, SN; Rainbow, PS (2008) Metal contamination in aquatic environments: science and lateral management. Cambridge, UK: Cambridge University Press.
- Pearl, J. (2009) Causality: models, reasoning, and inference, 2nd edition. Cambridge, UK: Cambridge University Press.
- Peters, GT; Cherry, DS; Cairns, J Jr. (1985) Response of *Isonychia bicolor* to alkaline pH: an evaluation of survival, oxygen consumption, and chloride cell ultrastructure. *Can. J. Fish. Aquat. Sci.* 42:1088–1095.
- Pond, GJ; Passmore, ME; Borsuk, FA; et al. (2008) Downstream effects of mountaintop coal mining: comparing biological conditions using family- and genus-level macroinvertebrate bioassessment tools. *J N Am Benthol Soc* 27(3):717–737.
- Puckett, RT; Cook, JL. (2004) Physiological tolerance ranges of larval *Caenis latipennis* (Ephemeroptera: *Caenidae*) in response to fluctuations in dissolved oxygen concentration, pH and temperature. *Texas J Sci* 56(2):123–130.
- Reimer, PS. (1999) Environmental effects of manganese and proposed freshwater guidelines to protect aquatic life in British Columbia. U. British Columbia, Vancouver, Canada. Available online at http://www.env.gov.bc.ca/wat/wq/wq_guidelines.html.
- Rowe, L; Hudson, J; Berrill, M. (1988) Hatching success of mayfly eggs at low pH. *Can J Fish Aquat Sci* 45: 1649–1652.
- Soucek, DJ; Cherry, DS; Zipper, CE. (2001) Aluminum-dominated acute toxicity to the cladoceran *Ceriodaphnia dubia* in neutral waters downstream of an acid mine drainage discharge. *Can J Fish Aquat Sci* 58:2396–2404.
- Stewart-Oaten, A. (1996) Problems in the analysis of environmental monitoring data. In: *Detecting environmental impacts*, Schmitt, RJ; Osenberg, CW; ed. New York, NY: Academic Press. p. 109–131.
- Swift, MC. (2002) Stream ecosystem response to, and recovery from, experimental exposure to selenium. *J Aqua Ecosyst Stress Rec* 9(3):159–184.
- U.S. EPA (Environmental Protection Agency). (1976) Quality criteria for water. Office of Water, Washington, DC; EPA/440/9-76/023. Available online at http://water.epa.gov/scitech/swguidance/waterquality/standards/current/upload/2009_01_13_criteria_redbook.pdf.
- U.S. EPA (Environmental Protection Agency). (1986) Quality criteria for water 1986. Office of Water, Washington, DC; EPA 440/5-86/001. Available online at http://water.epa.gov/scitech/swguidance/waterquality/standards/criteria/aqlife/upload/2009_01_13_criteria_goldbook.pdf.
- U.S. EPA (Environmental Protection Agency). (2004) Draft aquatic life water quality criteria for selenium. Washington, DC; EPA-822-D-04-001. Available online at <http://www.epa.gov/waterscience/criteria/selenium/pdfs/complete.pdf>.
- WVDEP (West Virginia Department of Environmental Protection). (2009) Selenium bioaccumulation among select stream and lake fisheries. West Virginia Department of Environmental Protection. Available online at http://www.dep.wv.gov/WVE/watershed/wqmonitoring/Documents/Selenium/Se_Fish_Tissue_Summary_Paper_final_Feb09.pdf.
- WVDEP (West Virginia Department of Environmental Protection). (2010) Selenium-induced developmental effects among fishes in select West Virginia waters. West Virginia Department of Environmental Protection. Available online at <http://www.dep.wv.gov/WVE/watershed/wqmonitoring/Documents/Selenium/Se%20Larvae%202010%20final.pdf>.

APPENDIX C
DATA SOURCES AND METHODS OF LAND USE/LAND COVER ANALYSIS USED
TO DEVELOP EVIDENCE OF SOURCES OF HIGH CONDUCTIVITY WATER

ABSTRACT

Potential sources of elevated conductivity were characterized for subwatersheds within the Coal, Upper Kanawha, Gauley, and New Rivers. From a large monitoring data set developed by the West Virginia Department of Environmental Protection (WVDEP), 190 <20-km² watersheds were found for which there was total maximum daily load (TMDL) and land cover information in southwestern West Virginia and macroinvertebrate samples identified to the genus level with at least one chemistry sample. Small <20-km² subwatersheds were selected to reduce confounding from multiple sources. Scatter plots of conductivity, SO₄²⁻, and Cl⁻, and alkalinity levels were generated for nine land cover classifications: open water, agriculture, forest, residential, barren, total mining, valley fill, abandoned mine lands, and mining excluding valley fill and abandoned mine lands. Conductivity was negatively correlated with the percentage of forest area and most strongly negatively associated with catchments with the greatest percentages of valley fills, and the HCO₃⁻ and SO₄²⁻ concentrations were greater than Cl⁻ concentration. Areas with more residences and farm buildings also had elevated conductivity but rarely exceeded 1,000 μS/cm, and Cl⁻ often exceeded SO₄²⁻ and HCO₃⁻ concentrations. These findings confirm sources of high conductivity waters that are used as evidence in the causal assessment that salts are a cause of impairment of aquatic macroinvertebrates in streams in West Virginia (see Section A.2.2.4).

C.1. INTRODUCTION

Analysis of land use and cover was used to determine if there was a source of high conductivity, to assess if land use was associated with conductivity levels, and to confirm the relative proportion of ions associated with land use and cover types reported in the literature for different sources. This information was used as evidence of preceding causation in the causal assessment described in Appendix A of this report.

C.2. METHODS

C.2.1. General Approach

Small catchments were delineated, and the proportions of land covers were regressed against water quality parameters. Watershed size was limited to <20 km² to minimize the variety of land use and cover types within a single watershed, thereby providing a clearer signal for each potential source of salinity. However, because the region has a long history of mining and land

cover information may not include legacy mining, persistent effects of mining are potentially present even when there is no current record of past or present mining activity in the publically available land cover databases. Also, residences are present in areas where mining occurs. Therefore, there are potential influences from multiple sources in most of the 190 watersheds, but these are minimized by using small catchments.

The final data set consisted of 190 small watersheds for which macroinvertebrate samples were identified to genus, water chemistry was available from at least one sampling effort, subwatershed area was ≤ 20 km², and detailed land cover information was available. The 190 sites are located in the Coal, Upper Kanawha, Gauley, and New Rivers of Ecoregion 69D. Water quality parameters are from the WVDEP's Watershed Assessment Branch Data Base (WABbase). For each watershed, scatter plots for several parameters were generated for nine land cover classifications: open water, agriculture, forest, urban/residential, barren, total mining, valley fill, and mining excluding valley fill and abandoned mine lands.

C.2.2. General Geographical Information Systems (GIS) Data Descriptions

Numerous geographic information system (GIS) data sets are available for the State of West Virginia. One such repository for data, the WVGISTC (2011), maintains publicly available shapefiles. WVDEP (2011a) also maintains a publicly available repository of statewide GIS data sets (<http://gis.dep.wv.gov/>). All relevant GIS metadata are available for the data housed at each repository site. All GIS coverages used in this EPA study are in universal transverse mercator (UTM) 1983 Zone 17, and the units are in meters. Table C-1 describes some of the publicly available GIS shapefiles that were originally used to develop base files for WVDEP's TMDL program. These base files were the beginning point for determining the 190 stations selected for the analyses described in Section C.2.3 and were used to estimate land uses (see Table C-2). The area in valley fill was from a 2003 coverage developed by WVDEP.

Table C-1. Publicly available GIS data used to generate land cover estimates

Data information	Data description	Source
General sources of land use/land cover information		
West Virginia GIS Technical Center	General West Virginia Universities GIS data repository location	http://wvgis.wvu.edu/data/data.php (WVGISTC, 2011)
WVDEP GIS data sets	General WVDEP's GIS data repository location	http://gis.dep.wv.gov/ (WVDEP, 2011a)
Base Land use/land cover		
GAP	GAP land use	http://wvgis.wvu.edu/data/dataset.php?ID=62 (WVGISTC, 2002)
NLCD 2001	NLCD land use	http://wvgis.wvu.edu/data/dataset.php?ID=269 (WVGISTC, 2001)
Other files		
Watershed Boundary Data sets	USGS 8-digit Hydrologic Unit Code boundaries	http://wvgis.wvu.edu/data/dataset.php?ID=123 (WVGISTC, 2004)
NHD Streams	National Hydrography Data set Streams	http://wvgis.wvu.edu/data/dataset.php?ID=235 (WVGISTC, 2010)
Abandoned Mine Lines (AML-Highwalls) and Polygons (AML Areas)	West Virginia abandoned mine lands coverages. Highwall mine coverage and AML area	http://wvgis.wvu.edu/data/dataset.php?ID=150 (WVGISTC, 1996)
DMR Mining NPDES Permits and Outlets	WVDEP Office of Mining and Reclamation NPDES permit and outlet coverages	http://gis.dep.wv.gov/data/omr.html (WVDEP, 2011b)
Mining Related Fills, Southern West Virginia	WVDEP valley fills coverage from 2003	http://gis.dep.wv.gov/data/omr.html (WVDEP, 2011c)
Mining Permit Boundaries	WVDEP Mining permit boundaries	http://gis.dep.wv.gov/data/omr.html (WVDEP, 2011d) http://wvgis.wvu.edu/data/dataset.php?ID=149 (WVGISTC, 2011)
Roads Paved	2000 TIGER/Line GIS and WV_Roads shapefiles	http://www.census.gov/geo/www/tiger/tiger2k/tgr2000.html (U.S. Census Bureau, 2000a) http://wvgis.wvu.edu/data/data.php (WVGISTC, 2011)
Roads Unpaved	2000 TIGER/Line GIS shapefile and digitized from aerial photographs and topographic maps	http://www.census.gov/geo/www/tiger/tiger2k/tgr2000.html (U.S. Census Bureau, 2000b) http://wvgis.wvu.edu/data/data.php (WVGISTC, 2011)

GAP = Gap Analysis Program; GIS = geographic information system; NHD = National Hydrography Data Set; NLCD = National Land Cover Database; NPDES = National Pollutant Discharge Elimination System; DMR = Division of Mining Reclamation; USGS = U.S. Geological Survey; WVDEP = West Virginia Department of Environmental Protection.

Table C-2. Detailed WV TMDL land use category derivation and land use derivation used in Appendix A. Base land use categories are highlighted in grey.

Detailed WV TMDL land use category	Data source	Base land use from which new source area was subtracted	Land use categories used in scatter plots in Appendix A
Water	Water—base LU coverage	N/A	Water
Wetland	Wetland—base LU coverage	N/A	Water
Forest	Forest—consolidated all forested types from base LU coverage	N/A	Forest
Grassland	Grassland—base LU coverage	N/A	Agriculture
Cropland	Cropland—consolidated all cropland types from base LU coverage	N/A	Agriculture
Urban pervious	Urban—consolidated urbanized types from base LU coverage	N/A	Urban/residential
Urban impervious	Urban—consolidated urbanized types from base LU coverage	N/A	Urban/residential
Barren	Barren—base LU coverage	N/A	Barren
Pasture	Source tracking	New area subtracted from Grassland	Agriculture
Paved roads	Roads shapefiles	New area subtracted from Urban Impervious	Urban/residential
Unpaved roads	Roads shapefiles	New area subtracted from Urban Pervious	Urban/residential
Revoked mining permits	AML information	New area subtracted from Barren	AML
Abandoned mine land	AML shapefile	New area subtracted from Barren	AML
Quarry	Mining shapefile	New area subtracted from Barren	Mining
Highwall	AML shapefile	New area subtracted from Barren	Mining
Oil and gas	Oil and Gas shapefile	New area subtracted from Barren	Mining

Table C-2. Detailed WV TMDL land use category derivation and land use derivation used in Appendix A (continued)

Detailed WV TMDL land use category	Data source	Base land use from which new source area was subtracted	Land use categories used in scatter plots in Appendix A
Surface Mine Water Quality permits	Mining shapefile	New area subtracted from Barren	Mining
Surface Mine Technology permits	Mining shapefile	New area subtracted from Barren	Mining
Comingled mine deep ground gravity discharge	Mining shapefile	New area subtracted from Barren	Mining
Comingled mine deep ground pump discharge	Mining shapefile	New area subtracted from Barren	Mining
Undeveloped surface mine WQ permits	Mining shapefile	New area subtracted from Forest	Mining
Undeveloped surface mine technology permits	Mining shapefile	New area subtracted from Forest	Mining
Undeveloped comingled mine gravity discharge	Mining shapefile	New area subtracted from Forest	Mining
Undeveloped comingled mine pump discharge	Mining shapefile	New area subtracted from Forest	Mining
Burned Forest	Forestry Dept. information	New area subtracted from Forest	Barren
Harvested Forest	Forestry Dept. information	New area subtracted from Forest	Barren
Skid Roads	Forestry Dept. information	New area subtracted from Forest	Barren
TMDL land use considers Valley Fill ^a area as part of the Surface Mine Water Quality and Technology Permit information	WVDEP valley fills coverage from 2003	New area subtracted from Mining, Barren, and Forest, as appropriate	Valley fill

^aValley fill land use was not part of the base TMDL land use and was specifically incorporated into the detailed land use analysis for this EPA report. See Table 1 for the source file.

AML = Abandoned Mine Line, LU = Land use, TMDL = total maximum daily load, WQ = water quality, WV = West Virginia.

C.2.3. Selection of Catchments

Catchments with available data that met the needs of the analysis involved a six-step selection process that resulted in 190 catchments. The steps were performed in the following sequence:

- Select all WVDEP WAB stations located within Ecoregion 69D. This generated 2,151 stations.
- Select stations where a macroinvertebrate sample was collected and identified to the genus level. During this selection process, stations had to have both a WVSCI and a GLIMPSS score. At least one chemistry sample was required to be associated with the macroinvertebrate sample from the same station location. This narrowed the available stations to 825.
- Select stations with detailed TMDL-associated land use located within the Coal, Upper Kanawha, Gauley, and New River watersheds. This narrowed the selection to 382 stations.
- Eliminate stations if the detailed land use was not created during the TMDL process. This eliminated 38 stations for a total of 344 stations.
- Eliminate stations located on undelineated tributary streams contained within a larger mainstem subwatershed. This eliminated an additional 33 stations for a total of 311 stations.
- Select stations with a total watershed drainage area <20 km² (4,942.08 acres). The total number of remaining stations in TMDL watersheds within Ecoregion 69D after this last reduction was 190 (see Figure C-1), and the data from these stations were assembled from 1997 to 2007, with the majority of samples collected from 2001 to 2006.

C.2.4. Land Use Analysis

To create the land use for the 190 stations, the original TMDL land uses from the Coal, Upper Kanawha, Gauley, and New Rivers were used as the starting point. These land uses were originally created by consolidating the available base land use (Gap Analysis Program [GAP] 2000 or National Land Cover Data [NLCD] 2001) into more general categories and then adding more detailed source land use categories (e.g., mining, oil and gas, roads) from detailed source information. To add these new land use categories, GIS shapefiles were used to locate sources and assign areas. These areas were then subtracted from the category they most likely would be attributed to in the original base land use. For example, a disturbed mine site would likely be classified as barren in GAP, so any area assigned as mining would be subtracted from barren to keep the total land use area in the watershed the same. Table C-2 contains the WVDEP TMDL

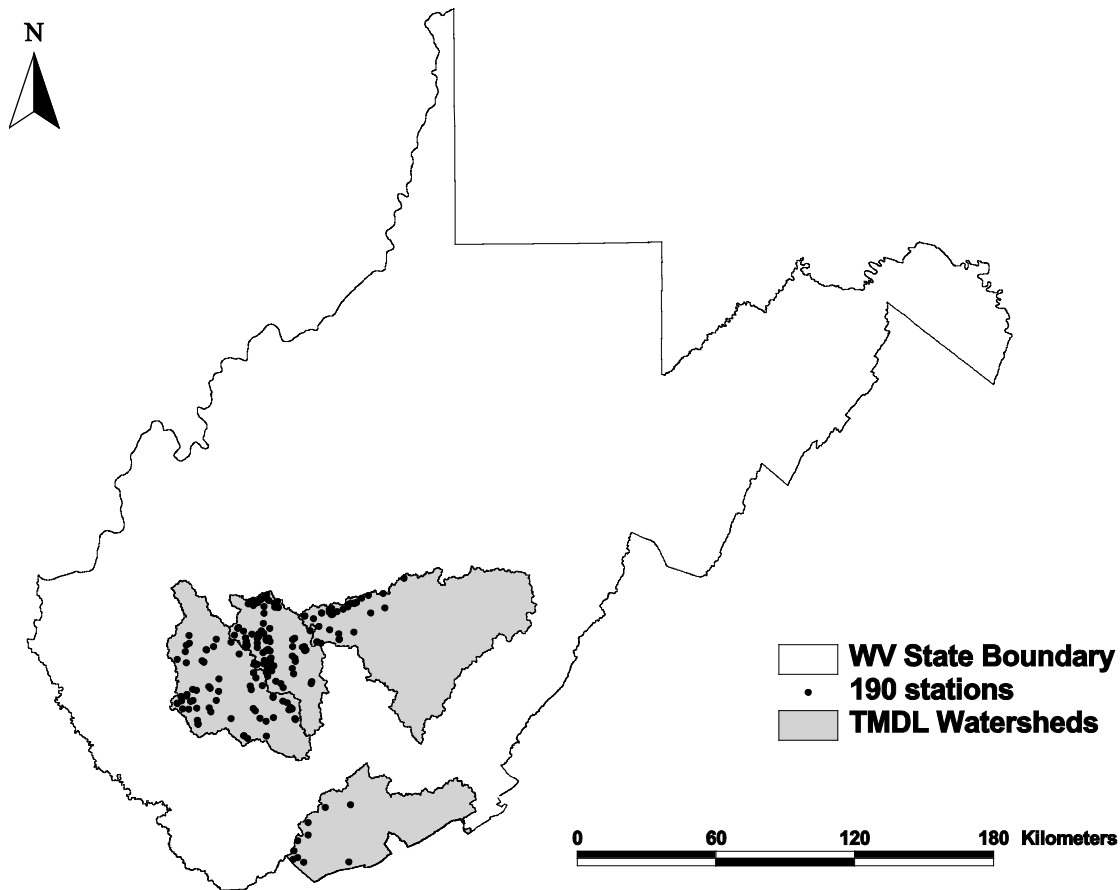


Figure C-1. Sampling locations used to develop evidence of sources of high conductivity inputs. The 190 stations (black dots) at the terminus of each >20-km² catchment are shown within the larger 8-digit HUCs in southwestern West Virginia.

land use categories, the data source from which the extent of the area and its location were determined, and the base land use from which any newly created land use categories were subtracted. In brief, nine land use categories were generated: total percentage area in mining (% Total Mining) which is the sum of % Abandoned Mine, % MTM-Valley Fill and % Mining; percentage in mountaintop mining valley fill (% MTM-Valley Fill); percentage of abandoned mine lands (% Abandoned Mine); percentage of mining (% Mining) excluding % MTM-Valley Fill and % Abandoned Mine; percentage barren land use (% Barren); percentage of residences, buildings, and roads (% Urban/residential); percentage in agriculture and pasture (% Agricultural); percentage in forest (% Forest), and percentage in open water (% Water).

Because the WVDEP TMDL land use manipulation process has undergone revisions and enhancements since the initiation of the TMDL program, WVDEP TMDL land use data sets for

the Upper Kanawha, Coal, Gauley, and New Rivers were manipulated to have equivalent land use when necessary and resulted in the consolidated land use for the 190 sampling stations. The land use representation used for more recently developed TMDLs is more detailed than that for TMDLs completed in earlier efforts. Therefore, consolidation of the detailed TMDL land use to seven basic land use categories was necessary. The valley fill GIS coverage was then incorporated into the TMDL land use by subtracting the valley fill acreage from Shank (2004) from the mining land use category. If more area was present in the valley fill coverage than was present in the TMDL mining area for each TMDL subwatershed, the remainder was subtracted from barren and then forest, respectively. The eight land use categories calculated for each of the 190 WAB sampling stations used seven categories consolidated from the TMDL land use (see Table C-2) and then included the addition of the valley fill area. The % Total Mining category is simply the sum of the % Mining, % MTM-Valley Fill, and % Abandoned Mine land categories. The % Mining land use represents all other types of mining activities except for abandoned mines and valley fill areas.

C.3. RESULTS

C.3.1. Characterization of Catchments and Ionic Matrix

The 190 small catchments used in the analysis are located near the borders of the 8-digit hydrologic unit codes (HUCS) where elevations are greater and headwaters of these small perennial streams are located (see Figure C-1). The ionic composition of these waters is not uniform, but bicarbonate and sulfate are usually greater than chloride (see Table C-3) (see also Table 1 and Table A-16, Pond et al., 2008). Because we were interested in all ions as well as the mixture, we did not exclude high Cl^- sites. Only one site, New West Hollow, had a conductivity measurement $>300 \mu\text{S}/\text{cm}$ and higher chloride (629 mg/L) than sulfate (89 mg/L). That watershed had the greatest area in residences, 16.4% urban, and a conductivity of 2,767 $\mu\text{S}/\text{cm}$. The potential presence of methane coal brine production was not ruled out.

C.3.2. Correlations with In-stream Biological and Water Quality Parameters

Pairs of land use and water quality parameters are listed in Table C-4 with at least one Pearson's correlation coefficient with an $r > |0.50|$, except for a few with spurious points or composed of only two points.

The two land use types that are most strongly and positively correlated with conductivity are percentage of mining and percentage of valley fill. Percentage of forest is negatively correlated with ion concentrations. Percentage of residential land use is not well correlated, and in this region, is somewhat confounded by mining land uses. Among the ions that are more strongly correlated, are total calcium and magnesium, also captured together as hardness,

Table C-3. Summary statistics of water quality parameters in the 190 catchments

Parameter	Units	Min	25 th centile	Median	75 th centile	Max	Mean	Valid N
Conductivity	µS/cm	6	254	474	851	3,964	445	1,671
Fecal	counts/ml	1	4	42	330	60,000	47.5	1,181
Alkalinity	mg/L	0.02	14.45	46	99.15	710	36.73	1,348
Hardness	mg/L	11.26	37.95	84.24	235.57	862.6	85.11	48
Sulfate	mg/L	5	84.55	192	358	2915	168.32	1,350
Chloride	mg/L	1	2.0	3.8	12.3	629	5.5	45
TSS	mg/L	0.3	3	5	8	1217	5.97	1,348
Al, total	mg/L	0.02	0.06	0.15	0.59	23.6	0.21	1,342
Al, dissolved	mg/L	0.02	0.02	0.05	0.07	23.5	0.063	1,335
Ca, total	mg/L	1.93	7.63	22.5	49.43	184	17.716	50
Cu, total	mg/L	0.001	0.003	0.004	0.005	0.014	0.004	24
Cu, dissolved	mg/L	0.001	0.003	0.003	0.004	1.91	0.004	40
Fe, total	mg/L	0.02	0.09	0.21	0.51	32.8	0.24	1,341
Fe, dissolved	mg/L	0.02	0.02	0.03	0.07	13.1	0.045	1,329
Mg, total	mg/L	1.28	4.3	8.0	26.3	97.9	9.97	49
Mn, total	mg/L	0.003	0.025	0.10	0.40	27.3	0.116	1,340
Se, total	mg/L	0.001	0.005	0.005	0.005	1.26	0.005	436
Se, dissolved	mg/L	0.001	0.001	0.003	0.006	1.26	0.003	23
Zn, total	mg/L	0.005	0.009	0.01	0.021	0.18	0.014	25
Zn, dissolved	mg/L	0.005	0.005	0.005	0.01	0.726	0.009	40
Flow	ft ³ /s	0.004	0.41	1.45	4.545	63.01	1.25	839
Temperature	°C	0.05	8.705	12.65	17.77	30.72	13.139	1,672
pH	standard units	3.03	7.105	7.6	7.97	12.99	7.355	1,671
DO	mg/L	1.22	9.26	10.43	11.75	11.81	11.23	1,666

TSS = Total suspended solids, Mean is geometric mean except for temperature, pH, and DO =dissolved oxygen.

Table C-4. Correlation coefficients between pairs of land use and water quality parameters in the land use data set

Water quality parameter	% Valley fill	% Total mining	% Mining	% Forest
Conductivity	0.65	0.52	0.39	-0.54
Alkalinity	0.51	0.49	0.37	-0.51
Hardness	0.69	0.63	0.55	-0.63
Sulfate	0.64	0.52	0.39	-0.53
Calcium Total	0.67	0.61	0.52	-0.64
Magnesium Total	0.66	0.65	0.58	-0.59

Parameters yielding only $r < |0.50$ are not shown.

bicarbonate measured as alkalinity, and sulfate. Noticeably chloride is not strongly correlated, owing to fewer measurements of chloride, but also due to the low concentrations except at one site. Chloride was 629 mg/L chloride at the site with the greatest residential and mining land uses.

Individual scatter plots and associated correlation coefficients for conductivity can be found in Appendix A, Section 2.2.2.4 but are reproduced here for the convenience of the reader (see Figure C-2). At relatively low urban land use, the range of conductivity is highly variable. In contrast, there is a clear pattern of increasing conductivity as percentage of area in valley fill increases and of decreasing conductivity with increasing forest cover. When area in valley fill is subtracted from the total nonacid mining area, the correlation decreases by 25% (see Figure C-2d). The scatter plots illustrate that there are clear sources of increased conductivity, but that percentage area in valley fill has the strongest correlation with conductivity ($r = 0.65$), and percentage mining without a valley fill has a moderate correlation ($r = 0.39$).

Assuming that the lower conductivity values represent current best practices, we modeled the lower 25th quantile of the percent valley fill scatter plot (see Figure C-3).

From the 10th quantile regression, the intercept for 300 $\mu\text{S}/\text{cm}$ is 4% valley fill and the intercept for 500 $\mu\text{S}/\text{cm}$ is 8% valley fill. Using logistic regression at 300 and 500 $\mu\text{S}/\text{cm}$, the probability of impairment, based on a WVSCI score ≤ 68 , is around 0.59 and 0.72, respectively. At 300 $\mu\text{S}/\text{cm}$, 5% of genera are extirpated, and at 500 $\mu\text{S}/\text{cm}$, 17% of genera are extirpated (see Figure 9). Because these estimates do not take into account the volume of the fill, construction practices, distance from the fill, or dilution from tributaries, the estimate of conductivity associated with percent valley fill is useful as a general characterization but will vary for specific cases.

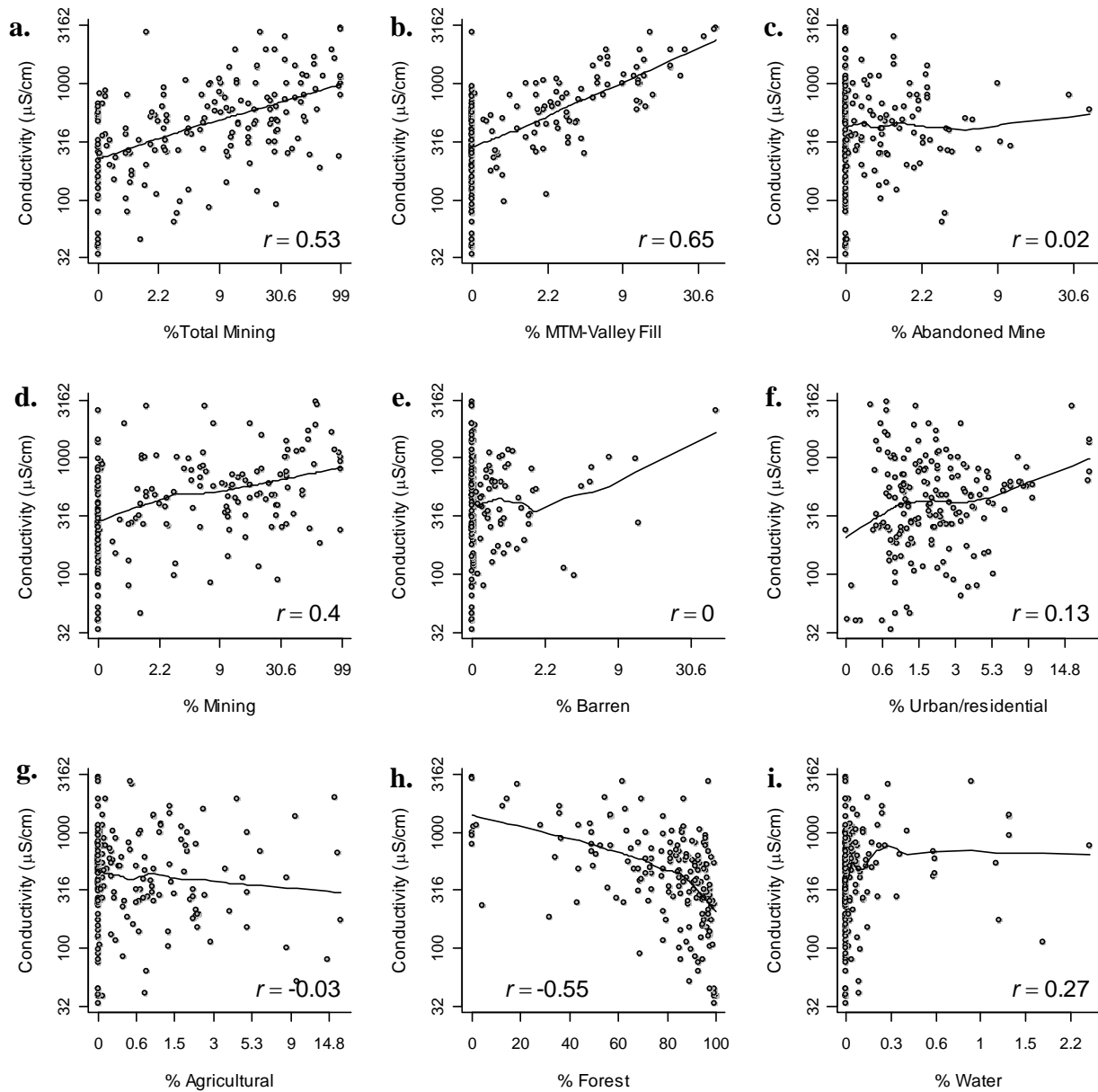


Figure C-2. Geometric mean conductivity associated with different land uses in 190 watersheds in Ecoregion 69D and Spearman's correlation coefficient.

Conductivity increases with increasing % MTM-Valley Fill and % Total Mining, and decreases with increasing % Forest, but there is less clear or no pattern with other land use. From left to right, they are (a) % Total Mining (percentage of deep, surface, quarry mining, MTM-Valley Fill, and abandoned mine land), (b) % MTM-Valley Fill (from mountaintop mining overburden), (c) % Abandoned Mine, (d) % Mining (inclusive of all types of mining except MTM-Valley Fill and Abandoned Mine), (e) % Barren, (f) % Urban/residential, (g) % Agricultural, (h) % Forest, and (i) % Water. Fitted LOWESS line with span set at 2/3.

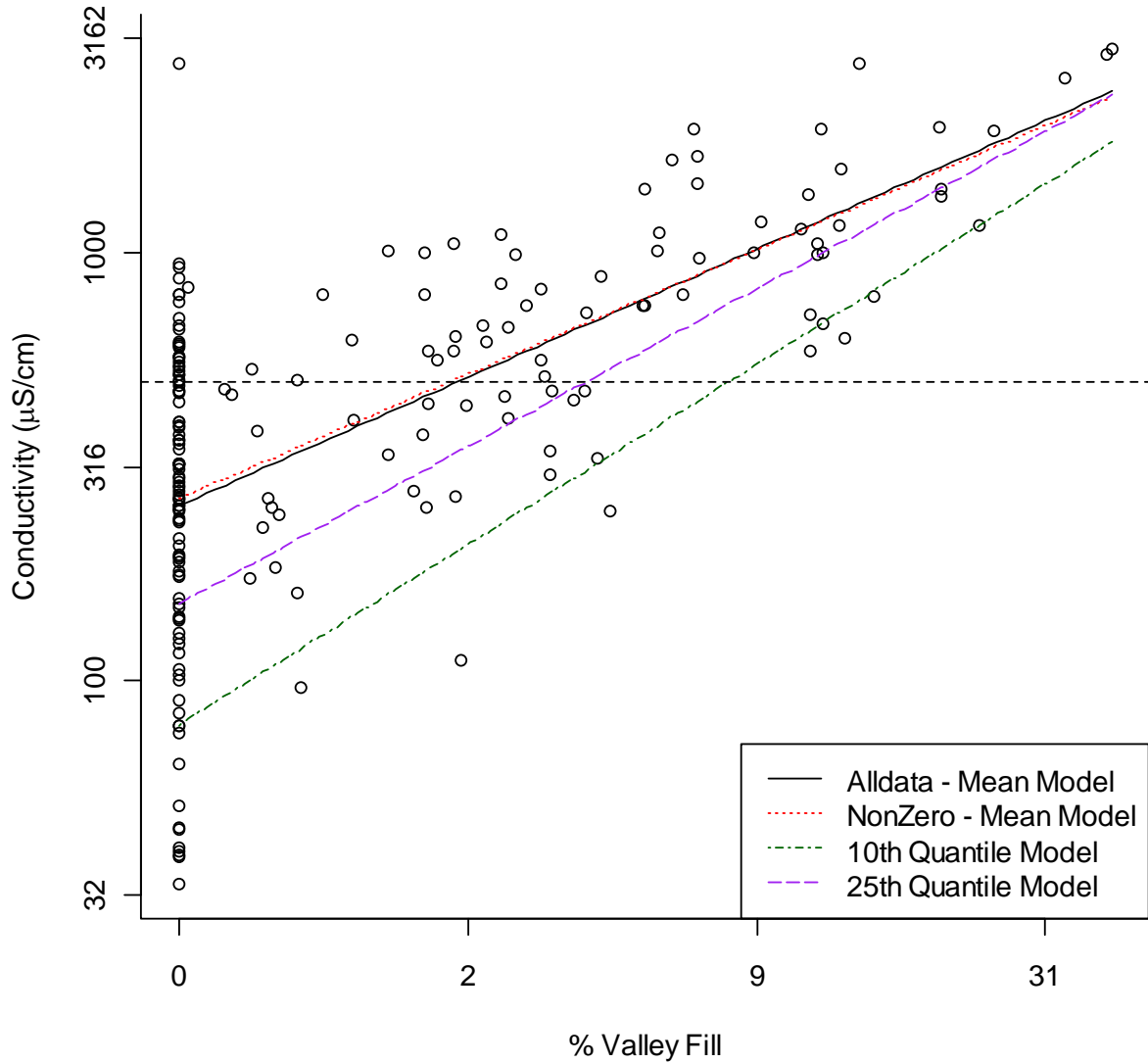


Figure C-3. Quantile regression of percentage of area in valley fill and conductivity in 190 small watersheds in Ecoregion 69D. Assuming the lowest conductivity points represent some of the best fill construction practices, the 10th and 25th quantile regression lines are shown. The intercepts for 500 $\mu\text{S}/\text{cm}$ (horizontal dashed line) are approximately 4% and 8% valley fill and for 300 $\mu\text{S}/\text{cm}$ are 1.5% and 3.9% valley fill for the 10th and 25th quantiles, respectively. The mean model based on samples minus those with zero percent valley fill shows that the relationship is unaffected by the removal of sites without valley fills.

C.4. CONCLUSIONS

Of the land uses in the small watersheds analyzed, only mining associated with valley fills are significant sources of the salts that are measured as conductivity. Disturbances associated with agriculture and human habitation may also contribute, but the densities of agricultural and urban land cover are relatively low, and a clear pattern of increasing conductivity and increasing land use is not evident. Furthermore, natural background is exceedingly low. For Ecoregion 69, the 25th centile from a probability-based sample from the WABbase data set was 72 $\mu\text{S}/\text{cm}$, $N = 617$ (see Section 5.5).

Although conductivity typically increases with increasing land use (Herlihy et al., 1998), conductivity is highly variable at relatively low urban land use. This may be caused by unknown mine drainage, deep mine break-outs, road applications, poor infrastructure condition (e.g., leaking sewers or combined sewers), gas drilling, or other practices. In contrast, there is a clear pattern of increasing conductivity as the percentage of valley fill area increases and decreasing conductivity with increasing percentage of forest cover area. This is evidence of at least one strong source of high conductivity in the region (see Appendix A for causal assessment).

REFERENCES

- Herlihy, AT; Stoddard, JL; Johnson, CB. (1998) The relationship between stream chemistry and watershed land cover data in the mid-Atlantic region, U.S. *Water, Air Soil Pollut* 105(1–2):377-386.
- Pond, GJ; Passmore, ME; Borsuk, FA; et al. (2008) Downstream effects of mountaintop coal mining: comparing biological conditions using family- and genus-level macroinvertebrate bioassessment tools. *J N Am Benthol Soc* 27(3): 717–737.
- Shank, M. (2004). Development of a mining fill inventory from multi-date elevation data. Presented at the Advanced Integration of Geospatial Technologies in Mining and Reclamation Conference, December 7-9, 2004, Atlanta, GA.
- U.S. Census Bureau. (2000a) 2000 TIGER/Line GIS and WV roads shapefiles. Geography division. Available online at <http://www.census.gov/geo/www/tiger/tiger2k/tgr2000.html>.
- U.S. Census Bureau. (2000b). 2000 TIGER/Line GIS shapefiles aerial photographs and topographic maps. Geography division. Available online at <http://www.census.gov/geo/www/tiger/tiger2k/tgr2000.html>.
- WVDEP (West Virginia Department of Environmental Protection). (2011a) General WVDEP's data repository locations. WVDEP GIS data sets. State of West Virginia, Charleston, WV. Available online at <http://gis.dep.wv.gov/>.
- WVDEP (West Virginia Department of Environmental Protection). (2011b). Mining permits, point locations. Division of Mining and Reclamation Data. State of West Virginia, Charleston, WV. Available online at <http://gis.dep.wv.gov/data/omr.html>.
- WVDEP (West Virginia Department of Environmental Protection). (2011c) Mining related fills, Southern West Virginia as of September 2003. Division of Mining and Reclamation Data. State of West Virginia, Charleston, WV. Available online at <http://gis.dep.wv.gov/data/omr.html>.

WVDEP (West Virginia Department of Environmental Protection). (2011d) Mining permit boundaries. Division of Mining and Reclamation Data. State of West Virginia, Charleston, WV. Available online at <http://gis.dep.wv.gov/data/omr.html>.

WVGISTC (West Virginia GIS Technical Center). (1996) Mining – abandoned mine lands. University of West Virginia, Morgantown, WV. Available online at <http://wvgis.wvu.edu/data/dataset.php?ID=150>.

WVGISTC (West Virginia GIS Technical Center). (2001) Land cover (NLCD 2001). WV NLCD (National land cover database). University of West Virginia, Morgantown, WV. Available online at <http://wvgis.wvu.edu/data/dataset.php?ID=269>.

WVGISTC (West Virginia GIS Technical Center). (2002) Land cover (GAP). WV GAP analysis program. University of West Virginia, Morgantown, WV. Available online at <http://wvgis.wvu.edu/data/dataset.php?ID=62>.

WVGISTC (West Virginia GIS Technical Center). (2004) Watershed boundary data sets (8, 10, and 12 digits). University of West Virginia, Morgantown, WV. Available online at <http://wvgis.wvu.edu/data/dataset.php?ID=123>.

WVGISTC (West Virginia GIS Technical Center). (2010) Streams – national hydrography data sets (24k). University of West Virginia, Morgantown, WV. Available online at <http://wvgis.wvu.edu/data/dataset.php?ID=235>.

WVGISTC (West Virginia GIS Technical Center). (2011) General West Virginia Universities GIS data repository locations. University of West Virginia, Morgantown, WV. Available online at <http://wvgis.wvu.edu/data/data.php>.

APPENDIX D
EXTIRPATION CONCENTRATION VALUES FOR
GENERA IN THE WEST VIRGINIA DATA SET

ABSTRACT

The purpose of Appendix D is to provide the reader with a list of the extirpation concentration (XC_{95}) values used to develop the species sensitivity distribution and the hazardous concentration (HC_{05}). Genera are ordered alphabetically (see Table D-1). The numbers of occurrences in the data set and at West Virginia Department of Environmental Protection (WVDEP) reference sites are noted in the right-hand columns.

Not all 95th centiles correspond to extirpation, and some imprecisely estimate the extirpation threshold. The following rules were applied to the XC_{95} values using the fitted curve and the confidence bounds from the plots in Appendix E. If the generalized additive model (GAM) mean curve at maximum conductivity is approximately equal to 0 (defined as less than 1% of the maximum modeled probability), then the XC_{95} value is listed without qualification. If the GAM mean curve at maximum conductivity is >0 but the lower confidence limit is approximating to 0 ($<1\%$ of the maximum mean modeled probability), then the XC_{95} value is listed as approximate (\sim). If the GAM lower confidence limit is >0 , then the XC_{95} value is listed as greater than ($>$) the 95th centile. All model fits and scatter of points were also visually inspected for anomalies, and if the model poorly fit the data, the uncertainty level was increased to either (\sim) or ($>$).

The assignation of (\sim) and ($>$) does not affect the HC_{05} . They are provided to alert users to the uncertainty of some XC_{95} values for other uses such as comparison with toxicity test results or with results from other geographic regions.

Table D-1. Extirpation concentration and sample size from West Virginia data set. XC₉₅ values reported without a preceding symbol indicate evidence of extirpation within the tested range. XC₉₅ values preceded by a (~) or (>) indicate extirpation with greater uncertainty or extirpation at a level above the reported value.

Order	Family	Genus	XC ₉₅	N	N from reference locations
Diptera	Chironomidae	<i>Ablabesmyia</i>	>1,1646	162	5
Ephemeroptera	Baetidae	<i>Acentrella</i>	1,337	752	31
Ephemeroptera	Baetidae	<i>Acerpenna</i>	~649	27	3
Plecoptera	Perlidae	<i>Acroneuria</i>	>2,630	512	60
Trichoptera	Glossosomatidae	<i>Agapetus</i>	365	27	6
Plecoptera	Capniidae	<i>Allocapnia</i>	542	33	15
Plecoptera	Chloroperlidae	<i>Alloperla</i>	246	101	15
Ephemeroptera	Ameletidae	<i>Ameletus</i>	591	219	30
Plecoptera	Nemouridae	<i>Amphinemura</i>	812	589	42
Diptera	Culicidae	<i>Anopheles</i>	>2,768	26	2
Diptera	Tipulidae	<i>Antocha</i>	>6,468	565	18
Isopoda	Asellidae	<i>Asellus</i>	960	33	2
Diptera	Athericidae	<i>Atherix</i>	>1,1646	157	3
Diptera	Ceratopogonidae	<i>Atrichopogon</i>	>2,257	43	3
Ephemeroptera	Ephemerellidae	<i>Attenella</i>	~698	34	1
Ephemeroptera	Baetidae	<i>Baetis</i>	>1,395	1527	71
Diptera	Ceratopogonidae	<i>Bezzia</i>	380	62	2
Odonata	Aeshnidae	<i>Boyeria</i>	>7,340	175	5
Diptera	Tipulidae	<i>Brachypremna</i>	408	27	2
Diptera	Chironomidae	<i>Brillia</i>	>2,005	95	6
Isopoda	Asellidae	<i>Caecidotea</i>	>4,713	141	1
Ephemeroptera	Caenidae	<i>Caenis</i>	>3,923	552	8
Decapoda	Cambaridae	<i>Cambarus</i>	>1,274	472	44
Diptera	Chironomidae	<i>Cardiocladius</i>	>2,257	191	2

Order	Family	Genus	XC ₉₅	N	N from reference locations
Ephemeroptera	Baetidae	<i>Centroptilum</i>	1,092	90	6
Trichoptera	Hydropsychidae	<i>Ceratopsyche</i>	>6,468	909	27
Diptera	Chironomidae	<i>Chaetocladius</i>	>5,057	184	4
Diptera	Empididae	<i>Chelifera</i>	>3,341	152	9
Trichoptera	Hydropsychidae	<i>Cheumatopsyche</i>	>9,180	1665	57
Trichoptera	Philopotamidae	<i>Chimarra</i>	>3,972	516	11
Diptera	Chironomidae	<i>Chironomus</i>	>11,646	105	1
Diptera	Tabanidae	<i>Chrysops</i>	>11,646	76	1
Ephemeroptera	Heptageniidae	<i>Cinygmula</i>	230	90	15
Diptera	Chironomidae	<i>Cladotanytarsus</i>	>11,646	104	5
Diptera	Empididae	<i>Clinocera</i>	>4,713	61	6
Diptera	Chironomidae	<i>Conchapelopia</i>	546	135	7
Odonata	Cordulegastridae	<i>Cordulegaster</i>	>1,436	43	3
Megaloptera	Corydalidae	<i>Corydalus</i>	>11,227	317	1
Diptera	Chironomidae	<i>Corynoneura</i>	>2,006	149	4
Amphipoda	Crangonyctidae	<i>Crangonyx</i>	>2,169	105	7
Diptera	Chironomidae	<i>Cricotopus</i>	>11,227	617	21
Diptera	Chironomidae	<i>Cryptochironomus</i>	>3,489	287	3
Diptera	Ceratopogonidae	<i>Dasyhelea</i>	>3,341	66	3
Diptera	Chironomidae	<i>Demicryptochironomus</i>	322	81	6
Diptera	Chironomidae	<i>Diamesa</i>	>4,713	486	14
Diptera	Tipulidae	<i>Dicranota</i>	>7,010	355	43
Diptera	Chironomidae	<i>Dicrotendipes</i>	>11,646	197	1
Ephemeroptera	Baetidae	<i>Dipheter</i>	632	148	17
Trichoptera	Hydropsychidae	<i>Diplectrona</i>	>2,527	618	59
Plecoptera	Perlodidae	<i>Diploperla</i>	315	106	2
Diptera	Dixidae	<i>Dixa</i>	>704	70	16
Trichoptera	Philopotamidae	<i>Dolophilodes</i>	>863	356	46

Order	Family	Genus	XC ₉₅	N	N from reference locations
Ephemeroptera	Ephemerellidae	<i>Drunella</i>	297	176	18
Coleoptera	Elmidae	<i>Dubiraphia</i>	>7,370	144	3
Plecoptera	Perlidae	<i>Eccoptura</i>	497	65	6
Coleoptera	Psephenidae	<i>Ectopria</i>	>1,380	324	32
Ephemeroptera	Heptageniidae	<i>Epeorus</i>	307	414	53
Ephemeroptera	Ephemeridae	<i>Ephemera</i>	696	148	20
Ephemeroptera	Ephemerellidae	<i>Ephemerella</i>	299	405	38
Diptera	Chironomidae	<i>Eukiefferiella</i>	>1,876	519	28
Ephemeroptera	Ephemerellidae	<i>Eurylophella</i>	490	189	19
Amphipoda	Gammaridae	<i>Gammarus</i>	>4,713	216	10
Trichoptera	Glossosomatidae	<i>Glossosoma</i>	>1,652	157	7
Trichoptera	Goeridae	<i>Goera</i>	~738	25	4
Plecoptera	Chloroperlidae	<i>Haploperla</i>	418	253	27
Diptera	Chironomidae	<i>Heleniella</i>	>1,697	62	7
Coleoptera	Dryopidae	<i>Helichus</i>	~11,646	333	18
Diptera	Empididae	<i>Hemerodromia</i>	>9,790	615	8
Ephemeroptera	Heptageniidae	<i>Heptagenia</i>	326	68	3
Diptera	Tipulidae	<i>Hexatoma</i>	>9,790	846	65
Coleoptera	Dytiscidae	<i>Hydroporus</i>	822	32	1
Trichoptera	Hydropsychidae	<i>Hydropsyche</i>	>7,010	999	21
Trichoptera	Hydroptilidae	<i>Hydroptila</i>	>11,227	281	4
Ephemeroptera	Isonychiidae	<i>Isonychia</i>	1,180	740	16
Plecoptera	Perlodidae	<i>Isoperla</i>	460	520	39
Diptera	Chironomidae	<i>Krenopelopia</i>	>2,320	62	2
Diptera	Chironomidae	<i>Krenosmittia</i>	~1,115	27	3
Odonata	Gomphidae	<i>Lanthus</i>	>2,087	66	7
Diptera	Chironomidae	<i>Larsia</i>	~2,630	96	3
Trichoptera	Lepidostomatidae	<i>Lepidostoma</i>	~121	91	12

