DRAFT
DO NOT CITE OR QUOTE

EPA/600/R-10/023A
March 2010
External Review Draft

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 information is distributed solely for the purpose of predissemination peer review under applicable information quality guidelines. It has not been formally disseminated by the U.S. EPA. It does not represent and should not be construed to represent any Agency determination or policy.

Preferred Citation:
U.S. EPA (Environmental Protection Agency). 2010. 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-10/023A.

CONTENTS

CONTENTS (continued)

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS AND ACRONYMS

CDF	cumulative distribution function
DC_x	depletion concentration
GLIMPSS	genus level index of most probable stream status
HC_x	hazardous concentration
KDOW	Kentucky Division of Water
RBP	rapid bioassessment protocol
SSD	species sensitivity distribution
TMDL	total maximum daily load
U.S. EPA	United States Environmental Protection Agency
WABbase	Watershed Assessment Branch Data Base
WVDEP	West Virginia Department of Environmental Protection
WVSCI	West Virginia Stream Condition Index
XC_x	extirpation concentration

AUTHORS, CONTRIBUTORS, AND REVIEWERS

AUTHORS

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

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

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

Lei Zheng, Ph.D.
Tetra Tech, Inc.
Owings Mills, MD 21117

CONTRIBUTORS

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

Jennifer Flippin, M.S.
Tetra Tech, Inc.
Owings Mills, MD 21117

Jeroen Gerritsen, Ph.D.
Tetra Tech, Inc.
Owings Mills, MD 21117

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

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

Samuel P. Wilkes, M.S.
Tetra Tech, Inc.
Charleston, WV 25301

REVIEWERS

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

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

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

Chuck Hawkins, Ph.D.
Western Center for Monitoring and Assessment of Freshwater Ecosystems
Department of Watershed Sciences
Utah State University
Logan, UT 84322

REVIEWERS (continued)

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

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

ACKNOWLEDGMENTS

Greg Pond, Susan B. Norton, Teresa Norberg-King, Peter Husby, Peg Pelletier, Treda Grayson, Amy Bergdale, Candace Bauer, Brooke Todd, Lana Wood, Heidi Glick, Cris Broyles, Linda Tackett, Stacey Lewis, Bette Zwayer, 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 may be applied to waters in the Appalachian Region that are dominated by salts of SO_4^{2-} and HCO_3^- at 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 U.S. EPA's standard methodology for deriving water quality criteria. 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 in that it used the 5th percentile of a species sensitivity distribution (SSD) as the benchmark value. SSDs represent the response of aquatic life as a distribution with respect to exposure. It is implicitly assumed that if the exposure level is kept below the 5th percentile of the SSD, at least 95% of species will be protected. Data analysis followed the standard methodology in aggregating species to genera and using interpolation to estimate the percentile. It differs primarily in that the points in the SSDs are extirpation concentrations (XCs) rather than median lethal concentrations (LC_{50}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 percentile of the distribution of the probability of occurrence of a genus with respect to conductivity was used as a 95th percentile 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 percentile 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 rare species or waters designated by state and federal agencies as exceptional.

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 or 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 is a cause of impairment and the relationship between conductivity and biological responses apparently is not significantly confounded.

The chronic aquatic life benchmark value for conductivity derived from all-year data from West Virginia is 300 μS/cm. It is applicable to parts of West Virginia and Kentucky. It is expected to be applicable to the same regions in Ohio, Pennsylvania, Tennessee, and Maryland, but data from those states have not been analyzed. It may also be appropriate for other nearby regions such as Ecoregions 67 but has only been validated for use in Ecoregions 68, 69, and 70 at this time. However, this level may not apply when the relative concentrations of dissolved ions are not dominated by salts of SO_4^{-2} and HCO_3^-.

1. INTRODUCTION

At the request of U.S. Environmental Protection Agency (U.S. EPA) Regions 3, 4, and 5, and the Office of Water, 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 SO_4^{2-} and HCO_3^- anions at 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 U.S. EPA's standard methodology for deriving water quality criteria (Stephen 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: Na^+, Ca^{2+}, Mg^{2+}, K^+
Anions: Cl^-, HCO_3^-, CO_3^{2-}, SO_4^{2-}

The salinity of water may be expressed in various ways, but the most common is *specific conductivity*. Specific conductivity (henceforth simply conductivity) is the ability of a material to conduct an electric current measured in microsiemens per centimeter (μS/cm) standardized to 25°C. (In this report we use "conductivity" to refer to the measurement and resulting data and "salinity" to refer 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. A measure such as conductivity is necessary because the effects of salts are a result of exposure to 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 criterion for an individual ion could be under-protective. The ionic composition of mixtures of salts affects its toxicity (Mount et al., 1997). Therefore, this aquatic life benchmark for conductivity is applicable for streams in the Appalachian Region where conductivity is dominated by salts of SO_4^{2-} and HCO_3^- at circum-neutral to mildly alkaline pH.

Salinity has numerous sources (Ziegler et al., 2007). Freshwater can become increasingly salty due to evaporation, which concentrates salts such as occurs with 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), and runoff from treating pavements for icy conditions (Environment Canada and Health Canada, 2001; Evans and Frick, 2000).

Exposure of aquatic organisms to salinity is direct. Fish, amphibians, mussels, and aquatic macroinvertebrates are exposed as they ventilate their gills or other respiratory surfaces in the course of taking up oxygen. The respiratory surfaces contain specific structures to actively take up nutrient ions and control the osmotic balance of organisms. However, these structures may only be able to operate 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, their cuticle is permeable to the uptake of salt, and they take up salt using specialized external chloride cells on their gills (Komnick, 1977). Also, 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).

1.2. APPROACH

The approach used to derive the benchmark is based on the standard method for the U.S. Environmental Protection Agency's (U.S. EPA's) published Section 304(a) Ambient Water Quality Criteria. Those criteria are the 5th percentiles of species sensitivity distributions (SSDs) based upon laboratory toxicity tests, such that the goal is to protect 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 responds at exposure level X_1, the second most sensitive species responds 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 (Posthuma et al., 2002). Centiles of the distribution can be derived using interpolation, parametric regression, or nonparametric regression.

For the conductivity benchmark, the SSDs are derived from field data. There are several reasons that 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. For example, traditional toxicity testing and the resulting criteria derivation procedures used by EPA do not allow for the study of a pollutant's effects on migration, predation, and other behaviors or for species interaction. Furthermore, toxicity tests are rarely completed for the most susceptible species and sensitive life stages, which are difficult to identify or to maintain and test in the laboratory. The result is that the criteria are derived based on toxicity tests conducted upon species that can be cultured in a laboratory setting, and these tests do not include a substantial fraction of the species inhabiting an ecosystem. In sum, SSDs based on laboratory studies cannot replicate the full range of effects or species interactions that could reasonably be expected to occur in the environment (Suter et al., 2002).

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 subject to random 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 regions and 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 bare rock. 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 circum-neutral to mildly alkaline pH.

2. DATA SETS

Data are required to develop the benchmark and to support it. 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). 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. Nevertheless, conductivity values from WVDEP's reference sites were low and similar in different years (see Figure 2), providing evidence that the sites were reasonable reference sites. The 75th percentiles were below 200 µS/cm in most years.

2.2. DATA SOURCES

All data used in this study were taken from the WVDEP's in-house Watershed Assessment Branch Data Base (WABbase) 1999−2007. The WABbase contains data from Level III Ecoregions 66, 67, 69, and 70 in West Virginia (see Figure 1) (U.S. EPA, 2000; Omernik, 1987; Woods et al., 1996). Chemical, physical, and/or biological samples were collected from 3,286 distinct locations 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 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 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 U.S. EPA (WVDEP, 2008b). Quality assurance and standard procedures are described by WVDEP (2006, 2008a). All contracted analyses for chemistry and macroinvertebrate identification follow WV's internal quality control and quality assurance protocols. This is a well-documented, regulatory database. We judged the quality assurance to be excellent based on the database itself, supporting documentation, and experience of EPA Region 3 personnel.

Background 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 and from the U.S. EPA's Ecological Toxicology Database. (2) Information on the effects of dissolved salts on freshwater invertebrates was taken from standard texts and other physiological reviews. (3) The original data for Table 3 in Pond et al. (2008) were obtained from the authors to evaluate the relative contribution of different ions in drainage from valley fills of large scale surface mining. (4) The constituent ions for Marcellus Shale brine were provided by EPA Region 3 based on analyses by drilling operators.

2.3. DATA SET CHARACTERISTICS

Biological sampling usually occurred once per sampling period (March through October) with the WVDEP (1996−2007) sampling protocol. Repeat biological samples from the same location were minimal and not excluded from the data set. They represented approximately 4% of the sampled sites; therefore, no correction was made for pseudoreplication. Summary statistics for ion concentration and other parameters for the data set are provided in Table 1. The benchmark applies to waters with a similar composition.

Data from a sampling event at a site were excluded from calculations if they lacked a conductivity measurement, for obvious reasons. They were excluded if the samples were identified as being from a large river (>155 km^2), because the assemblages are not comparable with wadable streams (Flotemersch et al., 2001). They were excluded if the salt mixture was dominated by Cl^- rather than SO_4^{2-} (conductivity >1,000 µS/cm, SO_4 <125 mg/L, and Cl^- >250 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 Marcellas Shale brines.