Order	Family	Genus	XC ₉₅	N	N from reference locations
Ephemeroptera	Leptophlebiidae	<i>Leptophlebia</i>	251	87	8
Ephemeroptera	Heptageniidae	<i>Leucrocuta</i>	424	225	29
Plecoptera	Leuctridae	<i>Leuctra</i>	>2,087	1199	84
Diptera	Tipulidae	<i>Limnophila</i>	~1,503	54	10
Diptera	Chironomidae	<i>Limnophyes</i>	>5,120	88	1
Diptera	Tipulidae	<i>Limonia</i>	>5,057	62	1
Isopoda	Asellidae	<i>Lirceus</i>	~1,323	72	6
Ephemeroptera	Heptageniidae	<i>Maccaffertium</i>	~1,035	214	13
Coleoptera	Elmidae	<i>Macronychus</i>	>1,890	44	4
Plecoptera	Perlodidae	<i>Malirekus</i>	>904	27	6
Coleoptera	Elmidae	<i>Microcylloepus</i>	>3,341	94	2
Diptera	Chironomidae	<i>Micropsectra</i>	>6,468	227	24
Diptera	Chironomidae	<i>Microtendipes</i>	>3,489	532	33
Hemiptera	Veliidae	<i>Microvelia</i>	>2,523	46	3
Diptera	Tipulidae	<i>Molophilus</i>	~2,169	28	2
Diptera	Chironomidae	<i>Natarsia</i>	>1,842	54	1
Trichoptera	Uenoidae	<i>Neophylax</i>	316	166	35
Megaloptera	Corydalidae	<i>Nigronia</i>	>9,790	746	36
Diptera	Chironomidae	<i>Nilotanytus</i>	>2,266	112	3
Ephemeroptera	Heptageniidae	<i>Nixe</i>	319	77	3
Trichoptera	Hydroptilidae	<i>Ochrotrichia</i>	>2,791	32	1
Coleoptera	Elmidae	<i>Optioservus</i>	>9,790	1471	63
Decapoda	Cambaridae	<i>Orconectes</i>	>3,162	205	2
Diptera	Chironomidae	<i>Orthocladius</i>	>3,427	277	9
Coleoptera	Elmidae	<i>Oulimnius</i>	>2,791	227	27
Diptera	Chironomidae	<i>Pagastia</i>	>1,800	46	2
Plecoptera	Capniidae	<i>Paracapnia</i>	334	37	13
Diptera	Chironomidae	<i>Parachaetocladius</i>	>1,147	169	27

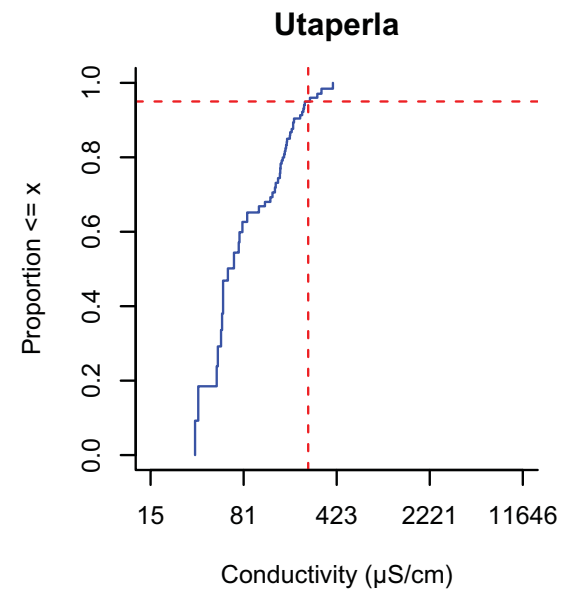
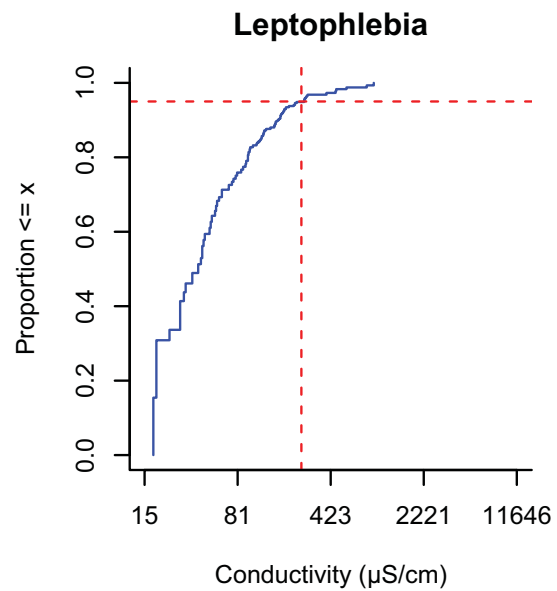
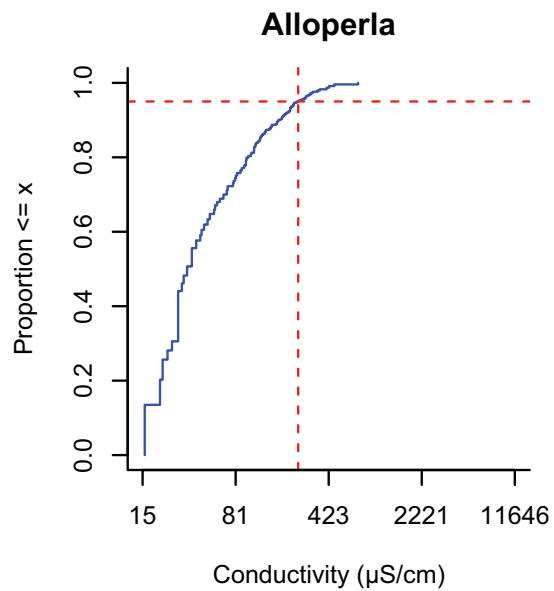
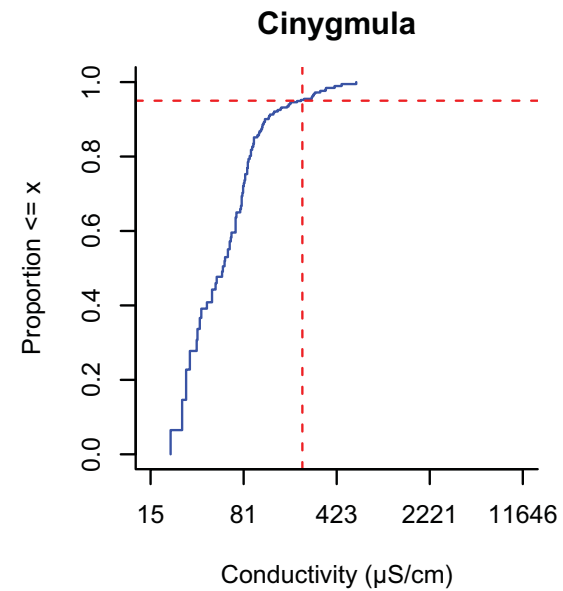
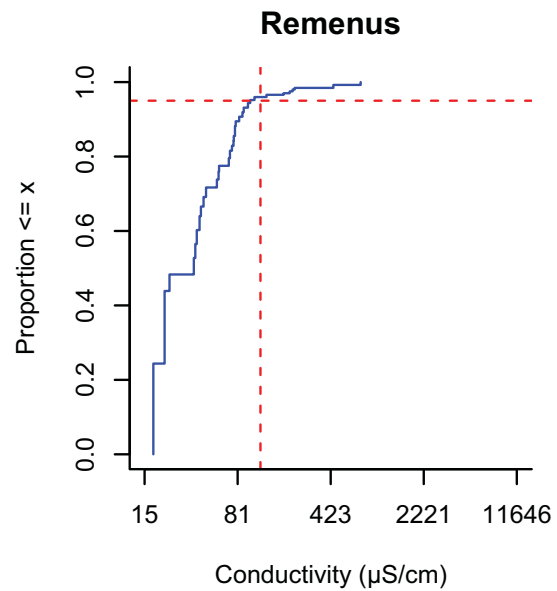
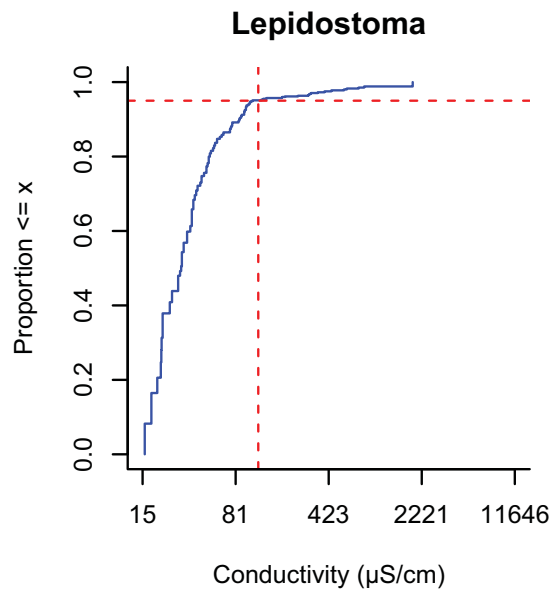
Order	Family	Genus	XC ₉₅	N	N from reference locations
Plecoptera	Perlidae	<i>Paragnetina</i>	2,087	40	3
Diptera	Chironomidae	<i>Parakiefferiella</i>	>1,757	75	2
Ephemeroptera	Leptophlebiidae	<i>Paraleptophlebia</i>	463	449	46
Diptera	Chironomidae	<i>Parametriocnemus</i>	>4,713	1501	72
Diptera	Chironomidae	<i>Paraphaenocladus</i>	>6,468	71	2
Diptera	Chironomidae	<i>Paratanytarsus</i>	>3,489	110	2
Plecoptera	Peltoperlidae	<i>Peltoperla</i>	>694	126	12
Plecoptera	Perlidae	<i>Perlesta</i>	3,314	315	8
Diptera	Chironomidae	<i>Phaenopsectra</i>	~2,332	89	1
Basommatophora	Physidae	<i>Physella</i>	>9,790	145	1
Veneroida	Pisidiidae	<i>Pisidium</i>	>1,795	34	2
Ephemeroptera	Baetidae	<i>Plauditus</i>	996	289	12
Trichoptera	Polycentropodidae	<i>Polycentropus</i>	>4,713	380	41
Diptera	Chironomidae	<i>Polypedilum</i>	>4,884	1648	70
Diptera	Chironomidae	<i>Potthastia</i>	>1,886	62	1
Diptera	Chironomidae	<i>Procladius</i>	>11,227	28	1
Ephemeroptera	Baetidae	<i>Procloeon</i>	702	78	3
Coleoptera	Elmidae	<i>Promoresia</i>	~672	79	5
Diptera	Simuliidae	<i>Prosimulium</i>	~531	106	20
Coleoptera	Psephenidae	<i>Psephenus</i>	>9,119	886	35
Diptera	Chironomidae	<i>Pseudochironomus</i>	>11,646	31	2
Diptera	Tipulidae	<i>Pseudolimmophila</i>	>1,357	135	11
Trichoptera	Psychomyiidae	<i>Psychomyia</i>	>1,131	39	3
Plecoptera	Pteronarcyidae	<i>Pteronarcys</i>	~634	113	25
Trichoptera	Limnephilidae	<i>Pycnopsyche</i>	295	44	10
Plecoptera	Perlodidae	<i>Remenus</i>	121	35	3
Hemiptera	Veliidae	<i>Rhagovelia</i>	>2,030	52	3
Diptera	Chironomidae	<i>Rheocricotopus</i>	>3,489	559	11

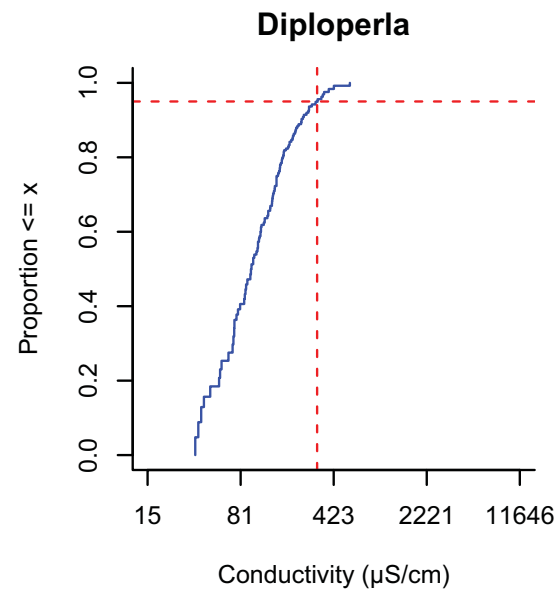
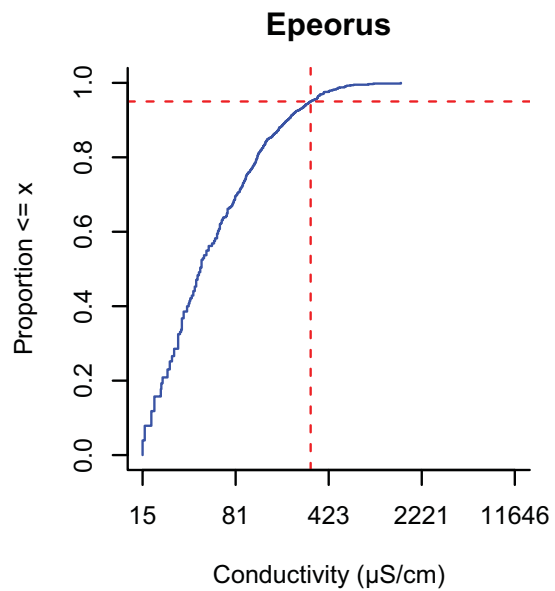
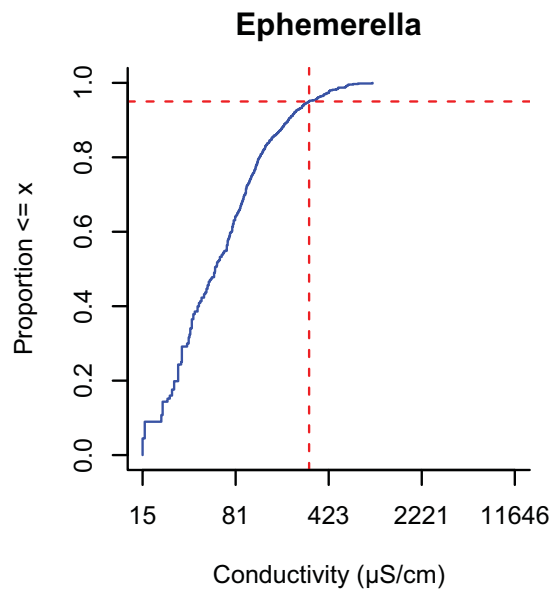
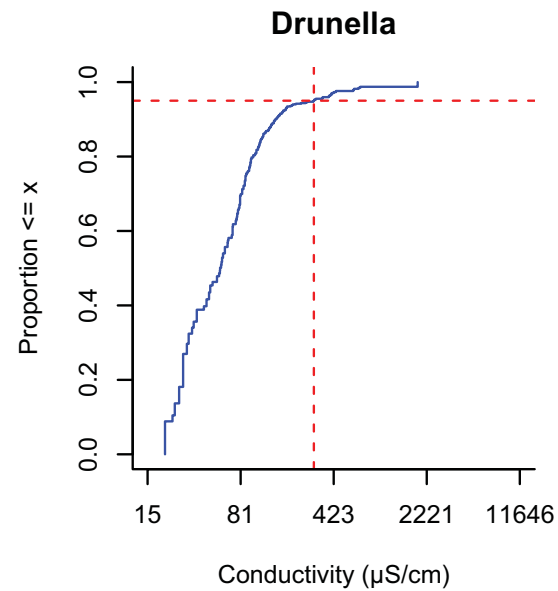
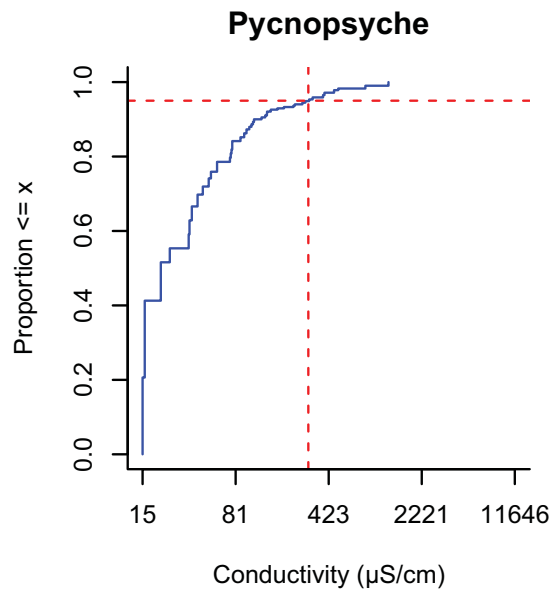
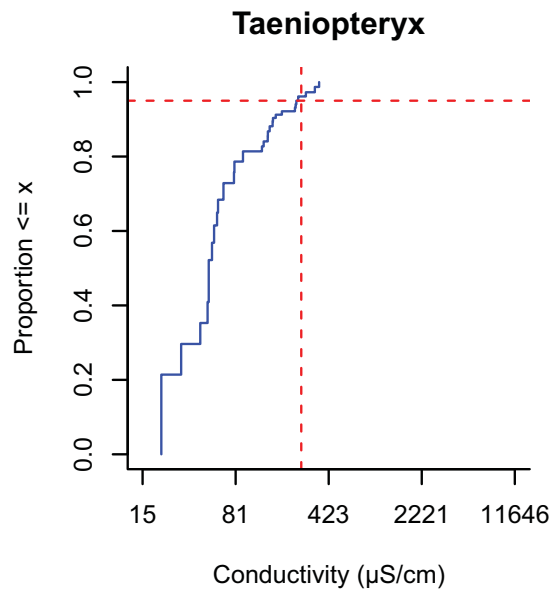
Order	Family	Genus	XC ₉₅	N	N from reference locations
Diptera	Chironomidae	<i>Rheopelopia</i>	~1,457	126	4
Diptera	Chironomidae	<i>Rheotanytarsus</i>	>3,489	949	28
Trichoptera	Rhyacophilidae	<i>Rhyacophila</i>	>1,890	415	57
Ephemeroptera	Ephemerellidae	<i>Serratella</i>	535	49	2
Megaloptera	Sialidae	<i>Sialis</i>	>11,227	264	3
Diptera	Simuliidae	<i>Simulium</i>	>6,468	1095	26
Diptera	Chironomidae	<i>Stempellina</i>	644	35	8
Diptera	Chironomidae	<i>Stempellinella</i>	>927	309	26
Ephemeroptera	Heptageniidae	<i>Stenacron</i>	~782	258	15
Coleoptera	Elmidae	<i>Stenelmis</i>	>9,790	1232	26
Ephemeroptera	Heptageniidae	<i>Stenonema</i>	745	922	57
Odonata	Gomphidae	<i>Stylogomphus</i>	>6,468	118	1
Diptera	Chironomidae	<i>Sublettea</i>	>2,421	182	2
Plecoptera	Chloroperlidae	<i>Sweltsa</i>	~750	315	42
Diptera	Tabanidae	<i>Tabanus</i>	>9,790	61	1
Plecoptera	Taeniopterygidae	<i>Taeniopteryx</i>	260	30	12
Plecoptera	Peltoperlidae	<i>Tallaperla</i>	478	89	16
Diptera	Chironomidae	<i>Tanytarsus</i>	>9,180	1232	64
Diptera	Chironomidae	<i>Thienemanniella</i>	>9,790	395	9
Diptera	Chironomidae	<i>Thienemannimyia</i>	>6,468	1345	56
Diptera	Tipulidae	<i>Tipula</i>	>1,979	621	36
Diptera	Chironomidae	<i>Tvetenia</i>	>2,613	760	40
Plecoptera	Chloroperlidae	<i>Utaperla</i>	255	47	2
Trichoptera	Philopotamidae	<i>Wormaldia</i>	>1,553	79	8
Plecoptera	Perlodidae	<i>Yugus</i>	655	75	12
Diptera	Chironomidae	<i>Zavrelia</i>	413	81	6
Diptera	Chironomidae	<i>Zavrelimyia</i>	>2,768	244	11

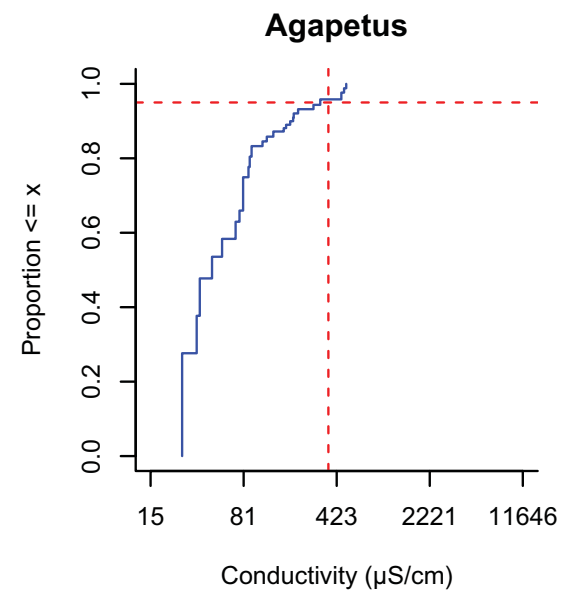
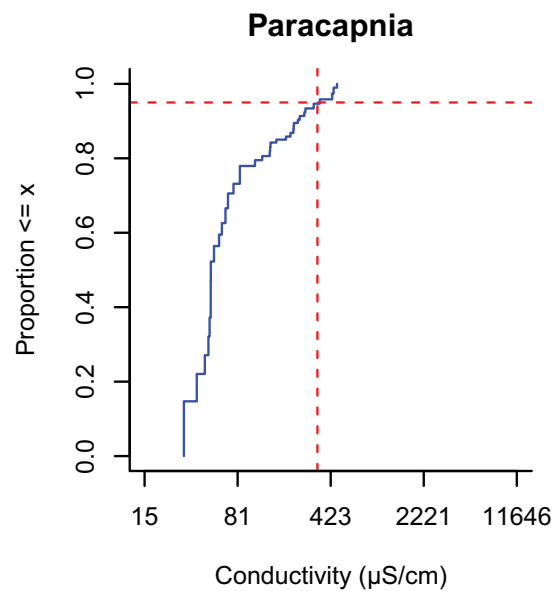
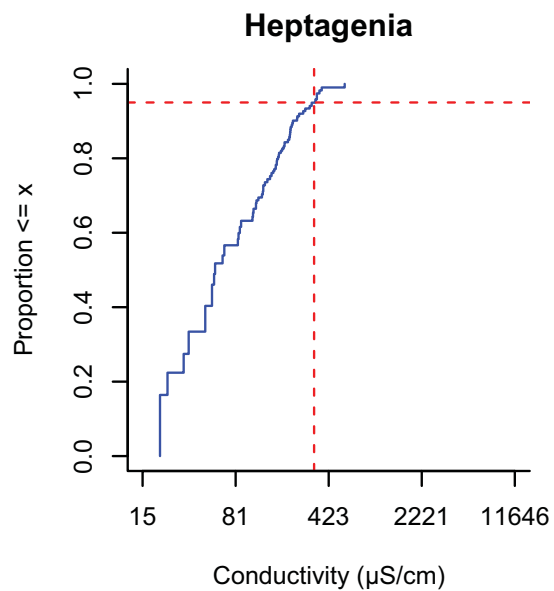
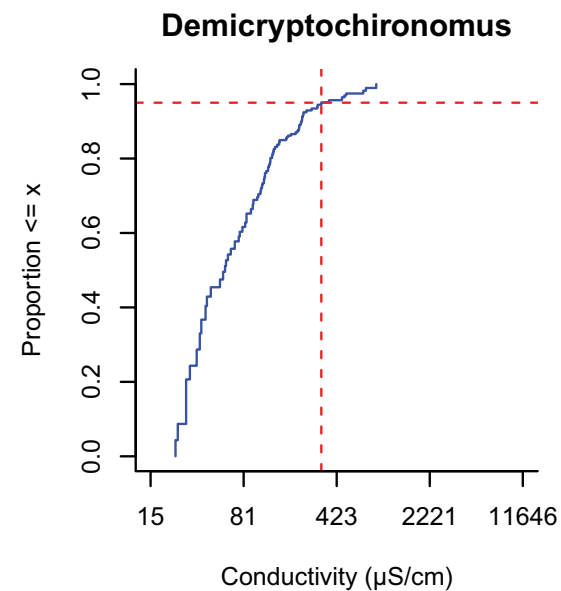
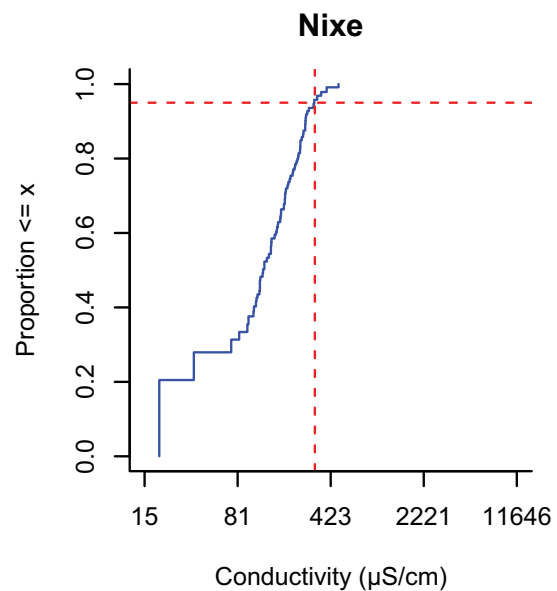
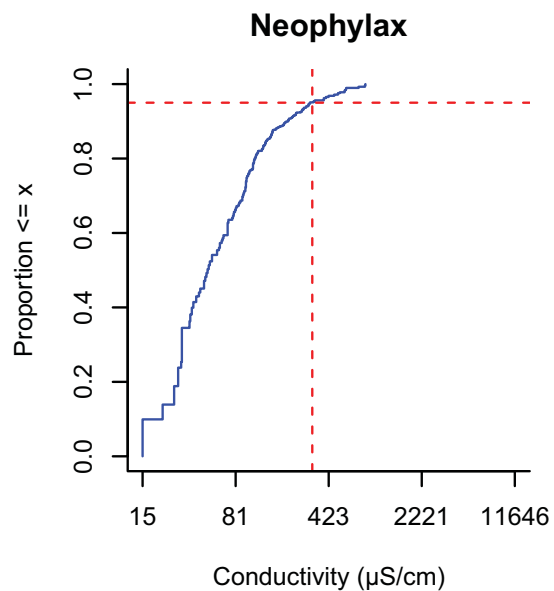
APPENDIX F
GRAPHS OF CUMULATIVE FREQUENCY DISTRIBUTIONS
FOR GENERA IN THE WEST VIRGINIA DATA SET

ABSTRACT

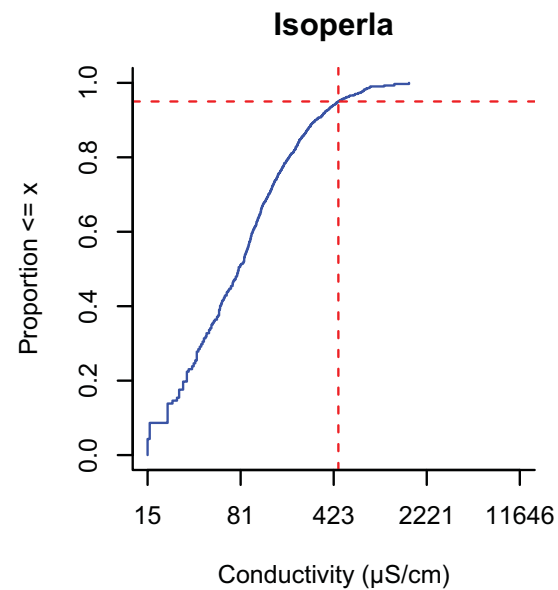
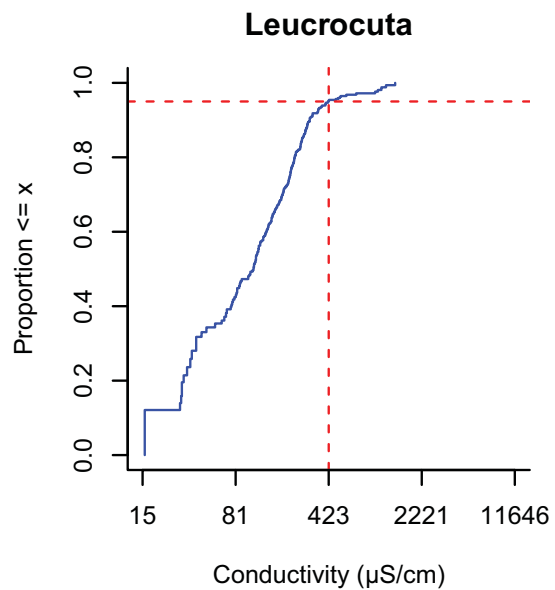
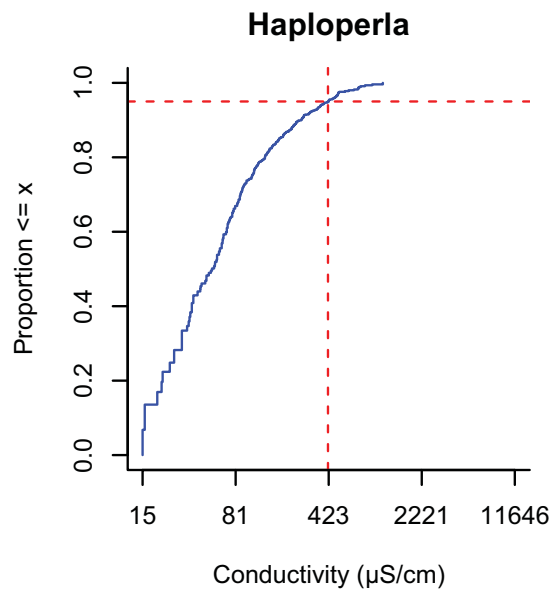
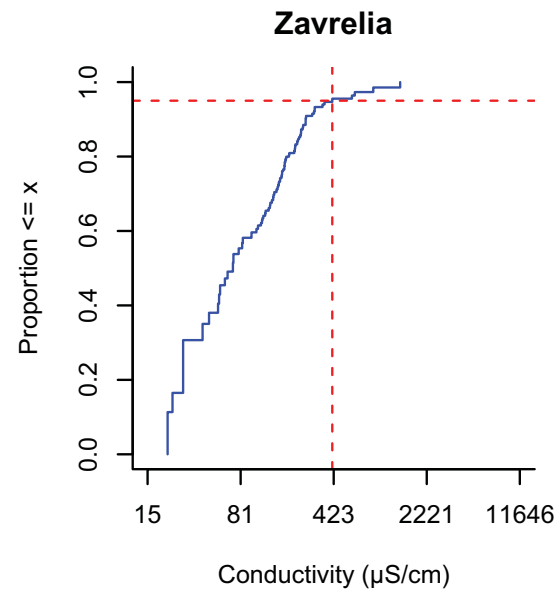
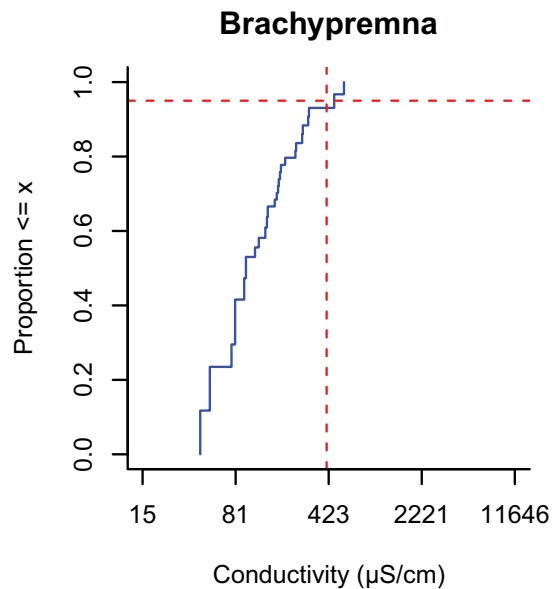
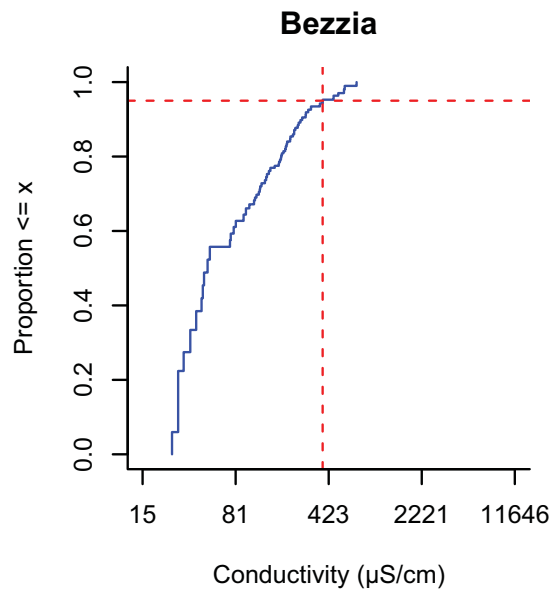
The purpose of Appendix F is to help the reader visualize the changes in the occurrence of each genus in the West Virginia data set as conductivity increases and understand how the extirpation concentration (XC_{95}) values are derived. Each plot contains the weighted cumulative distribution function (CDF) for the occurrence of a genus with respect to conductivity. For each genus, the points in the CDF represent the weighted proportions of occurrences of the genus in samples less than the indicated conductivity value ($\mu\text{S}/\text{cm}$), calculated using Equation 1. In a CDF, genera that are affected by increasing conductivity (e.g., *Drunella*) show a steep slope and asymptote well below the maximum conductivity, whereas genera unaffected by increasing conductivity (e.g., *Nigronia*) have a steady increase over the entire range of measured exposure and do not reach a perceptible asymptote. The 95th centile is found at the intersection of the dashed horizontal line with the CDF. The conductivity at the 95th centile is the XC_{95} value and is found at the intersection of the vertical line and the x-axis.

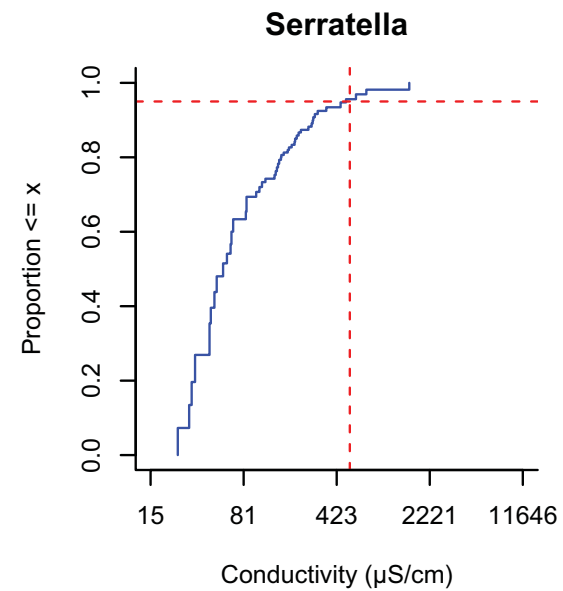
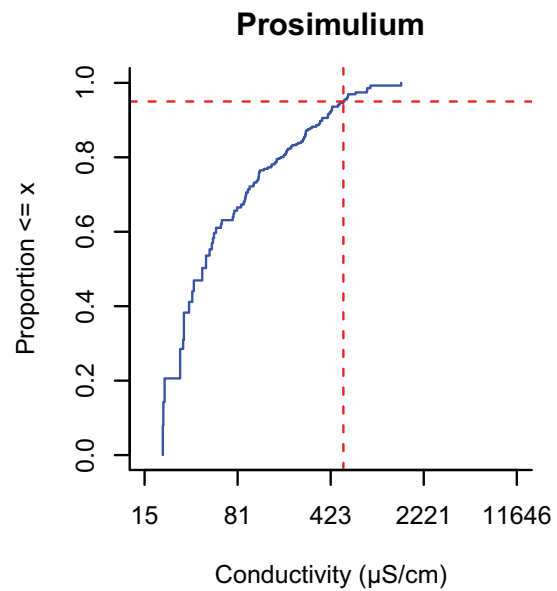
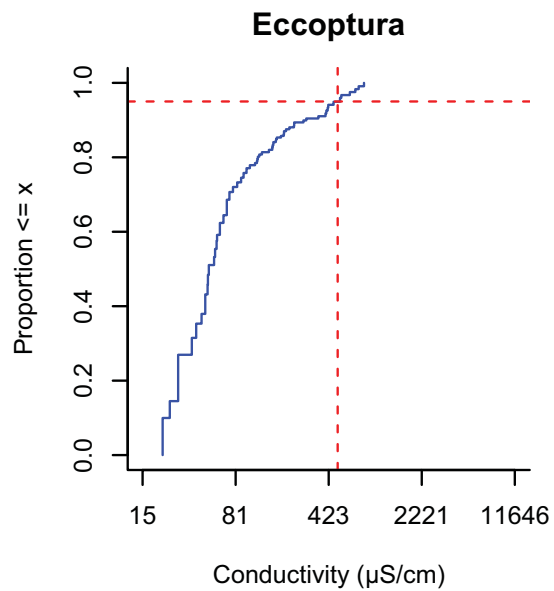
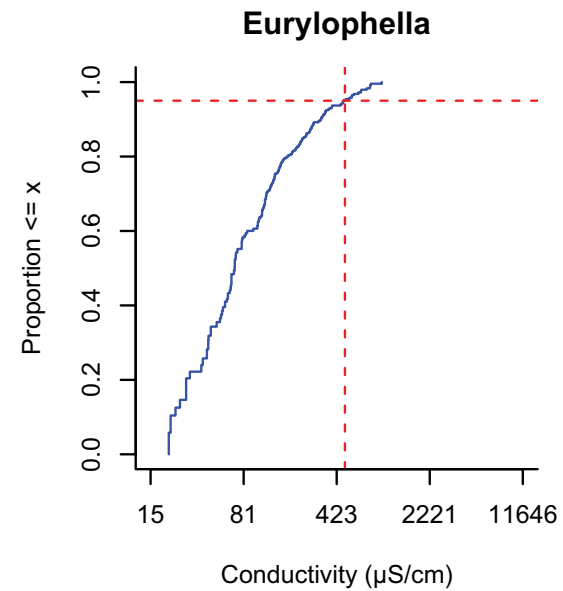
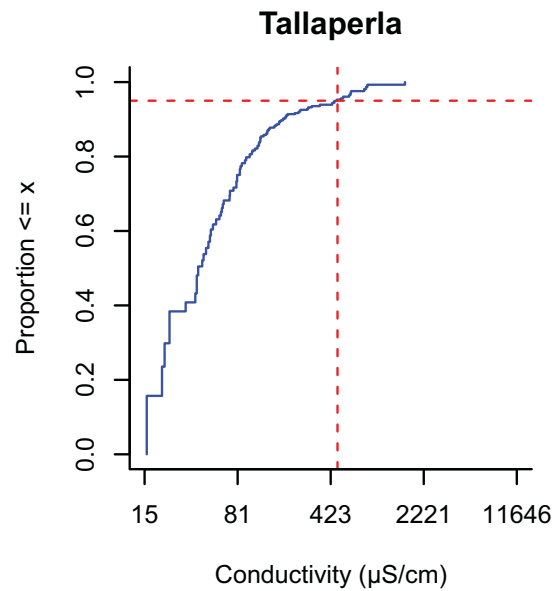
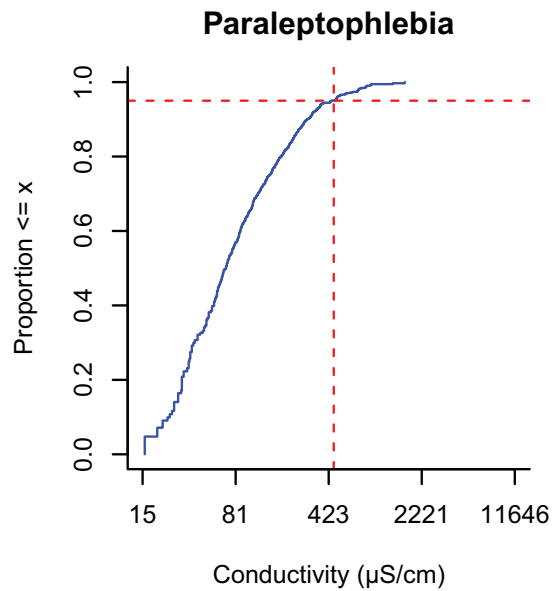


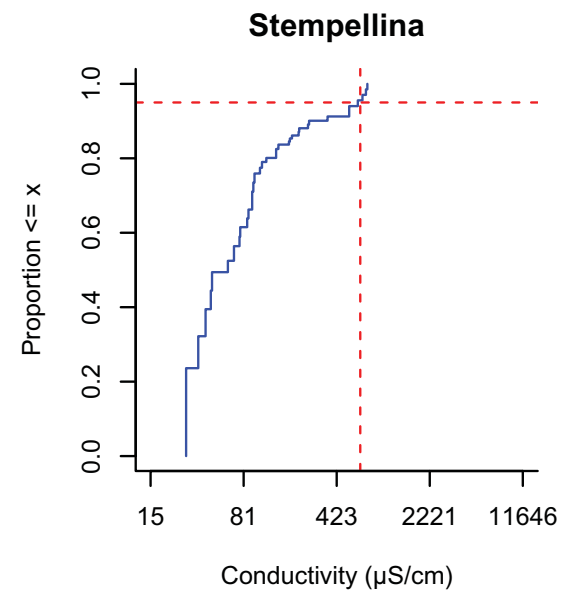
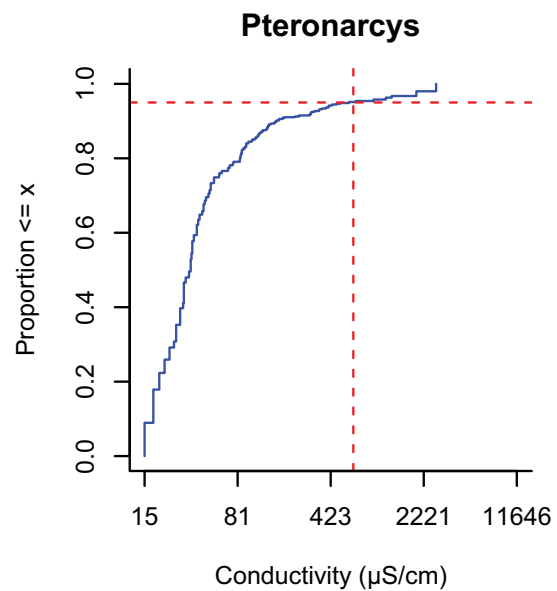
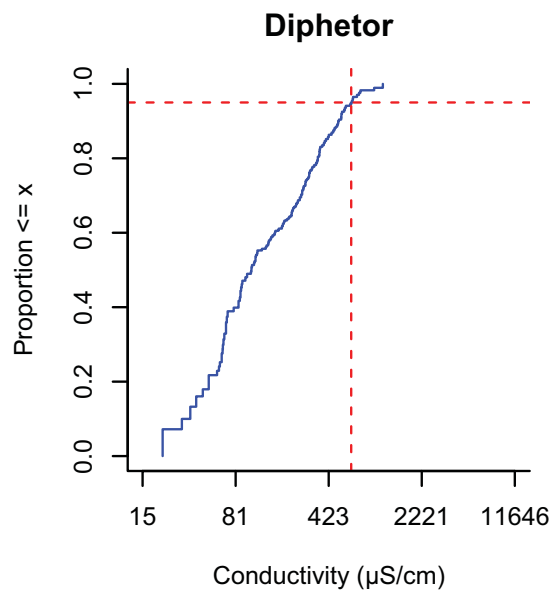
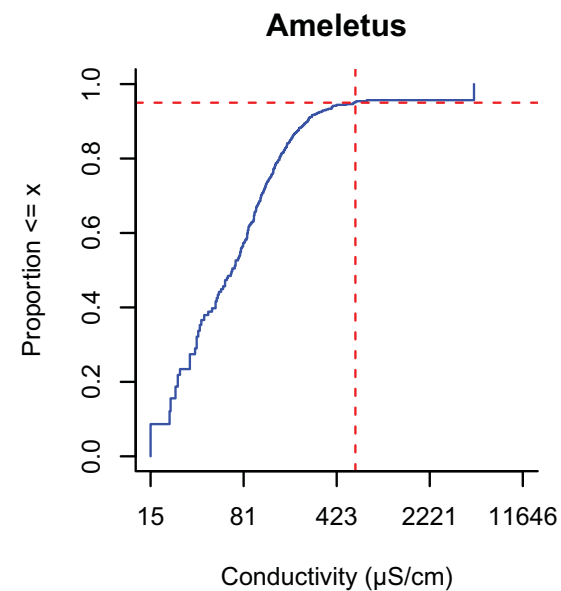
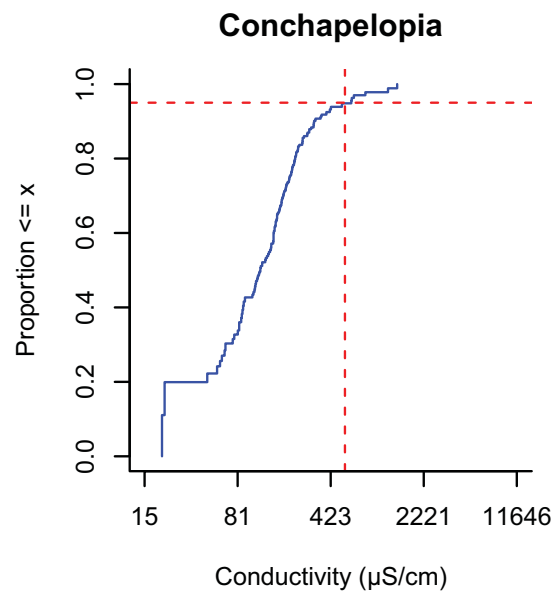
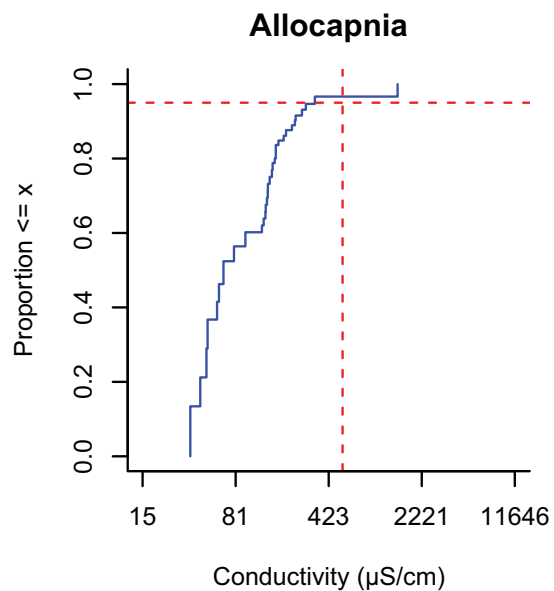


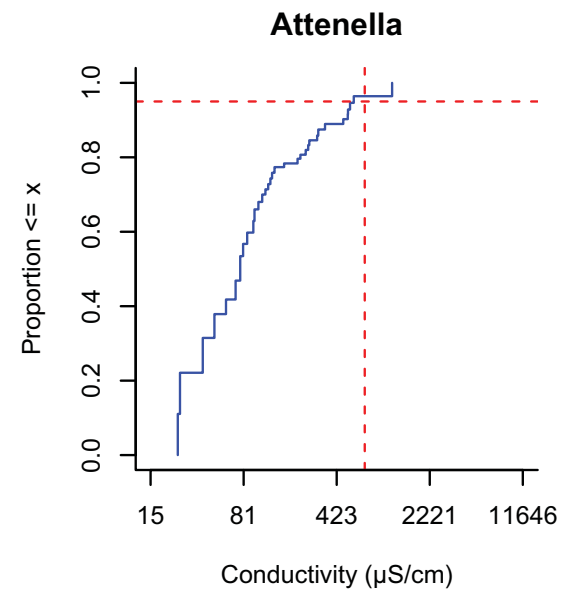
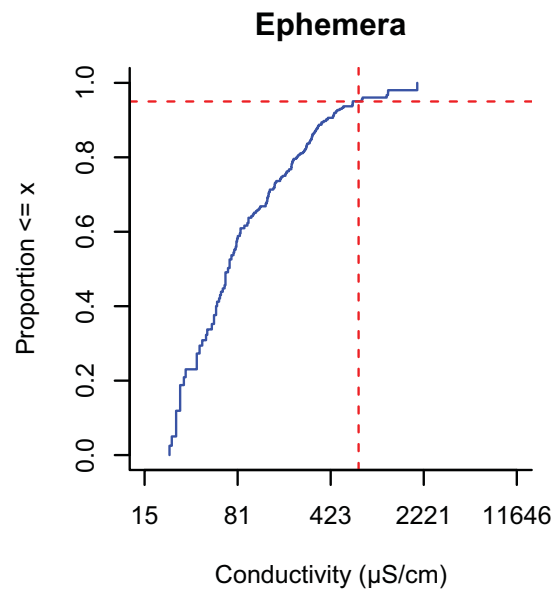
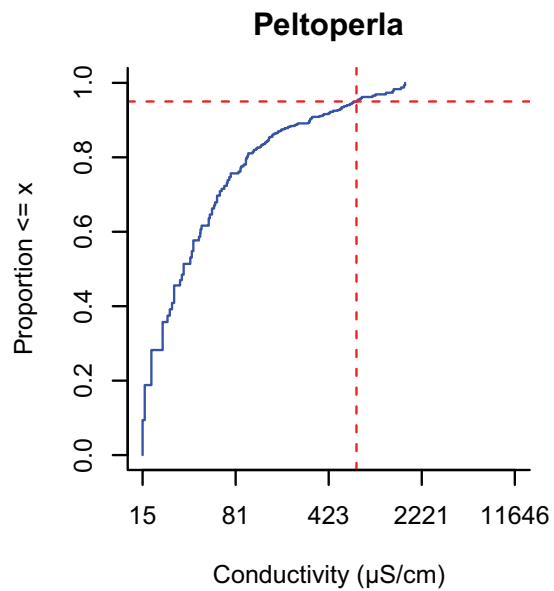
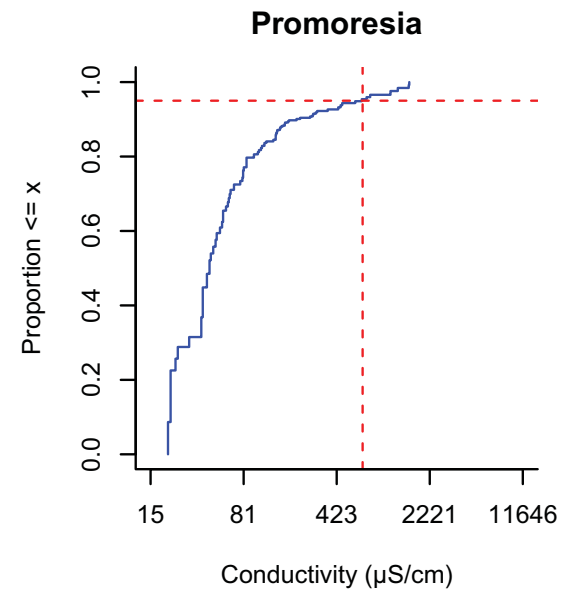
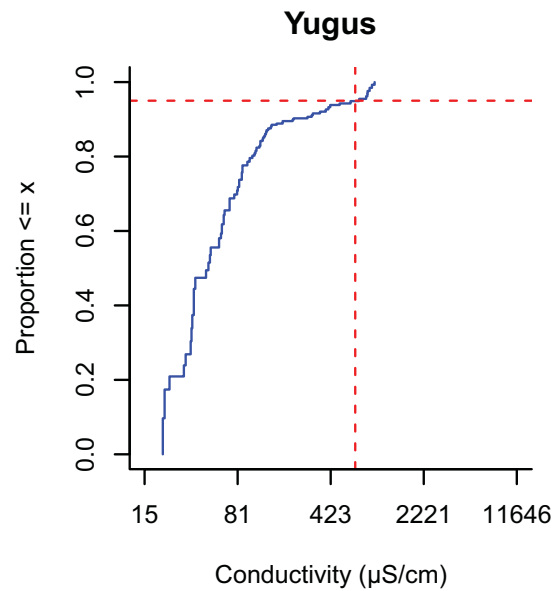
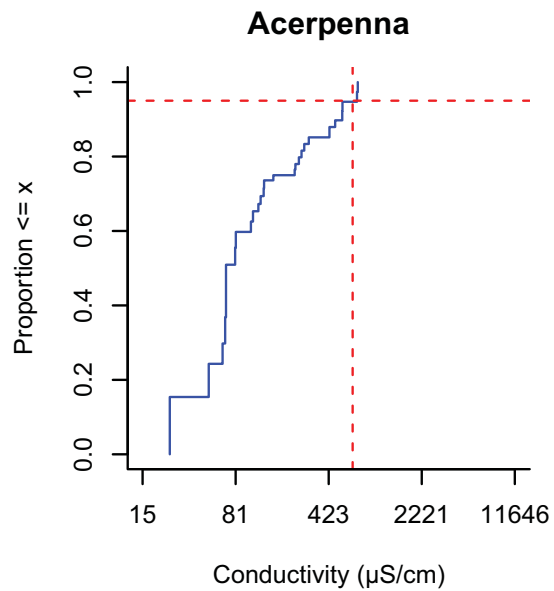


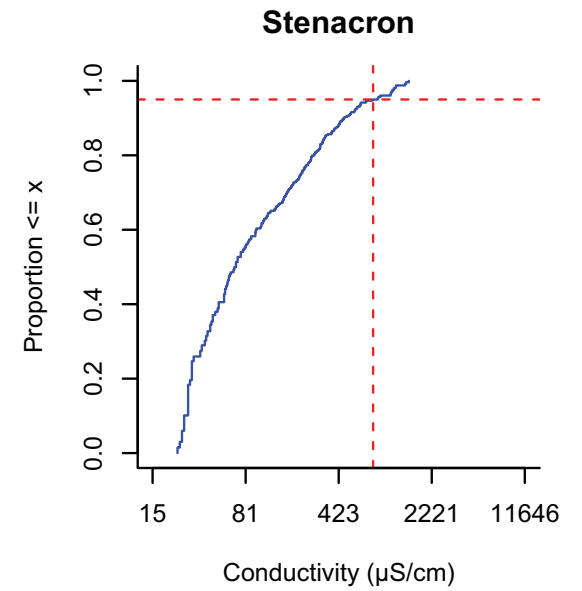
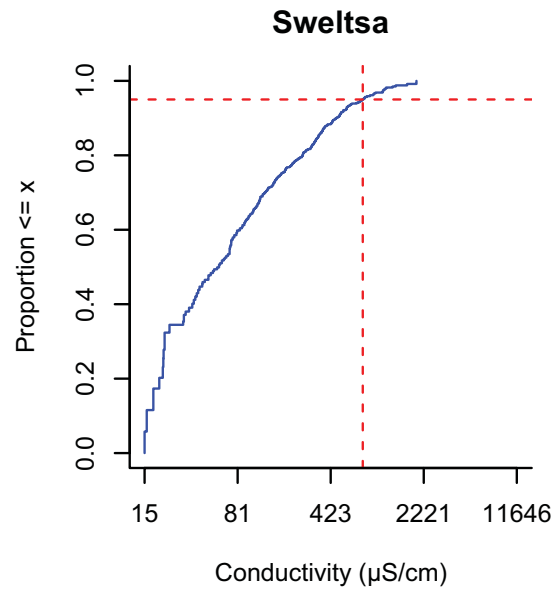
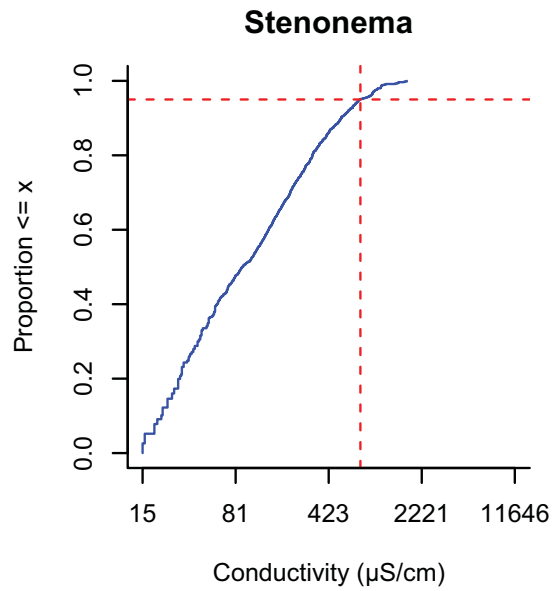
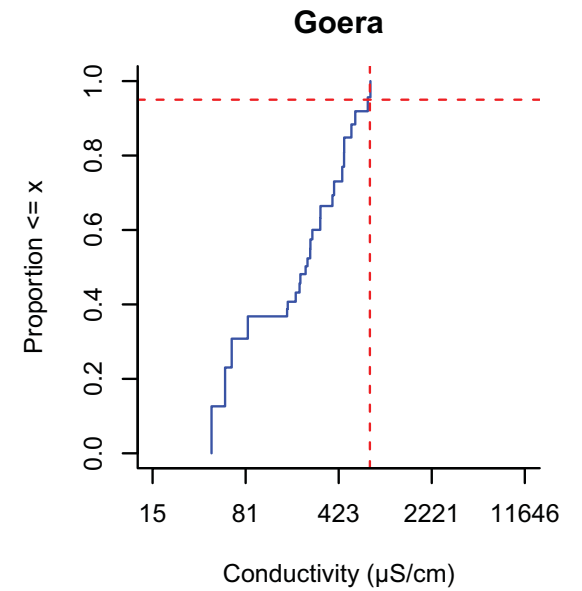
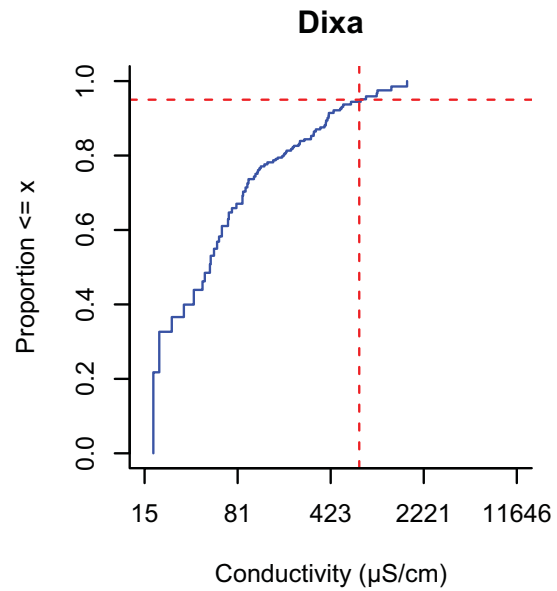
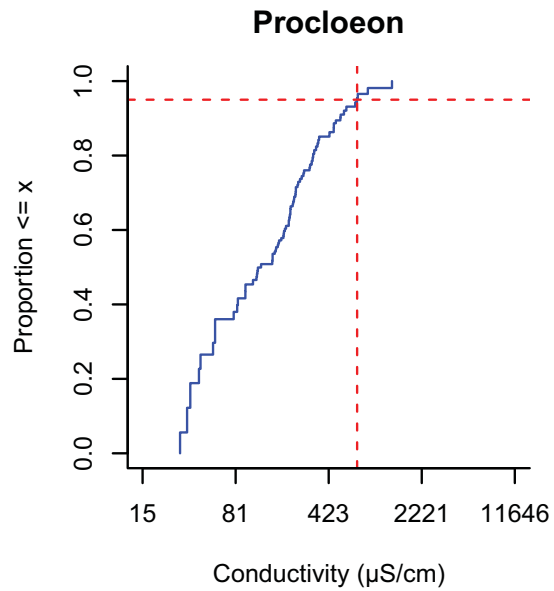
S-5

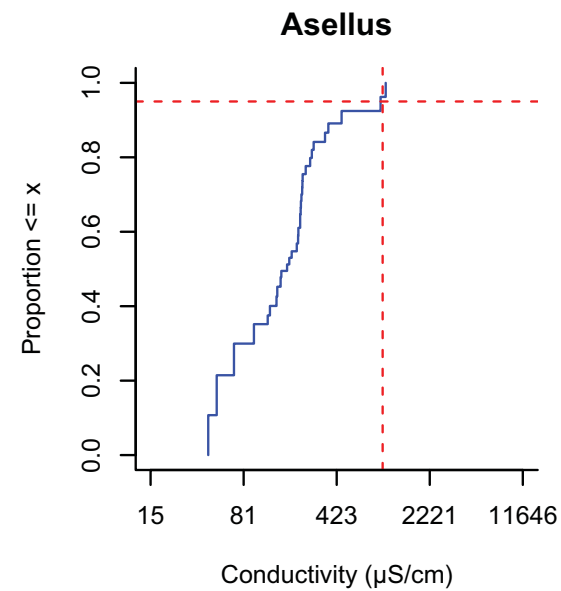
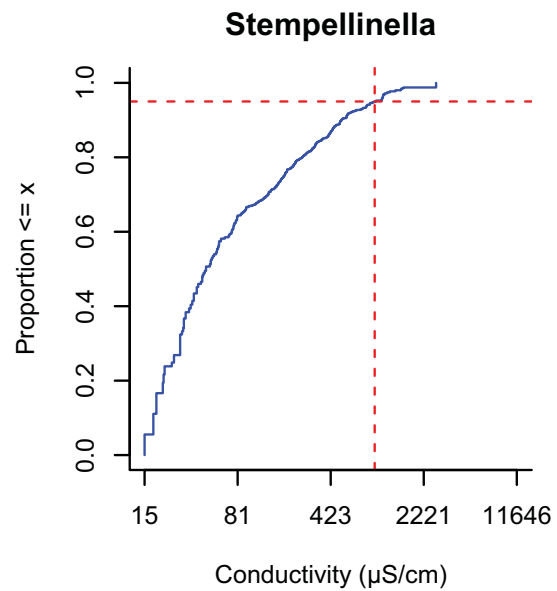
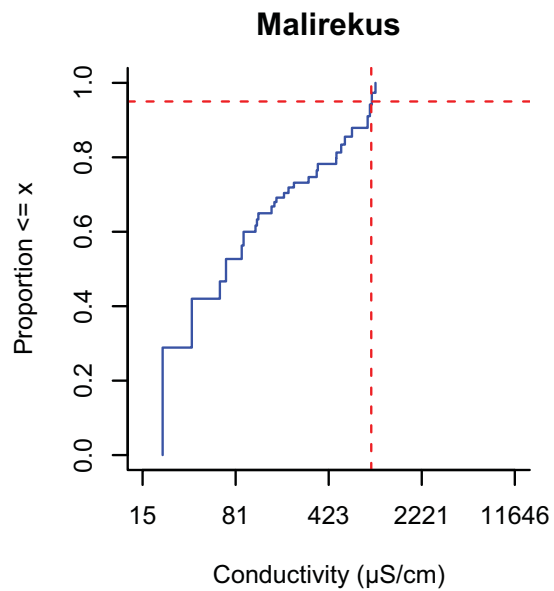
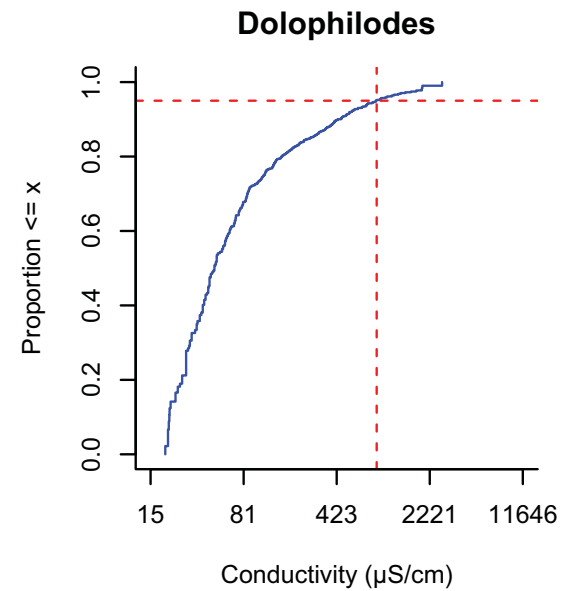
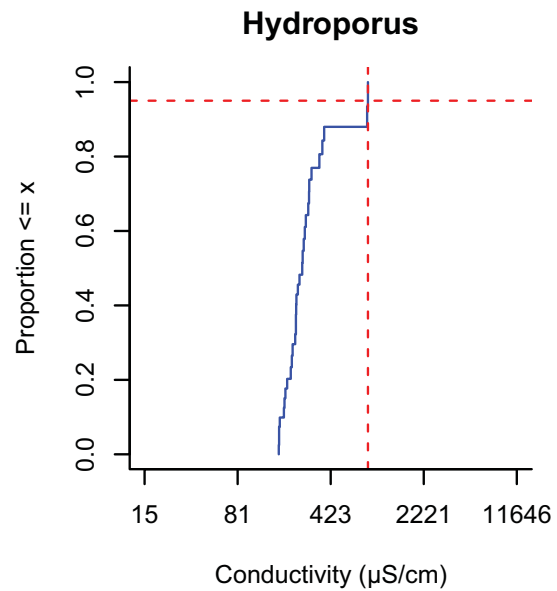
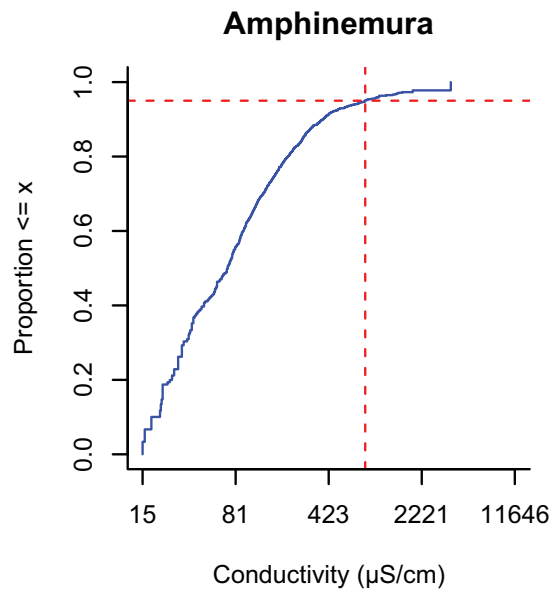


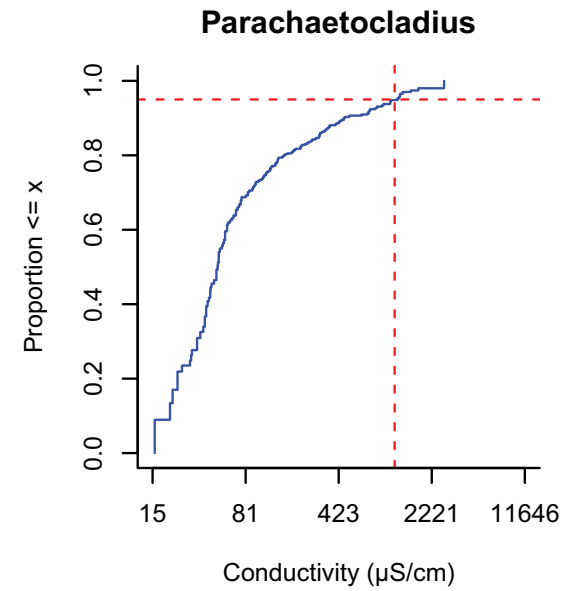
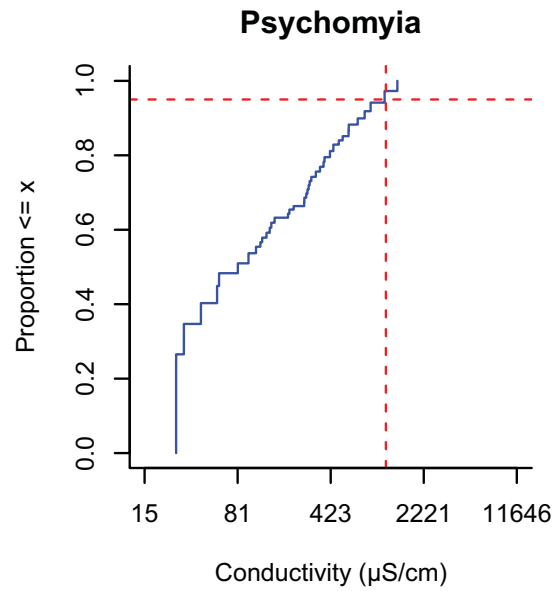
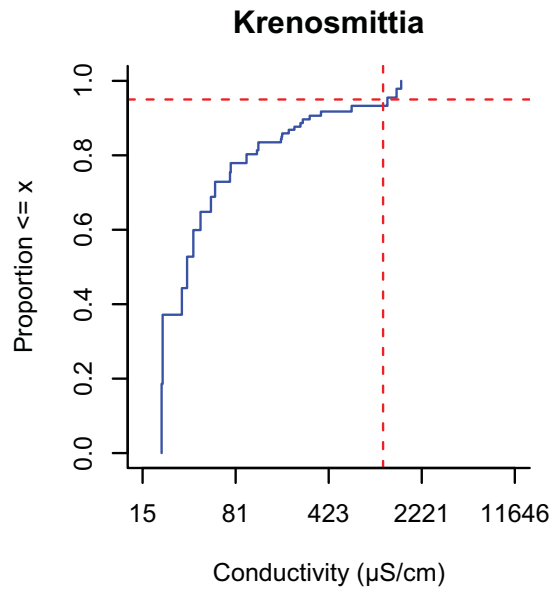
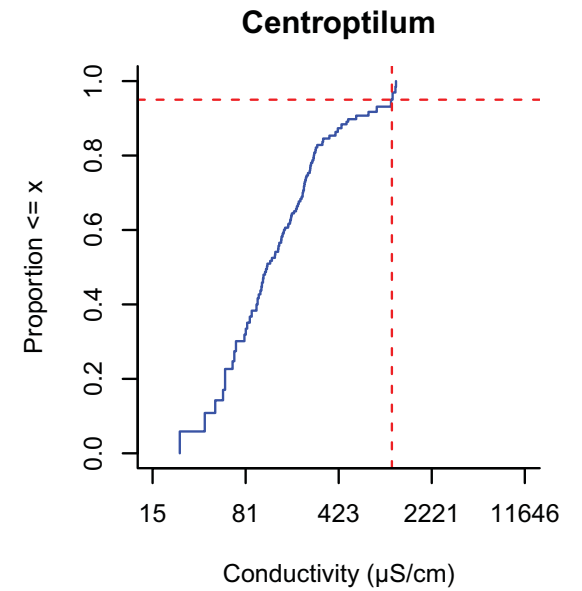
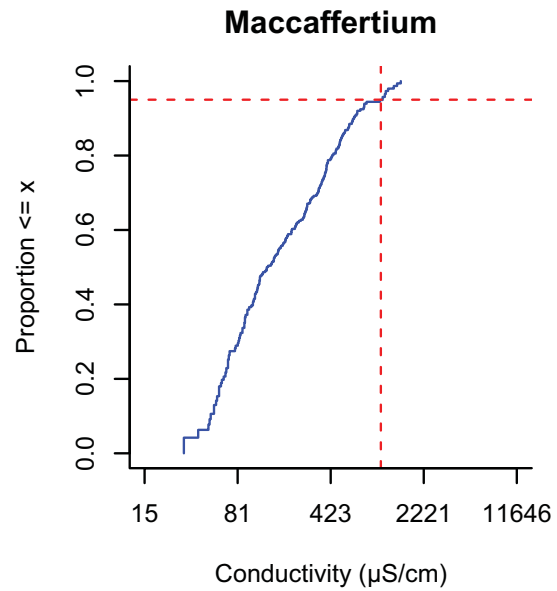
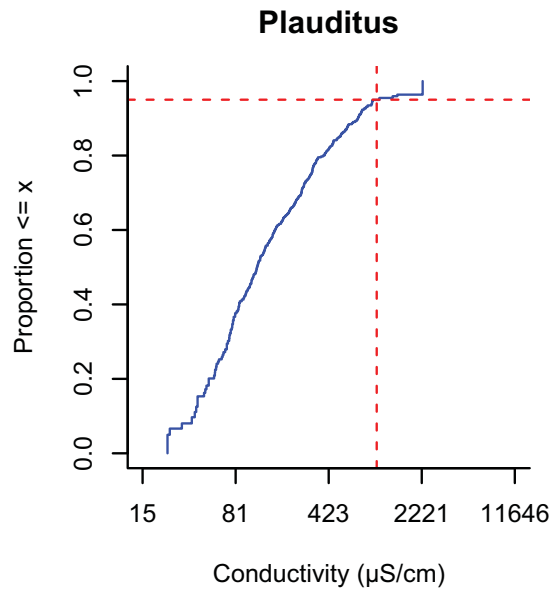


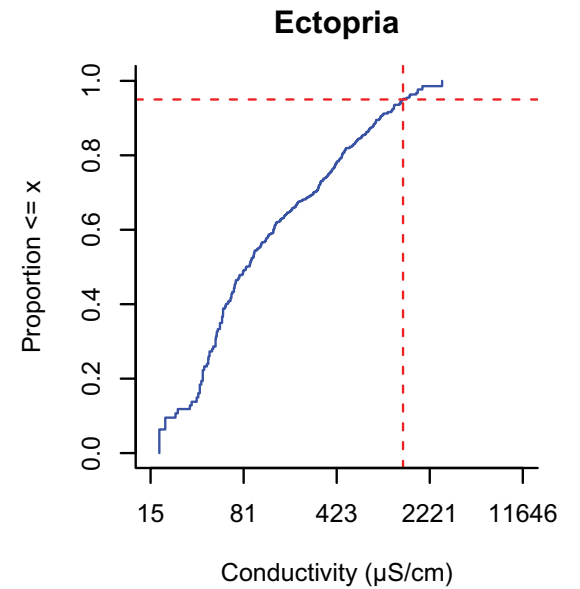
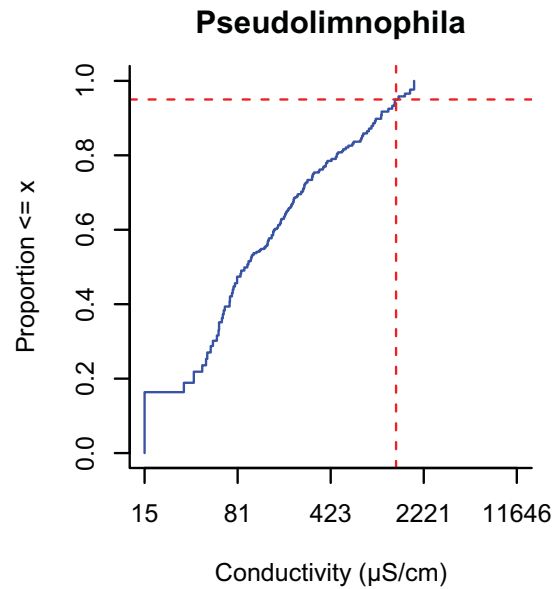
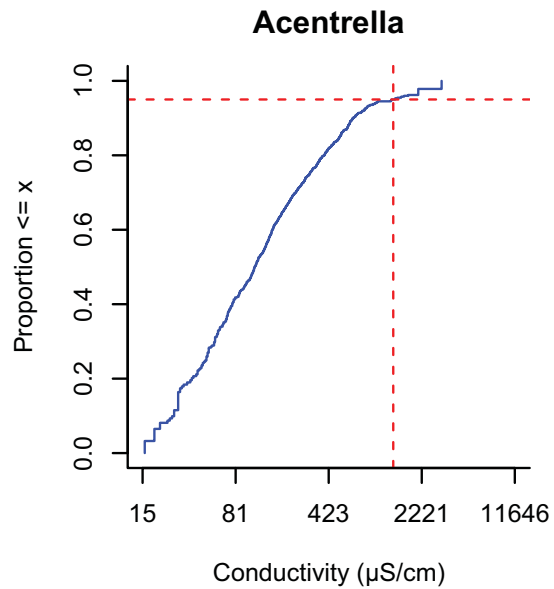
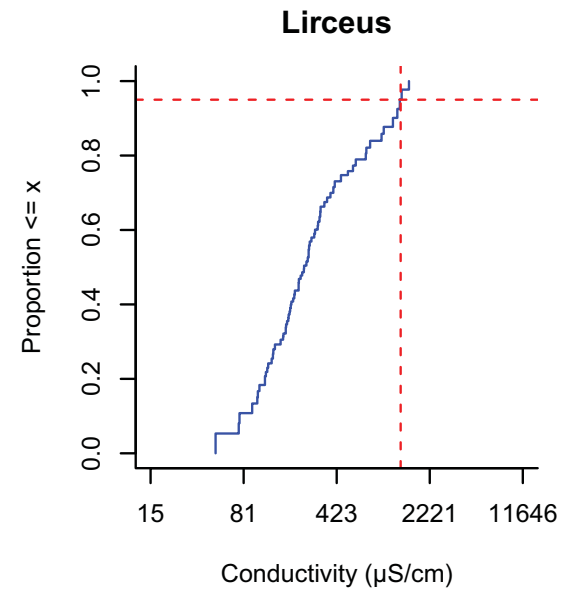
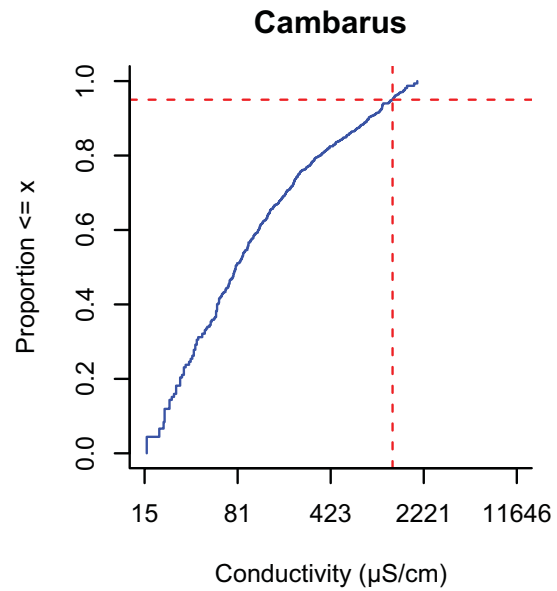
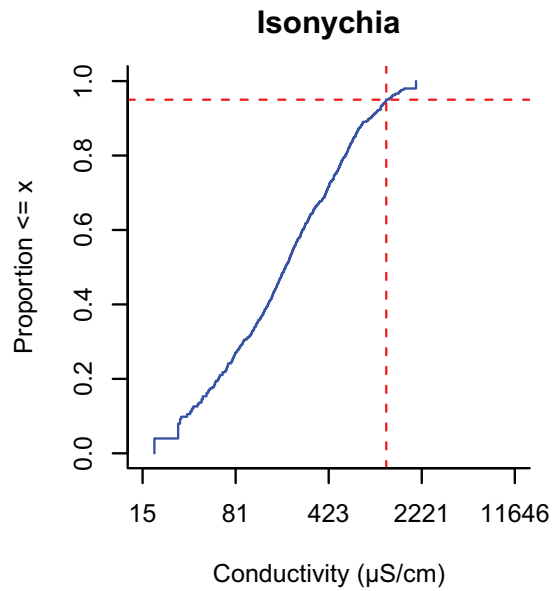


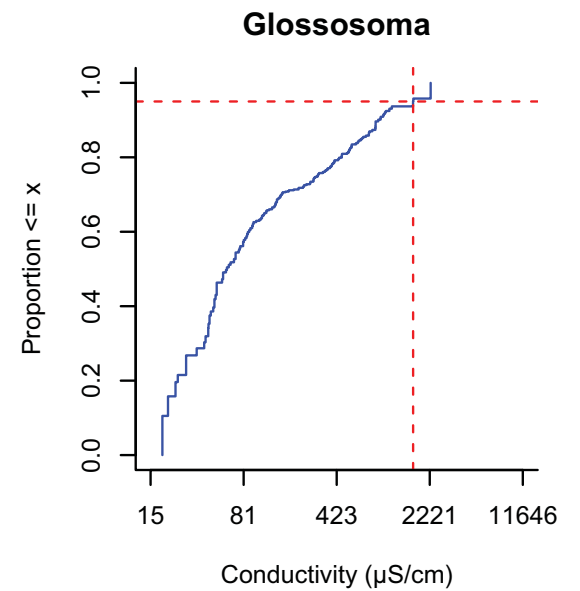
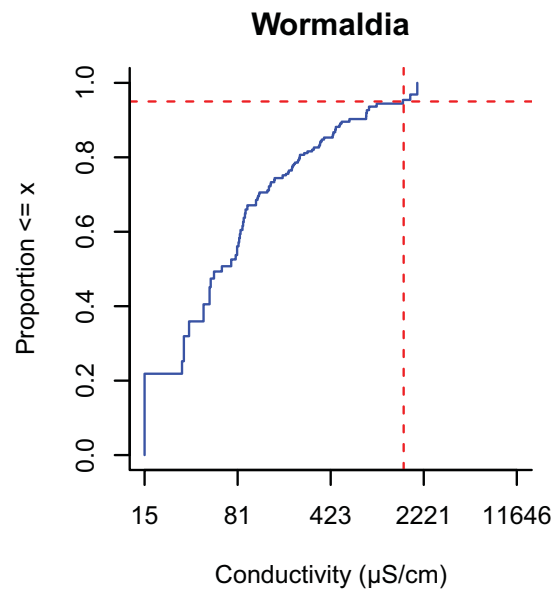
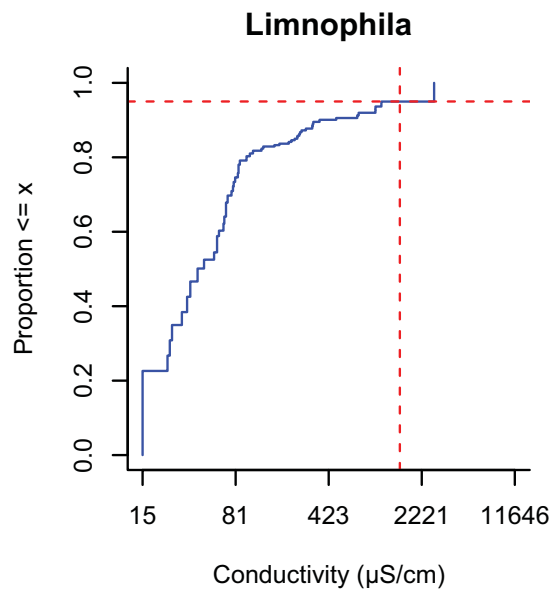
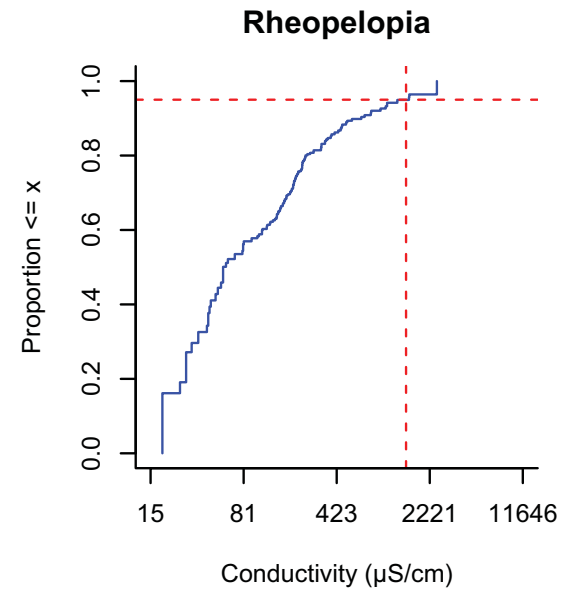
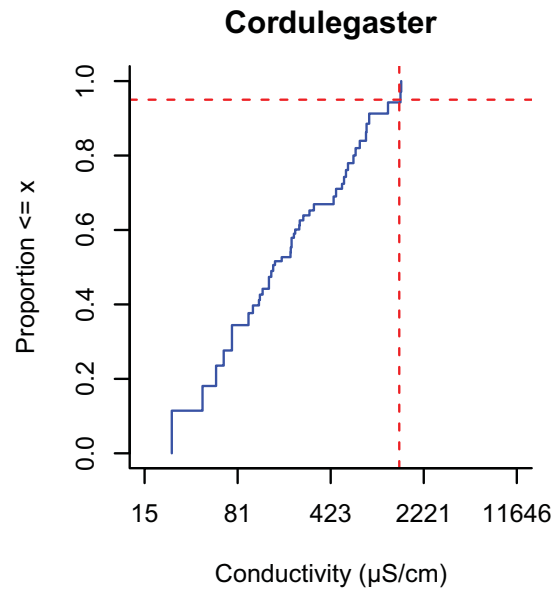
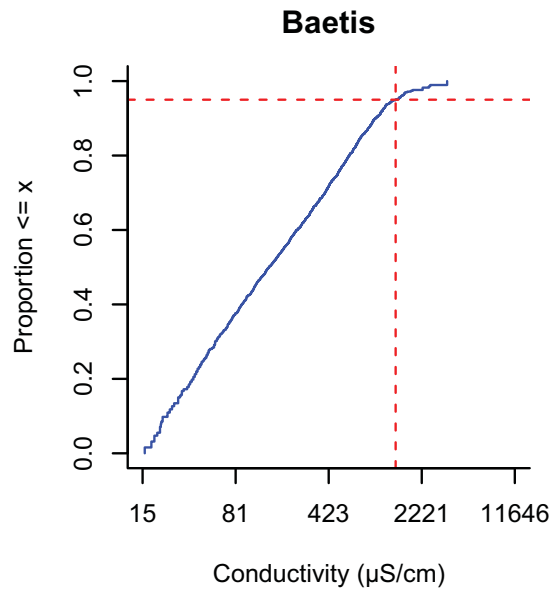


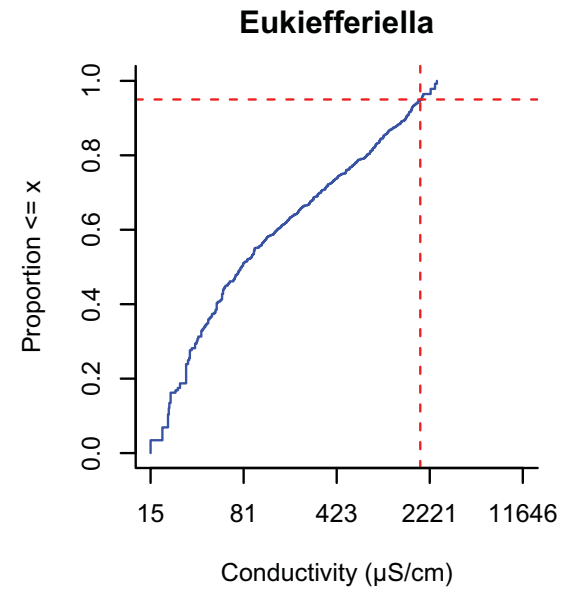
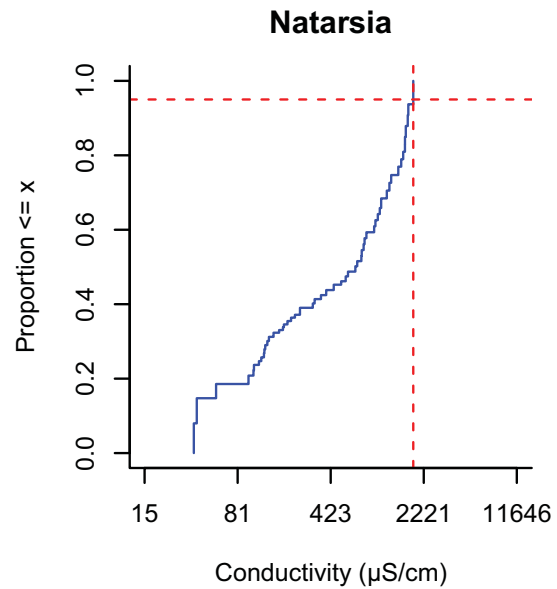
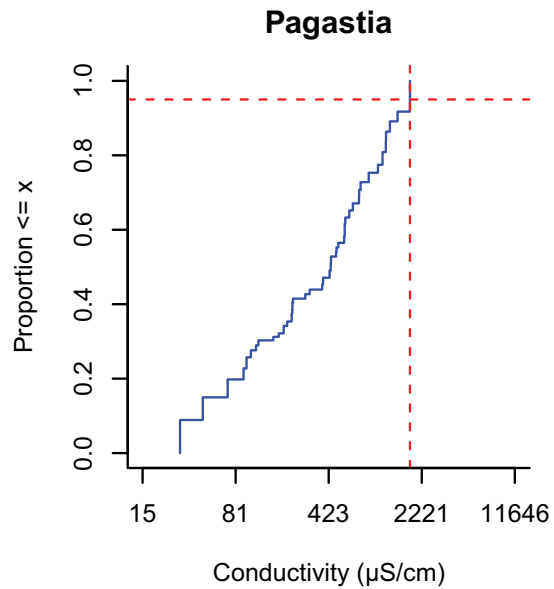
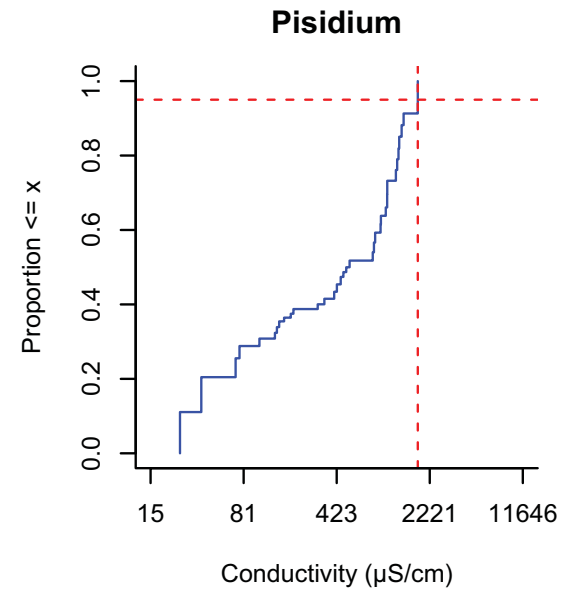
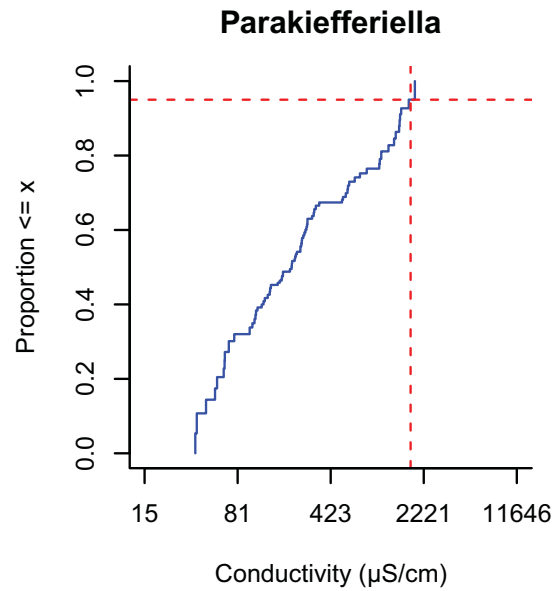
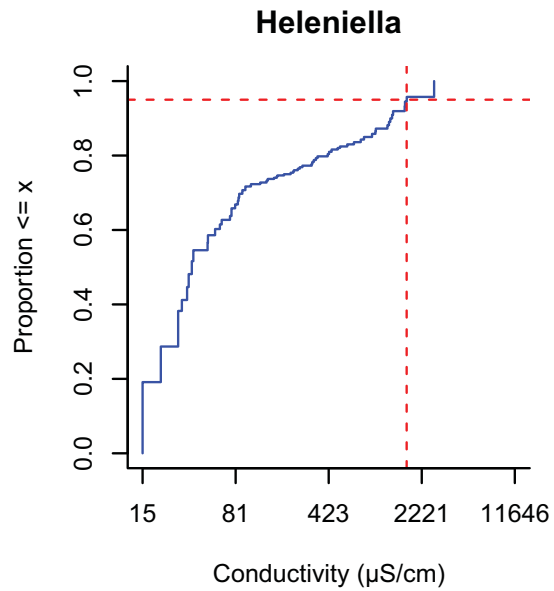


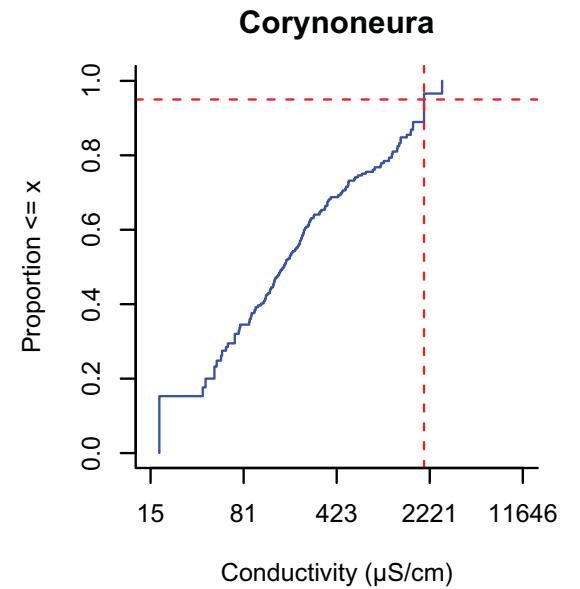
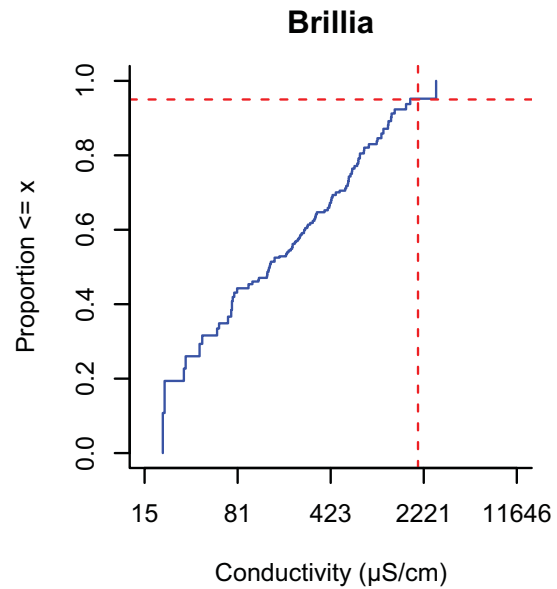
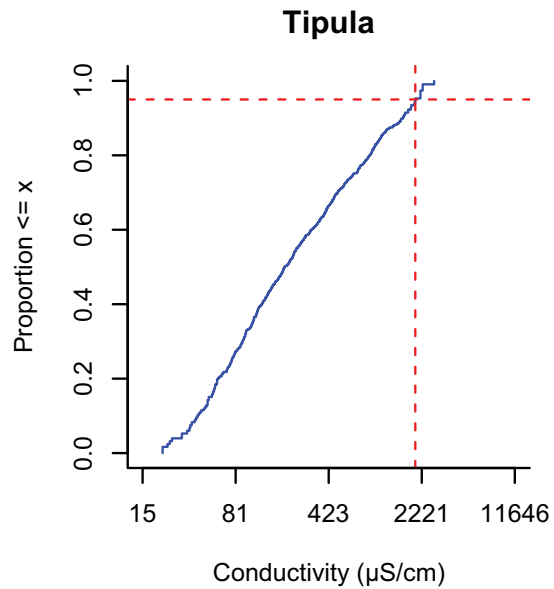
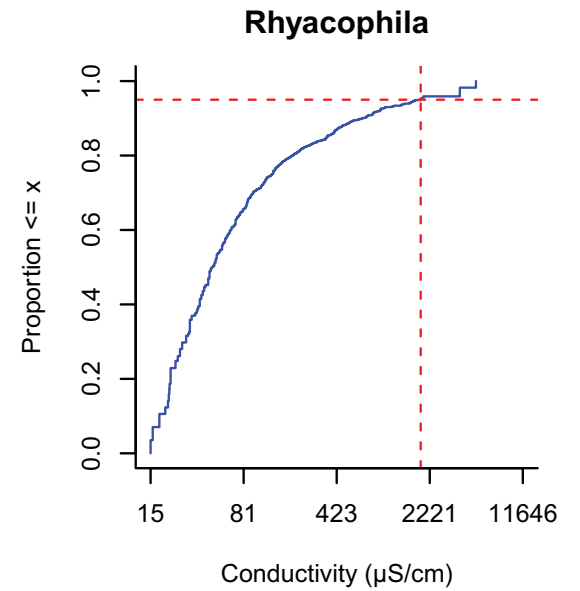
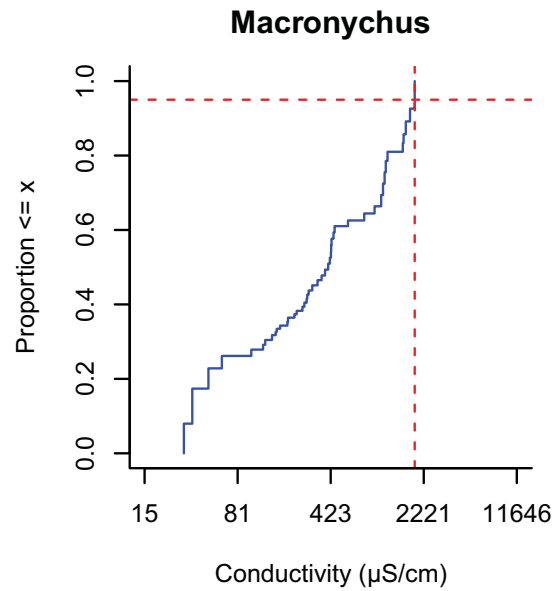
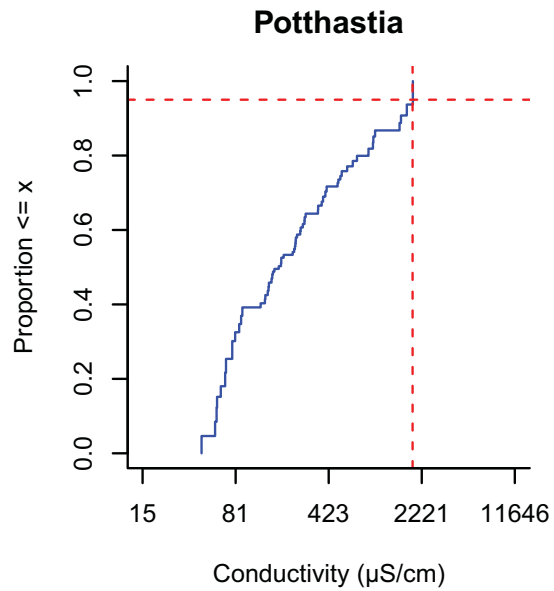