Data were excluded from calculations if the organisms were not identified to the genus level, and a genus was excluded if it was never observed at reference sites or it was observed at <30 sampling sites. Invertebrate genera, that did not occur at WVDEP tier 1 reference sites represented 7% of the spring genera and 8% of the summer genera, were excluded from the SSDs. They were excluded so that the data would be relevant to potentially unimpaired conditions and so as to not include opportunistic salt-tolerant organisms. The exclusion of genera that were observed at fewer than 30 sampling locations in the composited ecoregion ensured reasonable confidence in the evaluation of the relationship between conductivity and the presence and absence of a genus.

Before identifying the extirpation concentration (e.g., 95th percentile extirpation concentration [XC_{95}]) for each genus, we removed effects of low pH 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).

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.

In the WABbase, 498 benthic invertebrate genera were identified of which 213 genera occurred at the 75 reference sites in the two ecoregions (see Table 2). Genera that did not occur at reference sites were excluded from the SSD (see Table 3). Greater than 90% 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. Ecoregions 69 and 70 had 304 genera in common. Of the overall 498 genera, 170 occurred at >30 sampling locations in Ecoregions 69 and 70. Of the genera occurring at >30 sampling sites, 128 genera occurred in Ecoregions 69 and 129 in Ecoregion 70.

3. METHODS

The derivation of the benchmark for conductivity includes three steps: First, the benchmark values (XCs) for the invertebrate genera were derived. Second, the XC_{95} values were used to generate an SSD and the 5th percentile of the distribution, the 5th percentile hazardous concentration (HC_{05}). (The HC_X terminology for concentrations derived from SSDs is not in the 1985 U.S. EPA method, but 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.

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 endpoint because the field data set reflects exposure over the entire life cycle of the resident biota. The 95th percentile was selected because it is more stable than the maximum value, yet still represents the extreme of an organism's tolerance of conductivity.

3.1. EXTIRPATION CONCENTRATION DERIVATION

The XC_{95} is estimated as the 95th percentile of the cumulative distribution of probabilities of observing a genus at a site with respect to the concurrently measured conductivity at that site. The XC_{95} estimates a conductivity value above which very few, less than 5%, of the observations of a particular genus are likely to be found.

Observed conductivity values were nonuniformly distributed across a range of possible values (see Figure 3), 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, we used weighted cumulative distribution functions 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 0.048 log conductivity units wide, that spanned the range of observed conductivity values. We then calculated the number of samples that occurred within each bin (see Figure 3). 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:

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

where x_i 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, and 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 $F(x) = 0.95$. Eq. 1 is an empirical cumulative distribution function, and the output is the proportion of observations of the genus that occur at a given conductivity or lower. 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 4 for the mayfly, Drunella. The horizontal dashed red line indicates where $F(x) = 0.95$, and the conductivity value at the intersection of this line and the CDF is the XC_{95} value.

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 4). Therefore, it is necessary to identify those values that are actual extirpation values. We did this by examining plots of probabilities of occurrence, estimated as the proportion of samples within each bin in which the genus was observed. Examples are shown in Figure 5. The solid line is provided to help illustrate the association, and its position is calculated using a nonparametric smoothing spline fit with 3 degrees of freedom. These curves are not used in the calculation of the XC_{95}. The conductivity at the red, horizontal, dashed line is the estimated XC_{95} from the weighted cumulative distribution. The actual XC_{95} was greater than (>) the calculated XC_{95} if the trend was increasing or flat and the seven highest conductivity bins were not zeros (see Figure 5). For example, the XC_{95} for Cheumatopsyche (an extremely salt tolerant genus) is >9,180 µS/cm (see Appendices D-1 and D-2).

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. Because low pH waters are in violation of existing water quality criteria and because the data set was large, we excluded sites with pH <6 before identifying the XC_{95} for each genus.

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. Therefore, we used the SSD for the combined spring and summer samples, thus avoiding the need to apply judgment to define seasons that vary with longitude and elevation.

To further evaluate the effect of confounders on the HC_{05}, XC_{95} values for the full year were determined from a data set from which sites were removed as follows:

- pH of <6 (removes acidity and associated dissolved metals as a cause);
- rapid bioassessment protocol (RBP) score <135 (removes marginal habitat conditions as a cause). The RBP score is WV's composited index of qualitative measures of habitat parameters such as bank erosion, stream sinuosity, embeddedness, and cover; and
- fecal coliform >400 colonies/100 mL (removes sources of potential organic enrichment and potential toxicants from sewage treatment plants, failing septic tanks and livestock as causes).

XC_{95} values were recalculated with the trimmed data set and compared. That analysis found that the HC_{05} (300 μS/cm) was similar to the HC_{05} (297 μS/cm) for the data set with only low pH removed so the data set was not partitioned for either RBP score or fecal coliform when calculating the benchmark.

Other potential confounders were evaluated, but they were not partitioned from the data set prior to calculating the XC_{95} and HC_{05} values. Rather, we evaluated the potential magnitude of confounding by determining 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 Figure 6). The cumulative percentile 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 interpolation to estimate the percentile between the XC_{95} values bracketing $P = 0.05$ (i.e., the 5^{th} percentile of modeled genera). The benchmark is obtained by rounding the HC_{05} to two significant figures as directed by Stephen et al. (1985).

Exploratory SSDs were developed using different data sets to evaluate effects of potentially influential factors. The results of these exploratory analyses and other tests are discussed in the treatment of confounding factors (see Appendix B).

3.4. CONFIDENCE BOUNDS

The purpose of this analysis is to characterize the statistical uncertainty in the benchmark value by calculating confidence bounds on the HC_{05} values. Because the XC_{95} values were estimated from field data and then the HC_{05} values were derived from those XC_{95} 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.

Bootstrap estimates of the XC_{95} were derived for each genus used in the derivation of the benchmark by sampling with replacement from the data set used to derive the benchmark 2,145 times (the number of observations in the data set) (Efron and Tibshirani, 1993). For each bootstrap sample, the XC_{95} was calculated 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 XC_{95} 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_{95}s from the original data set, the mean XC_{95}s of the bootstrap distributions, and the confidence intervals are shown for the most sensitive 35 genera in Figure 7.

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 a two-tailed 95% confidence bounds on these bootstrap-derived values.

3.5. ESTIMATING BACKGROUND

In general, a benchmark should be greater than natural background. The background conductivities of streams were estimated using reference sites from the WABbase. The 75th percentile of this distribution in the summer index period (August−October), which is the period of greatest conductivity, is 100 µS/cm for Ecoregion 69 and 234 µS/cm for Ecoregion 70. The 75th percentile was selected because sites were among the least disturbed based on best professional judgment (U.S. EPA, 2000). We also estimated the background conductivity for the area using only probability samples from the WABbase, which do not rely upon any selection criteria other than representativeness of a stream order. The 25th percentile was selected because impaired sites are also included in the random sample (U.S. EPA, 2000). A total of 1,271 probability-based samples were collected from Ecoregions 69 and 70. The background values, based on the 25th percentile, were 72 µS/cm for Ecoregion 69 and 153 µS/cm for Ecoregion 70. The bases for these methods are explained in Section 5.5.

4. RESULTS

4.1. EXTIRPATION CONCENTRATIONS

The XC_{95} values are presented in Appendix C. Values are calculated for all macroinvertebrate genera that were observed at a minimum of 30 sampling sites in the two ecoregions. Distributions of occurrence with respect to conductivity are presented for each genus of macroinvertebrate in Figure D-1 and the CDFs used to derive the XC_{95} values are presented in Figure D-2.

4.2. SPECIES SENSITIVITY DISTRIBUTIONS

SSDs for invertebrates in spring, summer, and the entire sampling year (March through October) are derived from XC_{95} values of 150 genera (see Figure 6). 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 8 for better viewing of the plots near the 5^{th} percentile of genera.

4.3. HAZARDOUS CONCENTRATION VALUES AT THE 5TH PERCENTILE

The hazardous concentration values at the 5^{th} percentile of the SSDs are summarized in Table 4. The HC_{05} spring value is lower than the summer value and similar to the full year. The HC_{05} for year-long XC_{95} values are similar to spring values because the spring-only genera have low XC_{95} values (see Figure 6). Other seasonal differences result from exclusion of some taxa due to sample sizes less than 30 in spring or summer or seasonal differences in the ability to sample some genera (see Section 5.2). Rounding the HC_{05} for all year of 297 µS/cm to two significant figures yields a benchmark value of 300 µS/cm (see Figure 8).

4.4. UNCERTAINTY ANALYSIS

The following HC_{05}s resulted from the bootstrap-derived statistics, a lower confidence bound of 225 µS/cm and an upper confidence bound of 305 µS/cm. These confidence bounds are asymmetrical with respect to the point estimate of 297 µS/cm. In general statistical practice, confidence bounds around estimates are not infrequently asymmetric. In the case of bootstrap generated estimates as 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 WV 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 results of independent studies. One estimate of that larger uncertainty is provided by comparing the all-year HC_{05} values derived from West Virginia and Kentucky data. Even though the data were obtained in different areas by different agencies using different protocols, the values differ by only 7% (see Appendix E for details). In addition, the 95% confidence bounds on the HC_{05} 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 WV 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. SELECTION OF INVERTEBRATE GENERA

Selection of genera to model can affect the results. Using the data set of all taxa includes taxa 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 HC_{05} by only 2% in the full year data set.

Genera are also selected for statistical reasons. We restricted genera used in analysis to those recorded at a minimum of 30 sampling sites to reduce the chance that an apparent extirpation is due to sampling variance and to increase the likelihood that the models and exploratory analyses for potential confounding are reasonably strong.

5.2. 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.

Some invertebrate genera are observed only in the spring probably due to the size of the early life stages or diapausing eggs in summer months. Because many of these same genera may be among the most sensitive genera, they could have a strong influence on the HC_{05}. Therefore, we evaluated the summertime conductivities of the streams in which these genera were found in springtime. The conductivities of these streams were often stable throughout the year (see Figure 9), but in some streams that supported sensitive genera in spring, conductivities increased in the summer. This suggests that some genera might withstand higher conductivities that coincide with certain parts of a genus's lifecycle; however, these genera clearly do not tolerate those conductivities in the spring. Lower effects levels were not due to insufficient test range of conductivities because exposures as high as 5,200 µS/cm occurred in the spring samples. Furthermore, when we partitioned the data into spring and summer periods, the XC values are lower in the spring than in the summer (see Figures 6 and 8).

The effects of seasonality and life history were evaluated by comparing occurrences of individual invertebrate genera and XC_{95} and HC_{05} values partitioned for season (see Figures 6 and 8). The data set was partitioned into spring and summer based on seasonal patterns of conductivity at WVDEP reference sites (see Figure 2). The spring season is March through June. The summer season is July through October. The exposure is characterized by water quality parameters measured on the same date that a taxon is observed in the stream. Both high and low conductivity streams are represented in both spring and summer samples. However, the conductivity in certain streams may be three times greater in the summer than the spring.

We cannot be sure whether the greatest exposures in summer are tolerated by the spring-emergent genera. However, streams with conductivity <300 µS/cm in summer are also below the benchmark in spring 98% of the time (see Figure 10). For simplicity, we recommend the year-round value (see Section 6), but seasonal variation should be considered when planning monitoring of conductivity.

5.3. 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 should not be derived by this method in that region. In this case, WVDEP's reference sites were included as well as many probability sites with >90% forest cover, which are believed to be representative of good- to high-quality systems.

5.4. DEFINING THE REGION OF APPLICABILITY

If the method 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, >90% of genera in common, and a common dominant source of the stressor of concern.

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. In this case, background conductivities for Ecoregions 69 and 70 are 100 and 234 µS/cm, respectively, using 75th percentiles of reference sites in West Virginia. Reference sites are sites that are judged to be among the best within a category, but they are not necessarily pristine or representative of natural background. Some reference sites have unrecognized disturbances or have recognized disturbances that are less than most others in their category. Some 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 percentile is based on precedent and on the collective experience of EPA field ecologists (U.S. EPA, 2000).

The background values based on the 25th percentile range between 72 µS/cm for Ecoregion 69 (n = 617) and 153 µS/cm (n = 654) for Ecoregion 70 for probability samples in West Virginia. Samples from a probability design include all types of waters including impaired sites. In some regions there are no pristine streams. To characterize the best streams, the 25th percentile is commonly used by EPA field ecologists (U.S. EPA, 2000). None of these values exceed the HC_{05} values in Table 4.

5.6. INCLUSION OF OTHER TAXA

Fish were not included because their occurrence is affected by stream size making it difficult to determine XC_{95} values. 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 30 samples. Additional analyses may be necessary to ensure protection of federally or state listed rare, threatened, or endangered species of fish, amphibians, and mussels.

5.7. TREATMENT OF RARE 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, Alloperla, Pteronarcys, Cordulegaster, Ephemera, and Sweltsa, be identified to species in subsequent monitoring to evaluate the risk to these threatened taxa. 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 5^{th} percentile when necessary to protect threatened, endangered, or otherwise important species.

5.8. 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 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 species within that genus (i.e., it 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 may have evolved to partition niches afforded by naturally occurring causal agents such as conductivity. 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 Stephen et al. (1985) by using genera until such time that the advantages and disadvantages of using species can be more fully studied.

5.9. USE OF MODELED OR EMPIRICAL DISTRIBUTIONS

When deriving XC and HC values, one might use a percentile 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 percentiles 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 percentiles of interest. 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 this case, data are abundant, and either the empirical or modeled methods could work well. 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. Similarly, the HC_{05} is determined by interpolation of points on the empirical distributions of XC_{95} values as described in Stephan et al. (1985).

5.10. TREATMENT OF CAUSATION

Causation should not be an issue in laboratory toxicity tests, but, even with rigorous treatment of confounders, skeptics 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).

5.11. TREATMENT OF MIXTURES

In natural waters, salinity is a result of mixtures of ions. We use conductivity as a measure of the mixture. However, 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 SO_4^{2-} and HCO_3^{-} anions at circum-neutral to mildly alkaline pH. 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 were 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, ions of Na^{+} and Cl^{-} are the two most common in seawater and brines from Marcellus Shale drilling operations (Bryant et al., 2002). 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 11a−e).

6. AQUATIC LIFE BENCHMARK

The aquatic life benchmark of 300 µS/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 225 µS/cm and the upper bound is 305 µS/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, 319 µS/cm with a lower confidence bound of 180 µS/cm and an upper bound of 439 µS/cm (see Appendix E).

The method used to develop the benchmark is an adaptation of the standard method for deriving water quality criteria for aquatic life (Stephan et al., 1985), so it is supported by precedent. Because the organisms are exposed throughout their life cycle, this is a chronic value.

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 salts of SO_4^{2-} and HCO_3^- anions at circum-neutral pH. The aquatic life benchmark for conductivity is applicable to parts of West Virginia, which provided the data for its derivation, and Kentucky, which gave essentially the same result. It may be applicable to Ohio, Tennessee, Pennsylvania, and Maryland in Ecoregions 68, 69, and 70. (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 E]). The aquatic life benchmark may also be appropriate for other nearby regions. However, this level may not apply when the relative concentrations of dissolved ions are different (see Table 1 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.

Bear, JA; Cheng, HD; Sorek, S; et al. (1999) Seawater intrusion in coastal aquifers: concepts, methods and practices. Norwell, MA: Kluwer Academic Publishers.

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/MTMVFChemistryPart1.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/html/report.htm.

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, M; Frick, C. (2000) 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.

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.

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.

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.

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. Available online at http://www.water.ky.gov/NR/rdonlyres/ED76CE4E-F46A-4509-8937-1A5DA40F3838/0/coal_mining1.pdf.

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 27:717−737.

Posthuma, L; Suter, GW, II; Traas, TP; eds. (2002) Species sensitivity distributions in ecotoxicology. Boca Raton, FL: Lewis Publishers.

Rengasamy, P. (2002) Transient salinity and subsoil constraints to dryland farming in Australian sodic soil: an overview. Aust J Exp Agri 42:351−361.

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, Washington, D.C. PB85-227049. Available online at http://www.epa.gov/waterscience/criteria/library/85guidelines.pdf.

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, GW, Suter, II, Traas, T; eds. Species sensitivity distributions in ecotoxicology, L. Boca Raton: Lewis Publishers; pp 437−474.

Tarin, JJ, A Cano, JJTarbin.. (2000) Fertilization in protozoa and metazoan animals: Cellular and molecular aspects. Berlin: Springer-Verlag.

U.S. EPA (U.S. 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 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 ftp://ftp.epa.gov/wed/ecoregions/us/useco.pdf.

Werner, AD. (2009) A review of seawater intrusion and its management in Australia. Hydoegeol J Published on line http://www.springerlink.com/content/q75147m6j6766470/fulltext.pdf (accessed 10/14/2009).

WVDEP (West Virginia Department of Environmental Protection). (2006) Department of Water and Waste Management Division of Water and Waste Management quality assurance project plan for watershed branch monitoring activities. 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.wvdep.org/Docs/16495_WV_2008_IR_Supplements_Complete_Version_EPA_Approved.pdf.

WVDEP (West Virginia Department of Environmental Protection). (2008b) West Virginia Department of Environmental Protection Watershed Assessment Branch 2008 Standard Operating Procedures. Vo. 1 SOPP. 206.

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).

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/600R-96/077.

Ziegler, CR; Suter, GW, II; Kefford, BJ. (2007) Candidate Cause: Ionic Strength. Available online at www.epalgov/caddis (accessed 10/13/09).