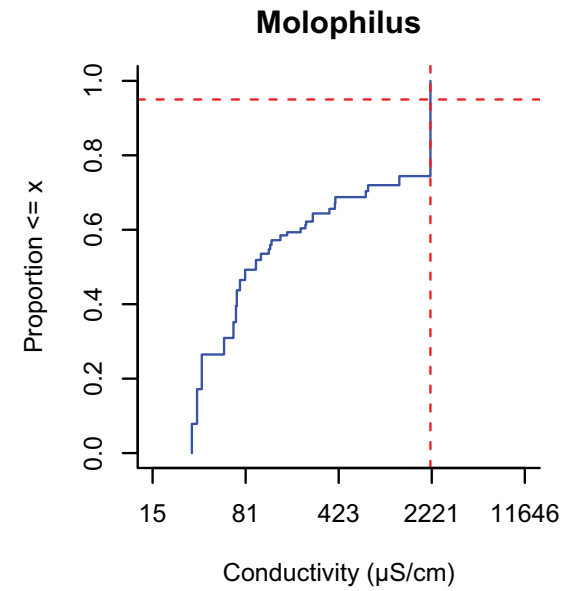
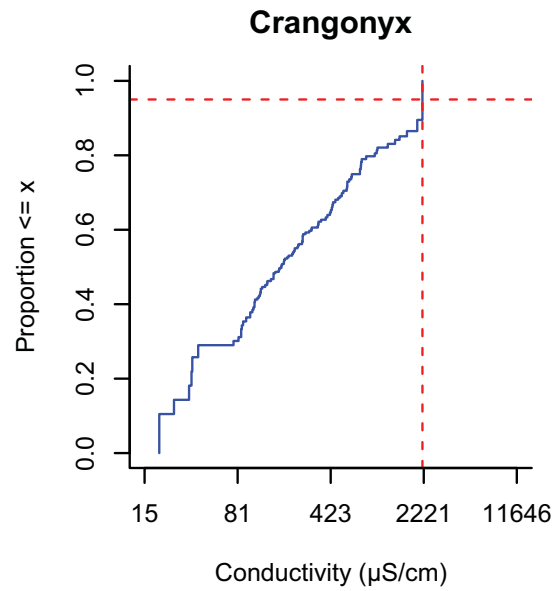
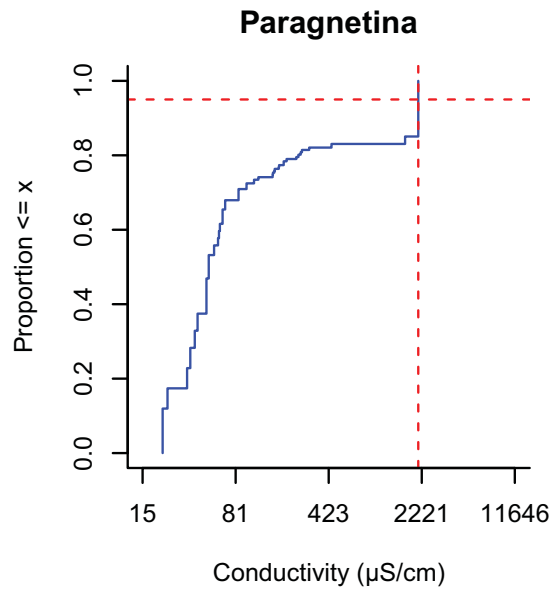
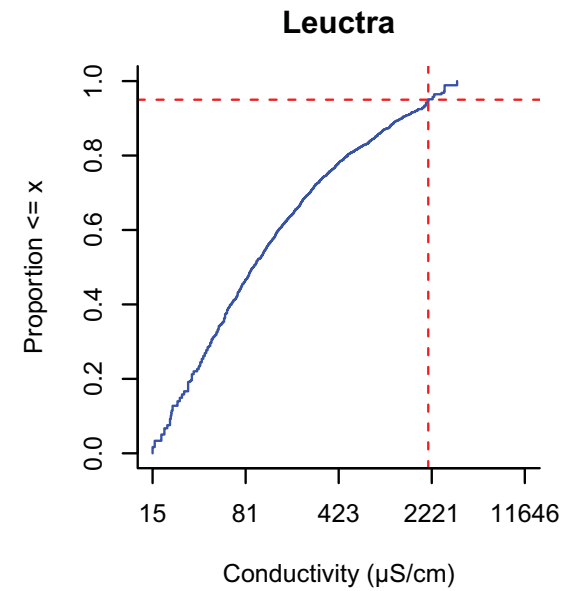
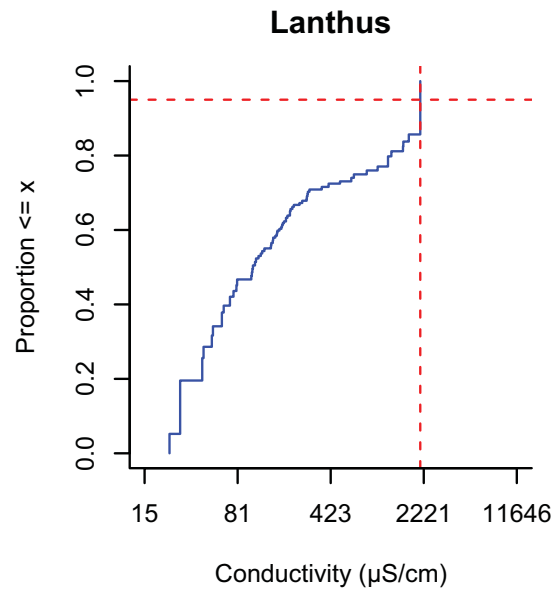
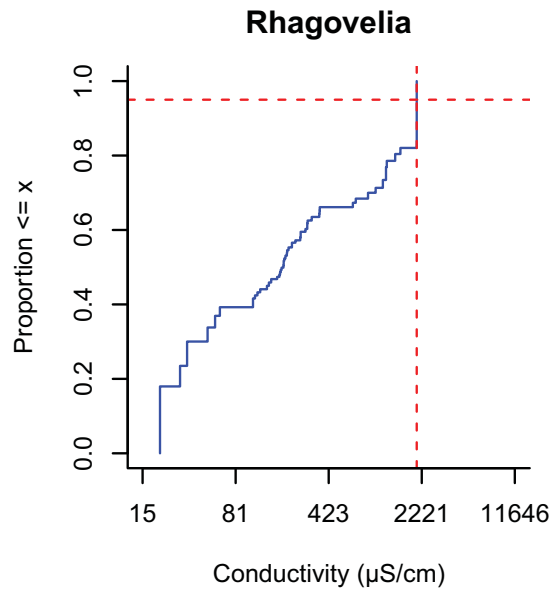


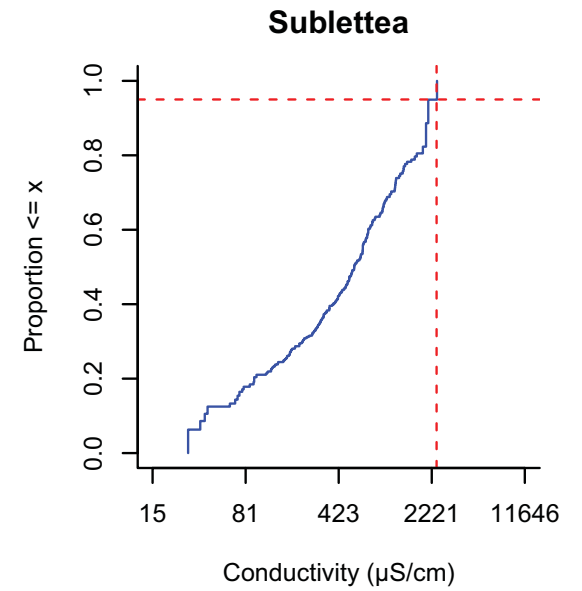
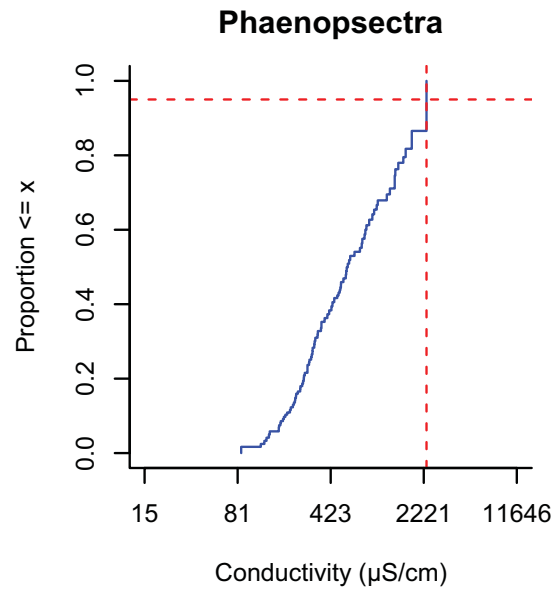
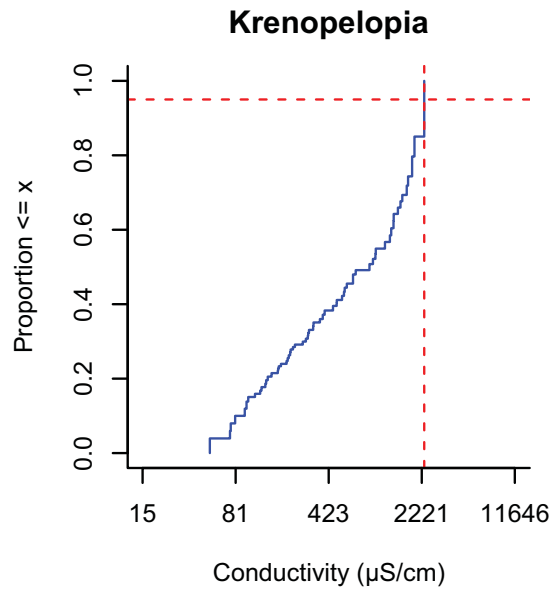
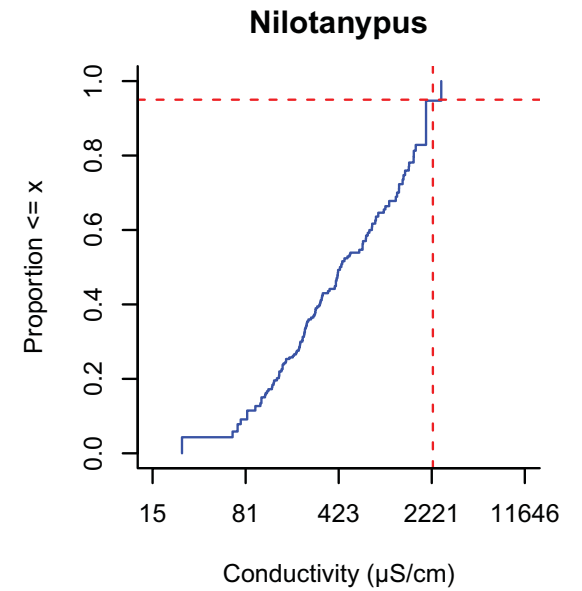
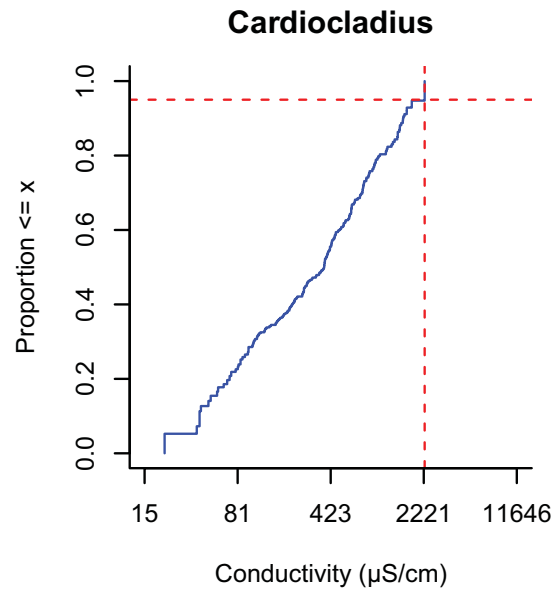
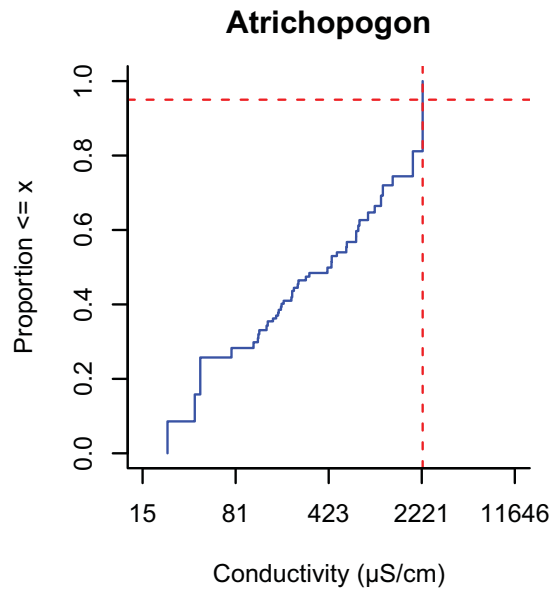


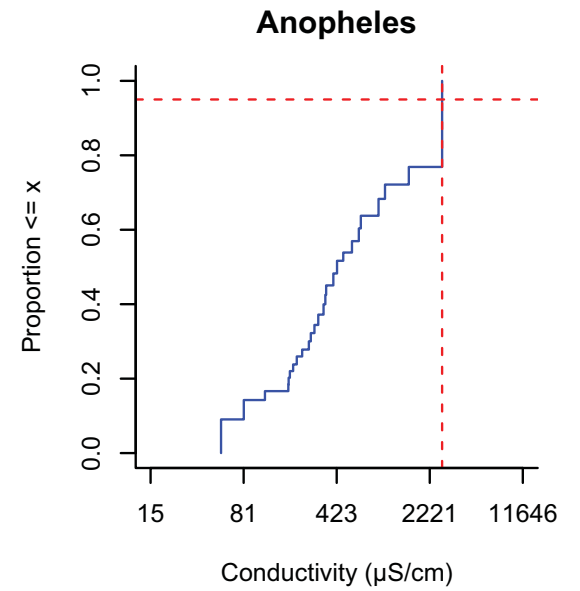
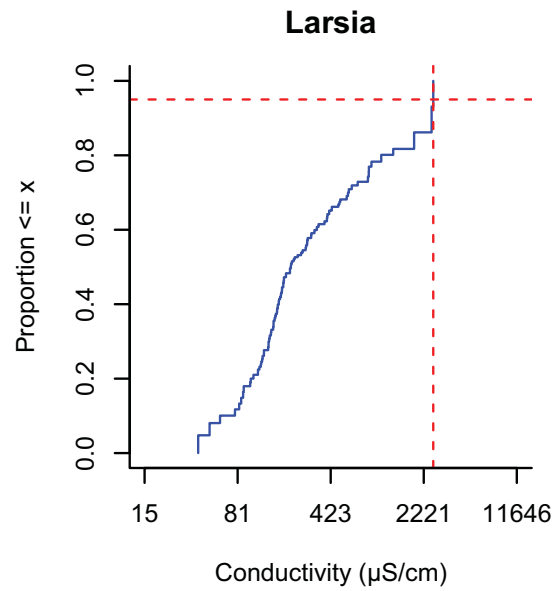
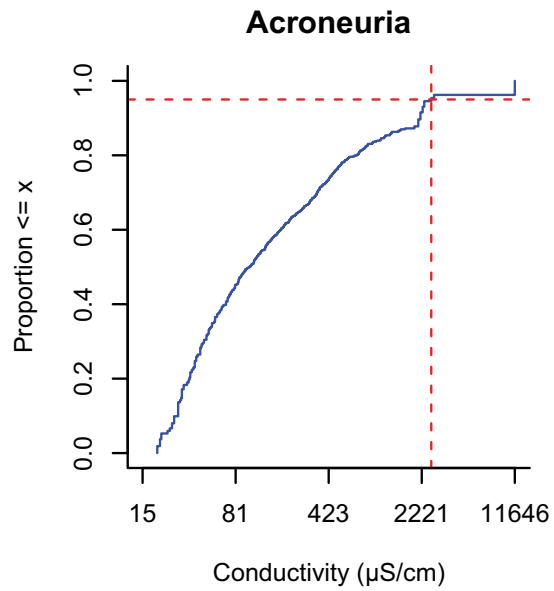
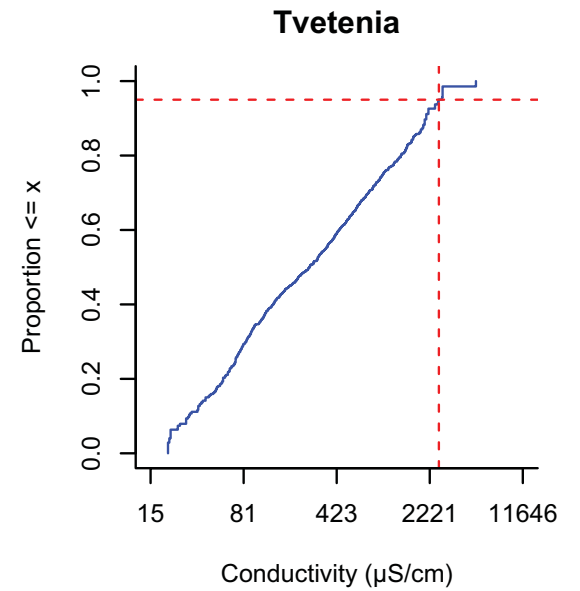
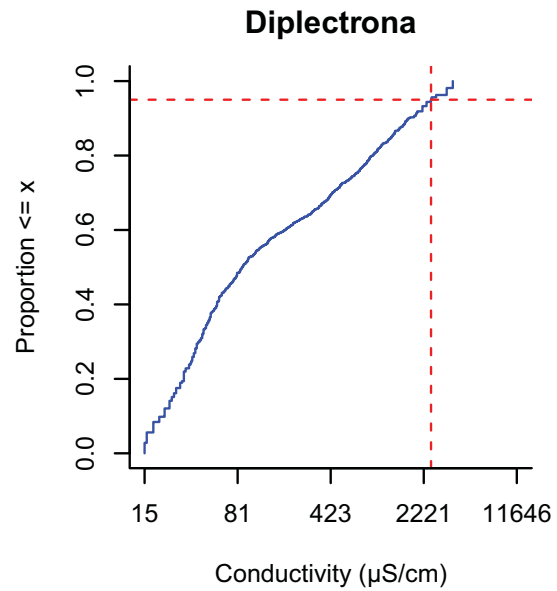
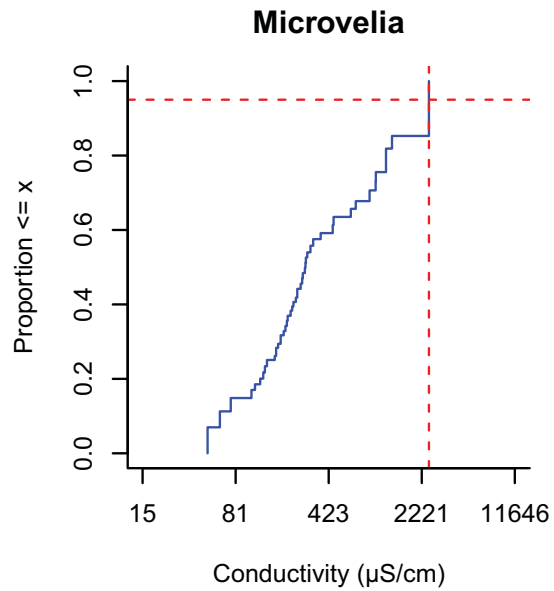


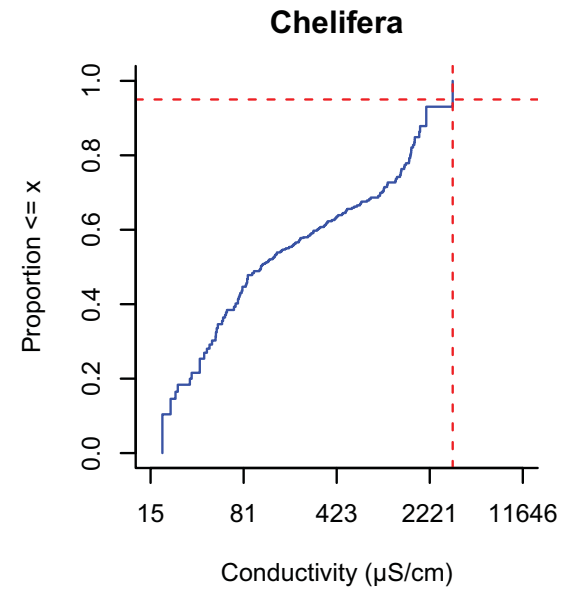
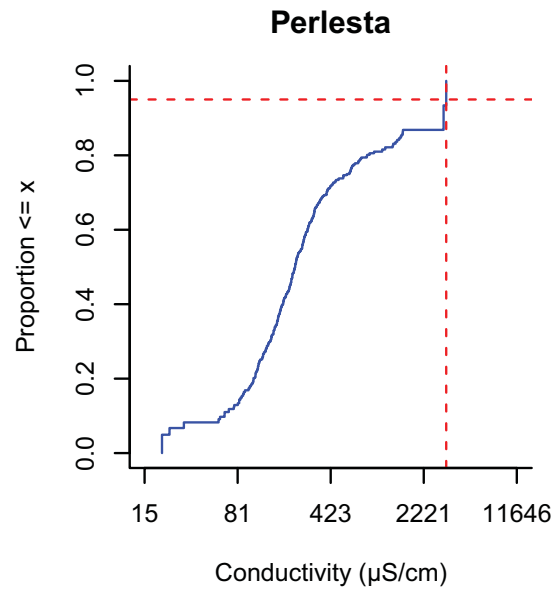
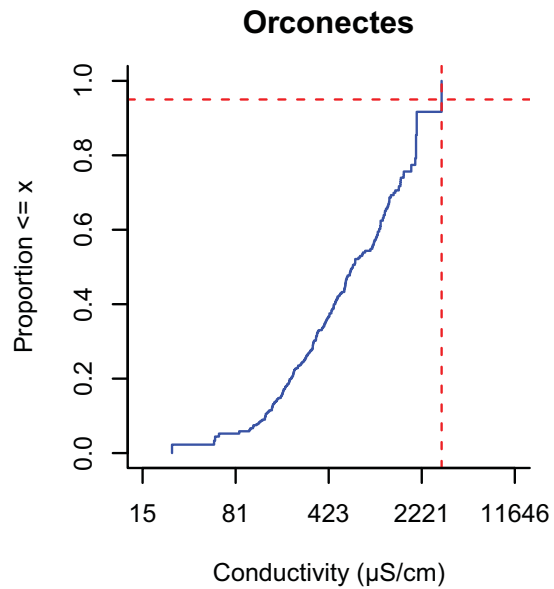
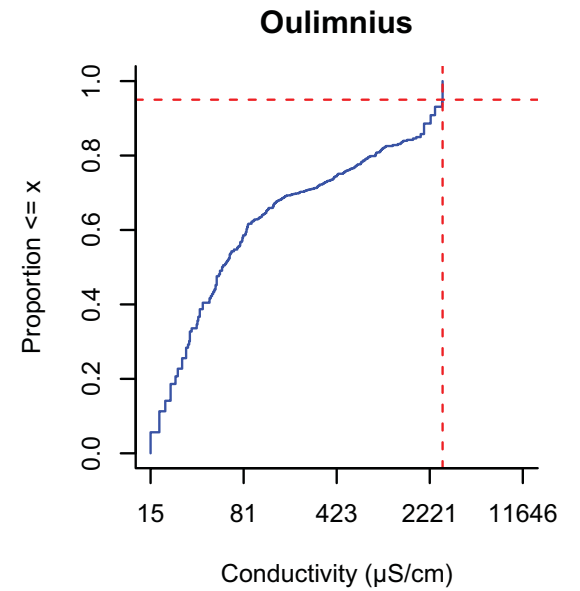
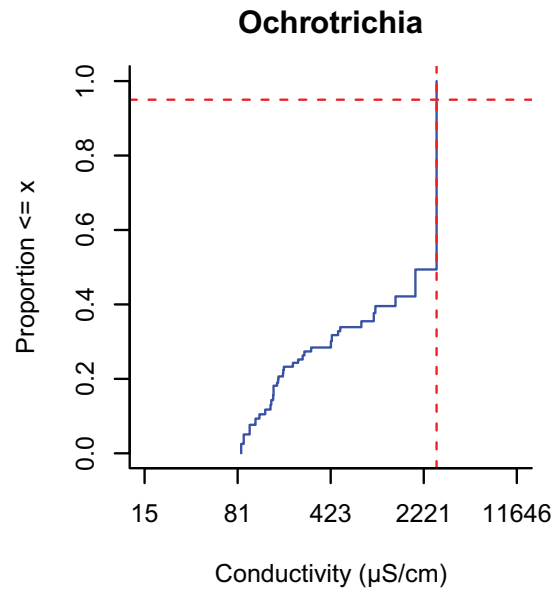
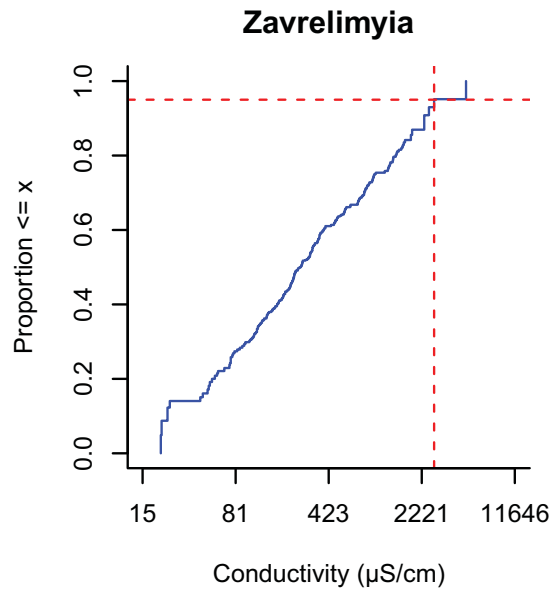


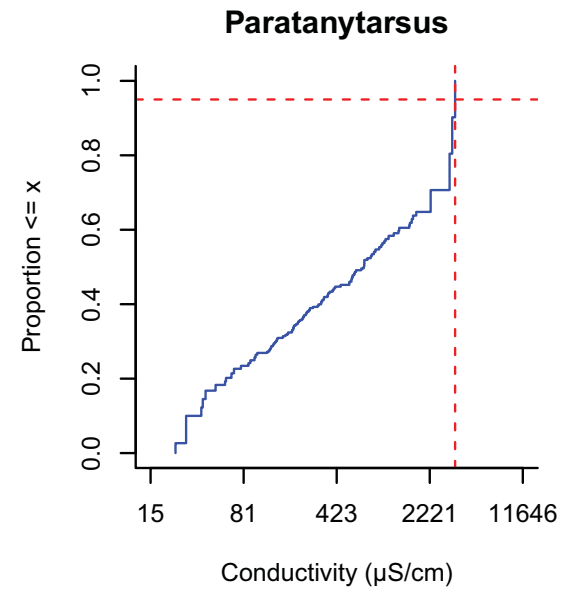
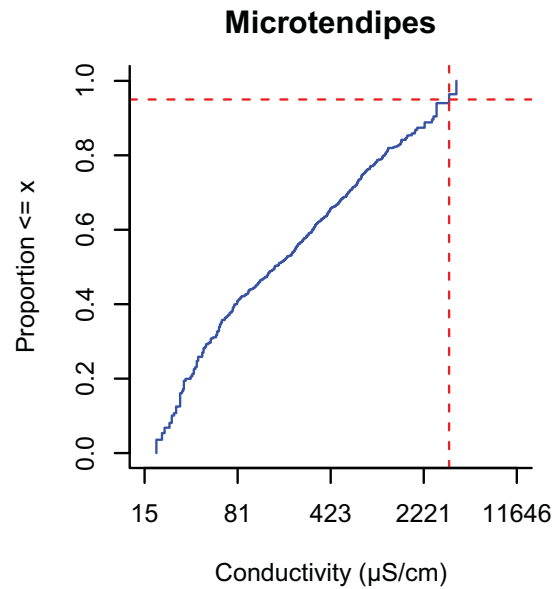
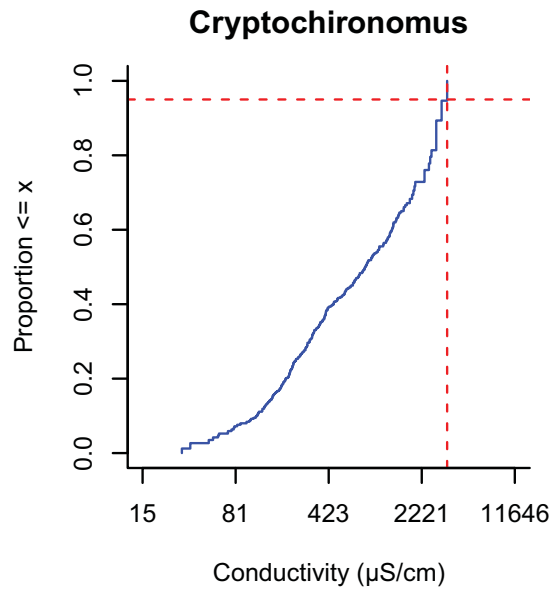
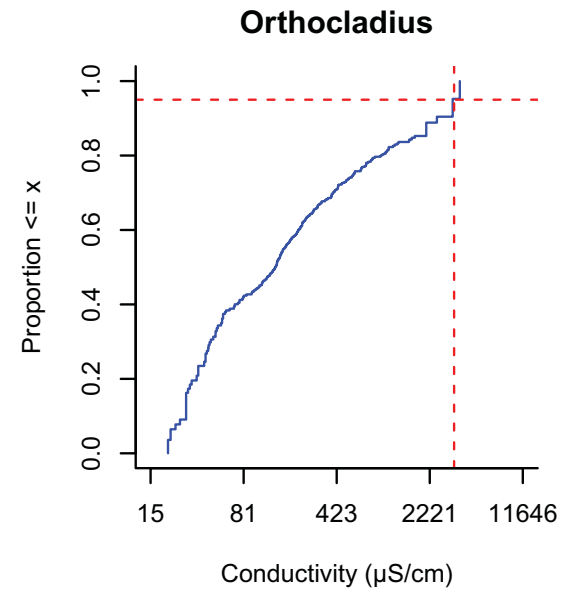
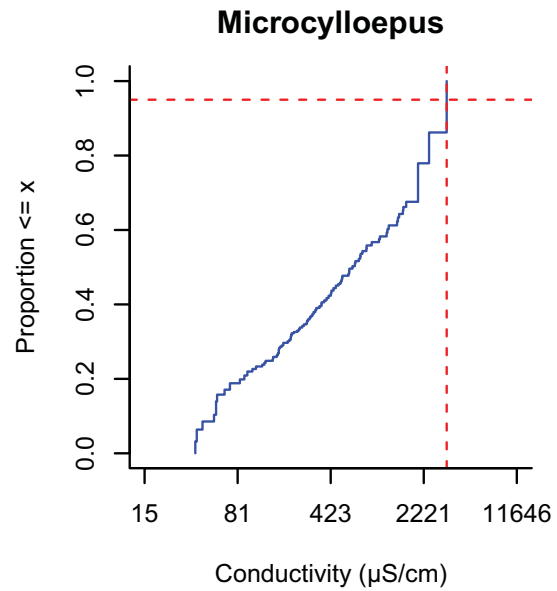
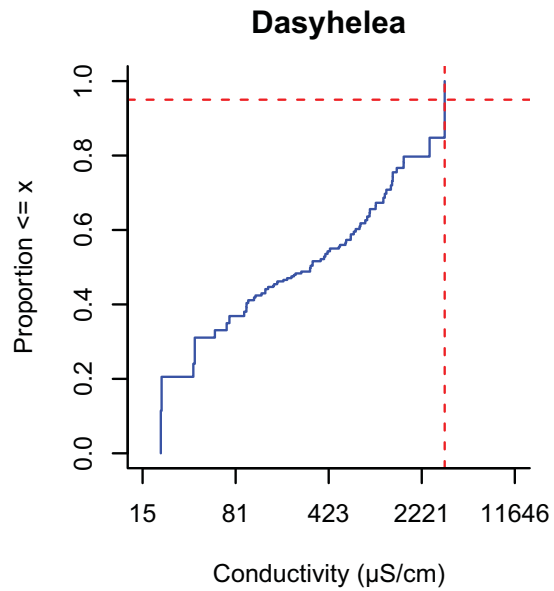


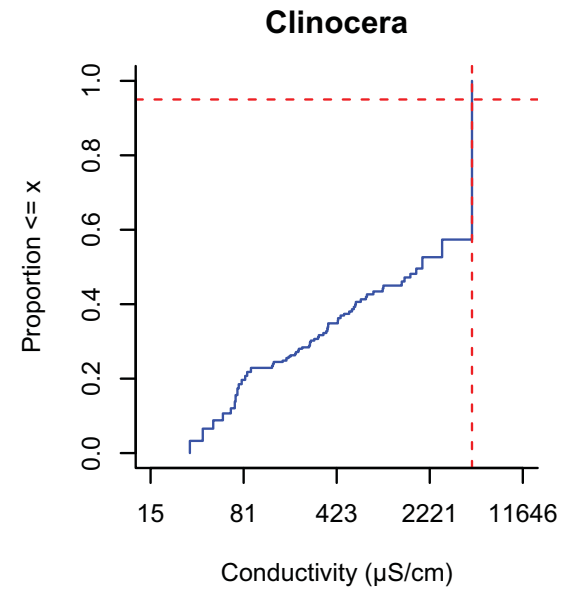
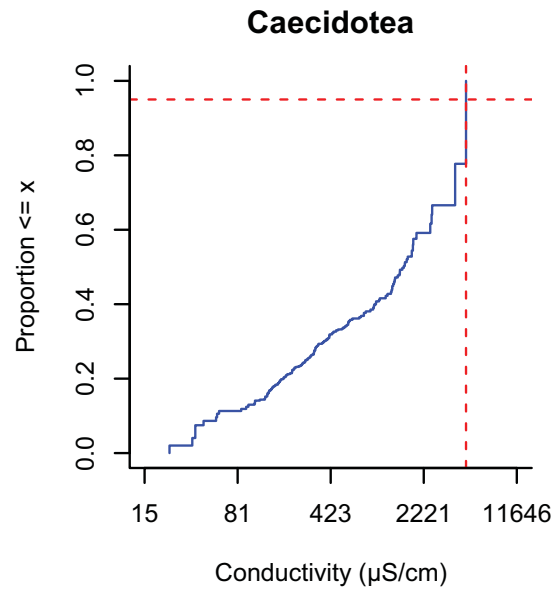
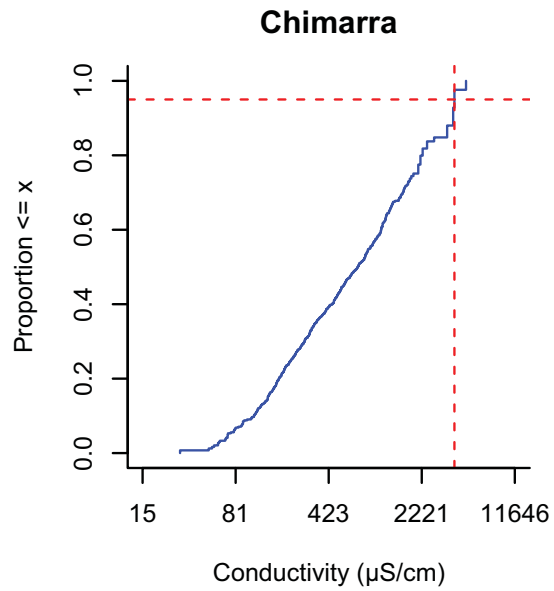
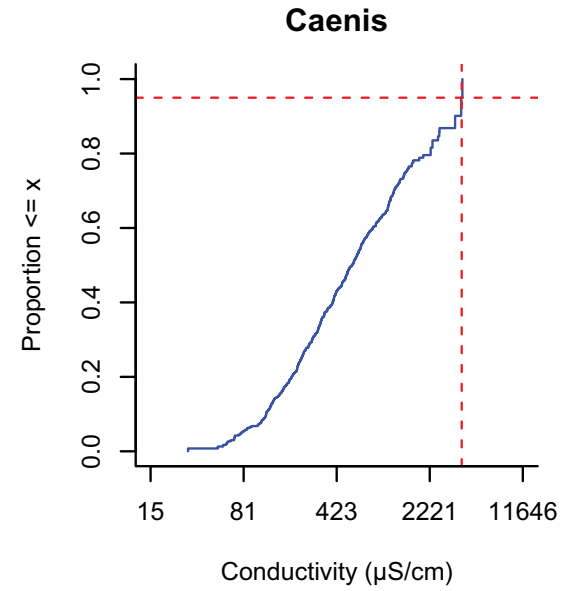
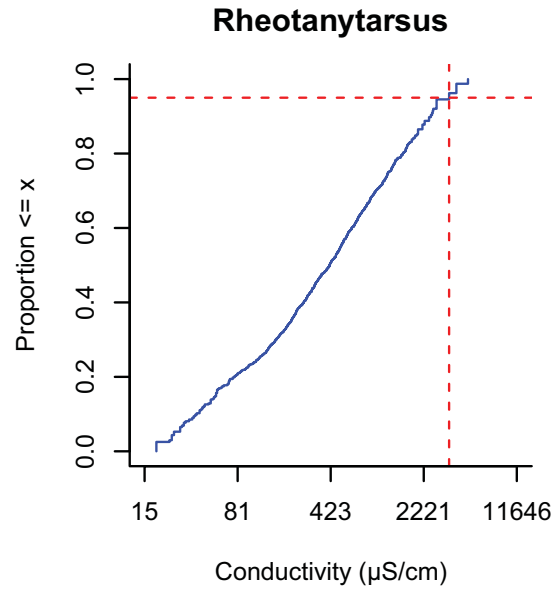
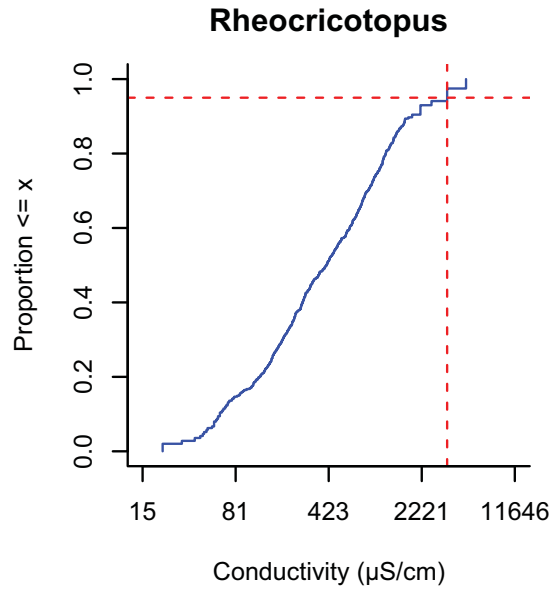


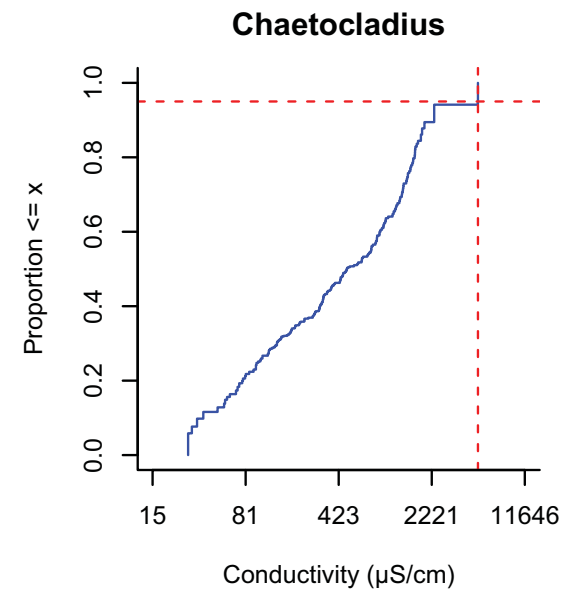
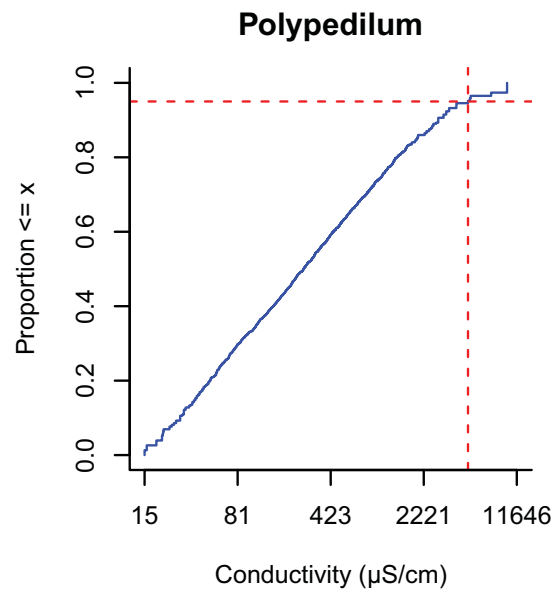
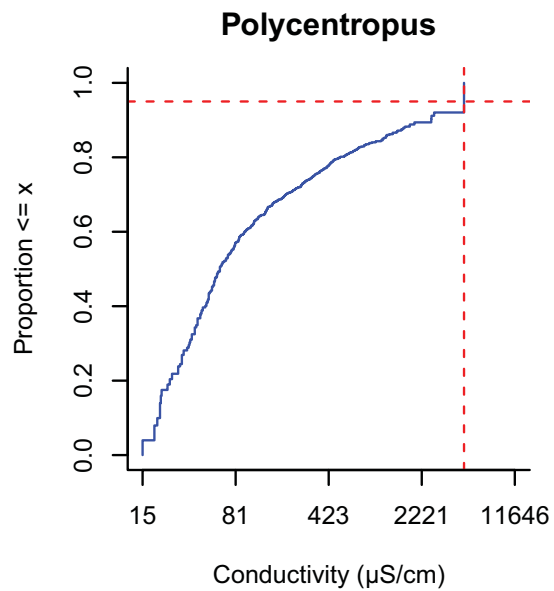
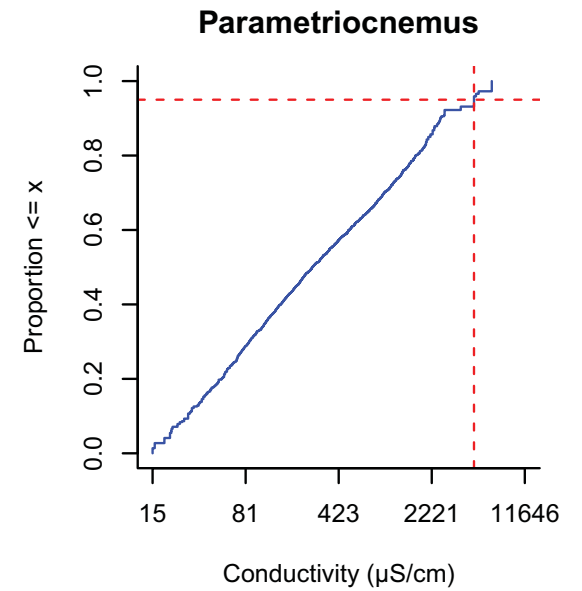
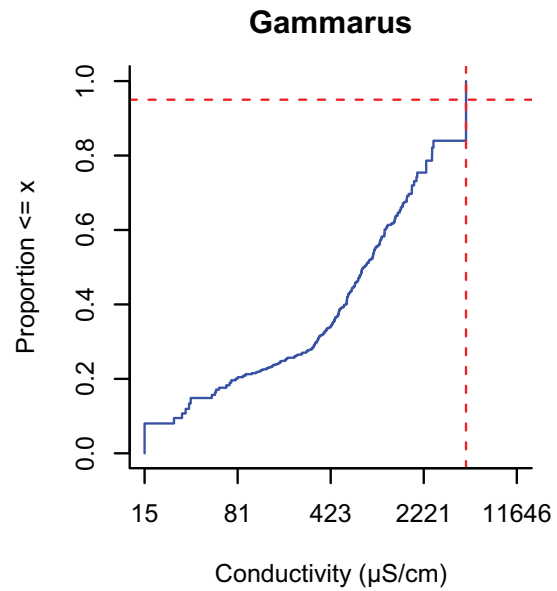
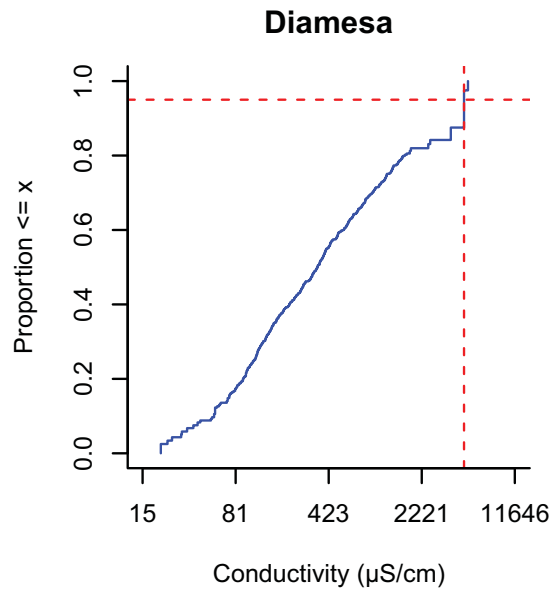


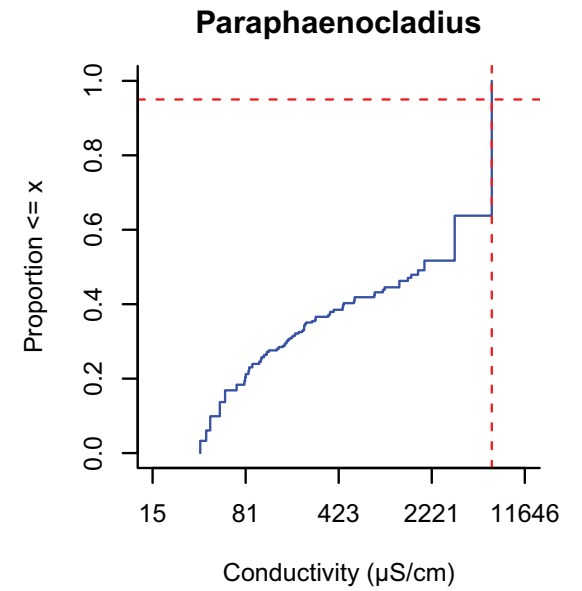
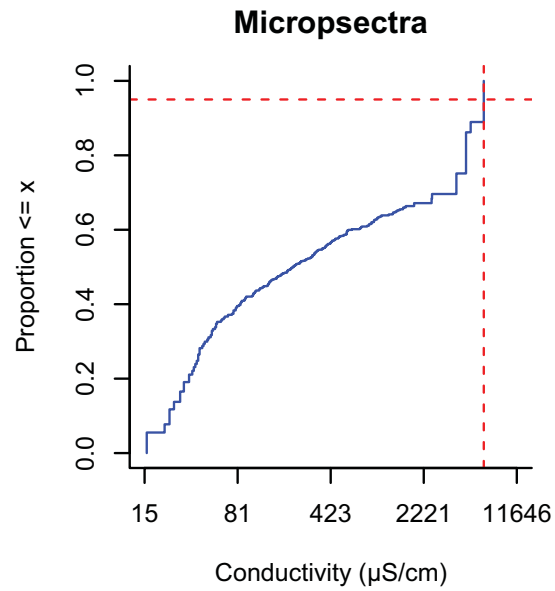
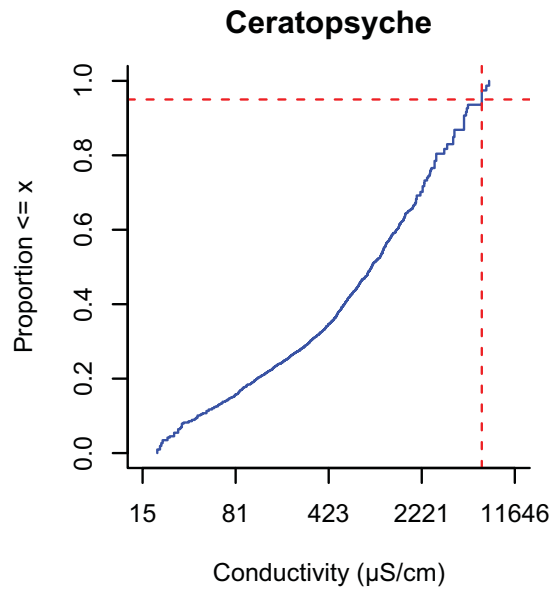
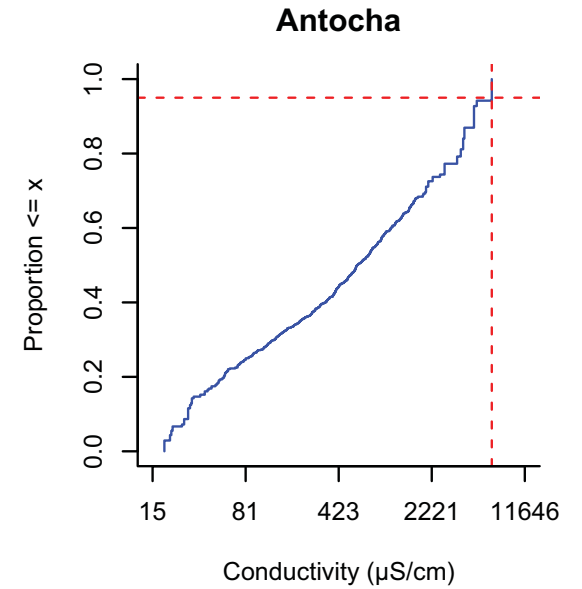
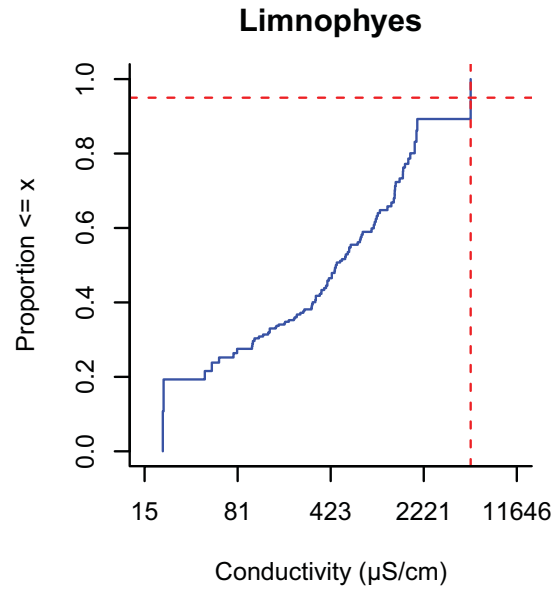
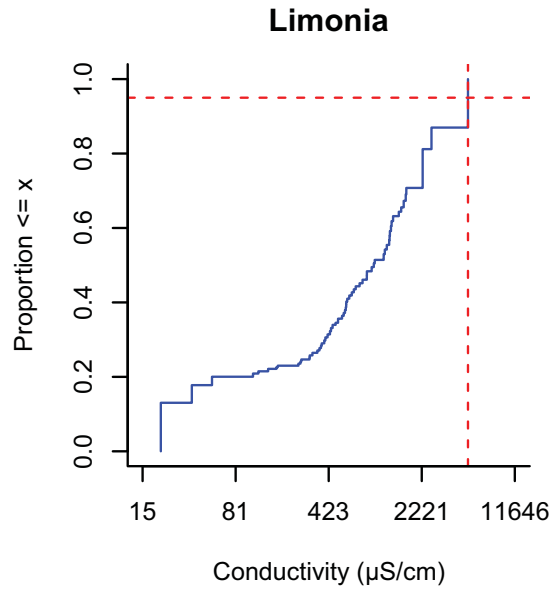