Table 1. Summary statistics of the measured water quality parameters

	Units	Min	25th percentile	Median	75th percentile	Max	Mean	Valid *N*
Conductivity	μS/cm	15.4	153	269.5	576	11,646	490.90	2,145
Hardness	mg/L	0.51	52.04	94.97	196.44	1,491.79	181.73	1,087
Alkalinity	mg/L	0.2	32.05	69.6	120	560	86.57	1,366
SO_4	mg/L	1	17.5	40	170.8	6,000	179.4	1,365
Ca, total	mg/L	0.002	14	26.2	51.1	430	48.89	1,091
Mg, total	mg/L	0.05	3.94	6.6	14.91	204	14.39	1,089
Chloride	mg/L	1	3	5.53	12.03	1,153	17.93	1,055
TSS	mg/L	1	3	4	6	190	6.5	1,380
Fe, total	mg/L	0.005	0.14	0.27	0.51	110	0.75	1,369
NO_2-NO_3	mg/L	0.01	0.1	0.21	0.38	30	0.47	1,113
Al, total	mg/L	0.01	0.09	0.12	0.24	12	0.27	1,372
Al, dissolved	mg/L	0.011	0.02	0.05	0.06	0.93	0.055	1,225
Fe, dissolved	mg/L	0.001	0.02	0.05	0.061	31.8	0.147	1,196
Mn, total	mg/L	0.003	0.02	0.042	0.105	7.25	0.145	1,367
Mn, dissolved	mg/L	0.01	0.03	0.07	0.22	1.06	0.16	19
Total phosphate	mg/L	0.01	0.02	0.02	0.03	2.36	0.039	1,116
Se, dissolved	mg/L	0.001	0.001	0.001	0.001	1.26	0.006	290

Table 1. Summary statistics of the measured water quality parameters (continued)

	Units	Min	25th percentile	Median	75th percentile	Max	Mean	Valid *N*
Se, total	mg/L	0.0003	0.001	0.001	0.005	1.26	0.006	472
Fecal coliform	counts/100 mL	0.19	40	175	600	250,000	1,515	1,998
DO	mg/L	1.0	8.2	9.2	10.2	18.4	9.2	2,118
pH	standard units	6.02	7.29	7.63	7.97	10.48	7.6	2,145
Catchment area	km^2	0.08	3.71	10.47	31.64	153.82	24.86	2,141

TSS = Total suspended solids

Note: K^+ and Na^+ not measured

Table 2. Number of samples with reported genera and conductivity meeting our acceptance criteria for calculating the benchmark value. Number of samples is presented for each month, ecoregion, and database

Region	Month								Total
	3	4	5	6	7	8	9	10	
69	1	63	188	103	79	267	232	58	987
70	4	186	232	179	194	237	118	8	1,158
Total	5	249	420	282	273	504	350	62	2,145

Table 3. Genera excluded from 95th percentile extirpation concentration calculation because they never occurred at reference sites

Argia	Baetisca	Calopteryx	Chironomus	Corbicula
Dineutus	Ferrissia	Fossaria	Palpomyia	Paratendipes
Nanocladius	Prostoma	Sphaerium	Stenochironomus	Stictochironomus
Tokunagaia	Tribelos	Tricorythodes		

Table 4. Hazardous concentration at the 5th percentile for invertebrates in Ecoregions 69 and 70

Season	HC_{05}
All year	297 µS/cm
Spring	322 µS/cm
Summer	479 µS/cm

Omernik Level III Ecoregions 69 and 70

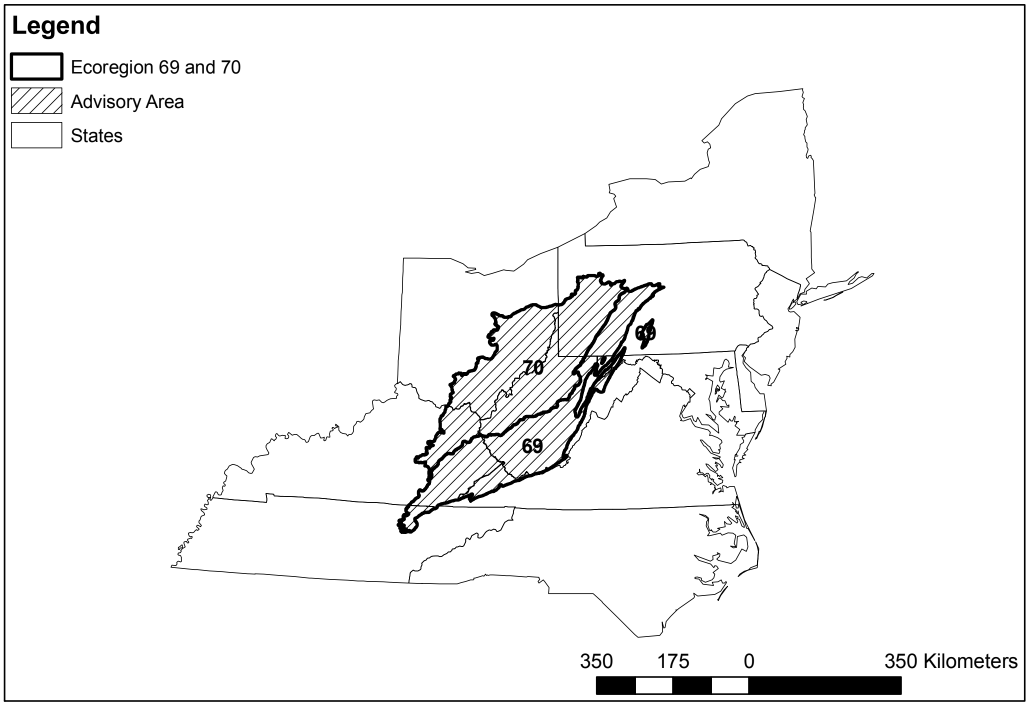

Figure 1. Data are from Tier III Ecoregions 69 and 70 spanning the states of Ohio, Pennsylvania, Kentucky, Tennessee, West Virginia, and Maryland.

Data source: State outlines from U.S. EPA Base Map Shapefile, Omernik Level III Ecoregions from National Atlas (National Atlas.gov) Projection NAD1983UTM17N.

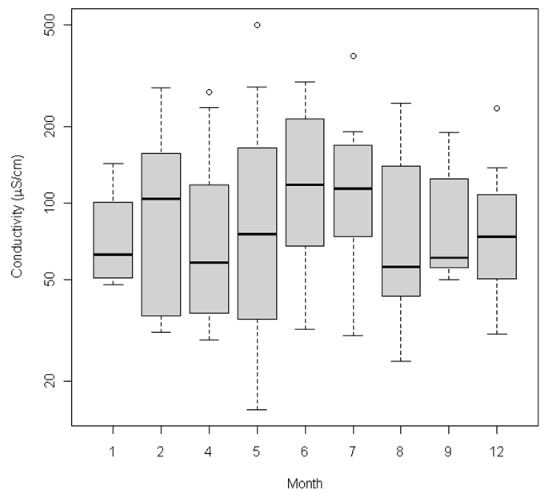

Figure 2. Box plot showing seasonal variation of conductivity (µS/cm) in the reference streams of Ecoregions 69 and 70 in West Virginia from 1999 to 2006. A total of 97 samples from 70 reference stations were used for this analysis. The 75th percentiles were below 200 µS/cm in most years.

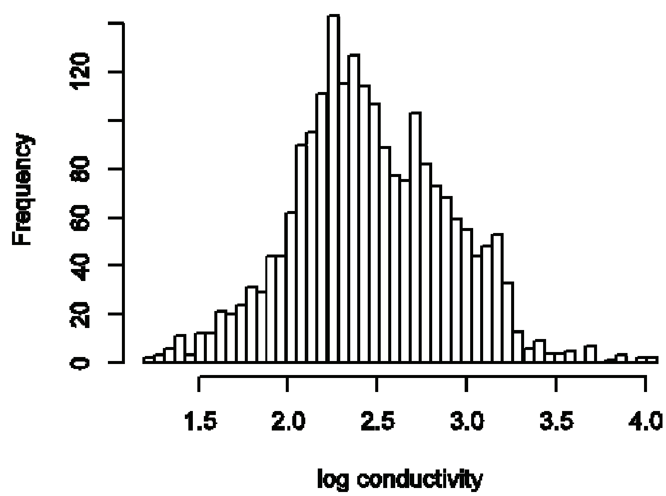

Figure 3. Histogram of the frequencies of observed conductivity values in samples from Ecoregions 69 and 70 from March to October. More of the sampled sites were in the midrange than in the extremes.

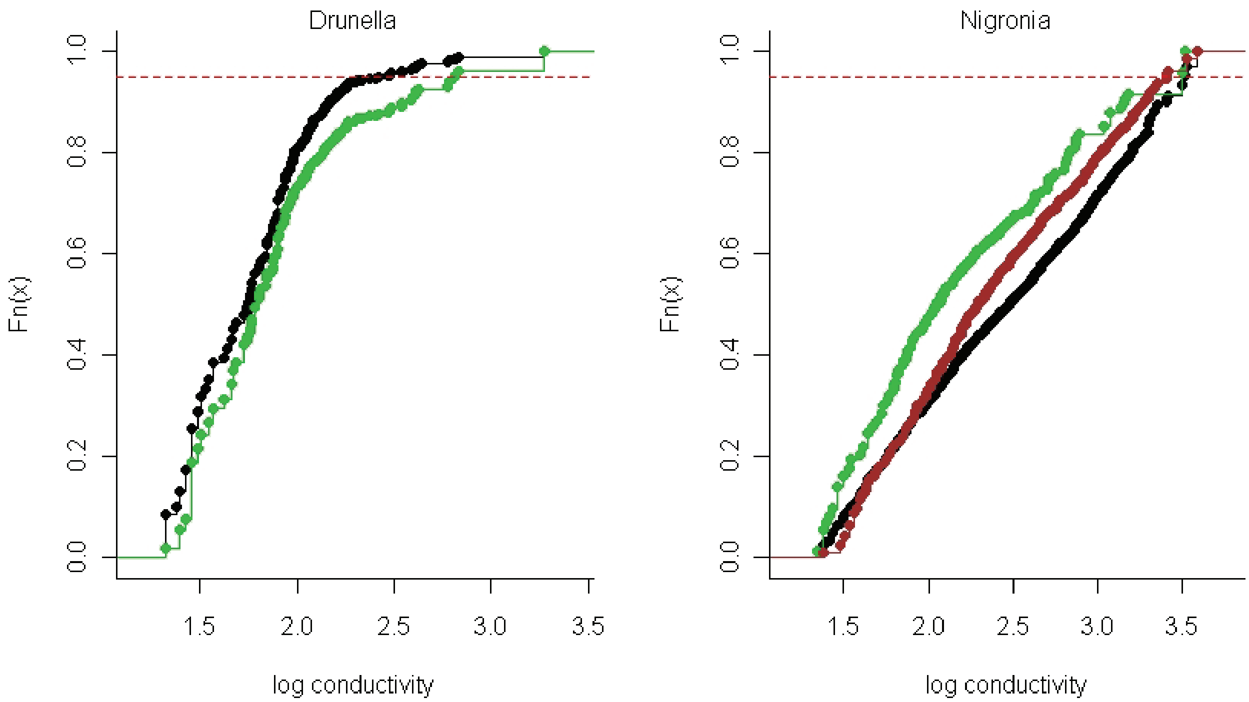

Figure 4. Example of a weighted CDF and the associated 95th percentile extirpation concentration value. Each point shows the weighted proportion of samples with Drunella or Nigronia present at (Fn(x)) the conductivity less than the indicated conductivity value (µS/cm). The XC_{95} is the conductivity at the 95th percentile of the cumulative distribution function (CDF) (horizontal dashed line). The CDF was calculated from observations from March through October (all year; black connected points) from March through June (spring; green connected points), and from July through October (summer; red connected points). As there were fewer than 30 observations of Drunella between July and October, no CDF was developed for the summer index period. In a CDF, genera that are affected by increasing conductivity (e.g., Drunella) show a steep slope and asymptote well below the measured range of 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 perceptible asymptote.

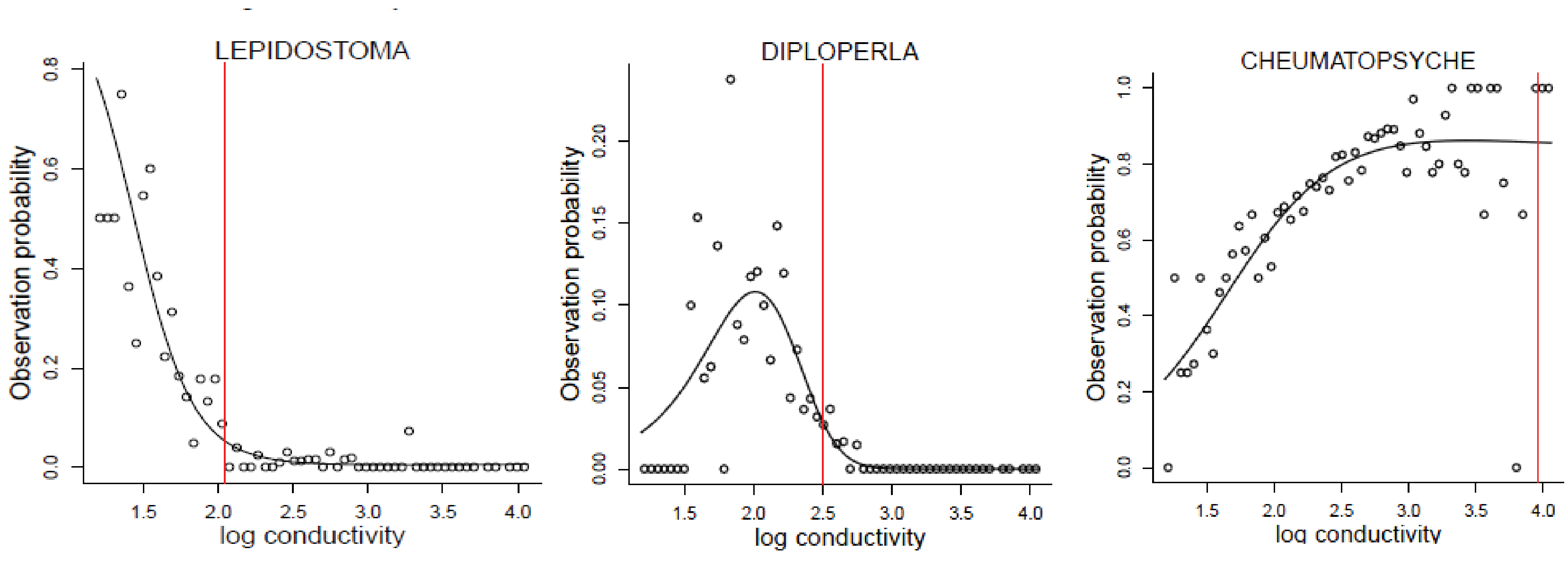

Figure 5. Three typical distributions of observation probabilities (March through October). Open circles are the probabilities of observing the genus within a range of conductivities. Circles at zero probability indicate no individuals at any sites were found at these conductivities. The line fitted to the probabilities is 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.

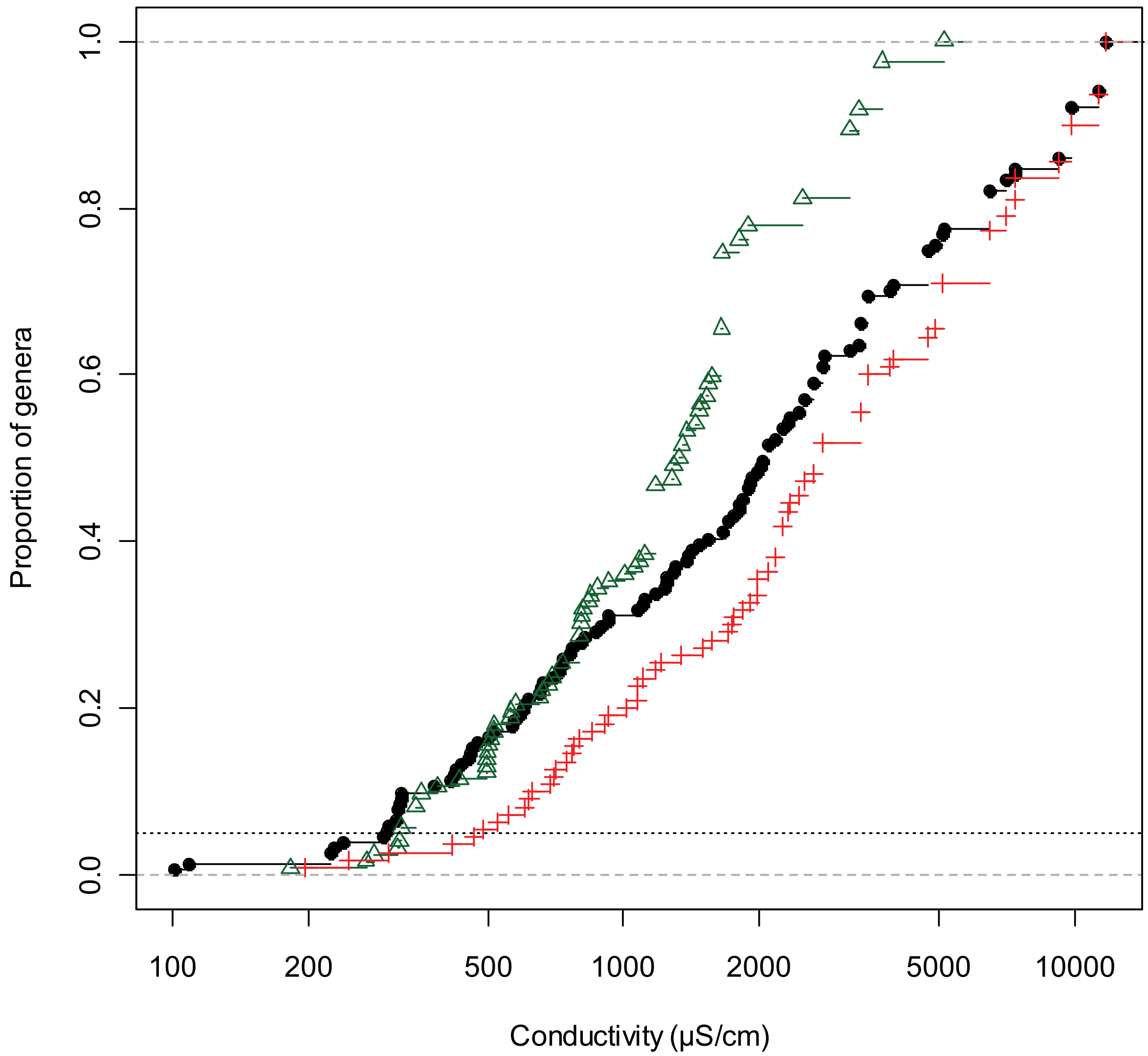

Figure 6. The species sensitivity distribution for all year (March through October [black circles], March through June [green triangles], and July through October [red +]). More than 100 genera are included. The HC_{05} is the conductivity at the intercept of the CDF with the horizontal line at the 5th percentile.