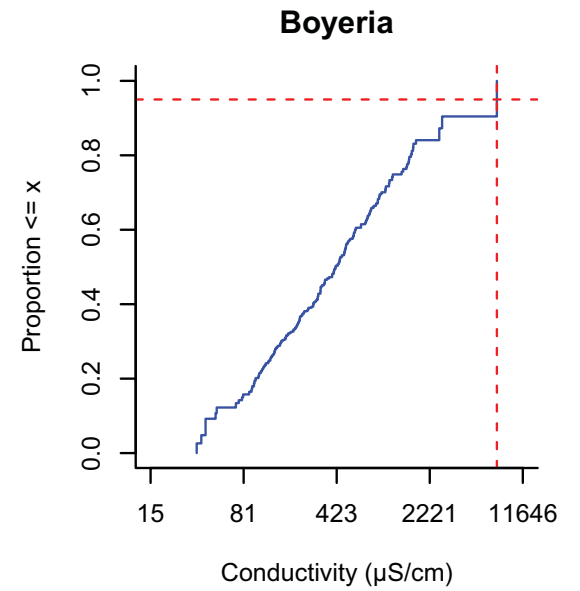
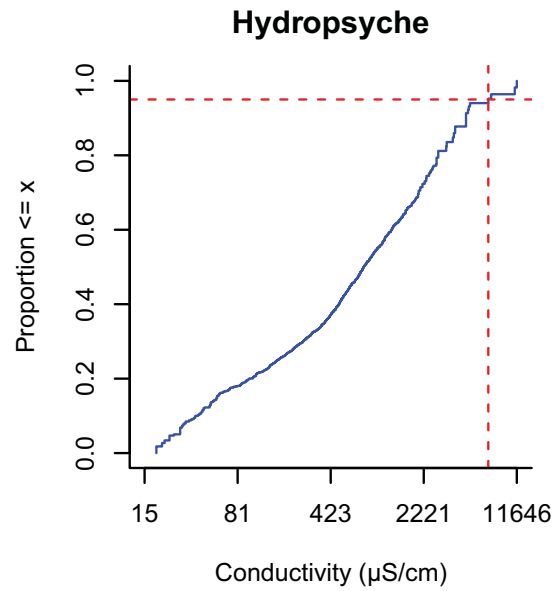
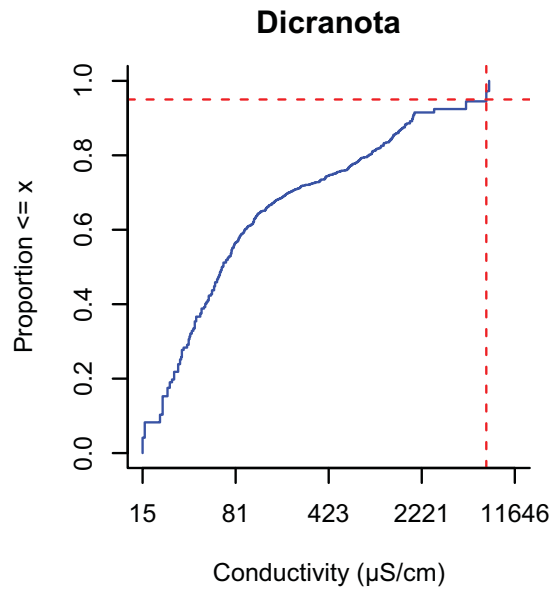
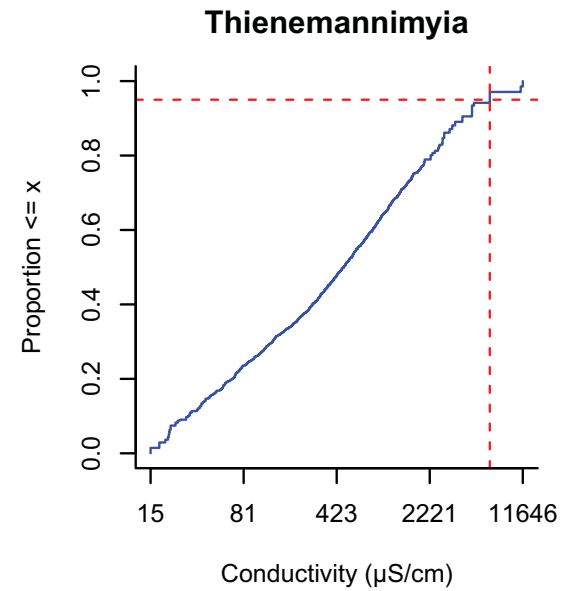
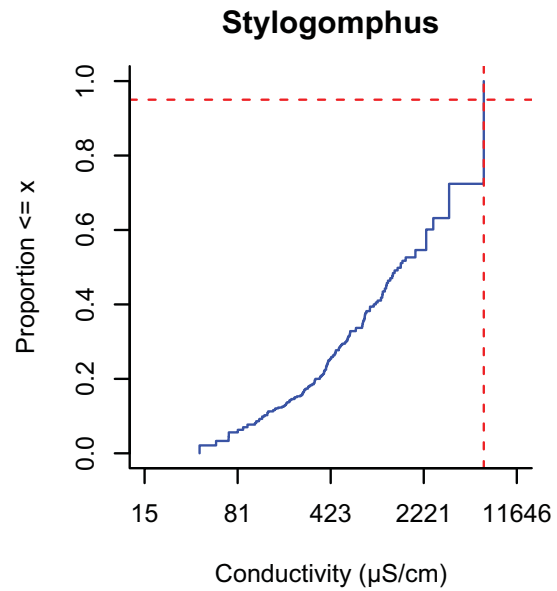
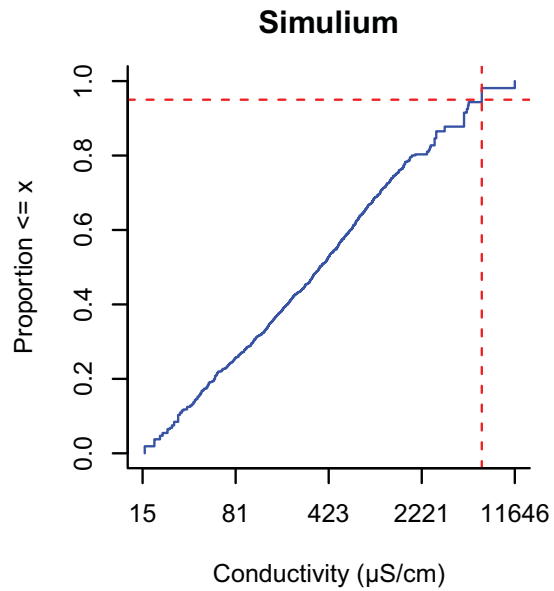


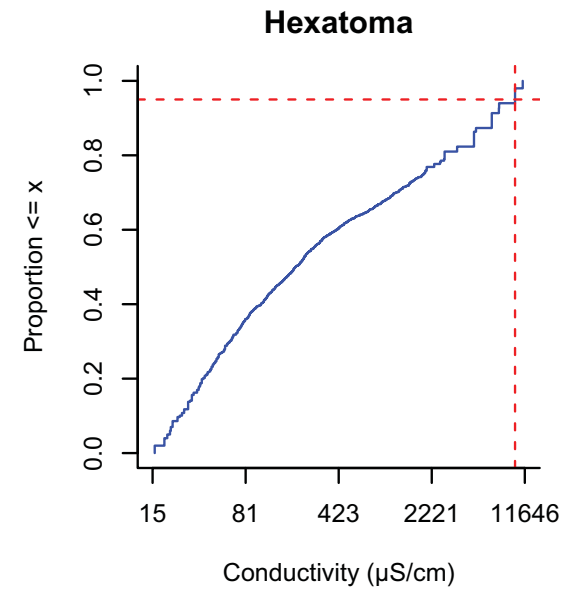
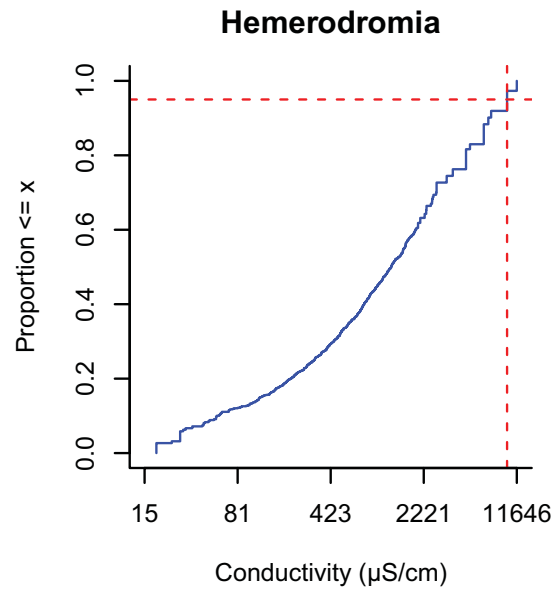
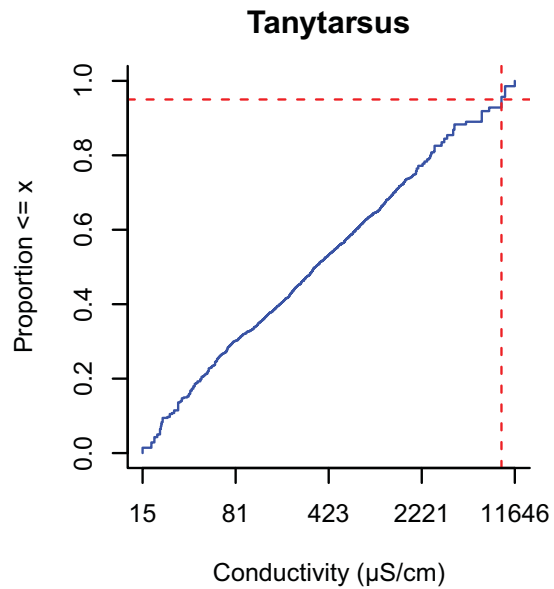
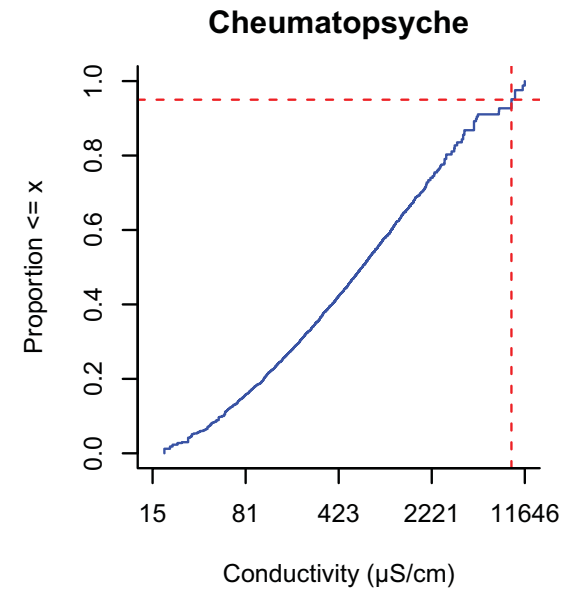
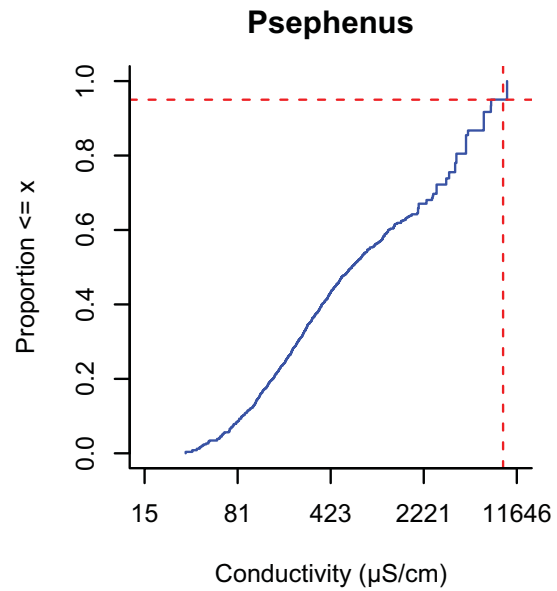
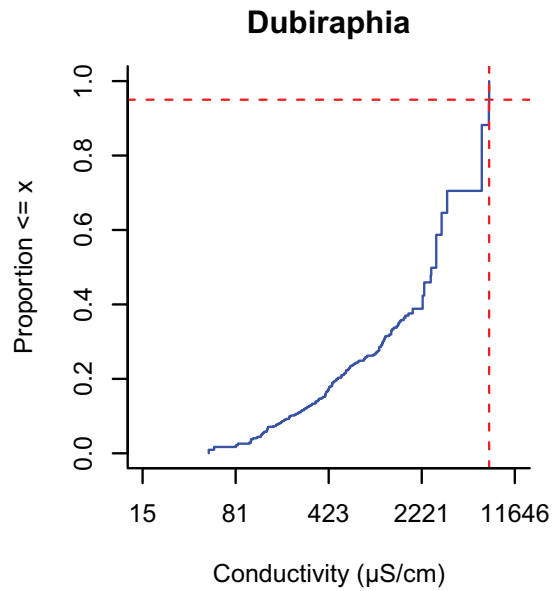


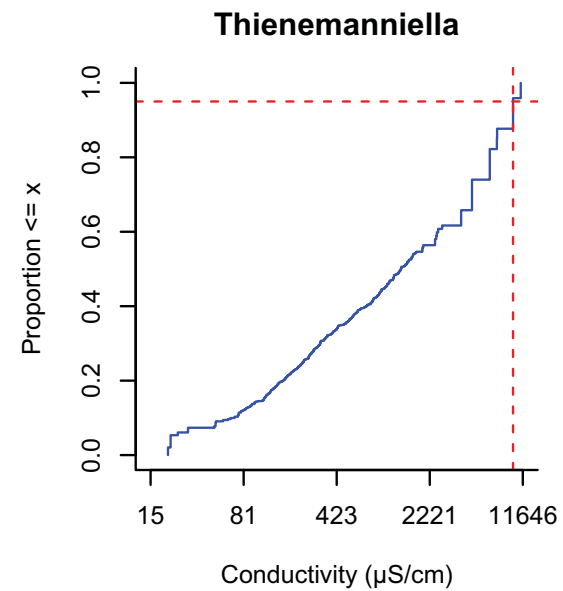
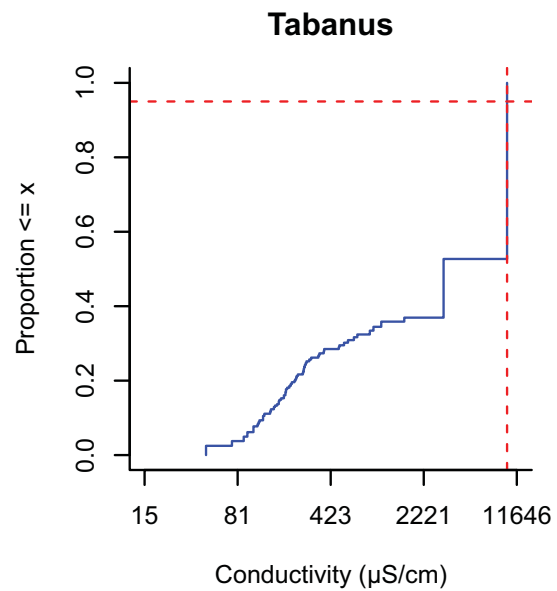
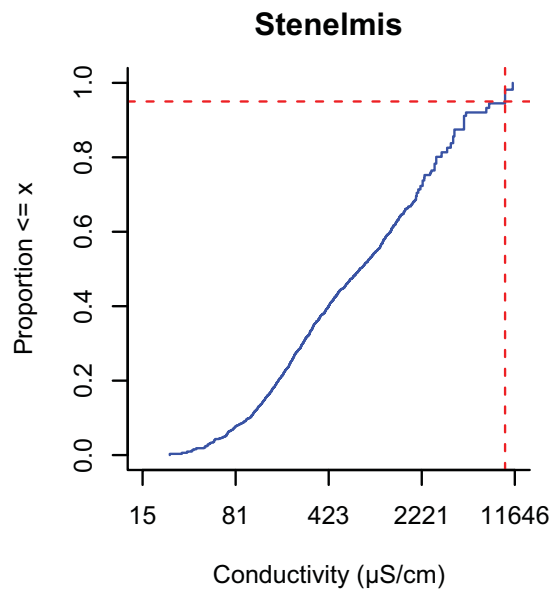
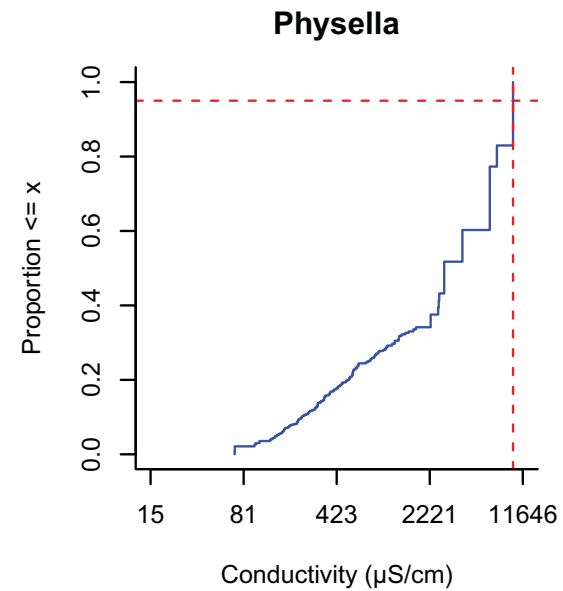
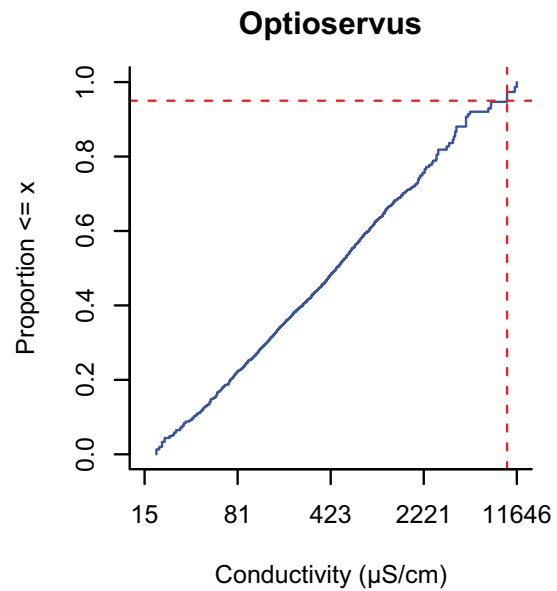
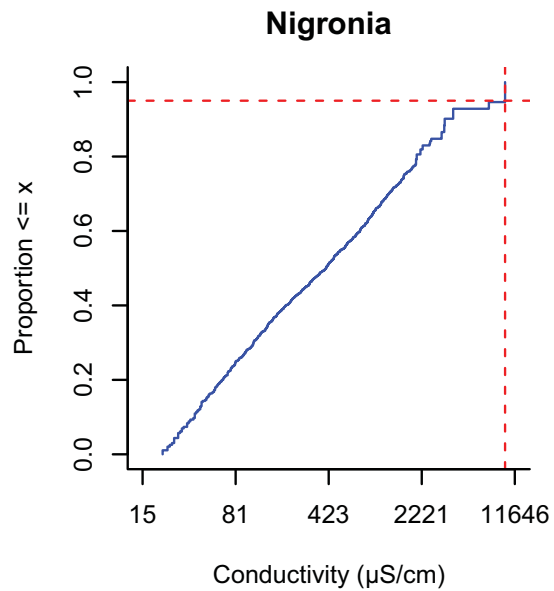


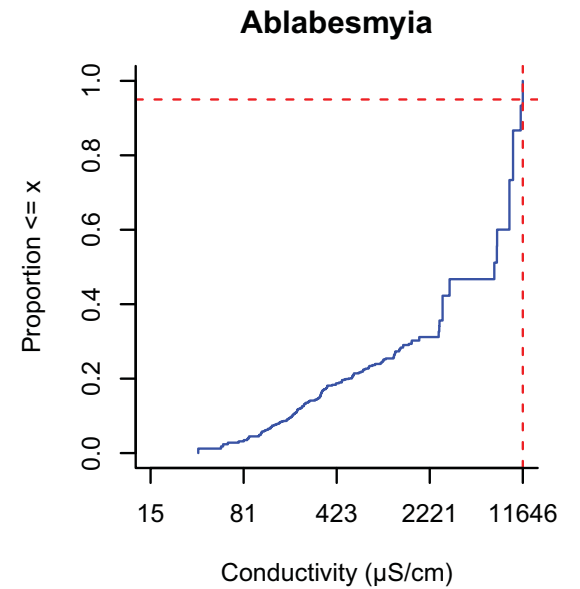
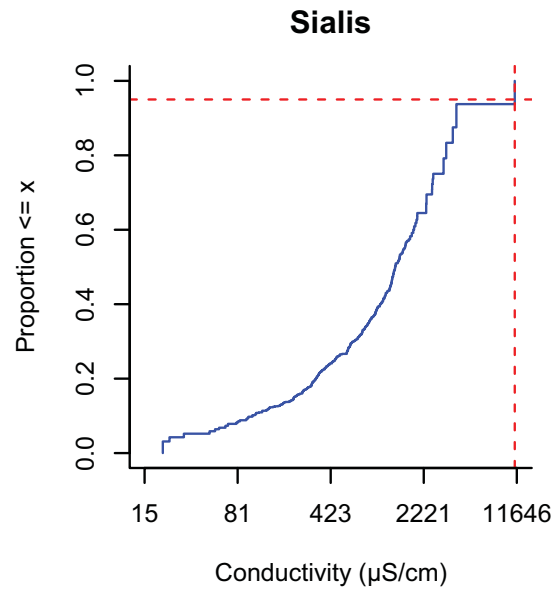
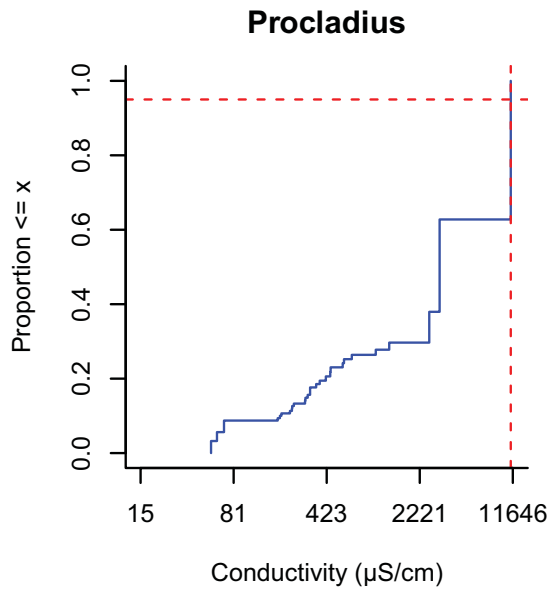
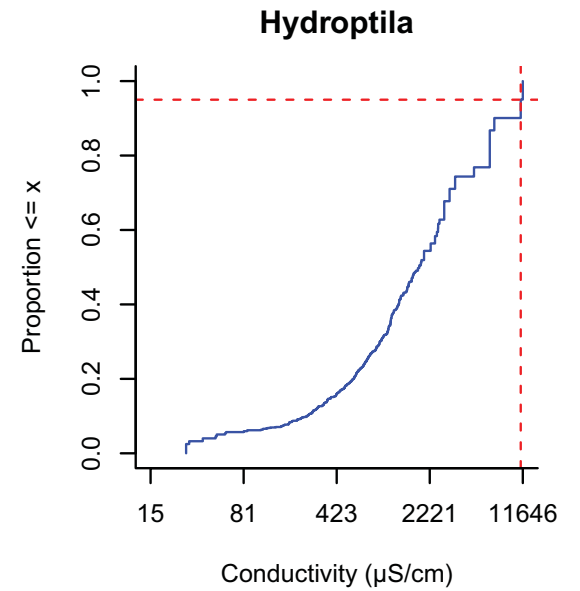
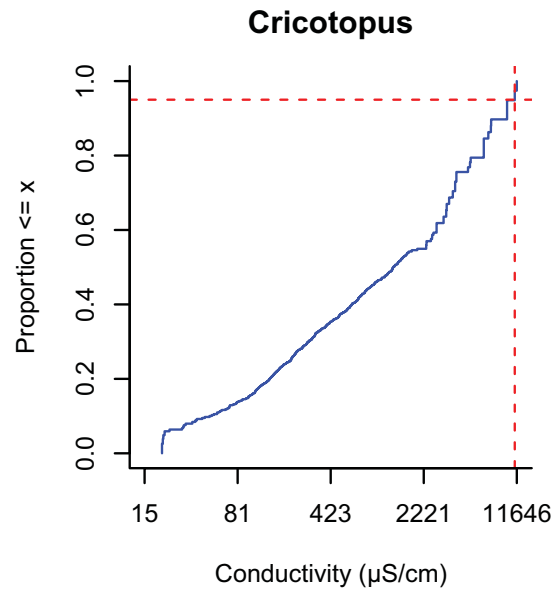
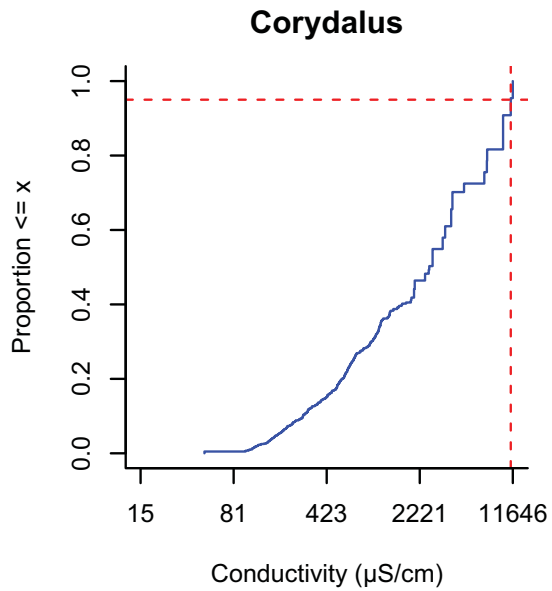


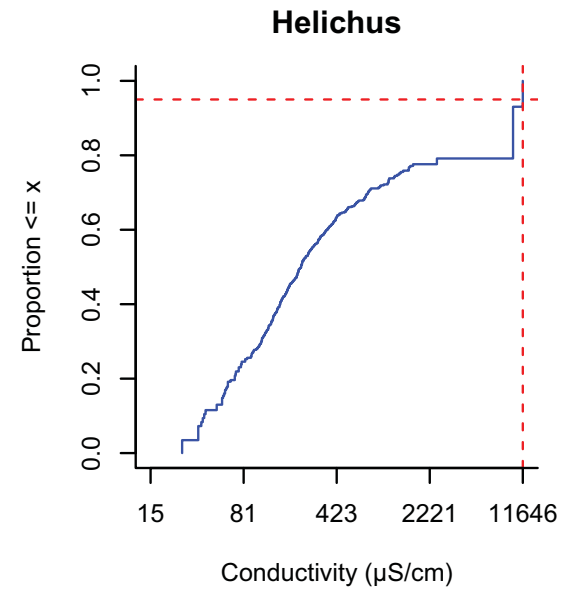
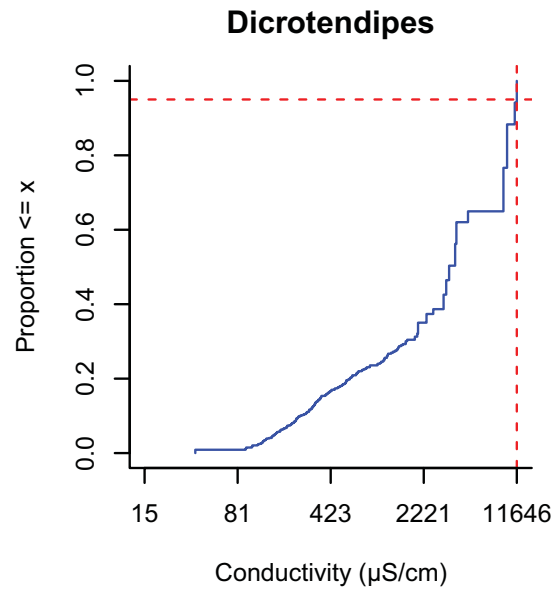
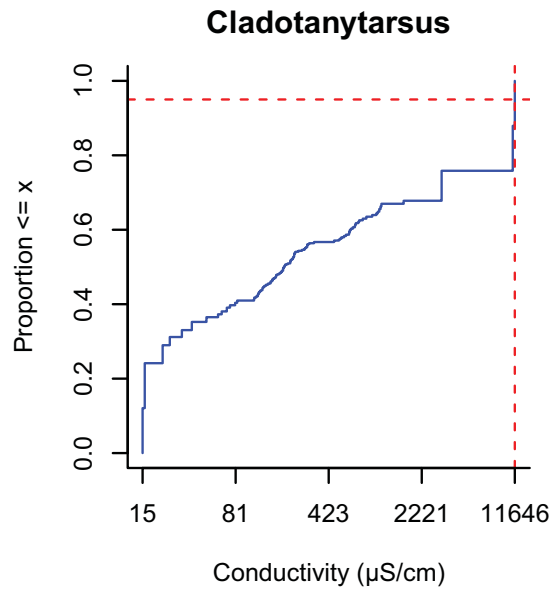
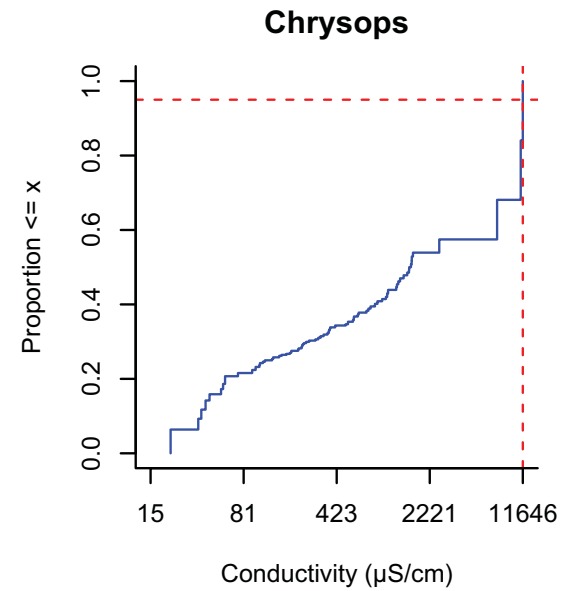
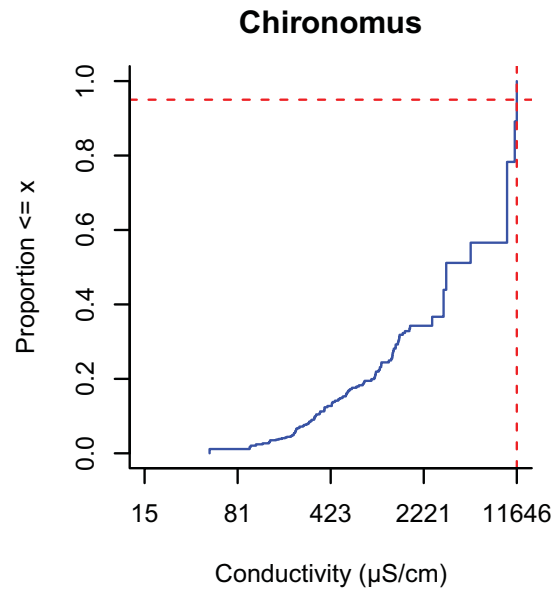
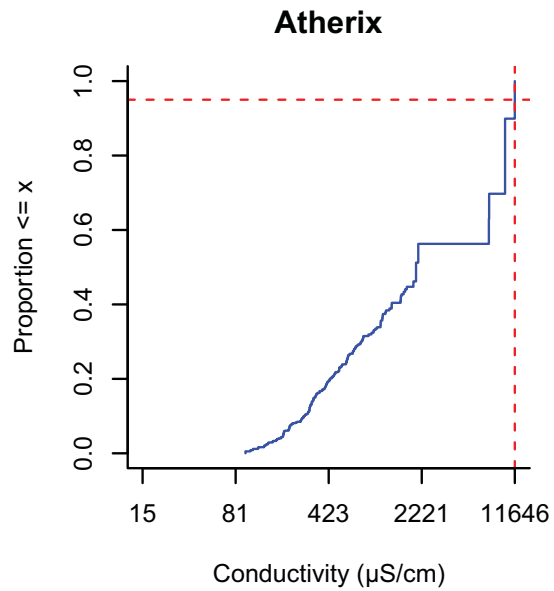


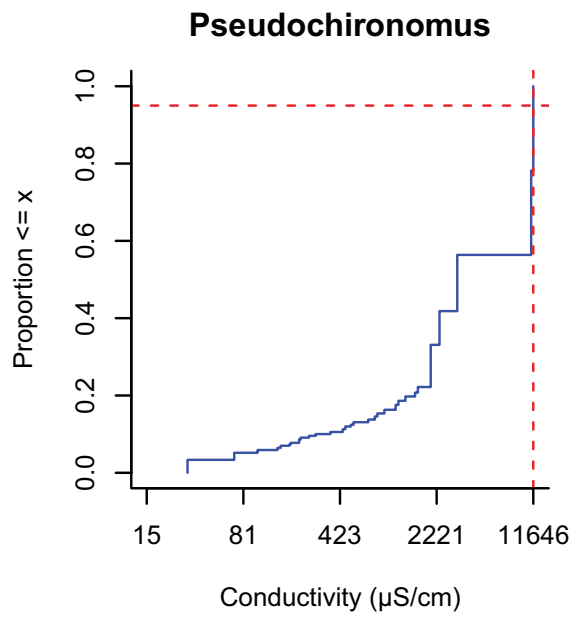












APPENDIX G
VALIDATION OF METHOD
USING FIELD DATA TO DERIVE AMBIENT WATER QUALITY BENCHMARK FOR
CONDUCTIVITY USING A KENTUCKY DATA SET

ABSTRACT

The method for developing the aquatic life benchmark for conductivity was validated by developing extirpation concentration (XC_{95}) and hazardous concentration (HC_{05}) values using a data set independently collected by the Kentucky Division of Water (KDOW) and comparing results with those found using the larger West Virginia database. Because samples were also drawn from the Central Appalachians (Ecoregion 69) and Western Allegheny Plateau (Ecoregion 70), the two data sets were expected to give similar results. Some differences were expected due to the different collection and taxa identification protocols, shorter sampling window, inclusion of the Southwestern Appalachians (Ecoregion 68), and the fewer number of samples in the Kentucky data set. Nevertheless, the HC_{05} value was 282 $\mu\text{S}/\text{cm}$ for the full Kentucky data set, which is very close to the West Virginia result.

G.1. DATA SET SELECTION

The Southwestern Appalachians (68), Central Appalachia (69), and Western Allegheny Plateau (70) ecoregions were selected for validation, because they are physiographically similar to Ecoregions 69 and 70 in West Virginia (U.S. EPA, 2000; Omernik, 1987; Woods et al., 1996) (see Figures G-1 and G-2). Although the Kentucky data set is smaller than the West Virginia data set, it was judged to be large enough for validation of the method (see Section 3.5). These regions have heavily forested areas as well as extensive areas developed for coal mining, and, as in West Virginia, conductivity has been implicated as a cause of biological impairment in the three Kentucky ecoregions, which were judged to be similar within the state of Kentucky in terms of water quality, resident biota, and sources of conductivity (Pond 2004, 2010). Background conductivity was not estimated due to the lack of designation of reference sites in the data set or a probabilistic sample of sufficient size. However, the 25th centile of the entire data set, which includes impaired sites, is 118 $\mu\text{S}/\text{cm}$ (see Figure G-3). Although not a background estimate, it does indicate conductivity levels are generally low in these ecoregions and is within the range of background values for West Virginia.

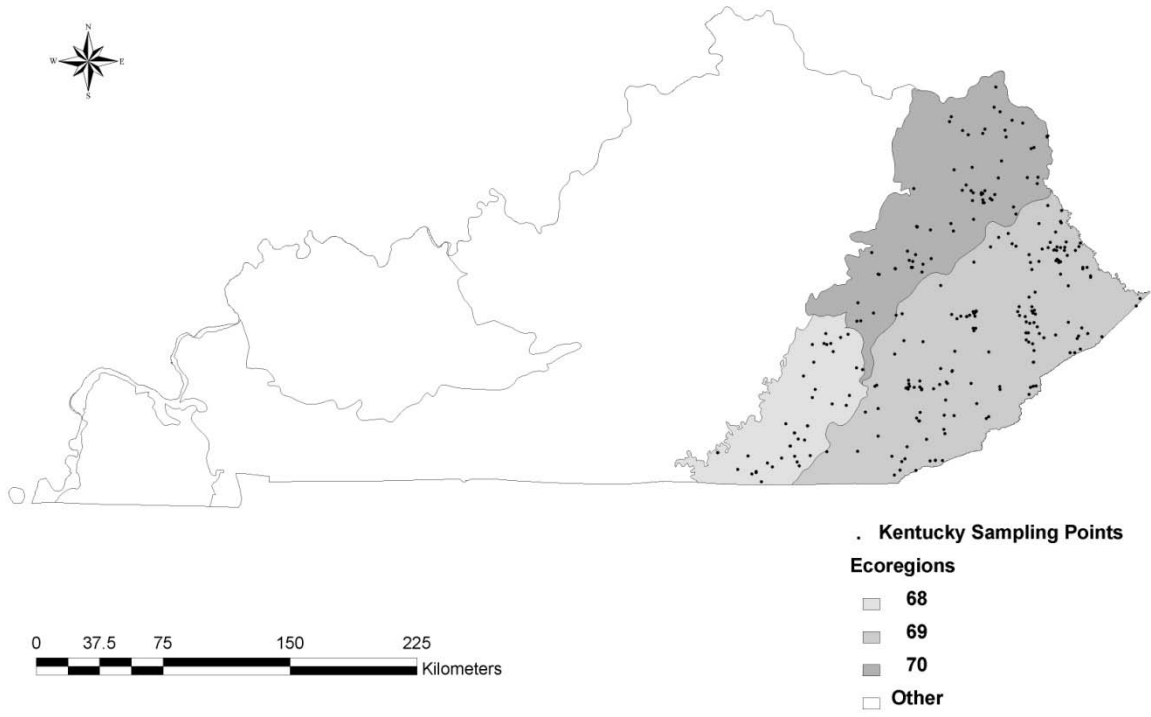


Figure G-1. Location of Southern Appalachia (68), Central Appalachia (69), and Allegheny Plateau (70) and sampling points.

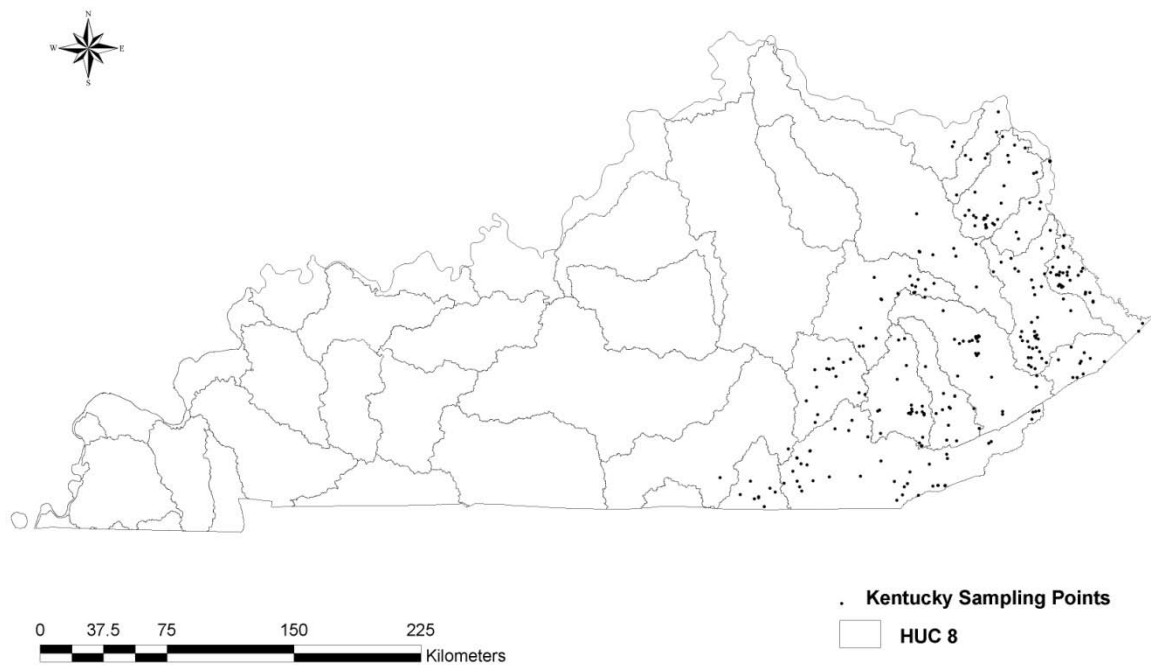


Figure G-2. Location of sampling points used to develop the Kentucky HC₀₅, shown with 8-digit HUC catchments.

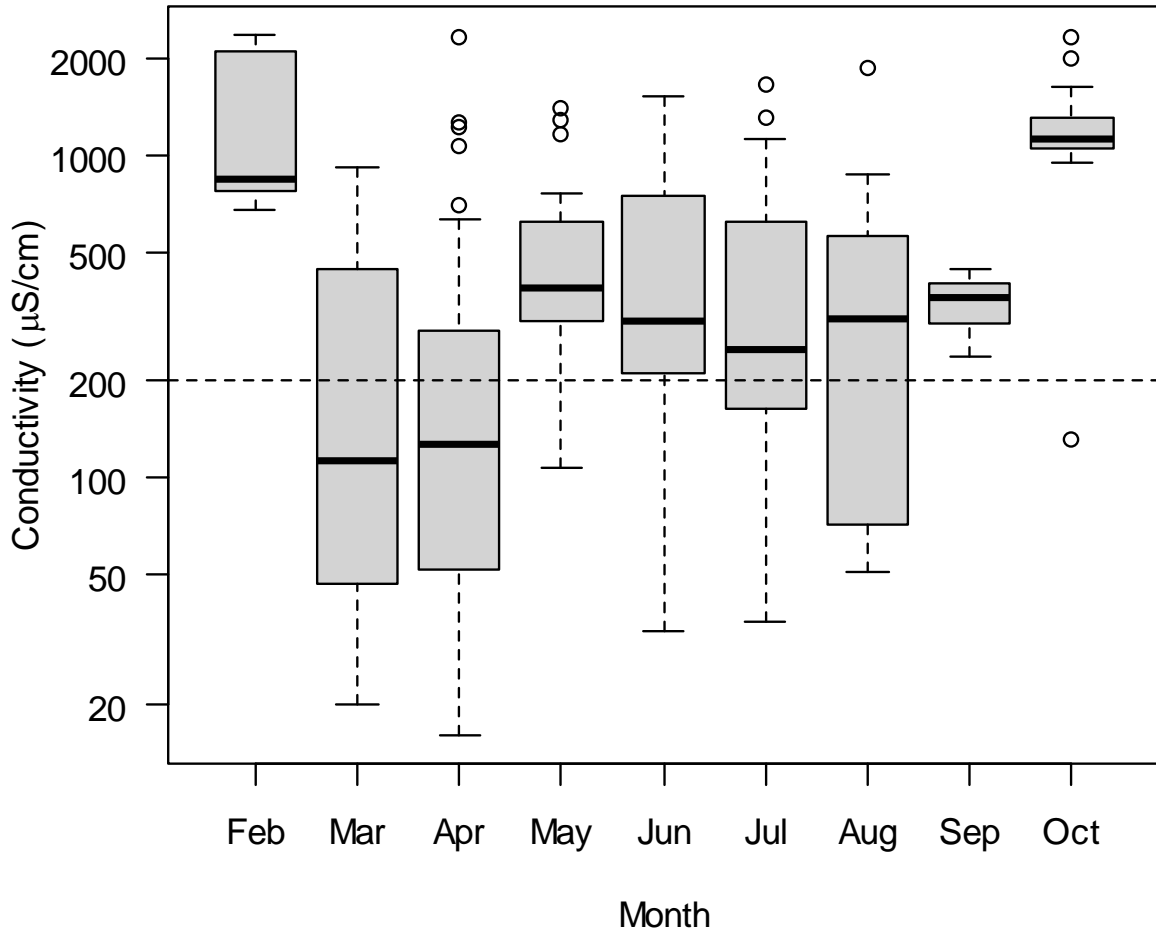


Figure G-3. Box plot showing seasonal variation of conductivity ($\mu\text{S}/\text{cm}$) from the data set used to develop the Kentucky HC_{05} . A total of 291 samples from 1998–2004 from Ecoregions 68, 69, and 70 in Kentucky are represented. Small sample sizes and targeted sampling could obscure seasonal patterns, if any.

G.2. DATA SOURCES

All data used in this study were taken from the Kentucky Division of Water, Water Quality Branch database, Ecological Data Application System (KY EDAS). Chemical, physical, or biological samples were collected from 274 distinct locations during February–October from 1998–2004 (see Table G-1). Like the West Virginia Department of Environmental Protection (WVDEP), the KDOW obtains biological data from both probability biosurvey and targeted ambient biological monitoring programs. The probability biosurvey program provides a condition assessment of the overall biological and water quality conditions for both basin and state levels. Targeted ambient biological monitoring involves intensive data-collection efforts

Table G-1. Number of samples with reported genera and conductivity. Number of samples is presented for each month and ecoregion

Ecoregion	Month												Total
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	
68		0	0	10	0	6	18	2	3	0			39
69		7	14	44	16	16	42	18	0	25			182
70		0	9	21	2	17	21	0	0	0			70
													291

for streams of interest as reference or impaired sites or for other reasons. Most sites have been sampled once during February to October. Quality assurance and standard procedures are described by KDOW (2008). Briefly, KDOW follows similar field and laboratory quality assurance methods as WVDEP. Macroinvertebrates are collected from mid-riffle/runs. Although KDOW also collects a separate sample from multihabitat survey, for consistency, only the riffle samples are quantified for this analysis. Four, 0.25-m² samples were collected mid-riffle in a 100-m sampling reach using a 1-m-wide, 600-µm mesh net. The four samples were composited, and all macroinvertebrates were removed and stored in 95%-ethanol. Samples were composited, and macroinvertebrates were identified to the lowest possible taxonomic level. A notable difference in the WVDEP and KDOW methods is that KDOW picks the entire sample in the laboratory, as opposed to WVDEP’s fixed count of 200 organisms.