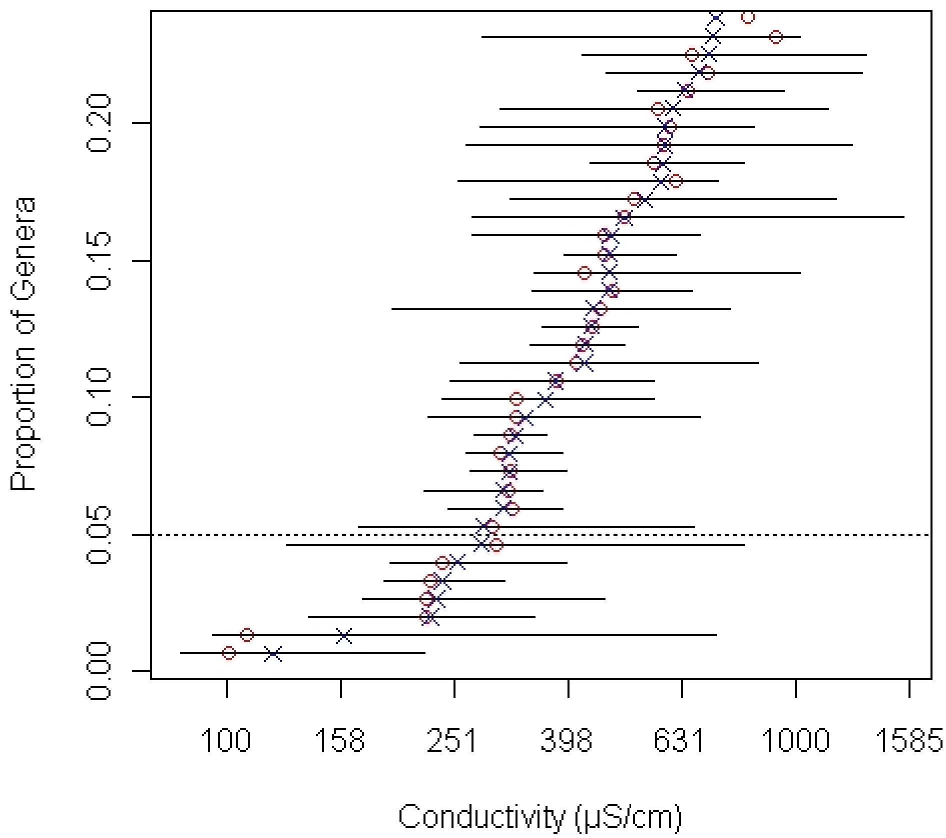

Figure 7. The cumulative distribution of XC_{95} values for the 35 most sensitive genera (red circles) and the bootstrap-derived means (blue x symbol) and two-tailed 95% confidence intervals (whiskers). The 5th percentile is shown by the dashed line.

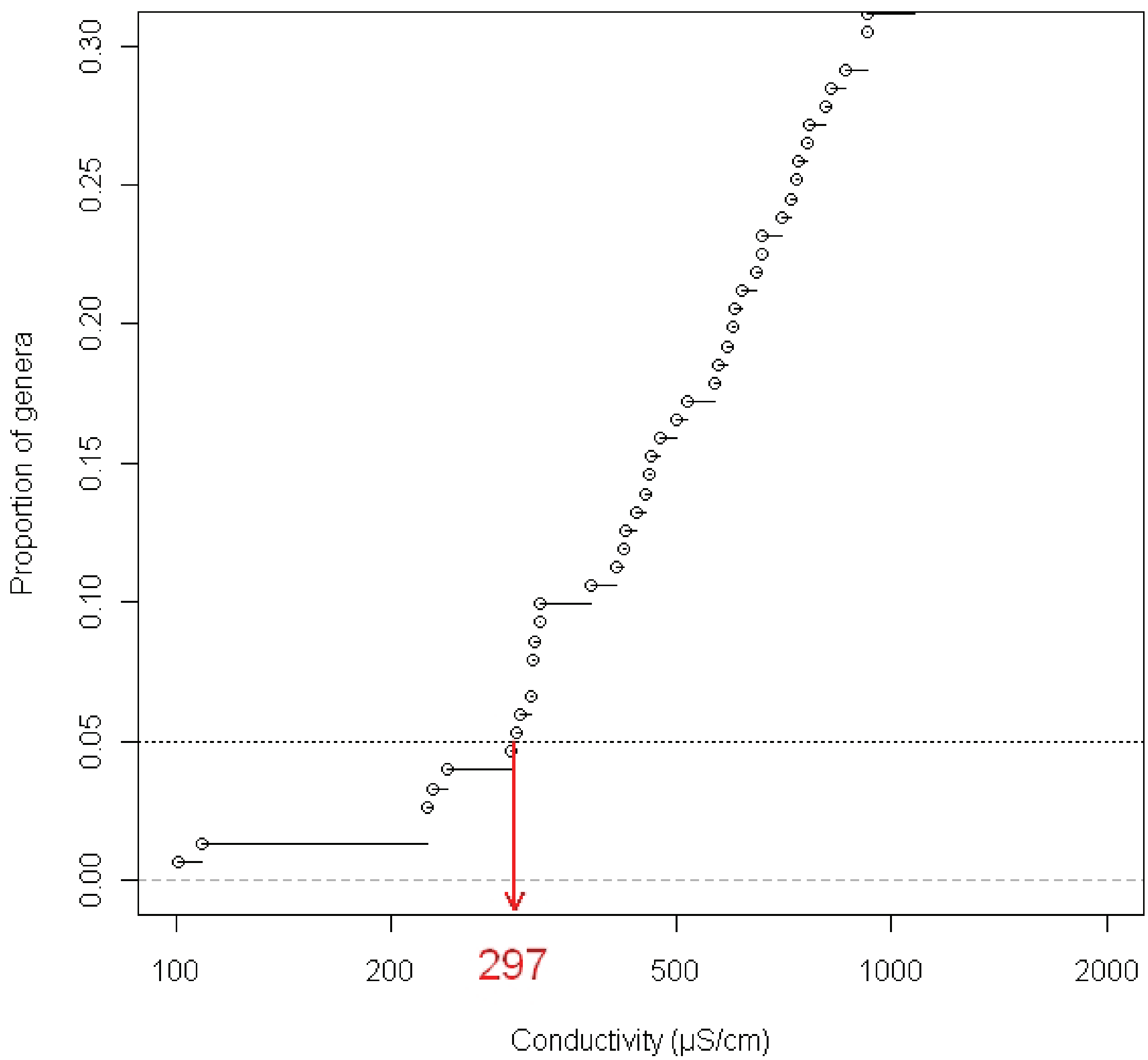

Figure 8. Species sensitivity distribution for all year. The dotted horizontal line is the 5th percentile. The vertical arrow indicates the HC_{05} of 297 µS/cm. Only the lower 50 genera are shown to better discriminate the points in the left side of the distribution.

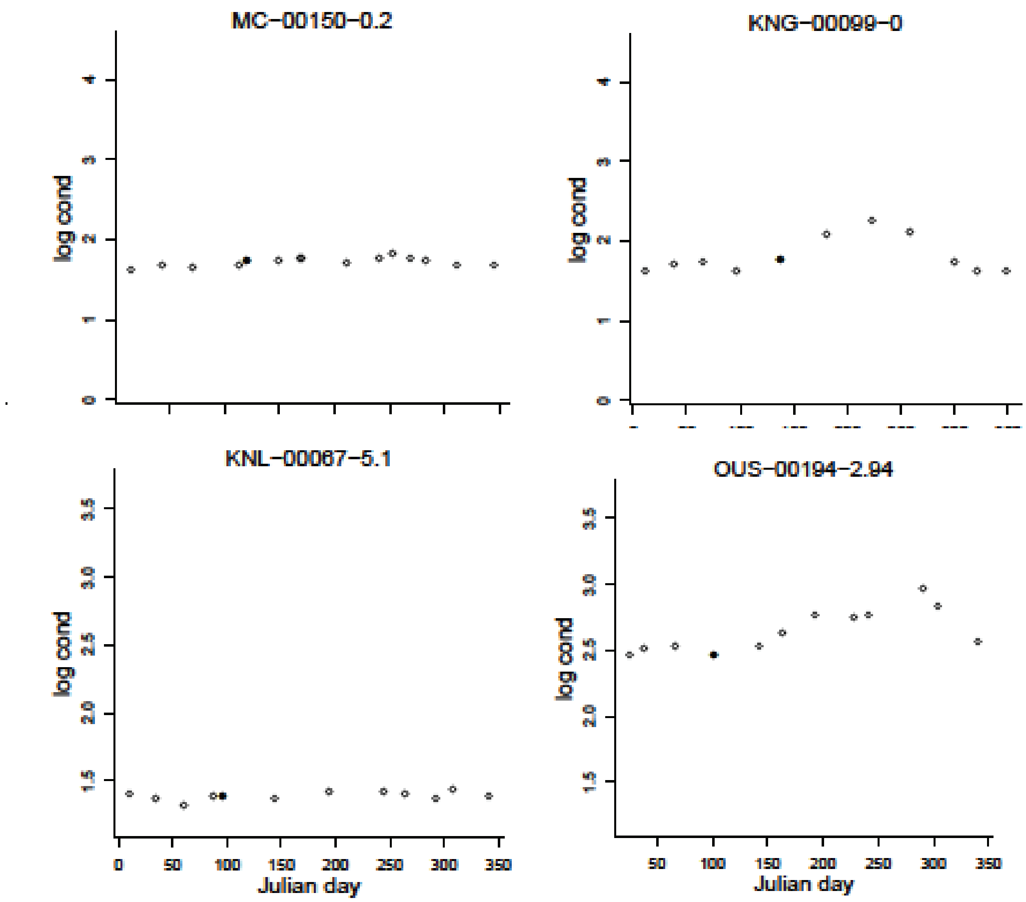

Figure 9. Examples of a monthly year-long stream conductivity record in a stream. Filled circles: events when macroinvertebrates were sampled concurrently with conductivity; open circles: physical-chemical sampling only. Some streams had a steady level of conductivity (plots on left); others increased in summer month (plots on right).

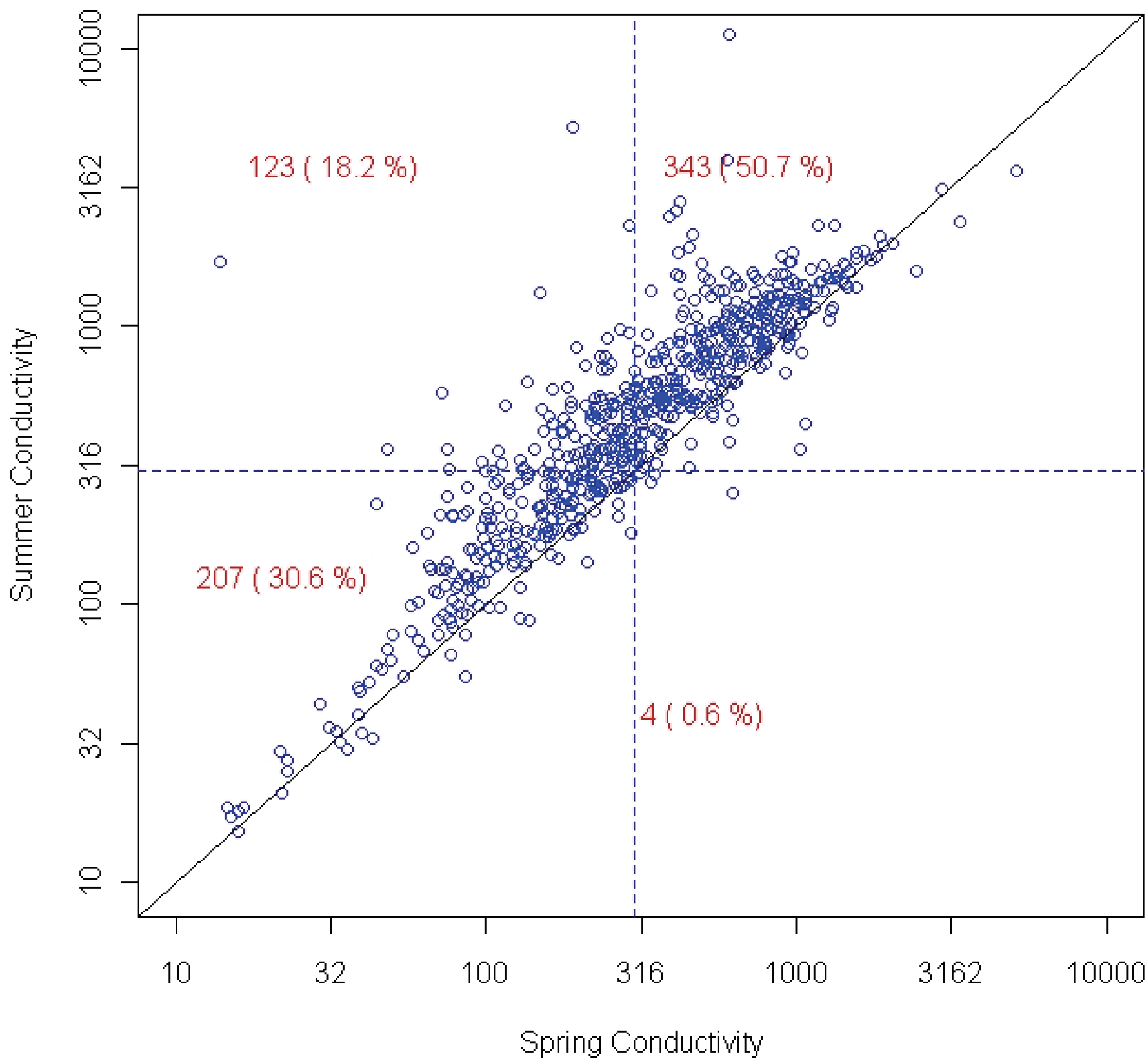

Figure 10. Correlation of conductivity values sampled from the same site in spring and summer. When conductivity is <300 µS/cm (broken lines) in March thru June, the conductivity is <300 µS/cm in the same stream 63% of the time July through October. When the conductivity is <300 µS/cm in July through October, the conductivity in the same stream March through June is <300 µS/cm 98% of the time.

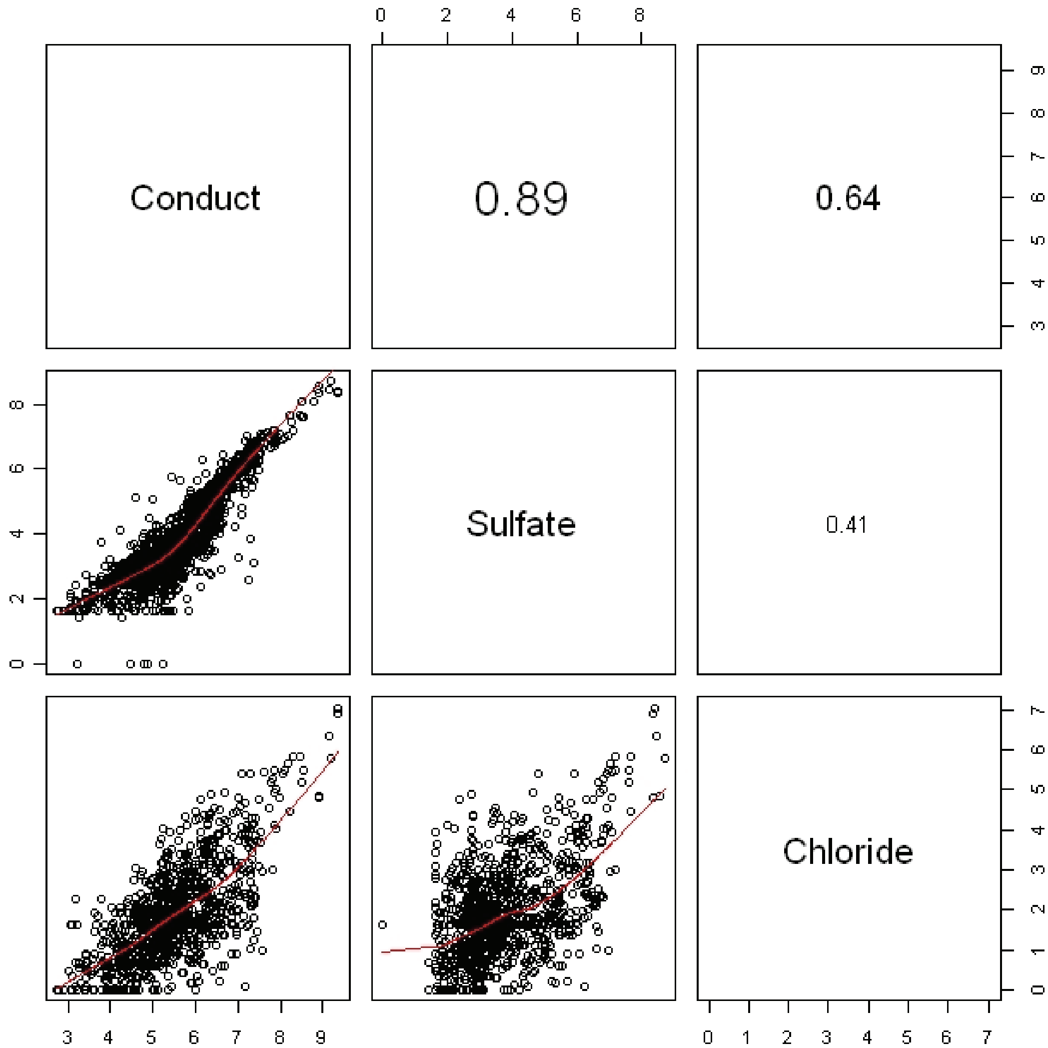

Figure 11a. Anions. Matrix of scatter plots and absolute Spearman correlation coefficients between conductivity (µS/cm), sulfate (mg/L), and chloride concentrations (mg/L) in streams of Ecoregions 69 and 70 in West Virginia. All variables are logarithm transformed. The smooth lines are the locally weighted scatter plot smoothing (LOWESS) lines (span = 2/3).