All contracted chemical analyses and macroinvertebrate identifications followed internal quality control and quality assurance protocols. This is a well-documented, regulatory database. The quality assurance was judged to be excellent based on the database itself, supporting documentation, and experience of EPA Region 4 personnel.

G.3. DATA SET CHARACTERISTICS

Biological sampling usually occurred once during February–October with the KDOW (1998–2004) wadeable sampling protocol. The Kentucky data set was treated in the same way as the West Virginia data used to derive the aquatic life benchmark for conductivity. A sample was excluded from calculations if (1) it lacked a conductivity measurement, or (2) the pH was low. XC₉₅ values were calculated for genera that occurred at ≥25 sampling locations. Organisms were not included unless identified to the genus level. Reference sites were not identified in the data set, so no genera were excluded in the species sensitivity distribution (SSD). Future analyses should identify invasive and opportunistic genera for a benchmark for

Kentucky. Repeat biological samples from the same location at the same time (or within a month) were excluded, but samples collected in different months/years were not excluded from the data set. These repeat biological samples from different years were retained and represented about 8% of the samples. All samples were from wadeable streams. No measures of individual ions were available, so no sites with high chloride and low sulfate were identified or removed from the Kentucky data set. Eighty-five percent of the 104 genera used to develop the SSD for Kentucky also occurred in the West Virginia SSD. Genera from both states were judged to be similarly susceptible to the effects of conductivity after exploratory analysis (see also Sections A.2.4.1 and A.2.1.2). Conductivity ranged from 16–2,390 $\mu\text{S}/\text{cm}$ for the Kentucky data set (see Table G-2, Figure G-4) and 15–11,646 $\mu\text{S}/\text{cm}$ for the West Virginia data set.

Table G-2. Summary statistics of the measured water quality parameters for the Kentucky data set

		Min	25 th	50 th	75 th	Max	Mean ^a
Specific conductance	291	16	118.5	272.1	674.4	2,390	265.4
pH	291	6.03	7.1	7.5	7.92	9.26	7.49
Total HAB score	291	56	115	138	161	191	136.3
Embeddedness	291	0	8	13	16	19	12

^aConductivity reported as geometric mean.

In the Kentucky database, 359 benthic invertebrate genera were identified. Of the 359 genera collected, 104 occurred in at least 25 sampling locations in Ecoregions 68, 69, and 70 (see Appendix H). All genera used to construct the SSD occurred in all three ecoregions.

Because of the data distributions, not all 95th centiles correspond to extirpation, and some imprecisely estimate the extirpation threshold. The following rules were applied to the XC_{95} values. If the generalized additive model (GAM) mean curve at maximum conductivity is approximately = 0 (<1% of the maximum modeled probability), then the XC_{95} is listed without qualification. If the GAM mean curve at maximum conductivity is >0 but the GAM lower confidence limit is approximating to 0, the value is listed as approximate (~). If the GAM lower confidence limit is >0, then the XC_{95} is listed as greater than (>) the 95th centile. All models fits and the scatter of points were also visually inspected for anomalies, and if the model poorly fit the data, the uncertainty level was increased to either (~) or (>). This procedure was applied to

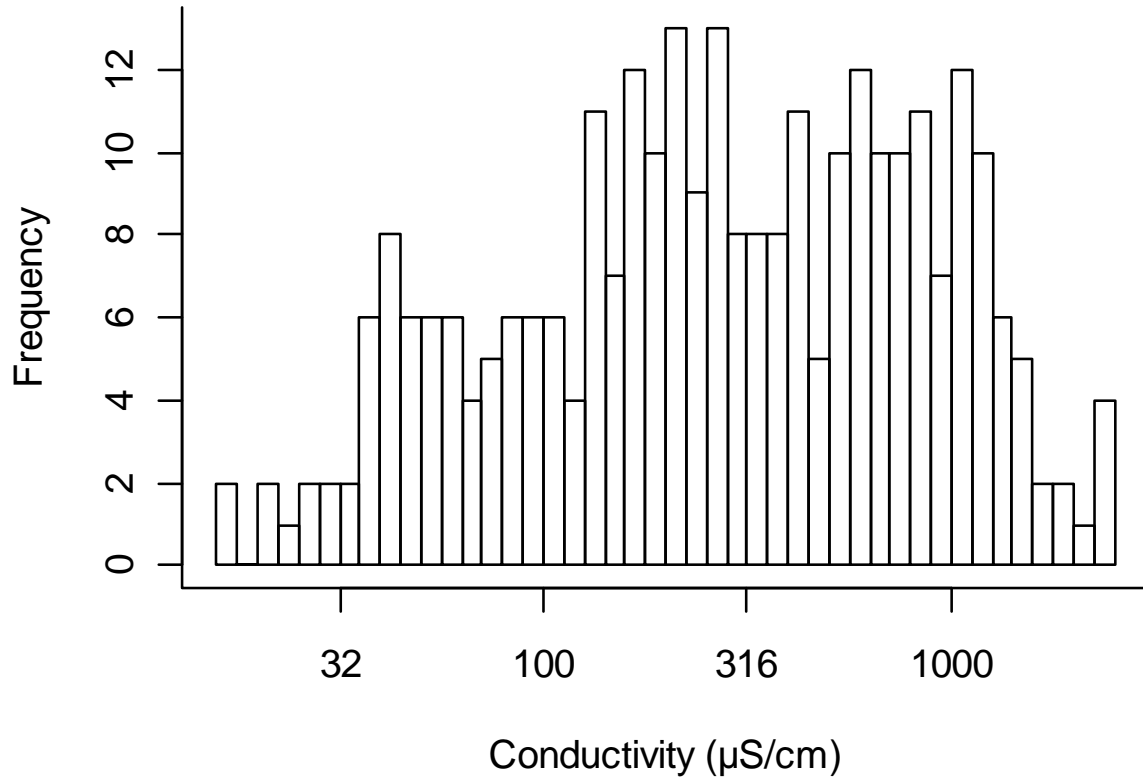


Figure G-4. Histograms of the frequencies of observed conductivity values in samples from Ecoregions 68, 69, and 70 in Kentucky sampled from 1998–2004.

plots in Appendix I, and the XC_{95} values appear in Appendix H. Also these models were used to evaluate when genera began to decline as evidence of alteration and sufficiency in Appendix A.

Many genera are marked as approximate because the Kentucky data set is small, the XC_{95} models are based on a smaller number of occurrences, and the maximum conductivity measured is lower than in the West Virginia data set. The assignment of (>) and (~) does not affect the HC_{05} but alerts users of the uncertainty of the XC_{95} values for other uses such as comparison with toxicity test results or with results from other geographic regions.

G.4. RESULTS

Appendix H lists the genera used to construct the SSD from the Kentucky sample and their corresponding XC_{95} values. The cumulative distribution functions (CDFs) used to develop them can be found in Appendix J. The full SSD is shown in Figure G-5, and an enlargement of the lower half of the model is shown in Figure G-6. Despite the differences in sampling method and geographic location, the HC_{05} values were similar: 282 $\mu\text{S}/\text{cm}$ for Kentucky compared to

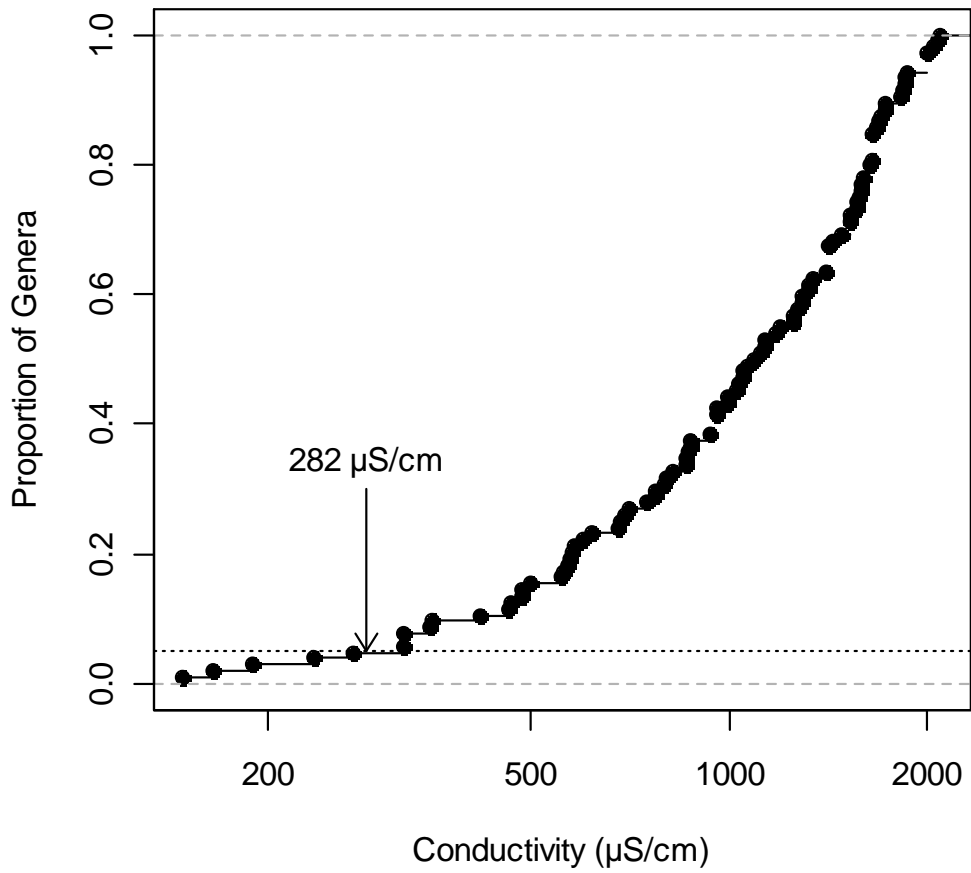


Figure G-5. Species sensitivity distribution (SSD) for Kentucky. A total of 104 genera are included in the SSD. The HC_{05} is the conductivity at the intercept of the CDF while the horizontal line at the 5th centile is 282 µS/cm.

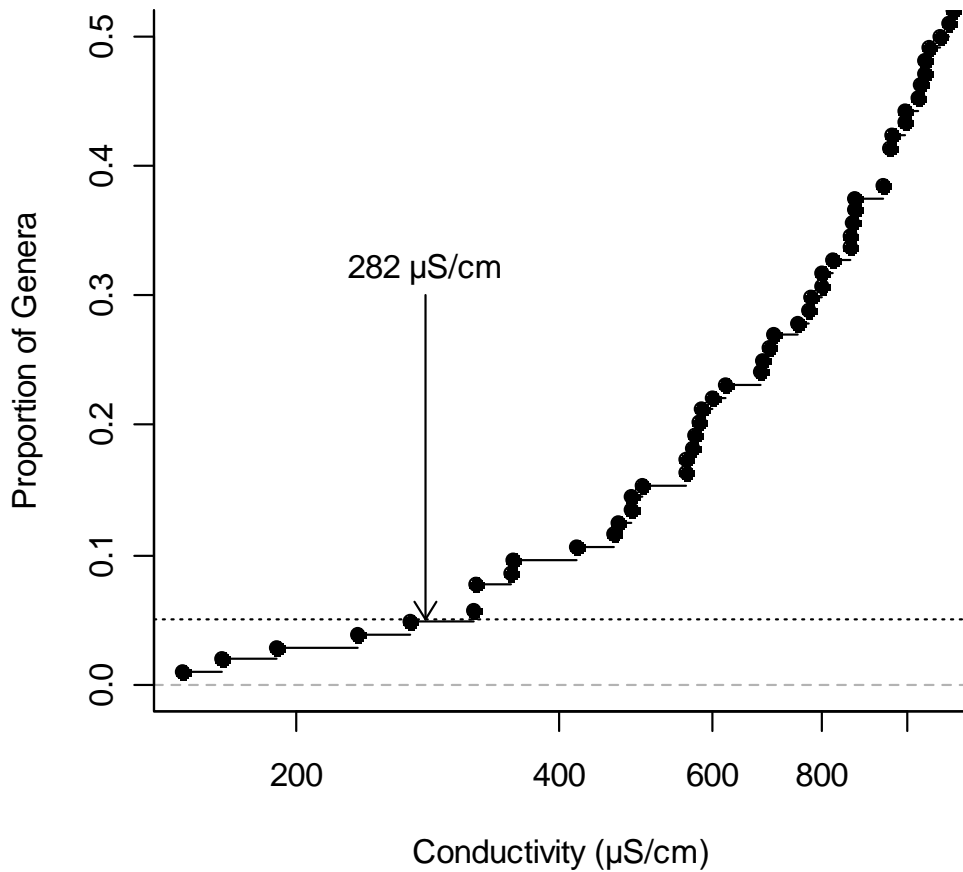


Figure G-6. Species sensitivity distribution (SSD) for Kentucky. Only the lower half of the total genera are shown to better discriminate the points in the left side of the SSD.

295 µS/cm for West Virginia (see Figures G-2 and G-3, Table G-3). The 95% confidence bounds for the Kentucky HC_{05} are 169 and 380 µS/cm, which overlap with the West Virginia data set's 95% confidence bounds of 228 and 303 µS/cm. Genera that exhibited a decreasing occurrence with increasing conductivity were among those with the lowest XC_{95} values in both states. Table G-4 shows the 10 lowest XC_{95} values for both West Virginia and Kentucky samples. The 5th centile occurs near the eighth genus for West Virginia samples and fifth genus for Kentucky samples.

Table G-3. HC₀₅ values for Kentucky and West Virginia data sets

	Kentucky	West Virginia
HC ₀₅	282 μ S/cm	295 μ S/cm
95% CI	169–380	228–303
Months represented	February–October	January–December
Sample	291	2,210
Genera in SSD	104	163

Table G-4. Comparison of the sensitive genera and XC₉₅ values

	West Virginia Genus	West Virginia XC₉₅	Kentucky Genus	Kentucky XC₉₅
1		~121		149
2		121		165
3		230		190
4		246		235
5		251		270
6		255		320
7		260		321
8		295		321
9		297		353
10		299		354

G.5. CONCLUSIONS

Based on the similar results, EPA judged the field-based method to be robust. The same aquatic life benchmark appears to be applicable to West Virginia and Kentucky streams in Ecoregions 68, 69, and 70. However, analysis of a larger statewide data set, removal of nonreference taxa, and verification of the basic water chemistry for the region are recommended.

REFERENCES

KDOW (Kentucky Division of Water). (2008) Standard methods for assessing biological integrity of surface waters in Kentucky. Commonwealth of Kentucky Environmental and Public Protection Cabinet Department for Environmental Protection, Division of Water February 2008, Revision 3. 120 pp. Available online at <http://www.water.ky.gov/sw/swmonitor/sop/> and at http://www.water.ky.gov/NR/rdonlyres/714984BB-54F8-46B9-AF27-290EF7A6D5CE/0/BiologicalSOPMainDocument03_08.pdf (accessed 12/19/2009).

Omernik, JM. (1987) Ecoregions of the conterminous United States. *Ann Assoc Am Geograph* 77:118–125.

Pond, GJ. (2004) Effects of surface mining and residential land use on headwater stream biotic integrity in the eastern Kentucky coalfield region. Kentucky Department of Environmental Protection, Division of Water, Frankfort, KY.

Pond, GJ. (2010) Patterns of Ephemeroptera taxa loss in Appalachian headwater streams (Kentucky, USA). *Hydrobiologia*. 641(1):185–201.

Stevens, DL, Jr.; Olsen, AR. (2004) Spatially balanced sampling of natural resources. *J Am Stat Assoc* 99(465):262–278.

U.S. EPA (U.S. Environmental Protection Agency). (2000) Nutrient criteria technical guidance manual: rivers and streams. Office of Water, Office of Science and Technology, Washington, DC. EPA/822/B-00/002. Available online at <http://www.epa.gov/waterscience/criteria/nutrient/guidance/rivers/rivers-streams-full.pdf>.

Woods, AJ; Omernik, JM; Brown, DD; et al. (1996) Level III and IV ecoregions of Pennsylvania and the Blue Ridge Mountains, the Ridge and Valley, and the Central Appalachians of Virginia, West Virginia, and Maryland. U.S. Environmental Protection Agency, National Health and Environmental Effects Research Laboratory, Corvallis, OR. EPA/600R-96/077. 50 pp.

APPENDIX H EXTIRPATION CONCENTRATION VALUES FOR GENERA IN A KENTUCKY DATA SET

ABSTRACT

The purpose of Appendix H is to provide the reader with a list of the extirpation concentration (XC_{95}) values used to develop the species sensitivity distribution and the hazardous concentration (HC_{05}) for Kentucky. Genera are ordered alphabetically (see Table H-1). The numbers of occurrences in the data set are noted in the right-hand column. Genera highlighted in gray do not occur at West Virginia reference locations, but were included in the Kentucky species sensitivity distribution (SSD). If they were removed the hazardous concentration, HC_{05} would be slightly lower.

Not all 95th centiles correspond to extirpation, and some imprecisely estimate the extirpation threshold. The following rules were applied to the XC_{95} values using the fitted curve and the confidence bounds from the plots in Appendix I. If the generalized additive model (GAM) mean curve at maximum conductivity is approximately equal to 0 (defined as less than 1% of the maximum modeled probability), then the XC_{95} value is listed without qualification. If the GAM mean curve at maximum conductivity is >0 but the lower confidence limit is approximating to 0 ($<1\%$ of the maximum mean modeled probability), then the XC_{95} value is listed as approximate (~). If the GAM lower confidence limit is >0 , then the XC_{95} value is listed as greater than ($>$) the 95th centile. All model fits and scatter of points were also visually inspected for anomalies, and if the model poorly fit the data, the uncertainty level was increased to either (~) or ($>$).

The assignation of (~) and ($>$) does not affect the HC_{05} . They are provided to alert users to the uncertainty of some XC_{95} values for other uses such as comparison with toxicity test results or with results from other geographic regions.

Table H-1. Extirpation concentration and sample size from Kentucky data set. Highlighted genera are not found at WV reference sites but were included in the SSD for Kentucky. XC_{95} values reported without a preceding symbol indicate evidence of extirpation within the tested range. XC_{95} values preceded by a (~) or (>) indicate extirpation with greater uncertainty or extirpation at a level above the reported value. Genera highlighted in gray do not occur at West Virginia reference locations.

Order	Family	Genus	XC_{95}	<i>N</i>
Diptera	Chironomidae	<i>Ablabesmyia</i>	>1,410	43
Ephemeroptera	Baetidae	<i>Acentrella</i>	>619	98
Plecoptera	Perlidae	<i>Acroneuria</i>	>697	105
Ephemeroptera	Ameletidae	<i>Ameletus</i>	>579	69
Plecoptera	Nemouridae	<i>Amphinemura</i>	>1,269	107
Coleoptera	Elmidae	<i>Ancyronyx</i>	798	30
Diptera	Tipulidae	<i>Antocha</i>	>958	49
Odonata	Coenagrionidae	<i>Argia</i>	>1,410	51
Diptera	Athericidae	<i>Atherix</i>	>1,650	61
Ephemeroptera	Baetidae	<i>Baetis</i>	>1,410	170
Odonata	Aeshnidae	<i>Boyeria</i>	>1,318	92
Ephemeroptera	Caenidae	<i>Caenis</i>	>1,410	85
Odonata	Calopterygidae	<i>Calopteryx</i>	>2,082	35
Decapoda	Cambaridae	<i>Cambarus</i>	>1,090	157
Trichoptera	Hydropsychidae	<i>Ceratopsyche</i>	>1,577	102
Trichoptera	Hydropsychidae	<i>Cheumatopsyche</i>	>1,630	230
Trichoptera	Philopotamidae	<i>Chimarra</i>	>2,000	90
Diptera	Chironomidae	<i>Chironomus</i>	>1,670	31
Ephemeroptera	Heptageniidae	<i>Cinygmula</i>	165	39
Veneroida	Corbiculidae	<i>Corbicula</i>	>1,829	84
Megaloptera	Corydalidae	<i>Corydalus</i>	>1,650	121
Diptera	Chironomidae	<i>Cricotopus</i>	>2,037	98
Diptera	Chironomidae	<i>Cryptochironomus</i>	>1,037	27
Diptera	Chironomidae	<i>Diamesa</i>	>2,074	54

Order	Family	Genus	XC₉₅	N
Diptera	Tipulidae	<i>Dicranota</i>	>484	25
Diptera	Chironomidae	<i>Dicrotendipes</i>	>1,437	29
Coleoptera	Gyrinidae	<i>Dineutus</i>	>874	45
Ephemeroptera	Baetidae	<i>Dipheter</i>	190	25
Trichoptera	Hydropsychidae	<i>Diplectrona</i>	>958	102
Plecoptera	Perlodidae	<i>Diploperla</i>	>997	35
Trichoptera	Philopotamidae	<i>Dolophilodes</i>	270	31
Ephemeroptera	Ephemerellidae	<i>Drunella</i>	321	37
Coleoptera	Elmidae	<i>Dubiraphia</i>	>1,650	86
Plecoptera	Perlidae	<i>Eccoptura</i>	>1,649	31
Lumbriculida	Lumbriculidae	<i>Eclipidrilus</i>	>1,294	92
Coleoptera	Psephenidae	<i>Ectopria</i>	>582	66
Neotaenioglossa	Pleuroceridae	<i>Elimia</i>	~1,131	33
Odonata	Coenagrionidae	<i>Enallagma</i>	>959	31
Ephemeroptera	Heptageniidae	<i>Epeorus</i>	321	65
Ephemeroptera	Ephemeridae	<i>Ephemera</i>	~559	42
Ephemeroptera	Ephemerellidae	<i>Ephemerella</i>	~467	70
Diptera	Chironomidae	<i>Eukiefferiella</i>	>1,842	54
Ephemeroptera	Ephemerellidae	<i>Eurylophella</i>	>499	84
Basommatophora	Ancylidae	<i>Ferrissia</i>	>872	29
Odonata	Gomphidae	<i>Gomphus</i>	>1,063	36
Plecoptera	Chloroperlidae	<i>Haploperla</i>	485	38
Coleoptera	Dryopidae	<i>Helichus</i>	>1,050	148
Diptera	Empididae	<i>Hemerodromia</i>	>2,000	123
Diptera	Tipulidae	<i>Hexatoma</i>	>1,134	106
Trichoptera	Hydropsychidae	<i>Hydropsyche</i>	>1,650	161
Trichoptera	Hydroptilidae	<i>Hydroptila</i>	>1,680	58
Ephemeroptera	Isonychiidae	<i>Isonychia</i>	>1,524	132
Plecoptera	Perlodidae	<i>Isoperla</i>	>1,176	81

Order	Family	Genus	XC₉₅	N
Odonata	Gomphidae	<i>Lanthus</i>	>1,564	34
Trichoptera	Lepidostomatidae	<i>Lepidostoma</i>	149	30
Ephemeroptera	Heptageniidae	<i>Leucrocuta</i>	>686	45
Plecoptera	Leuctridae	<i>Leuctra</i>	>1,029	131
Isopoda	Asellidae	<i>Lirceus</i>	~958	35
Odonata	Corduliidae	<i>Macromia</i>	~772	27
Coleoptera	Elmidae	<i>Macronychus</i>	>1,722	54
Diptera	Chironomidae	<i>Micropsectra</i>	~462	25
Diptera	Chironomidae	<i>Microtendipes</i>	>681	58
Diptera	Chironomidae	<i>Natarsia</i>	>1,630	45
Trichoptera	Uenoidae	<i>Neophylax</i>	353	73
Megaloptera	Corydalidae	<i>Nigronia</i>	>1,197	153
Trichoptera	Leptoceridae	<i>Oecetis</i>	>1,337	31
Coleoptera	Elmidae	<i>Optioservus</i>	>1,563	178
Decapoda	Cambaridae	<i>Orconectes</i>	>1,291	115
Diptera	Chironomidae	<i>Orthocladius</i>	>1,480	50
Coleoptera	Elmidae	<i>Oulimnius</i>	320	31
Ephemeroptera	Leptophlebiidae	<i>Paraleptophlebia</i>	~420	76
Diptera	Chironomidae	<i>Parametriocnemus</i>	>1,583	185
Plecoptera	Peltoperlidae	<i>Peltoperla</i>	>1,520	37
Plecoptera	Perlidae	<i>Perlesta</i>	>1,399	51
Basommatophora	Physidae	<i>Physella</i>	>1,856	52
Ephemeroptera	Baetidae	<i>Plauditus</i>	~703	55
Trichoptera	Polycentropodidae	<i>Polycentropus</i>	>570	82
Diptera	Chironomidae	<i>Polypedilum</i>	>1,251	158
Ephemeroptera	Baetidae	<i>Procloeon</i>	>800	42
Diptera	Simuliidae	<i>Prosimulium</i>	>866	54
Coleoptera	Psephenidae	<i>Psephenus</i>	>750	111
Ephemeroptera	Baetidae	<i>Pseudocloeon</i>	~861	36

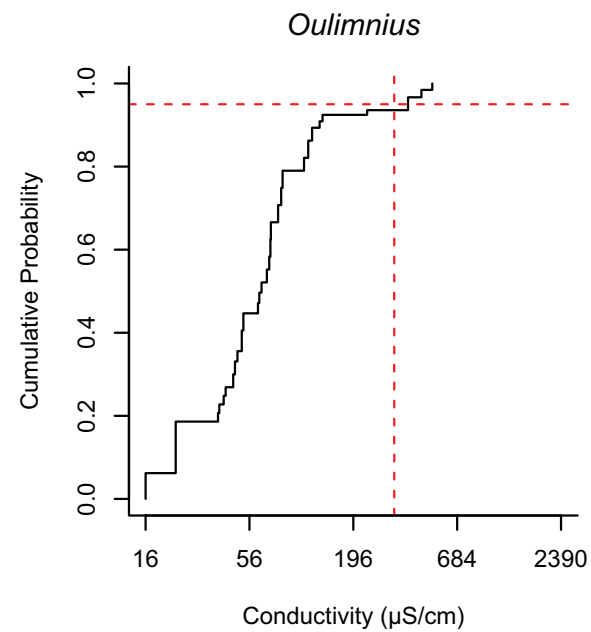
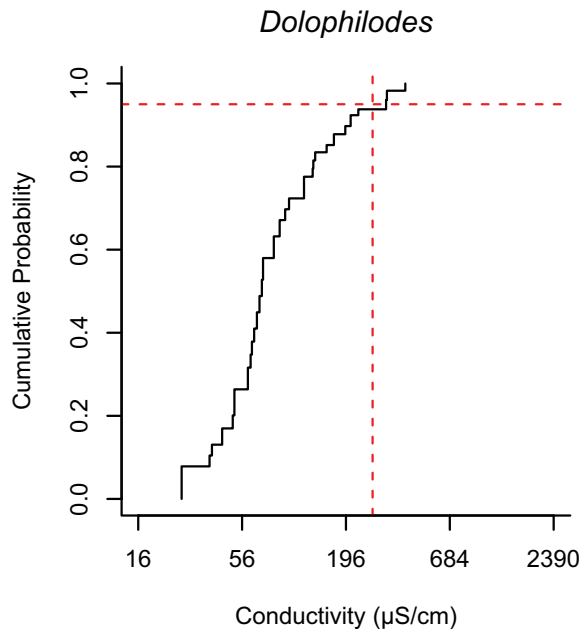
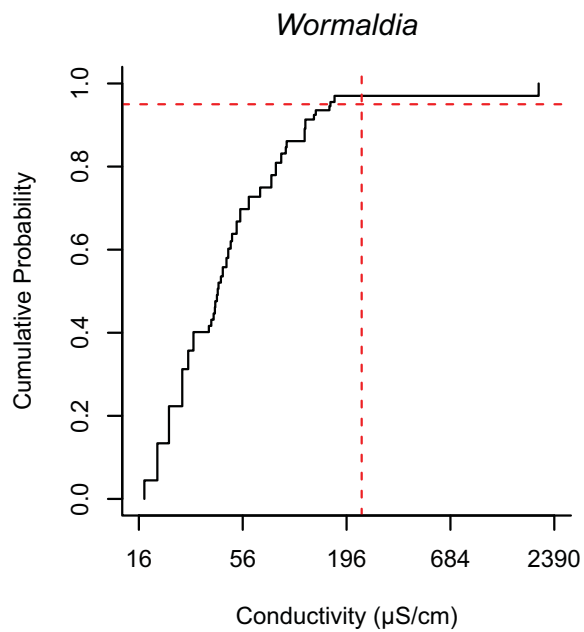
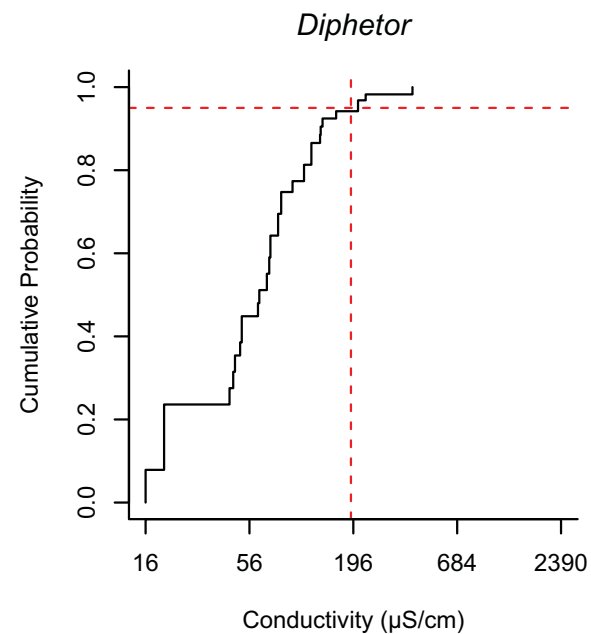
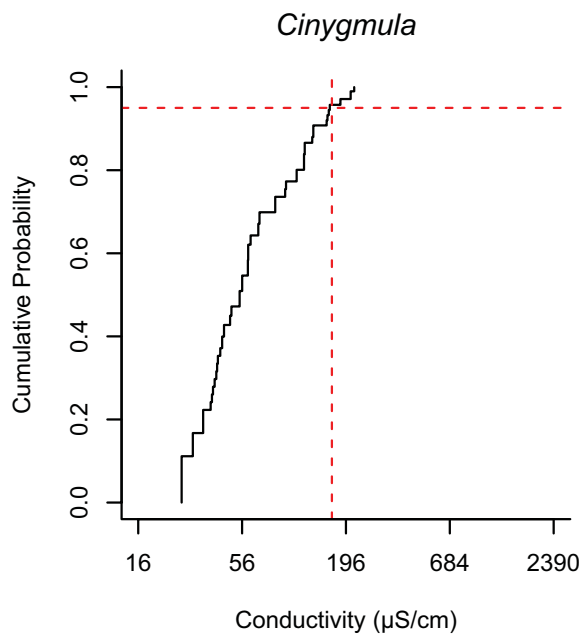
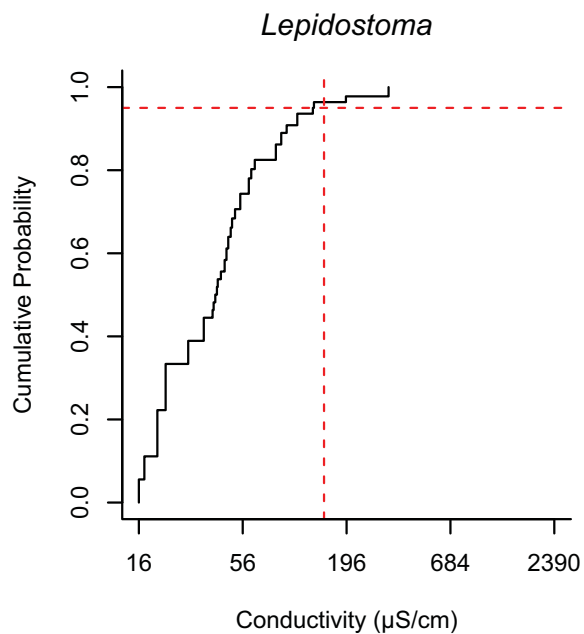
Order	Family	Genus	XC₉₅	N
Diptera	Tipulidae	<i>Pseudolimnophila</i>	>1,051	40
Trichoptera	Limnephilidae	<i>Pycnopsyche</i>	>775	64
Hemiptera	Veliidae	<i>Rhagovelia</i>	>600	27
Diptera	Chironomidae	<i>Rheocricotopus</i>	>1,117	51
Diptera	Chironomidae	<i>Rheotanytarsus</i>	>1,601	115
Trichoptera	Rhyacophilidae	<i>Rhyacophila</i>	>574	94
Megaloptera	Sialidae	<i>Sialis</i>	>1,843	64
Diptera	Simuliidae	<i>Simulium</i>	>1,580	179
Ephemeroptera	Heptageniidae	<i>Stenacron</i>	>862	90
Coleoptera	Elmidae	<i>Stenelmis</i>	>1,520	168
Diptera	Chironomidae	<i>Stenochironomus</i>	>824	35
Ephemeroptera	Heptageniidae	<i>Stenonema</i>	>993	178
Odonata	Gomphidae	<i>Stylogomphus</i>	>1,720	90
Plecoptera	Chloroperlidae	<i>Sweltsa</i>	558	55
Diptera	Chironomidae	<i>Tanytarsus</i>	>1,316	118
Diptera	Chironomidae	<i>Thienemannimyia</i>	>1,697	155
Diptera	Tipulidae	<i>Tipula</i>	>1,814	150
Trichoptera	Leptoceridae	<i>Triaenodes</i>	>938	31
Ephemeroptera	Leptohiphidae	<i>Tricorythodes</i>	>2,000	48
Diptera	Chironomidae	<i>Tvetenia</i>	>1,254	46
Trichoptera	Philopotamidae	<i>Wormaldia</i>	235	38
Plecoptera	Perlodidae	<i>Yugus</i>	354	25

APPENDIX J
GRAPHS OF CUMULATIVE FREQUENCY DISTRIBUTIONS
FOR GENERA IN A KENTUCKY DATA SET

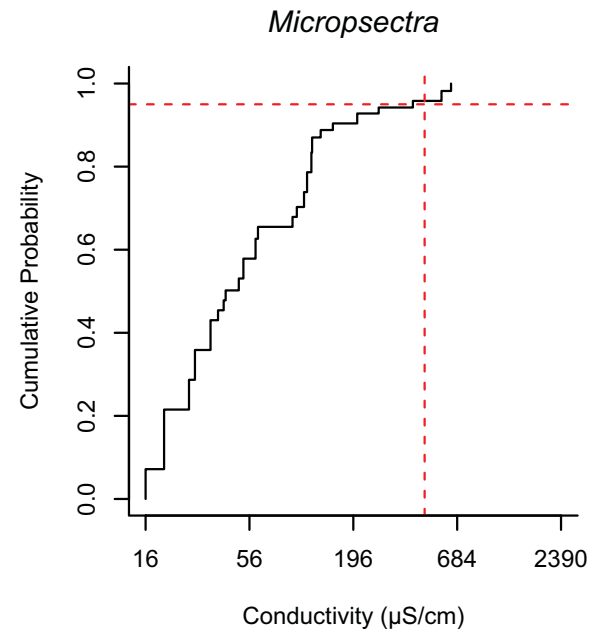
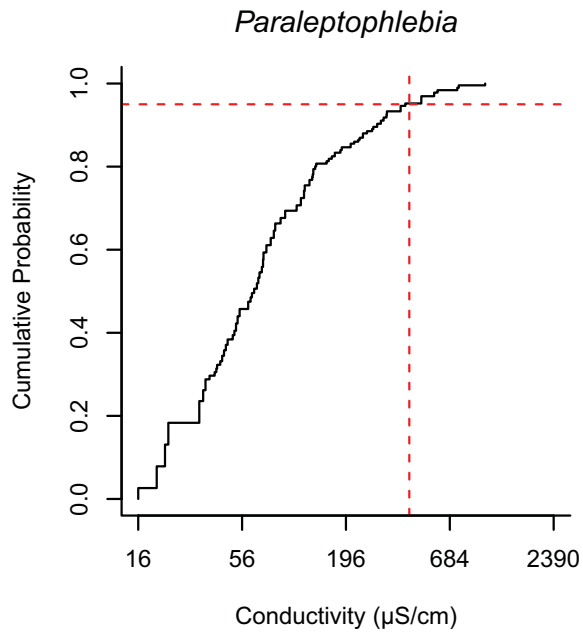
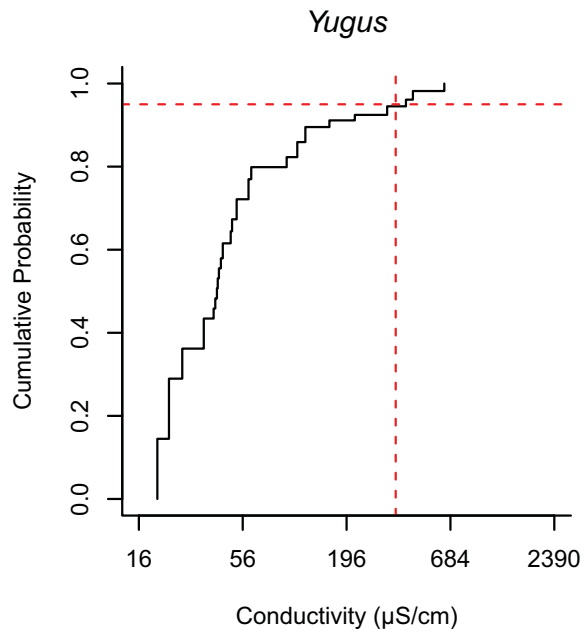
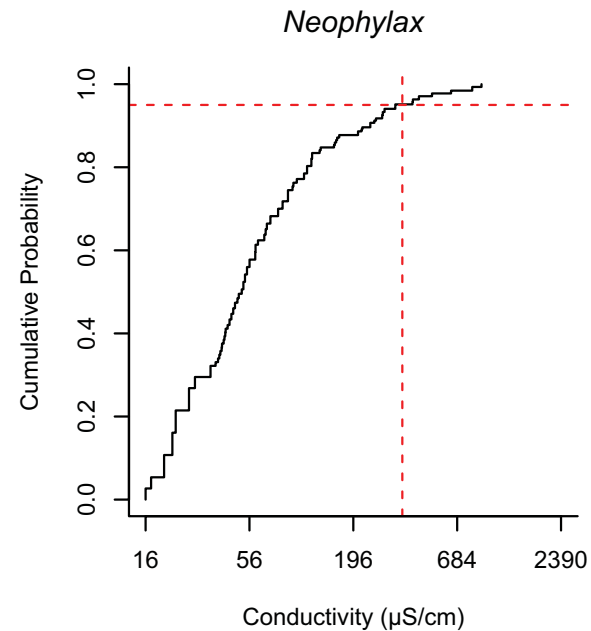
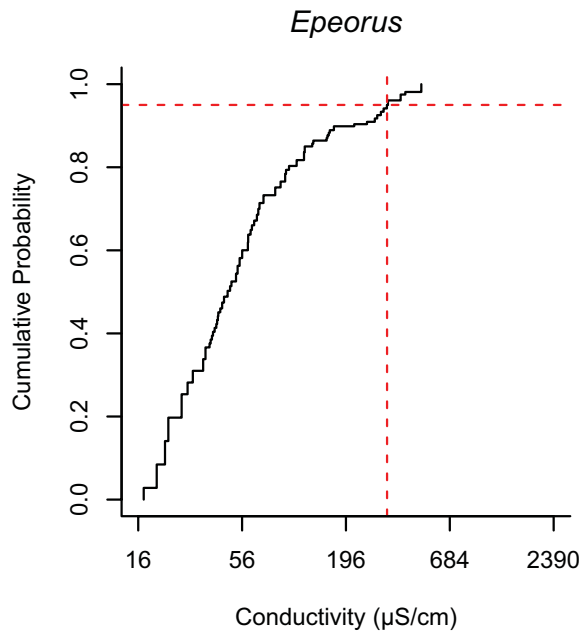
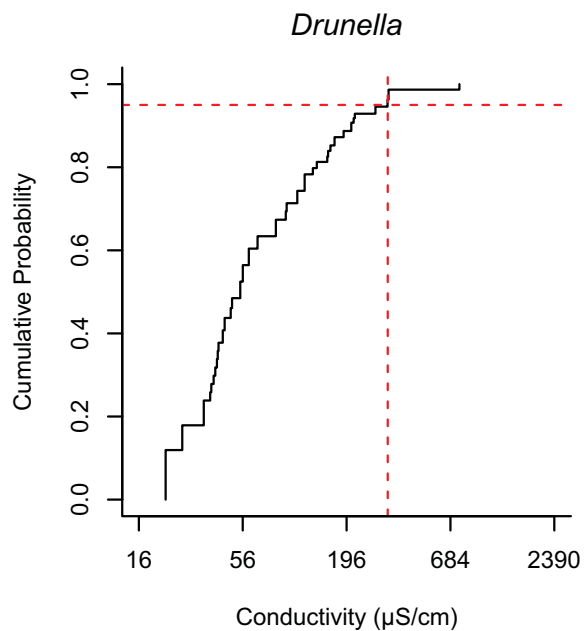
ABSTRACT

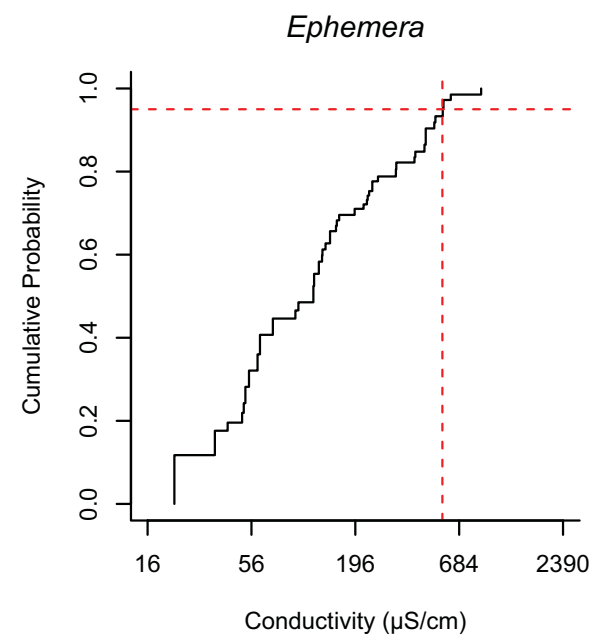
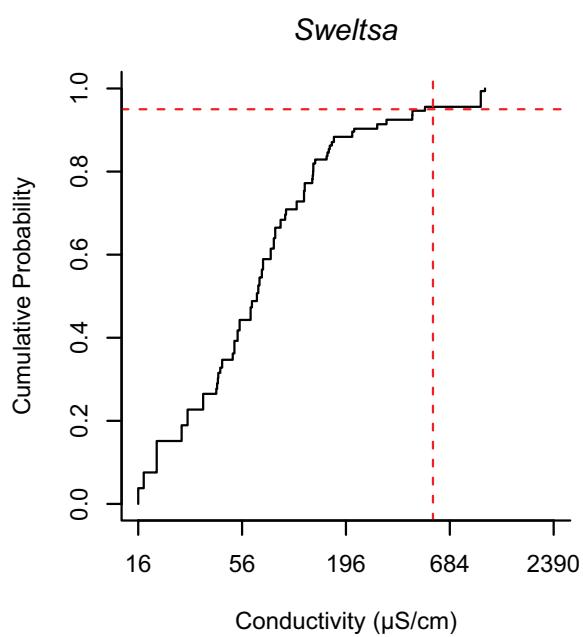
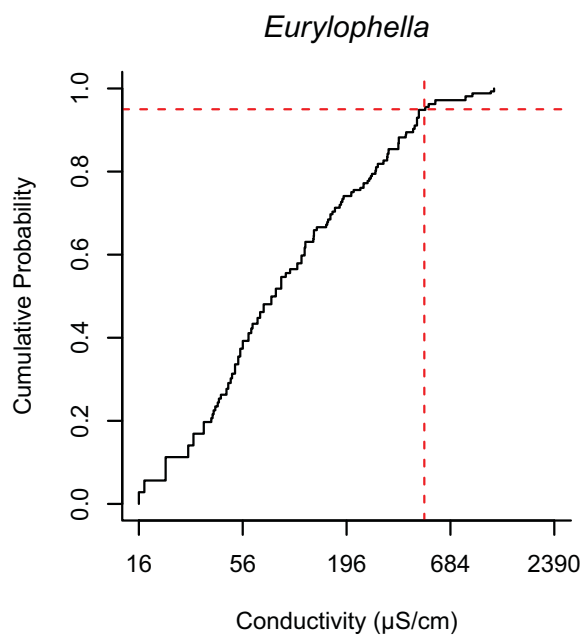
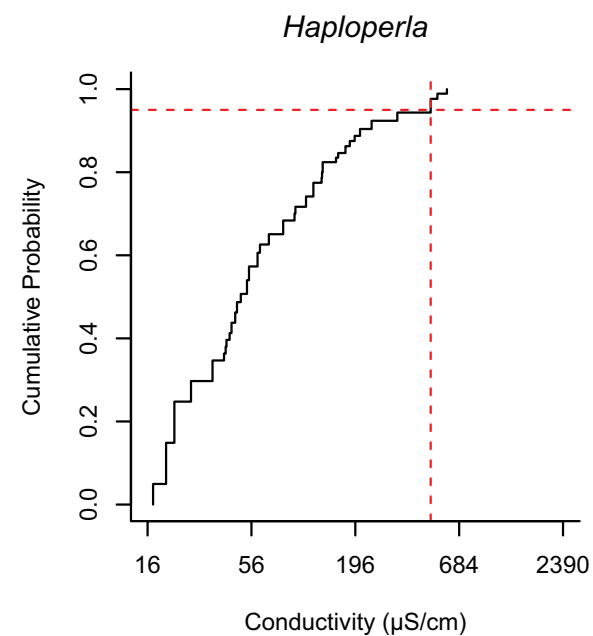
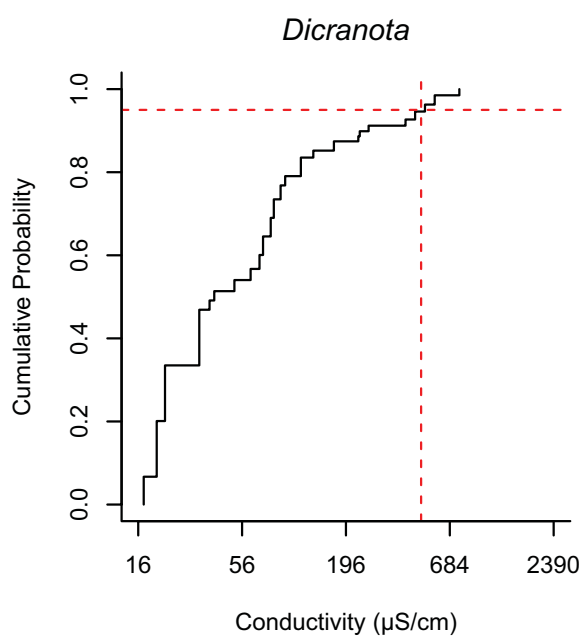
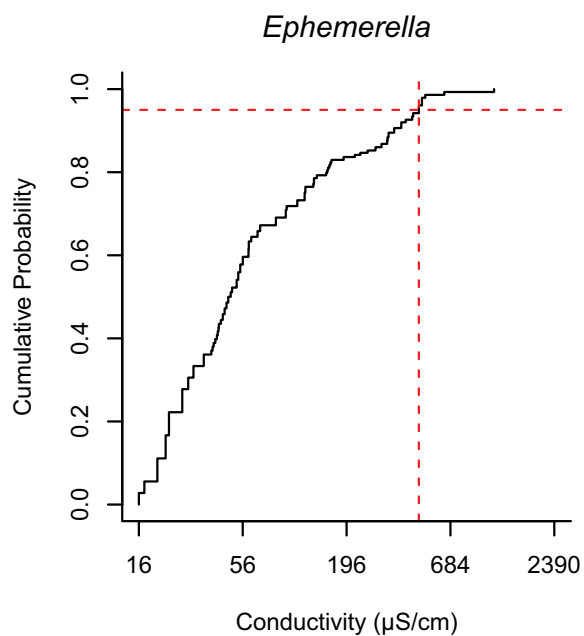
The purpose of Appendix J is to help the reader visualize the changes in the occurrence of each genus in the Kentucky data set as conductivity increases and understand how the extirpation concentration (XC_{95}) values are derived. Each plot contains the weighted cumulative distribution function (CDF) for the occurrence of a genus with respect to conductivity. For each genus, the points in the CDF represent the weighted proportions of occurrences of the genus in samples less than the indicated conductivity value ($\mu\text{S}/\text{cm}$), calculated using Equation 1. The 95th centile is found at the intersection of the dashed horizontal line with the CDF. The conductivity for the 95th centile is the XC_{95} value and is found at the intersection of the vertical line and the x-axis.

Z-f

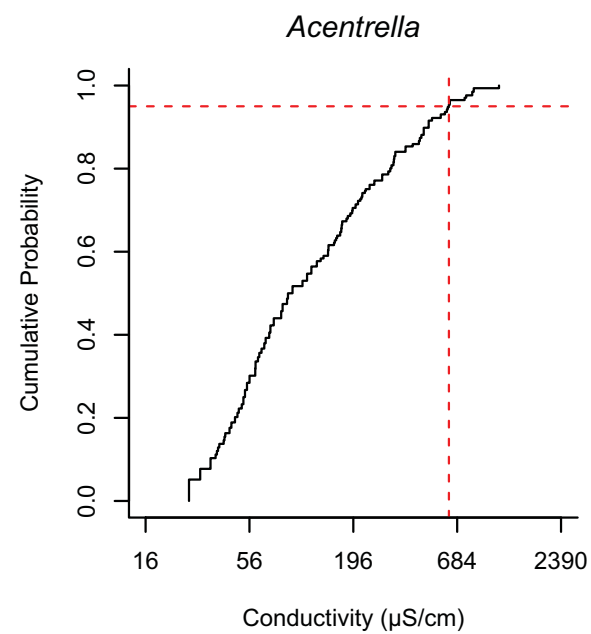
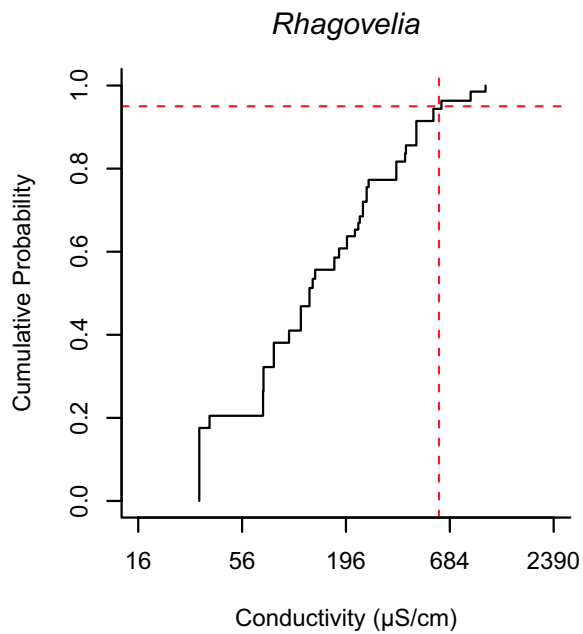
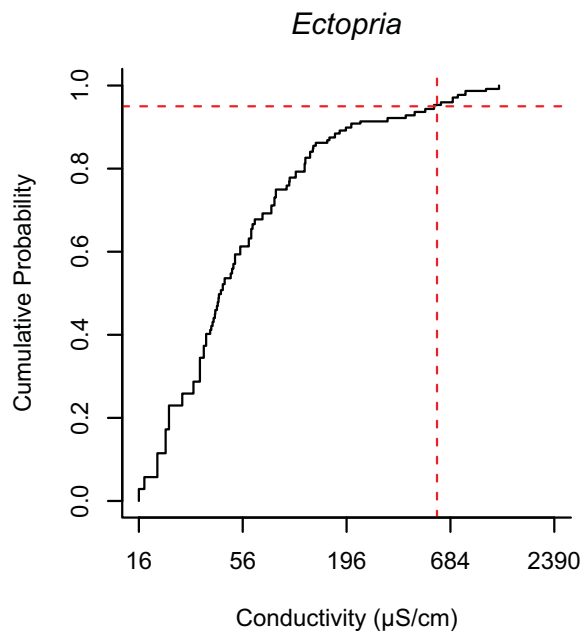
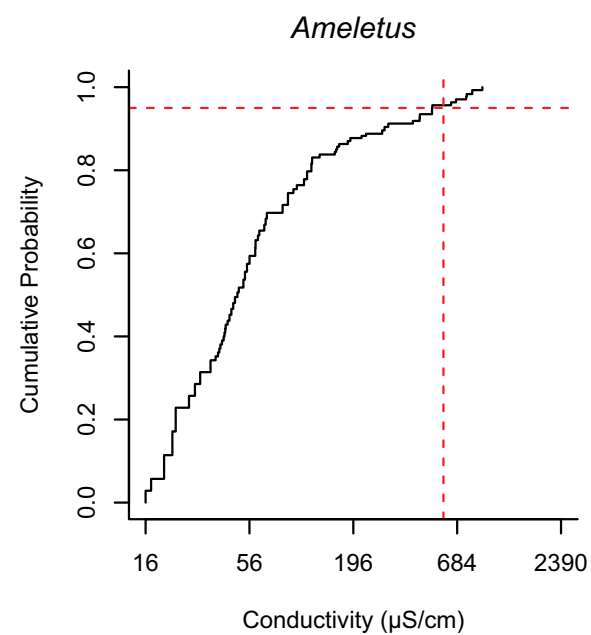
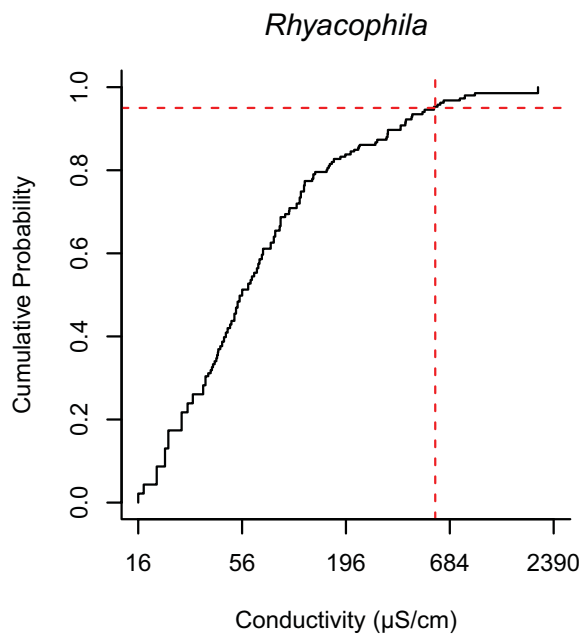
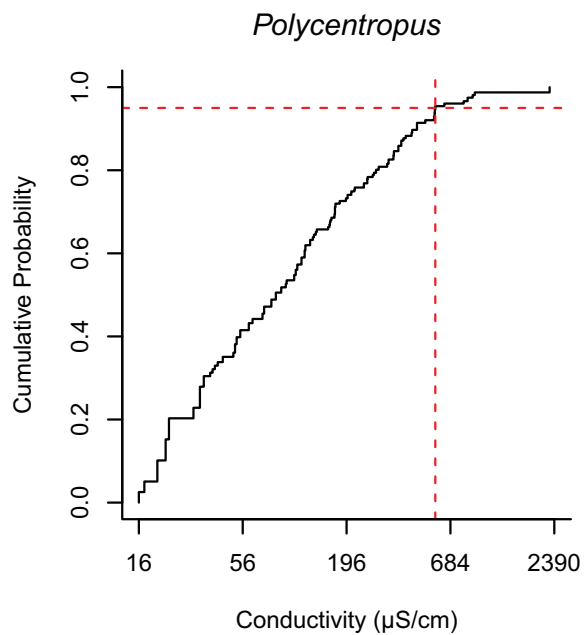


ξ -f

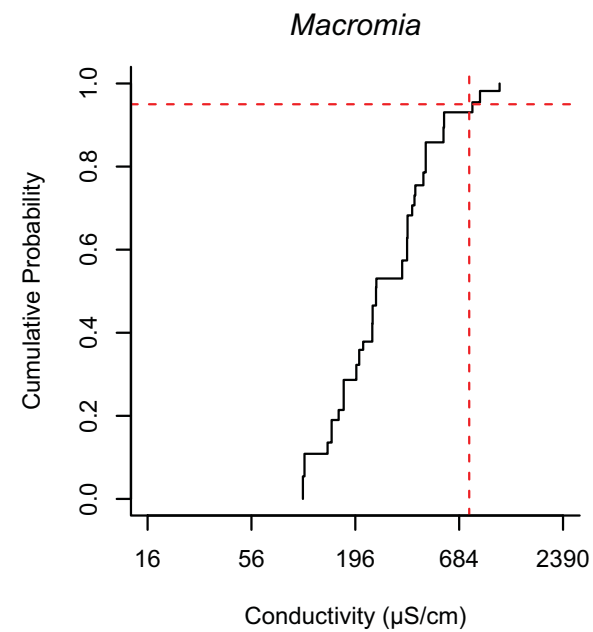
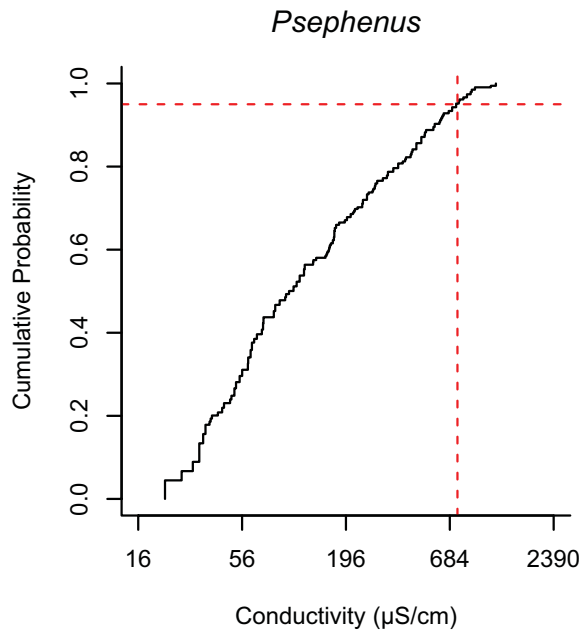
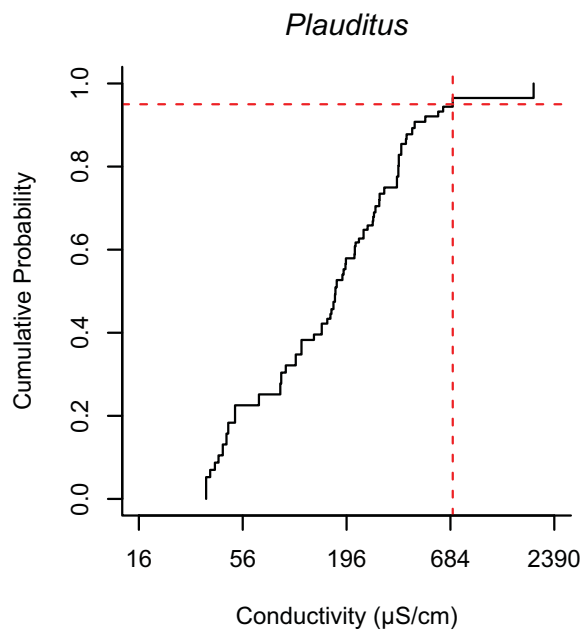
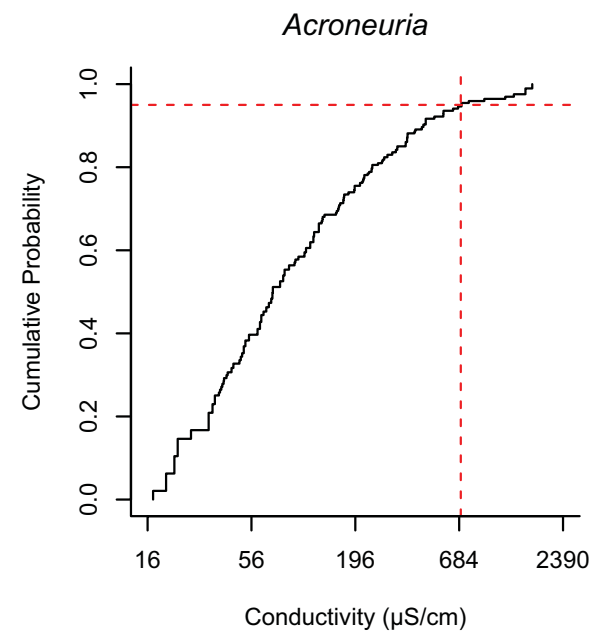
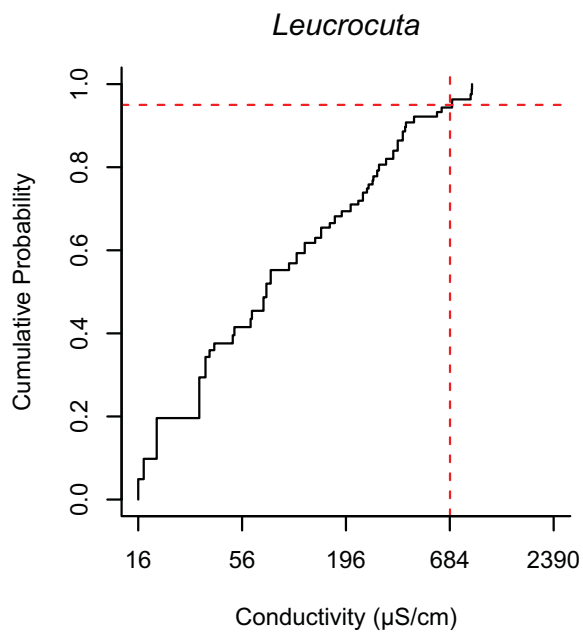
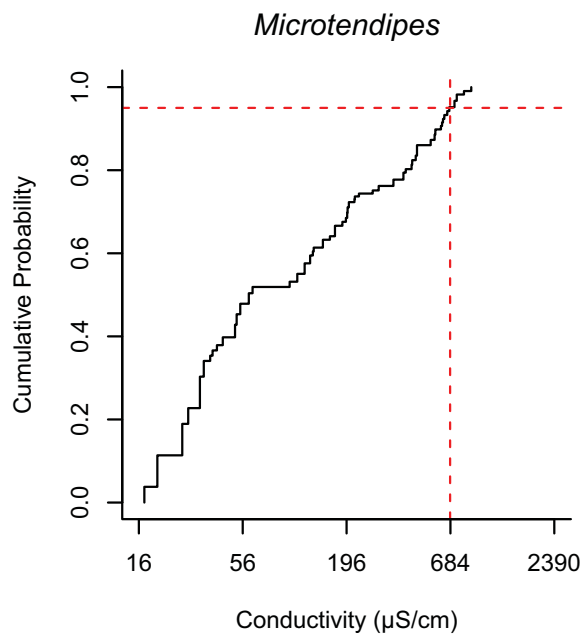




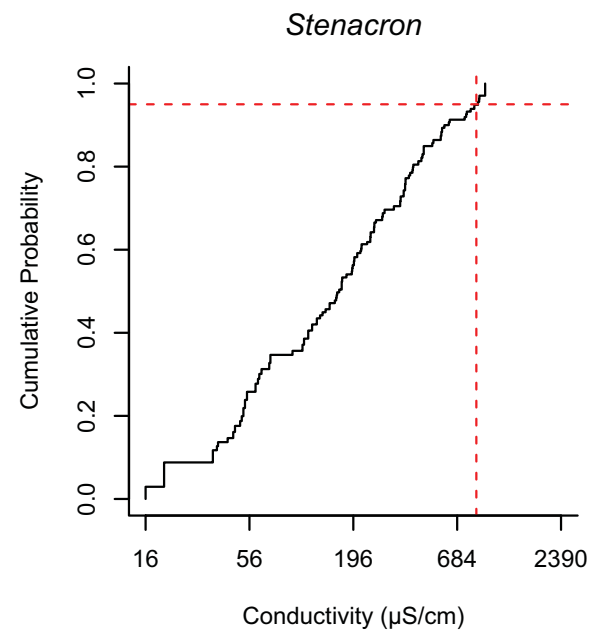
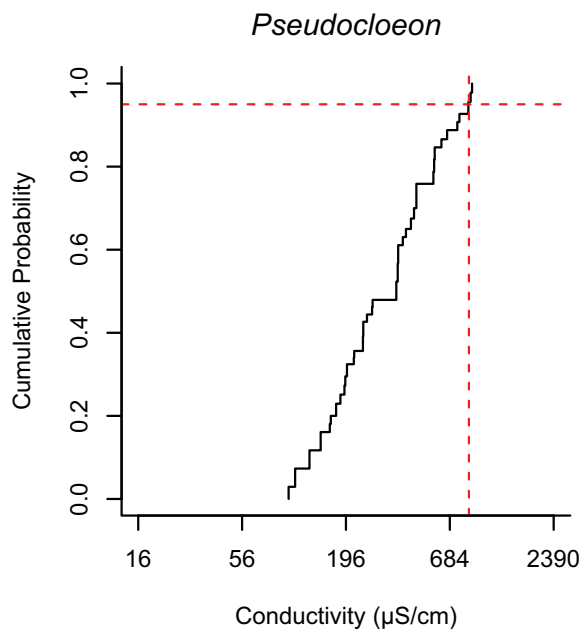
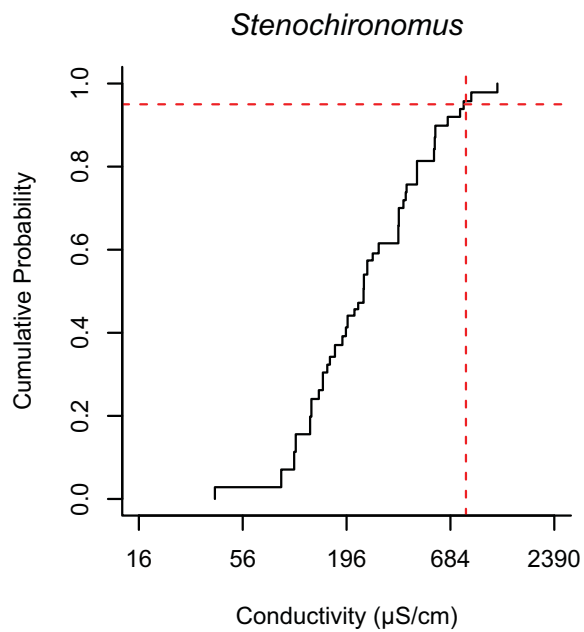
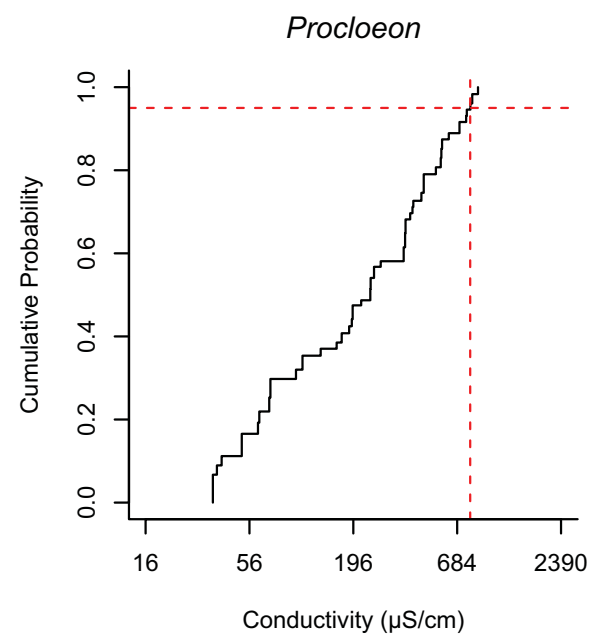
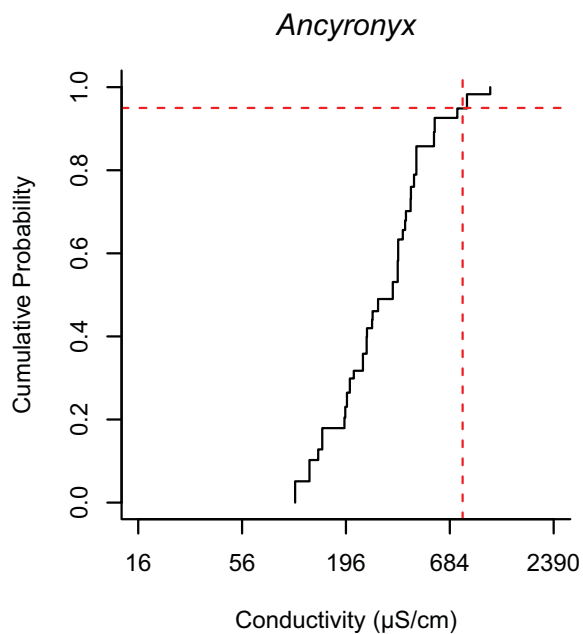
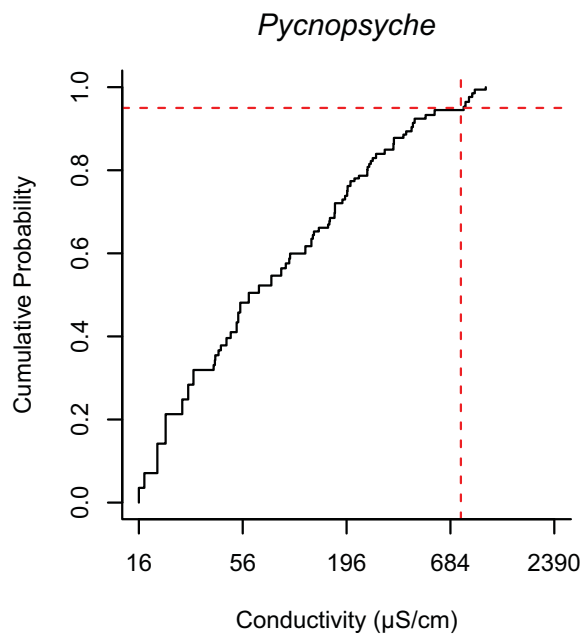
S-I



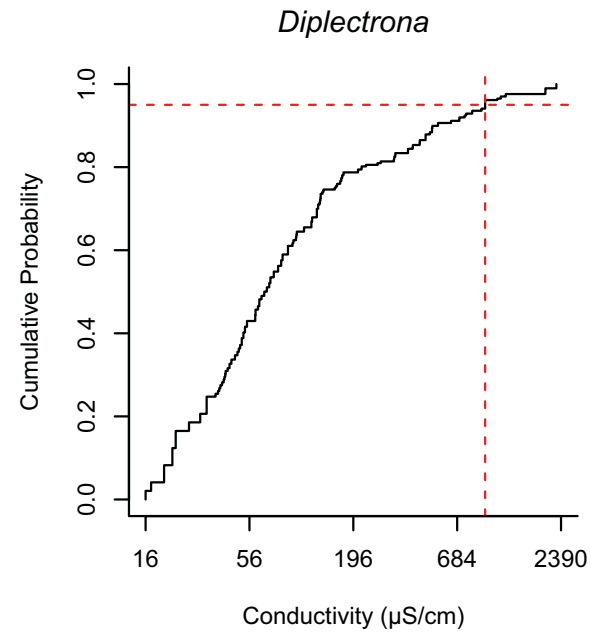
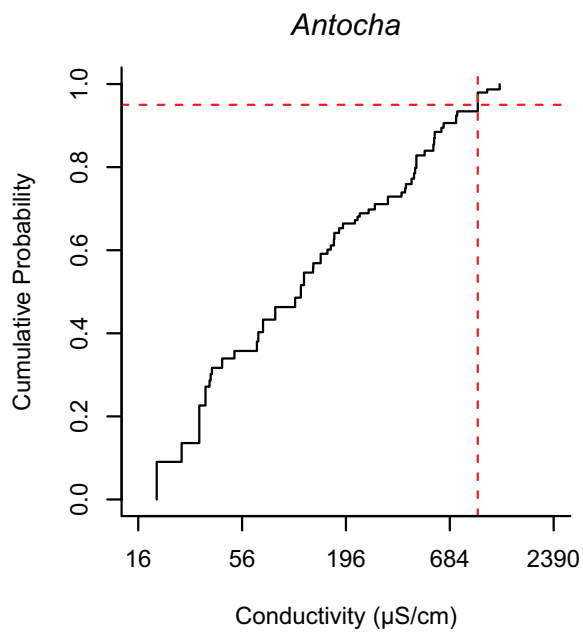
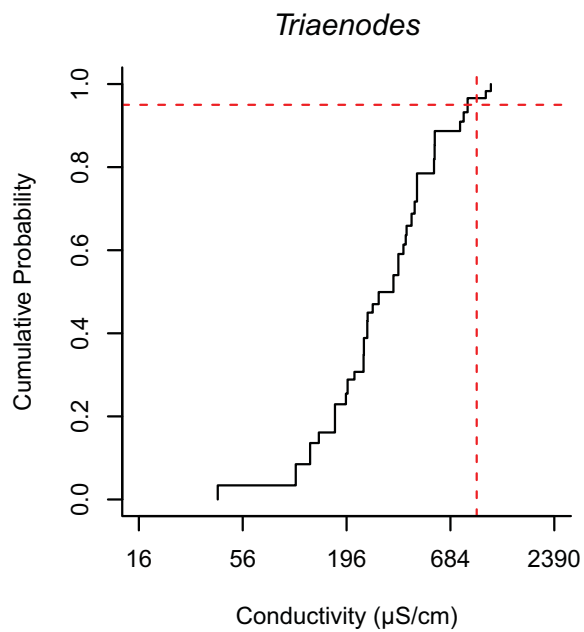
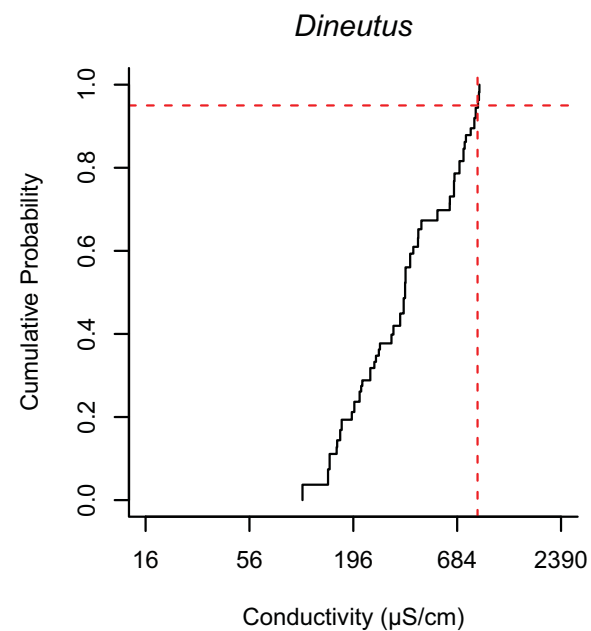
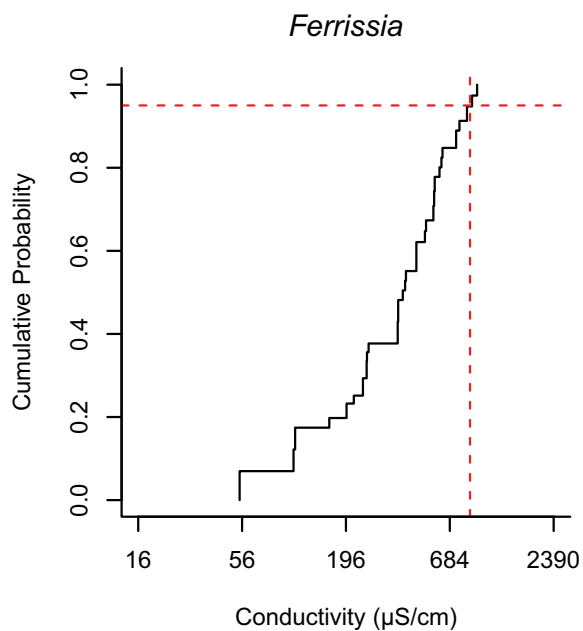
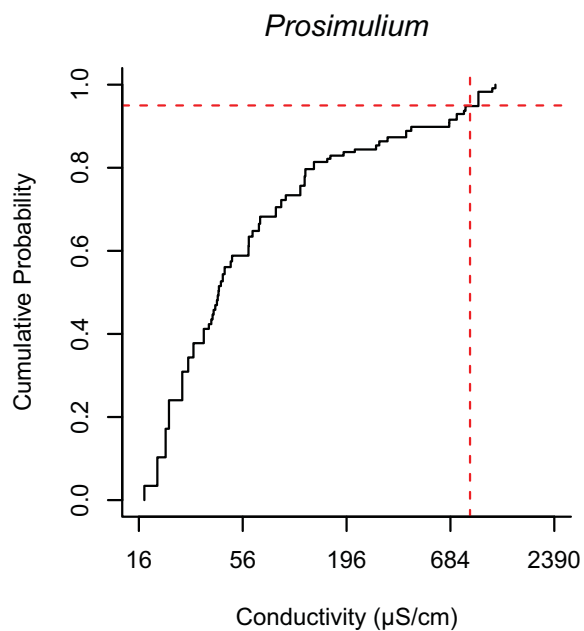
9-f



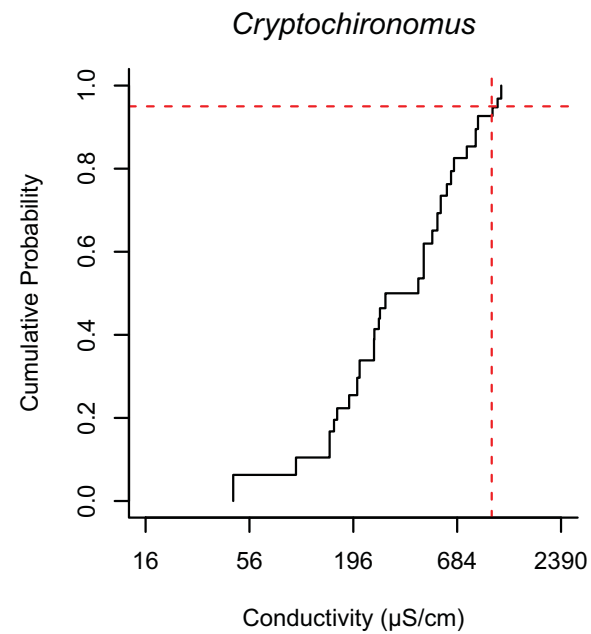
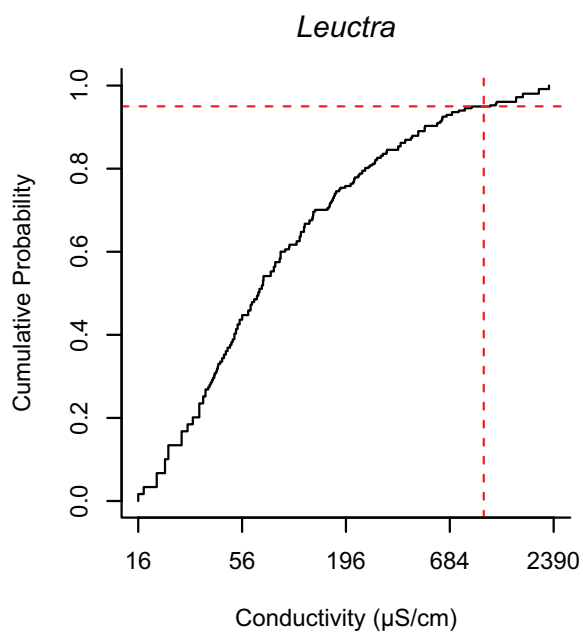
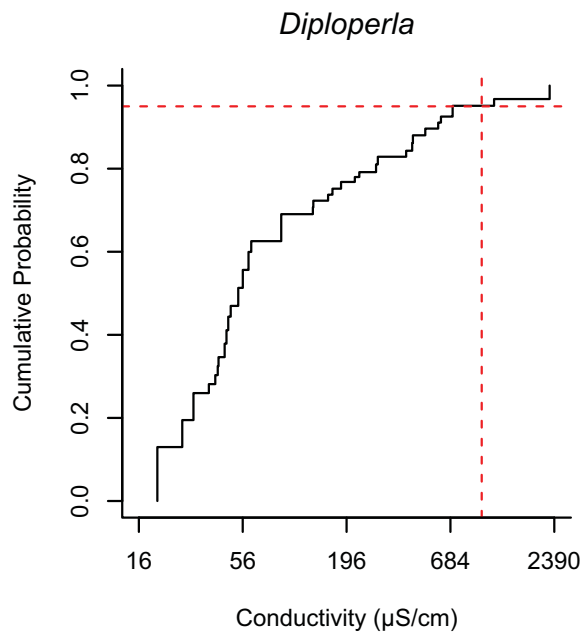
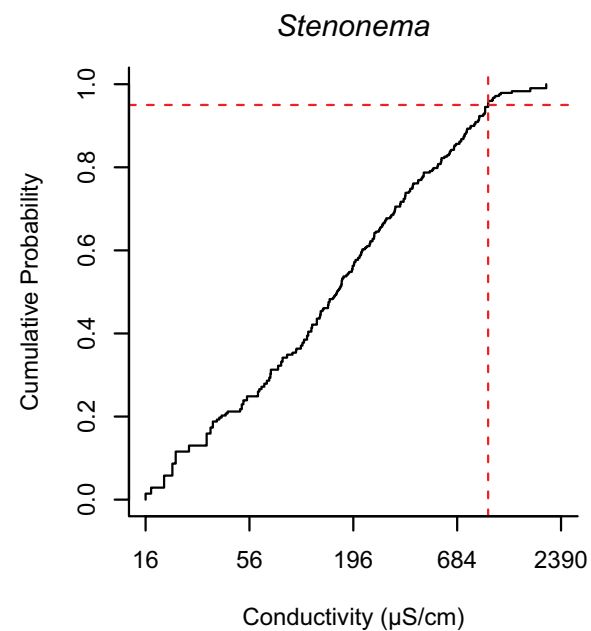
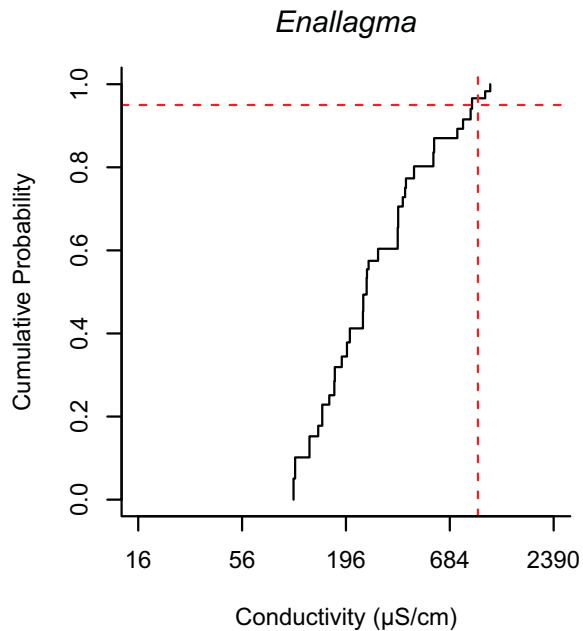
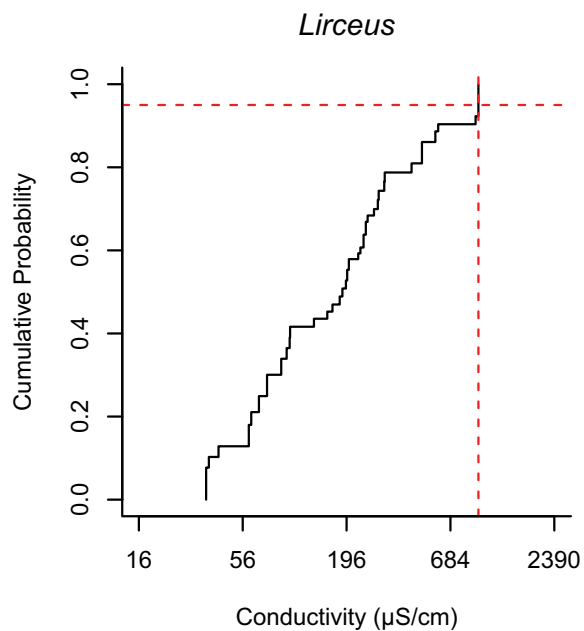
$L-f$

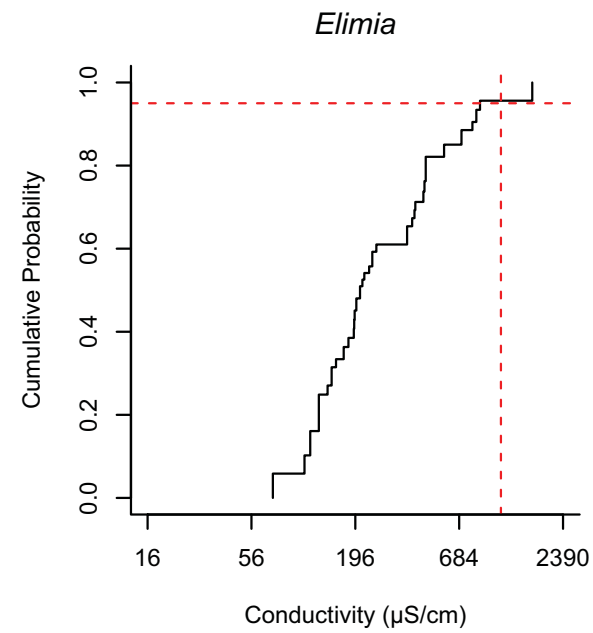
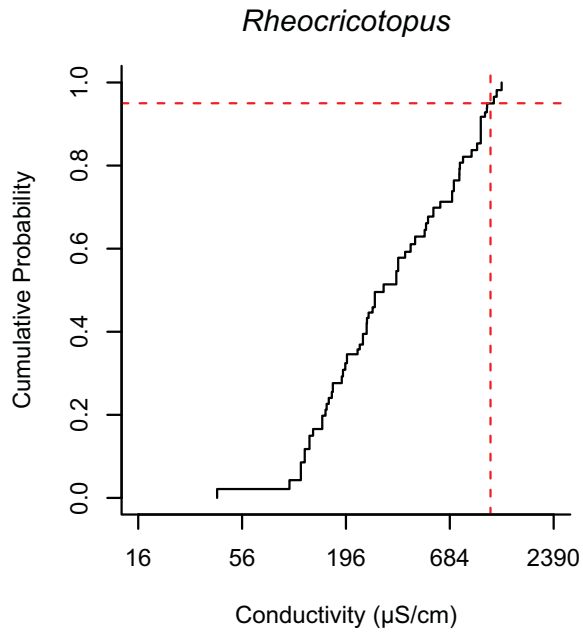
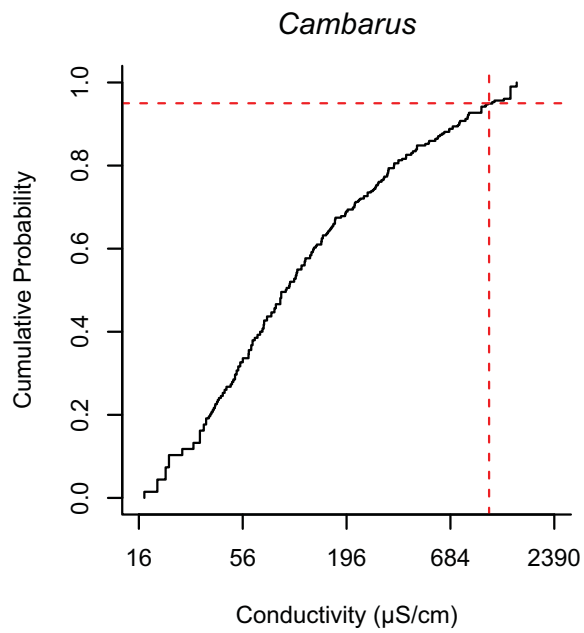
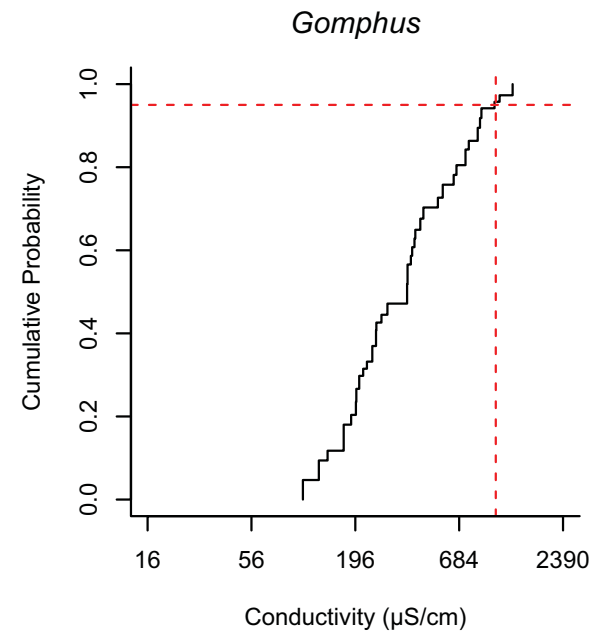
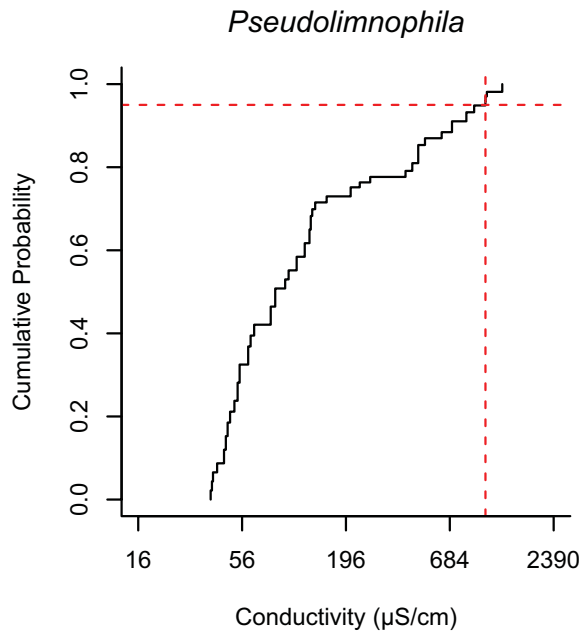
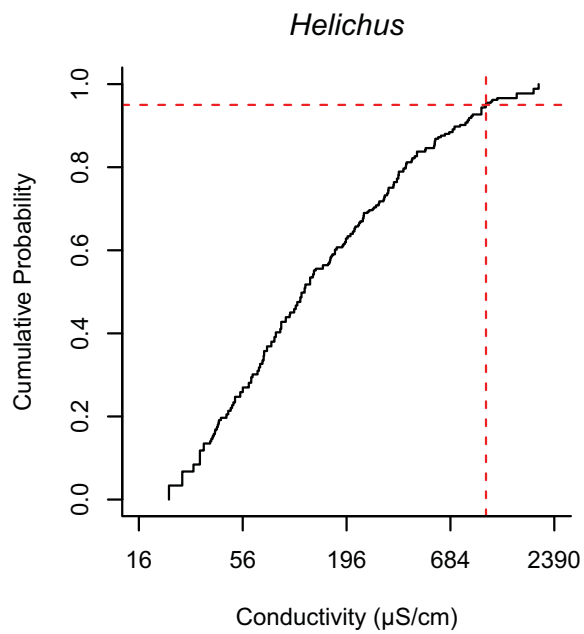


8-f

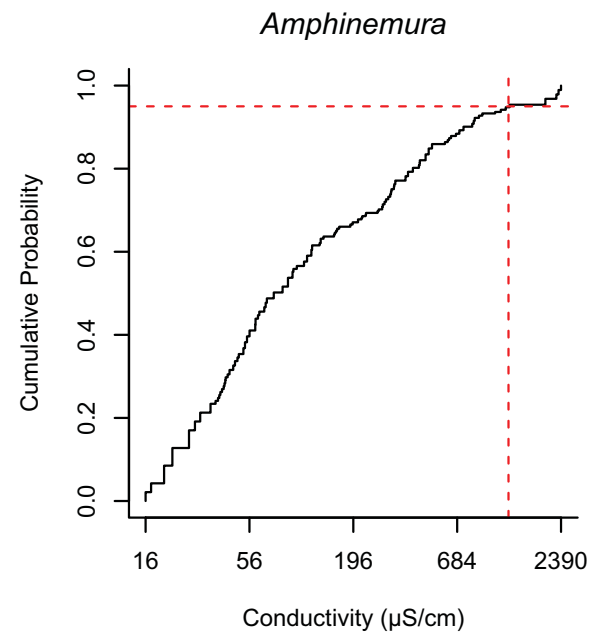
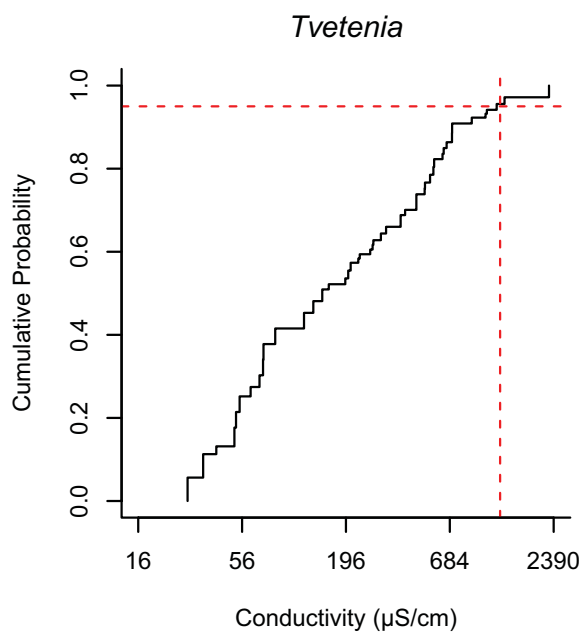
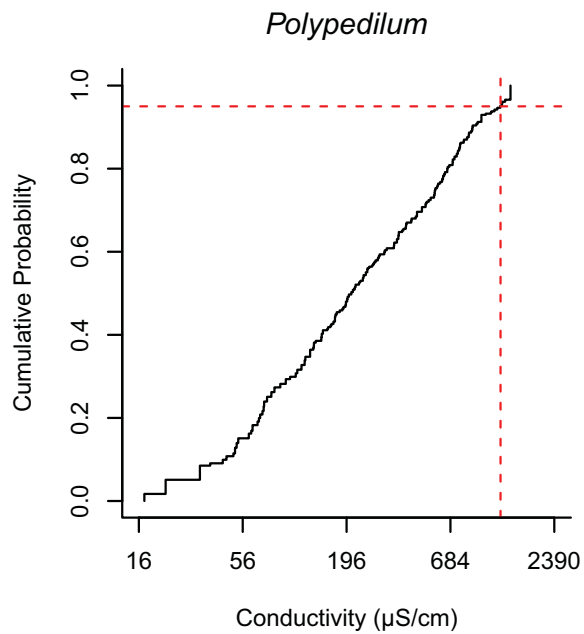
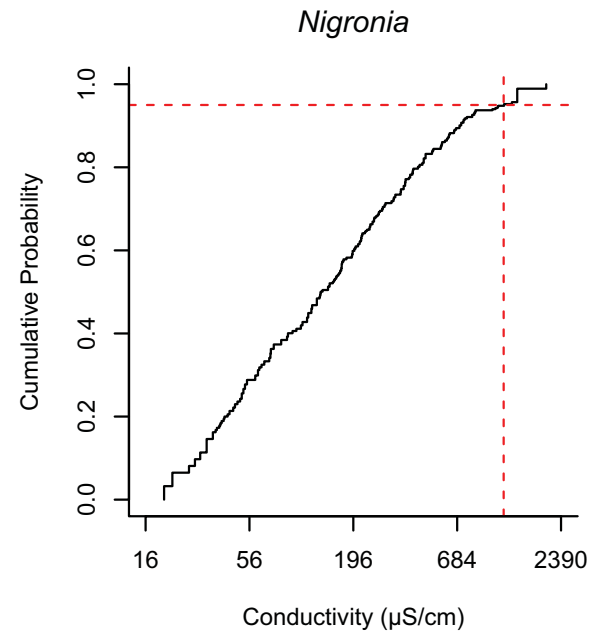
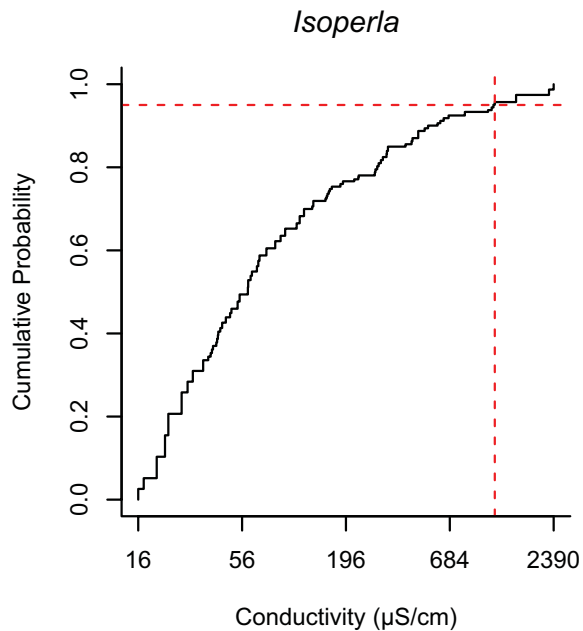
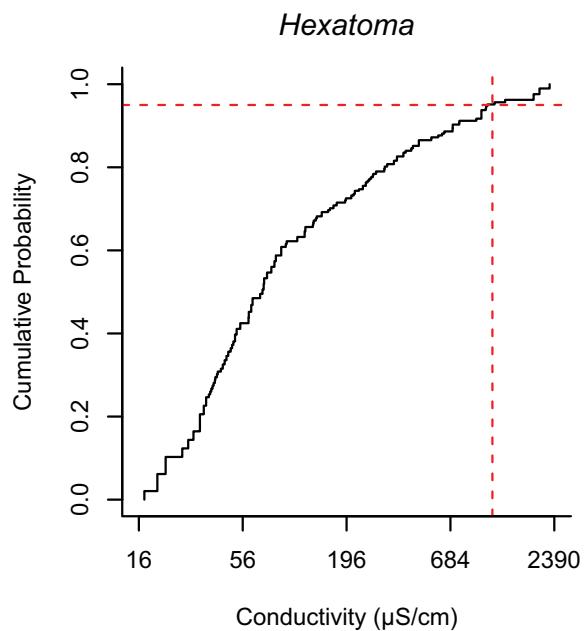


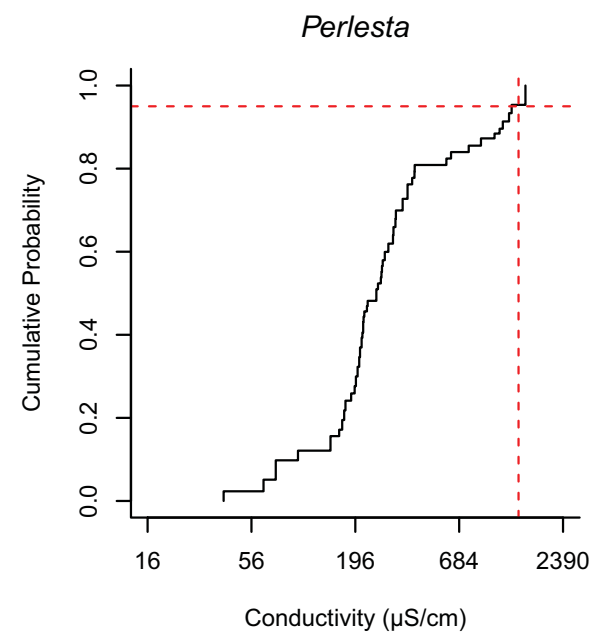
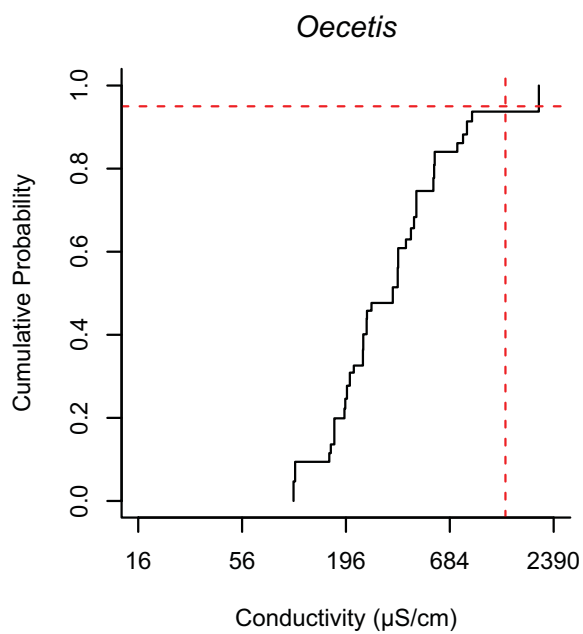
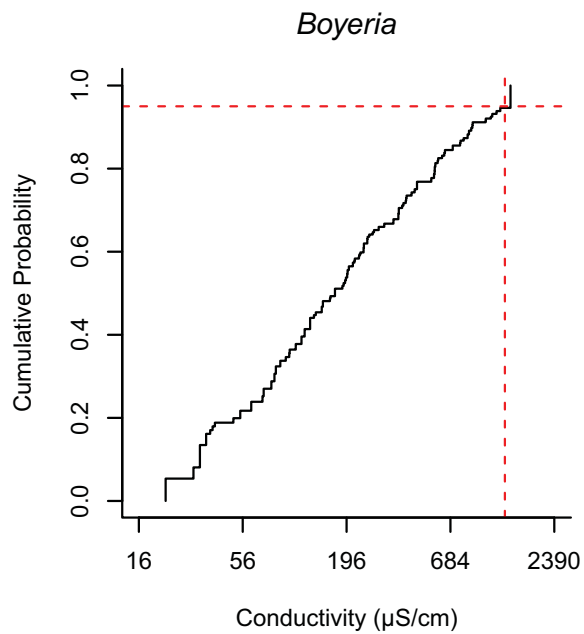
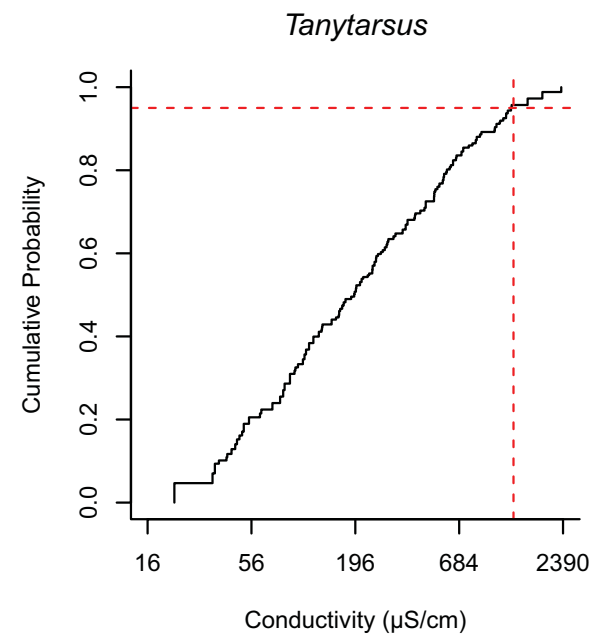
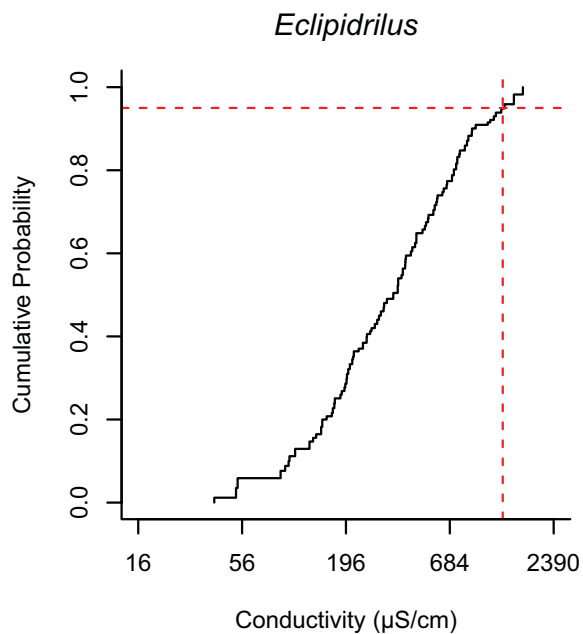
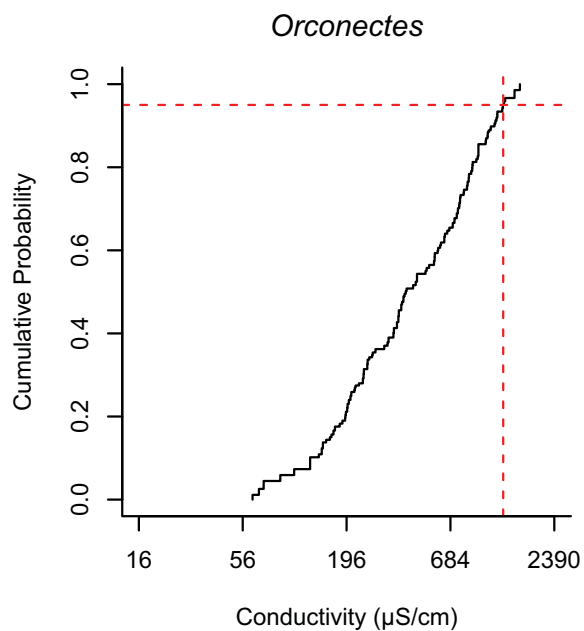
6-f



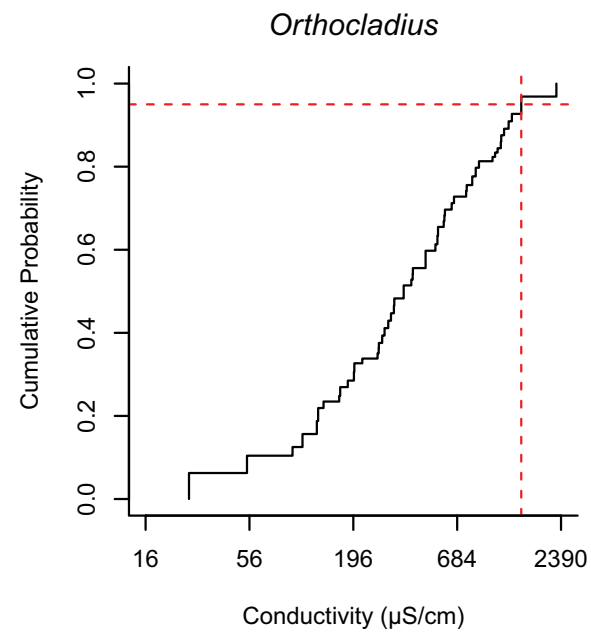
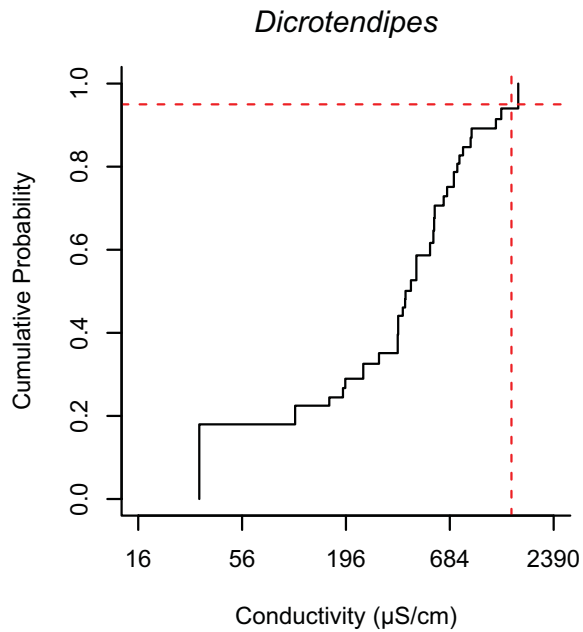
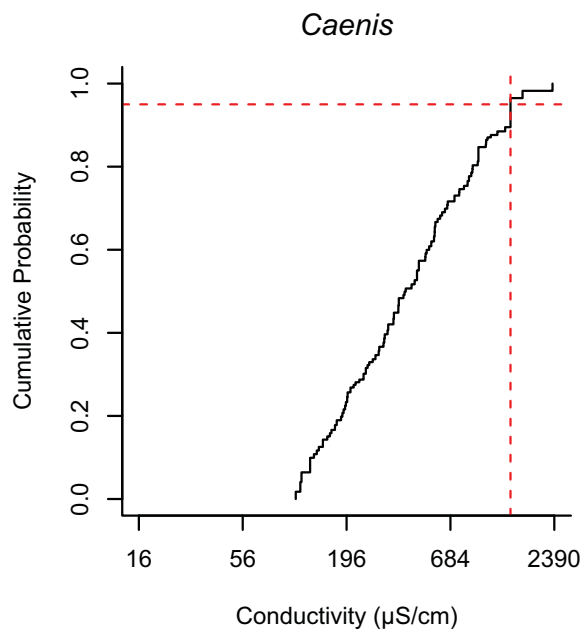
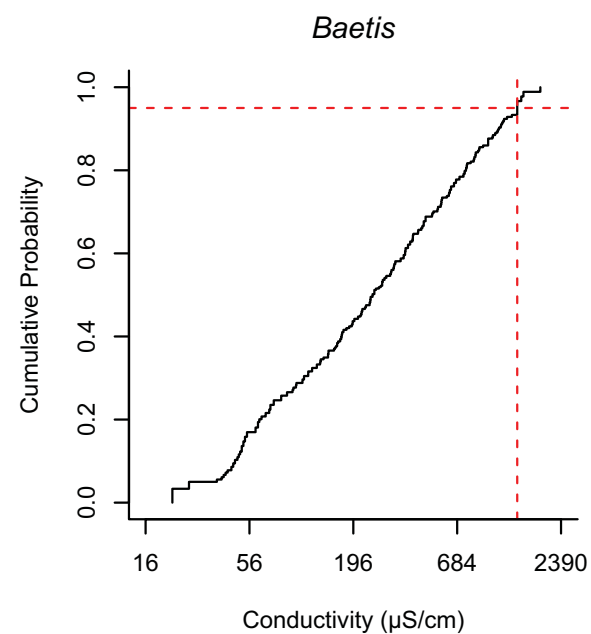
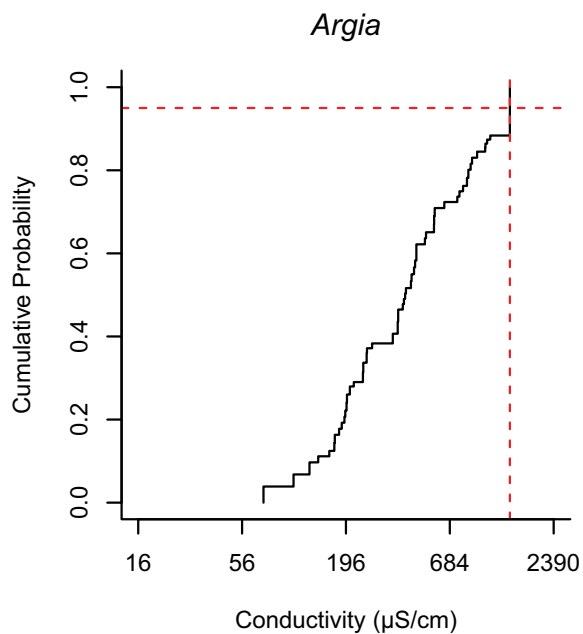
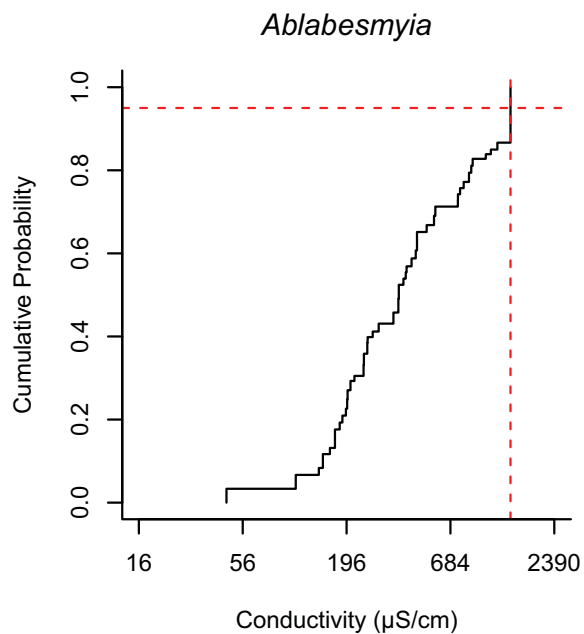


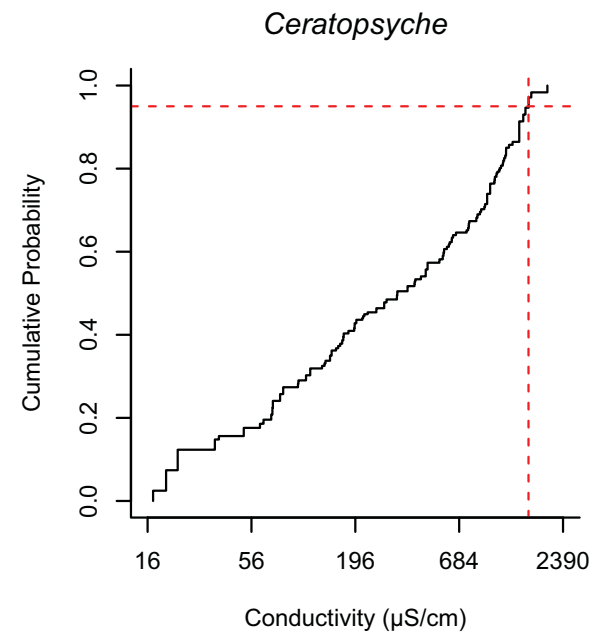
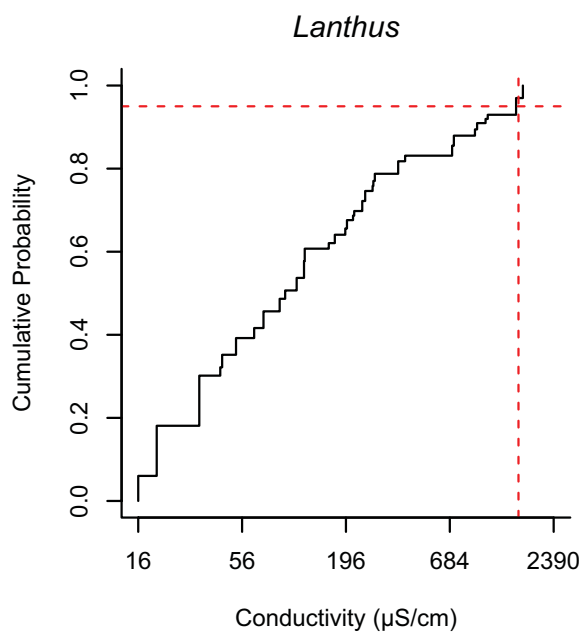
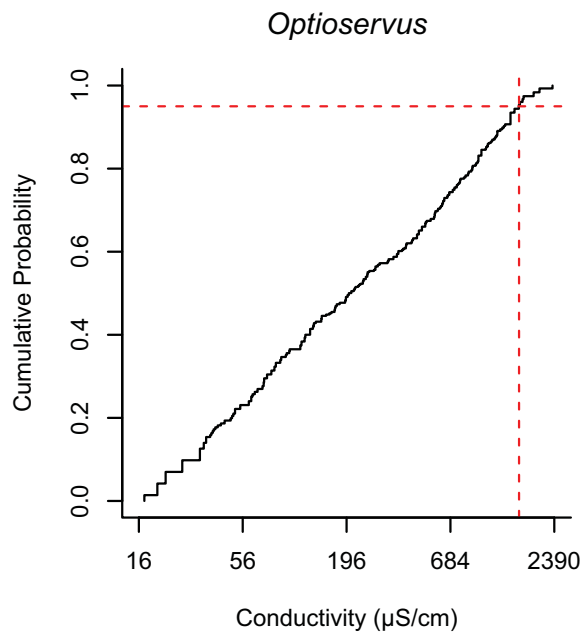
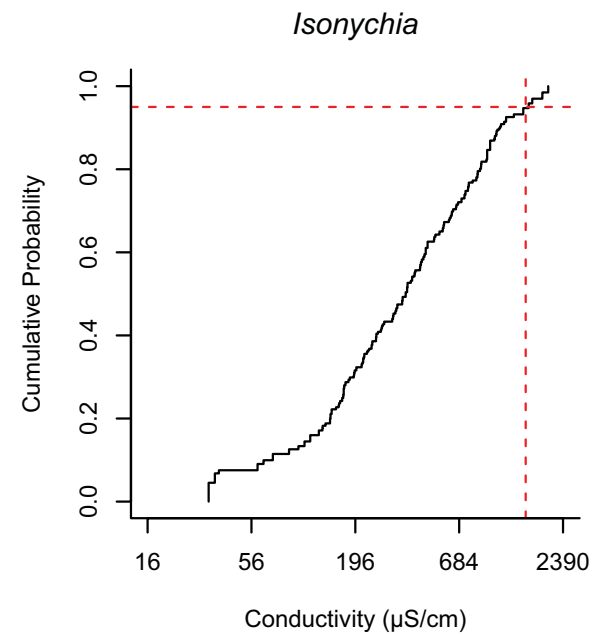
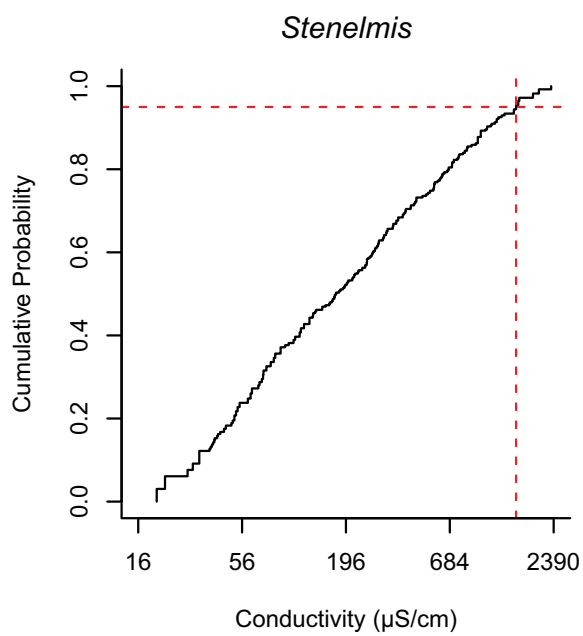
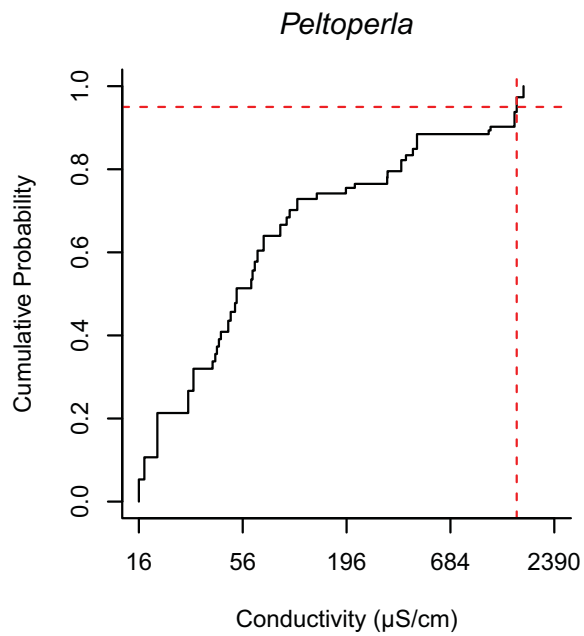
I I-f



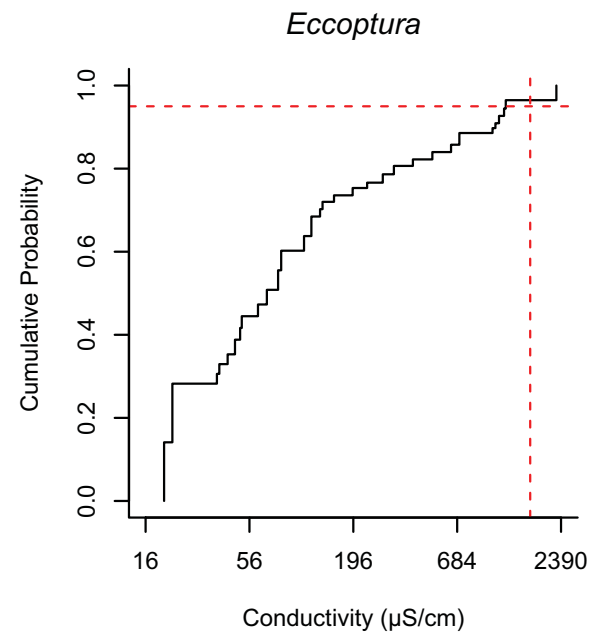
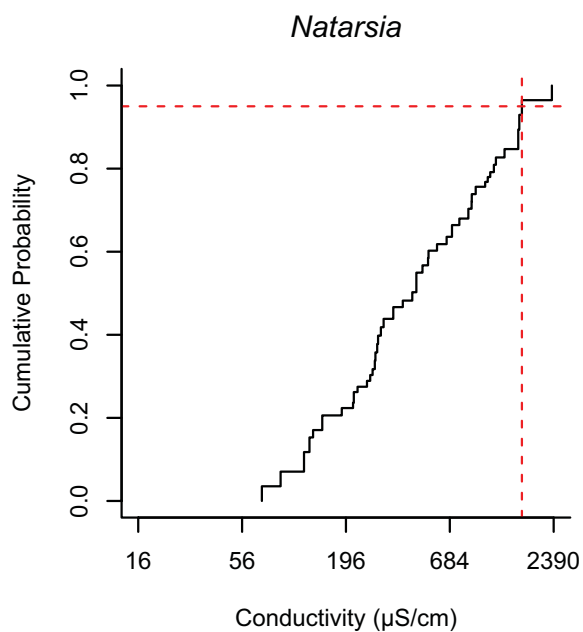
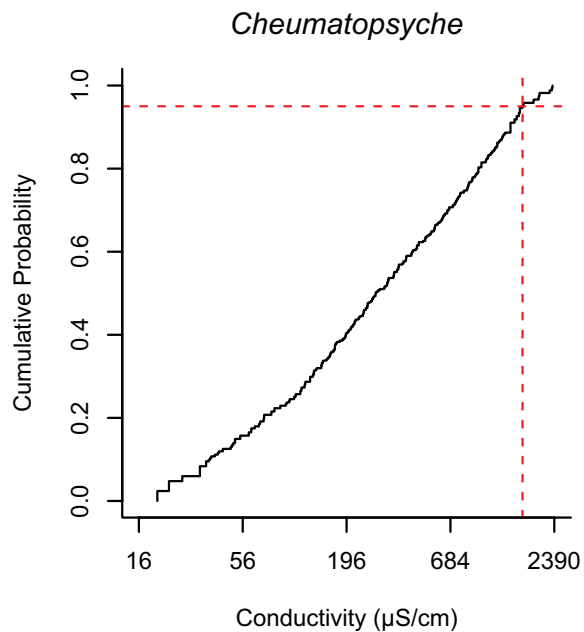
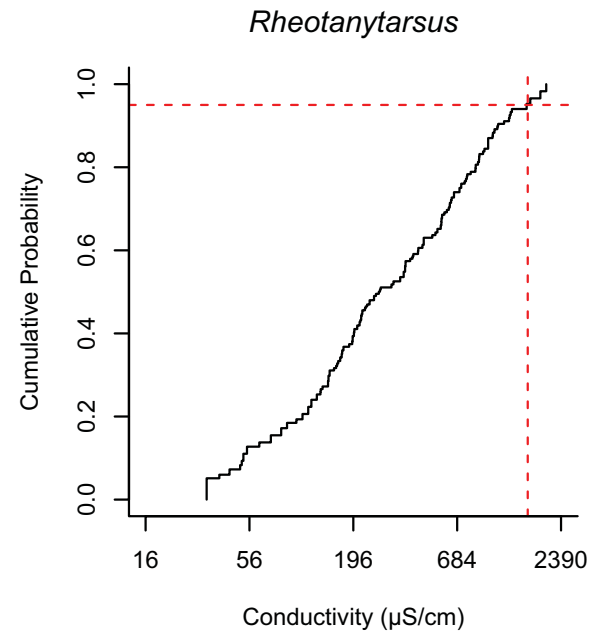
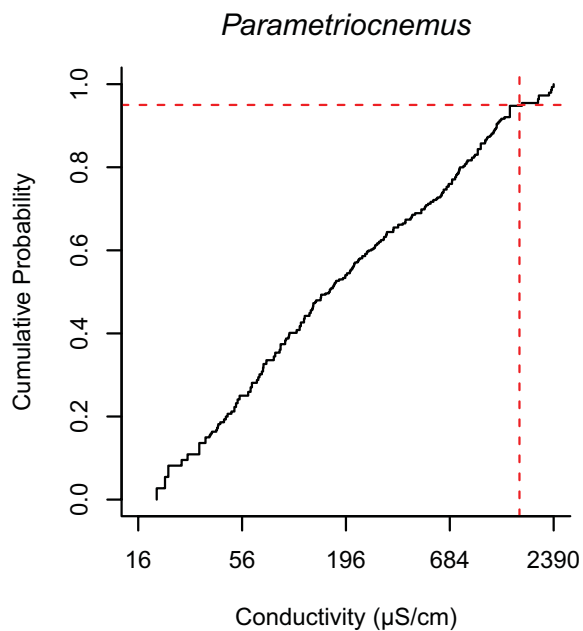
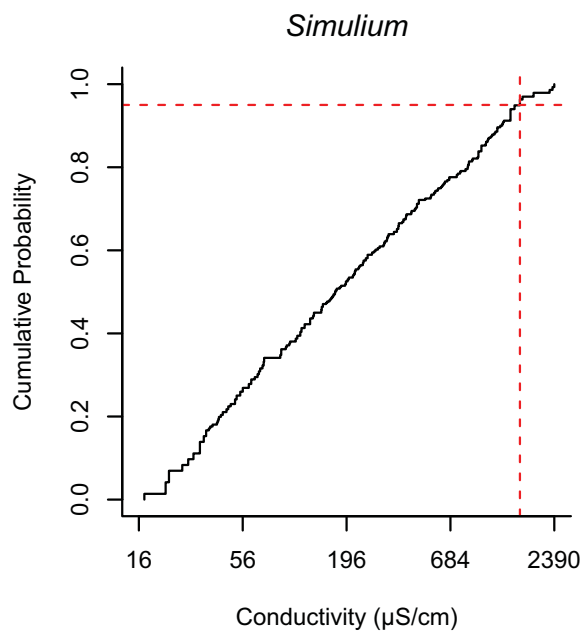


$\xi|f$

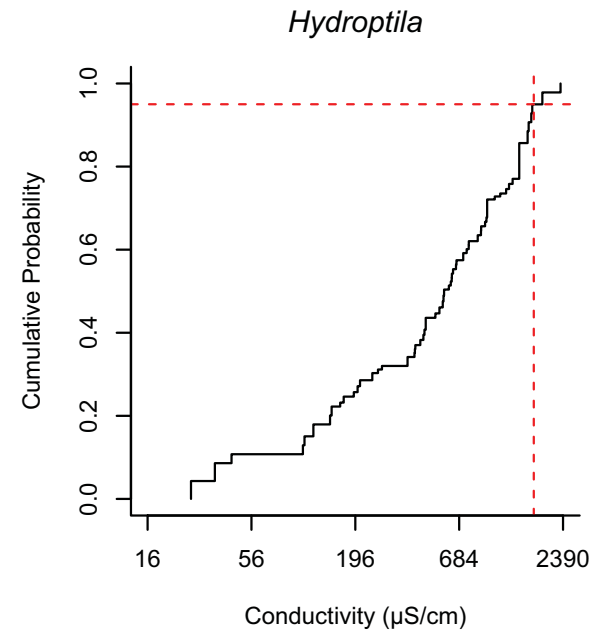
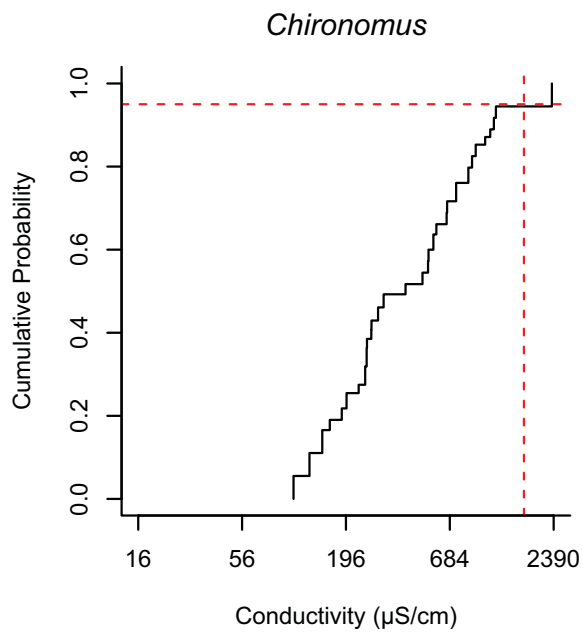
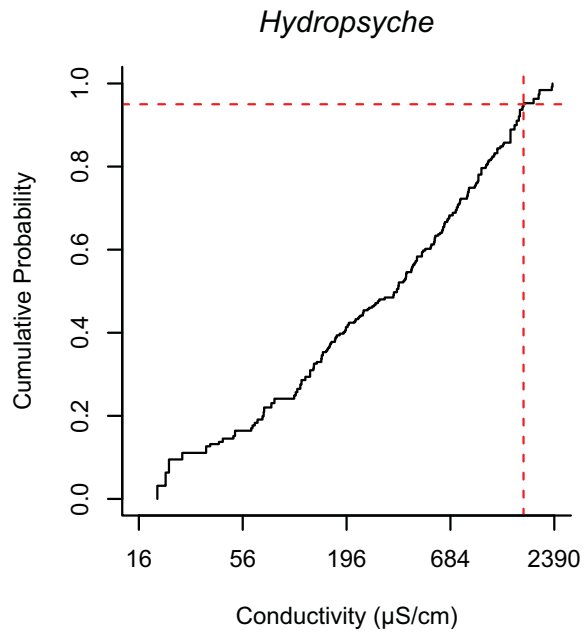
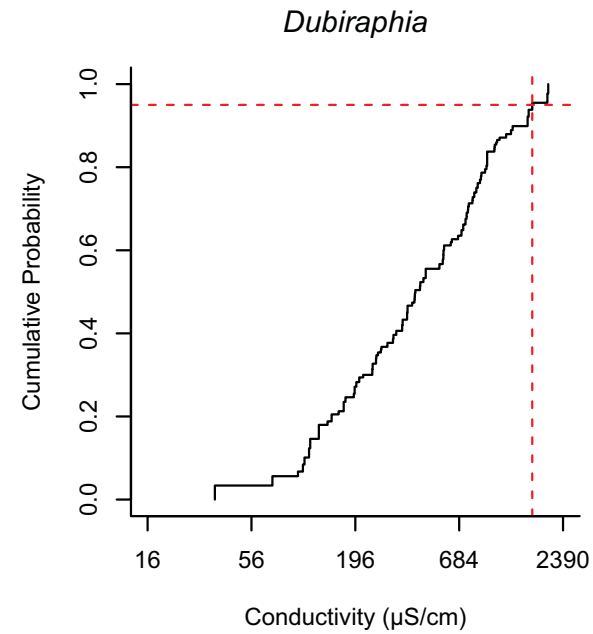
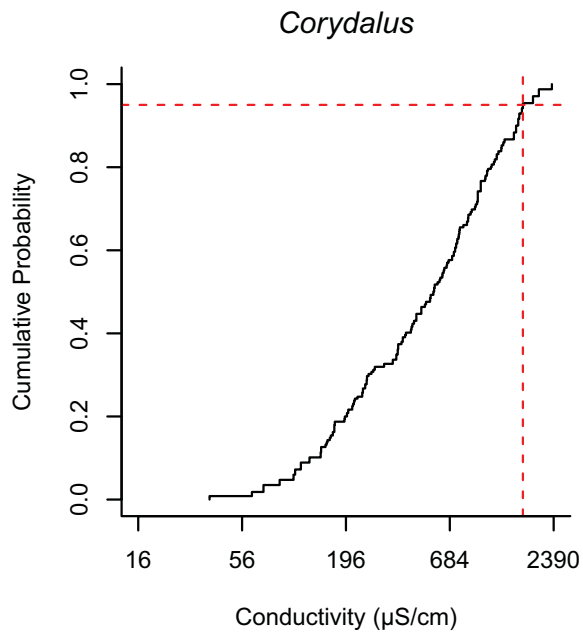
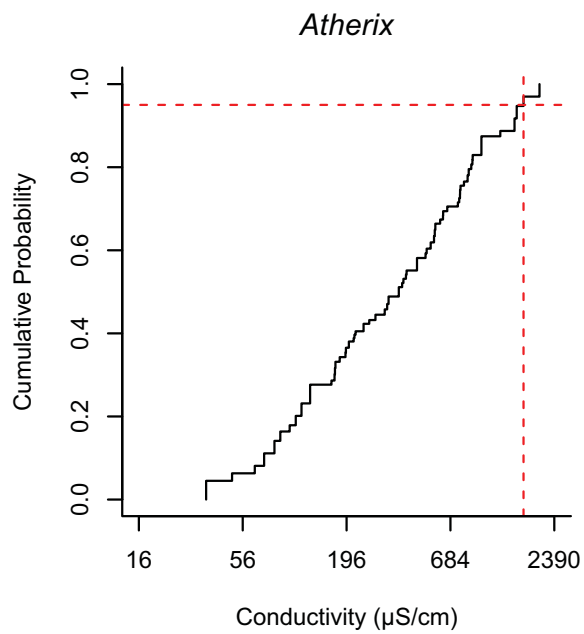




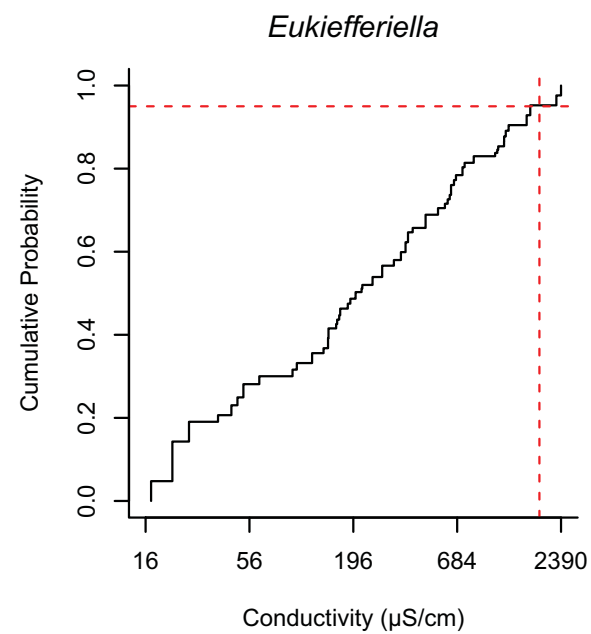
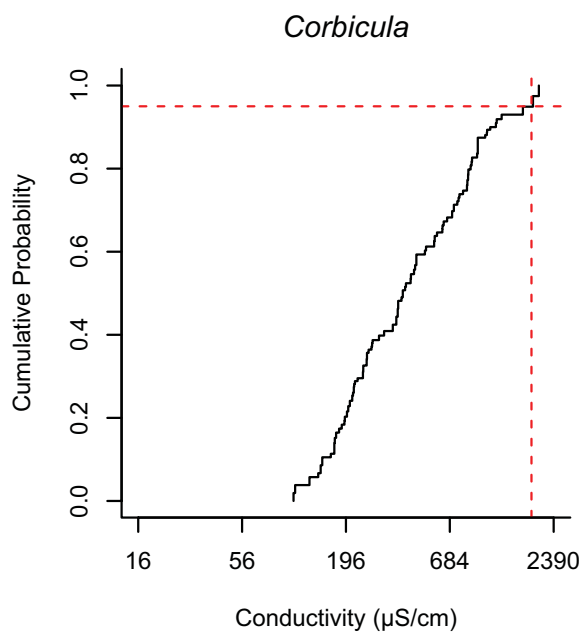
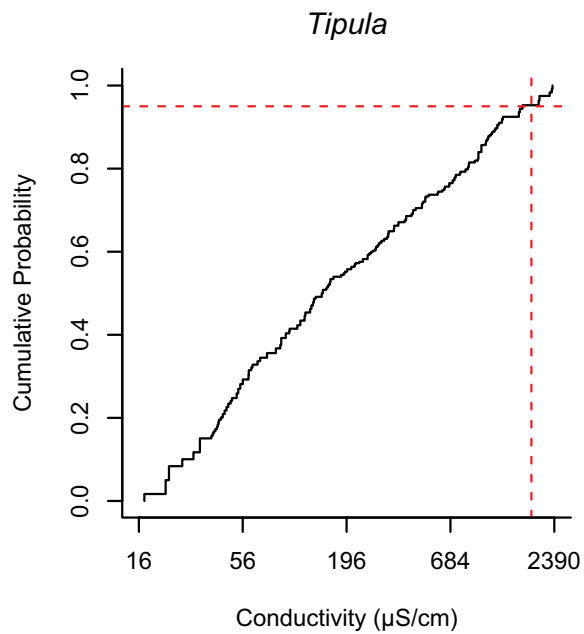
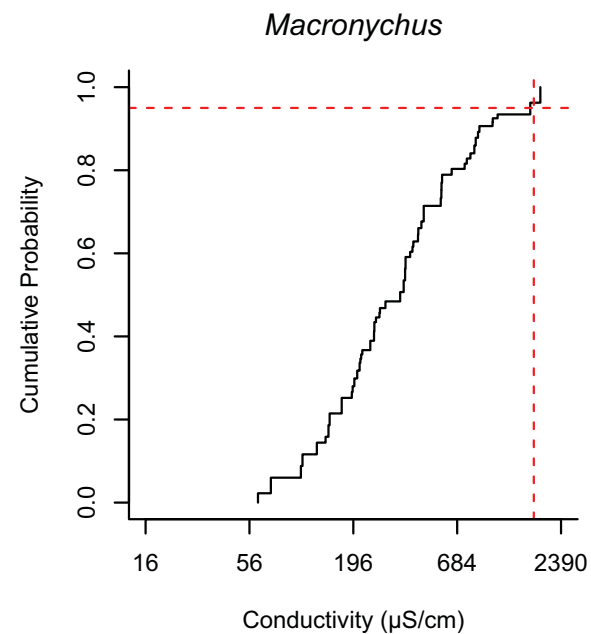
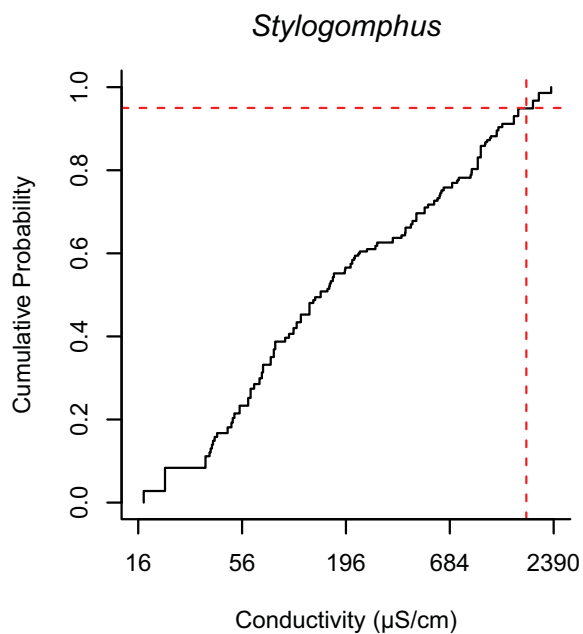
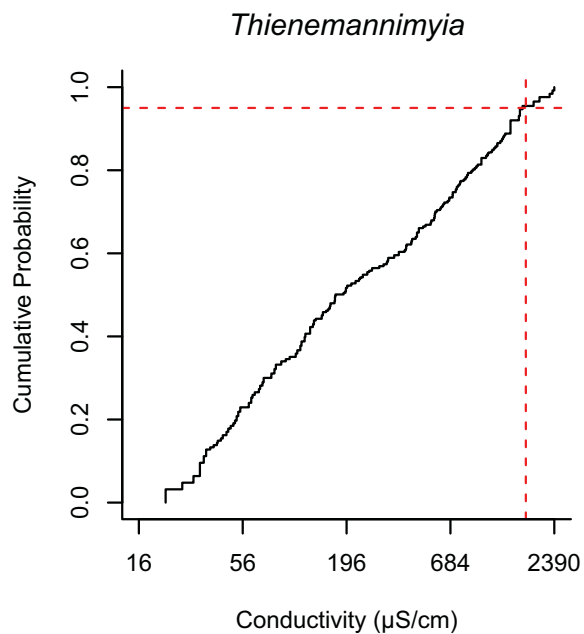
S1-f



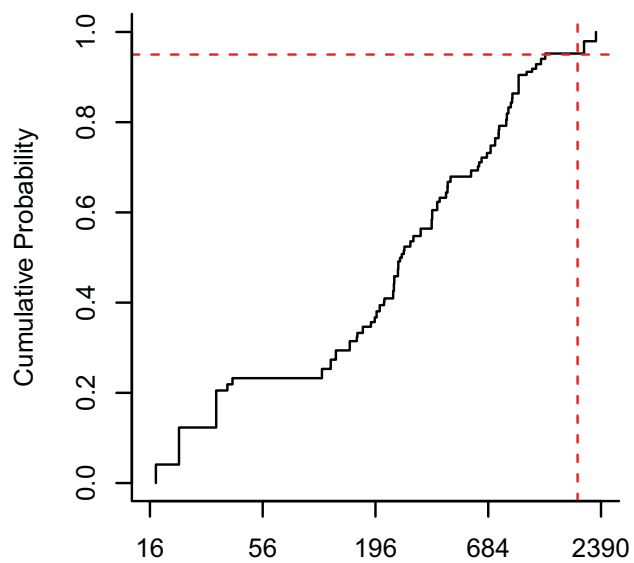
9|f



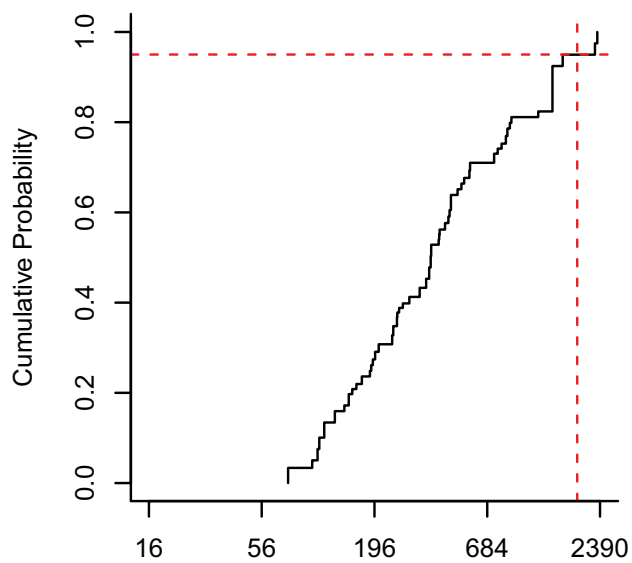
LI-f



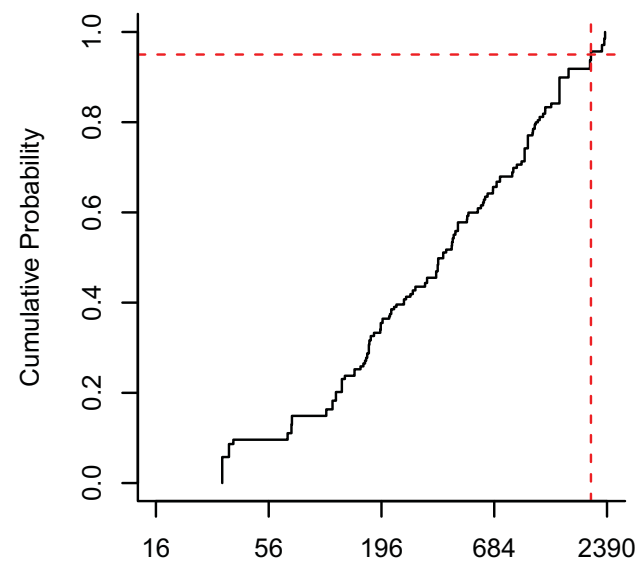
Sialis



Physella



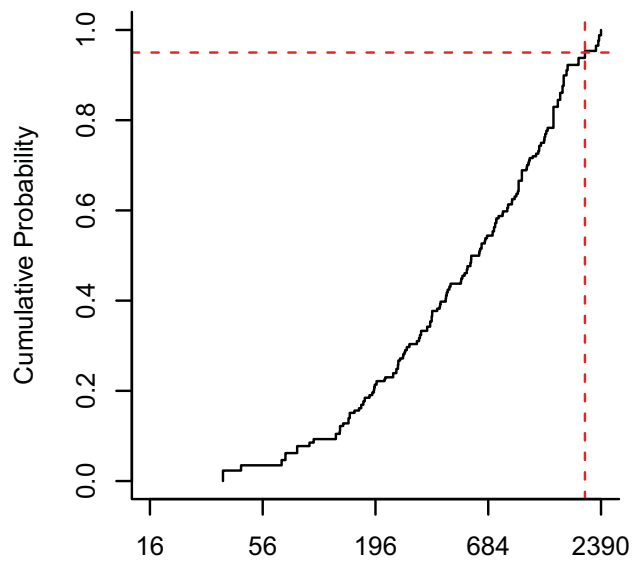
Chimarra



$8I-f$

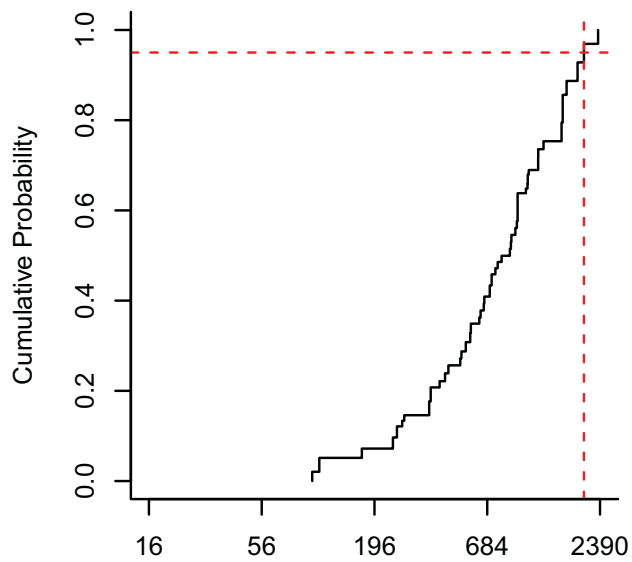
Conductivity ($\mu\text{S/cm}$)

Hemerodromia



Conductivity ($\mu\text{S/cm}$)

Tricorythodes



Conductivity ($\mu\text{S/cm}$)

Cricotopus