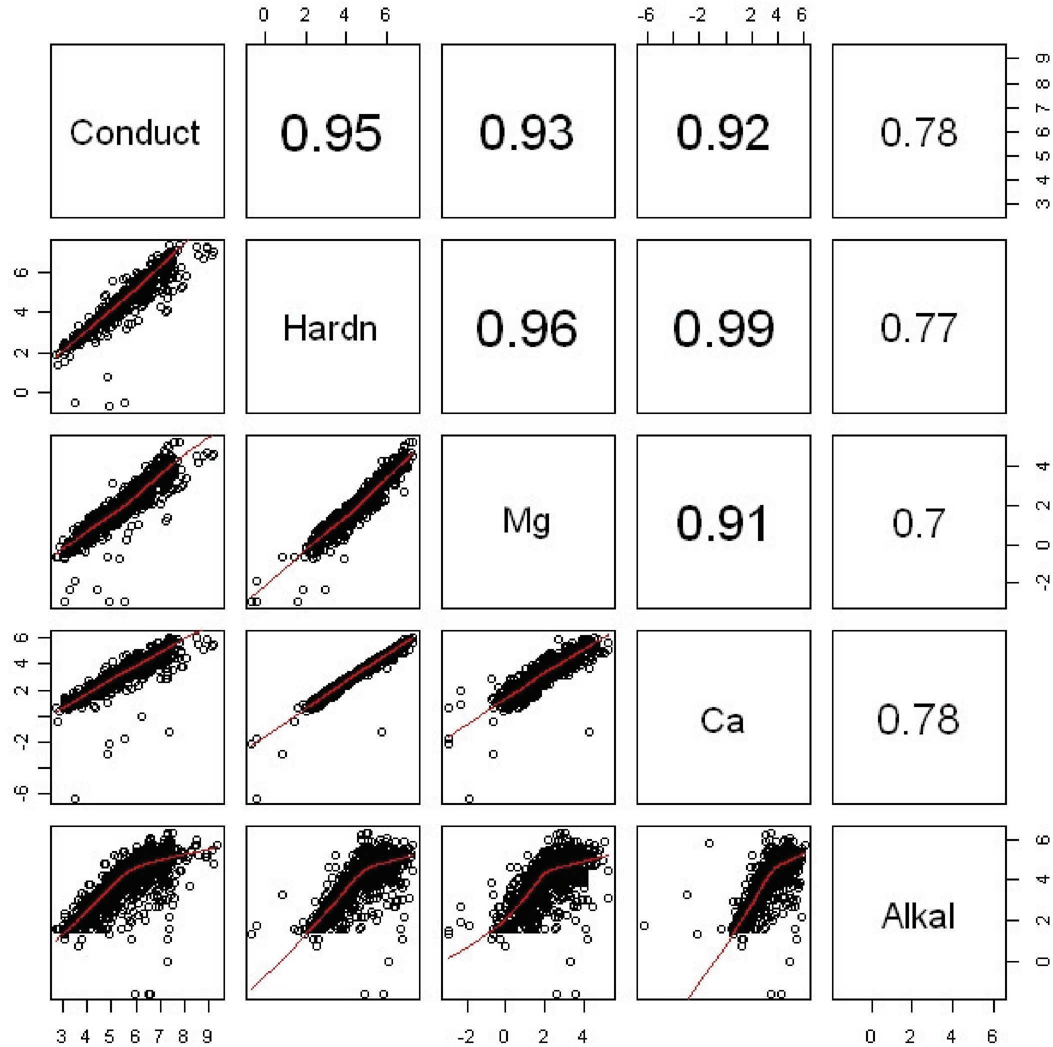

Figure 11b. Cations. Matrix of scatter plots and absolute Spearman correlation coefficients between conductivity (µS/cm), hardness (mg/L), Mg (mg/L), Ca (mg/L), and alkalinity (mg/L) in the streams of Ecoregions 69 and 70 in West Virginia. All variables are logarithm transformed. The smooth lines are the locally weighted scatter plot smoothing (LOWESS) lines (span = 2/3).

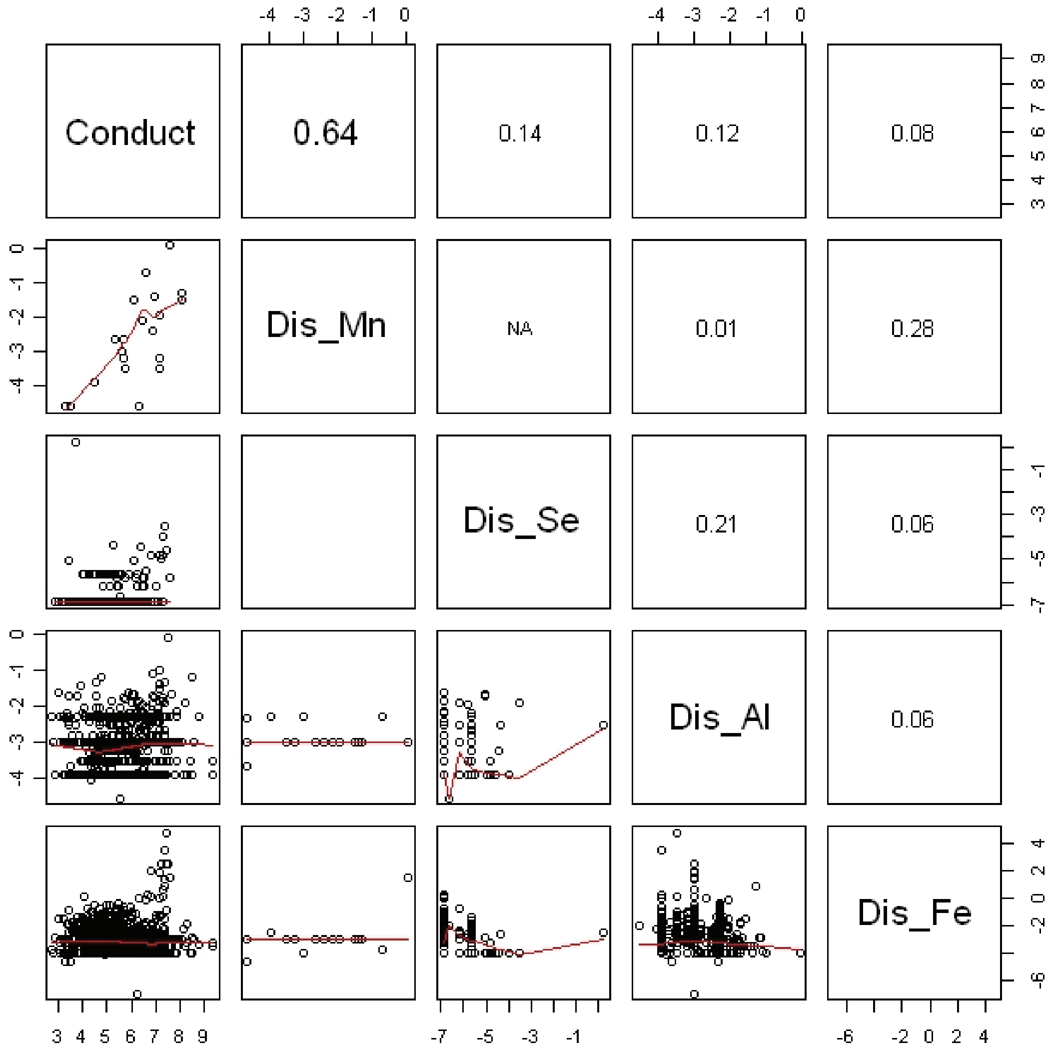

Figure 11c. Dissolved metals. Matrix of scatter plots and absolute Spearman correlation coefficients among conductivity (µS/cm) and dissolved metal concentrations (mg/L) in the streams of Ecoregions 69 and 70 in West Virginia. All variables are logarithm transformed. The smooth lines represent the locally weighted scatter plot smoothing (LOWESS) lines (span = 2/3).

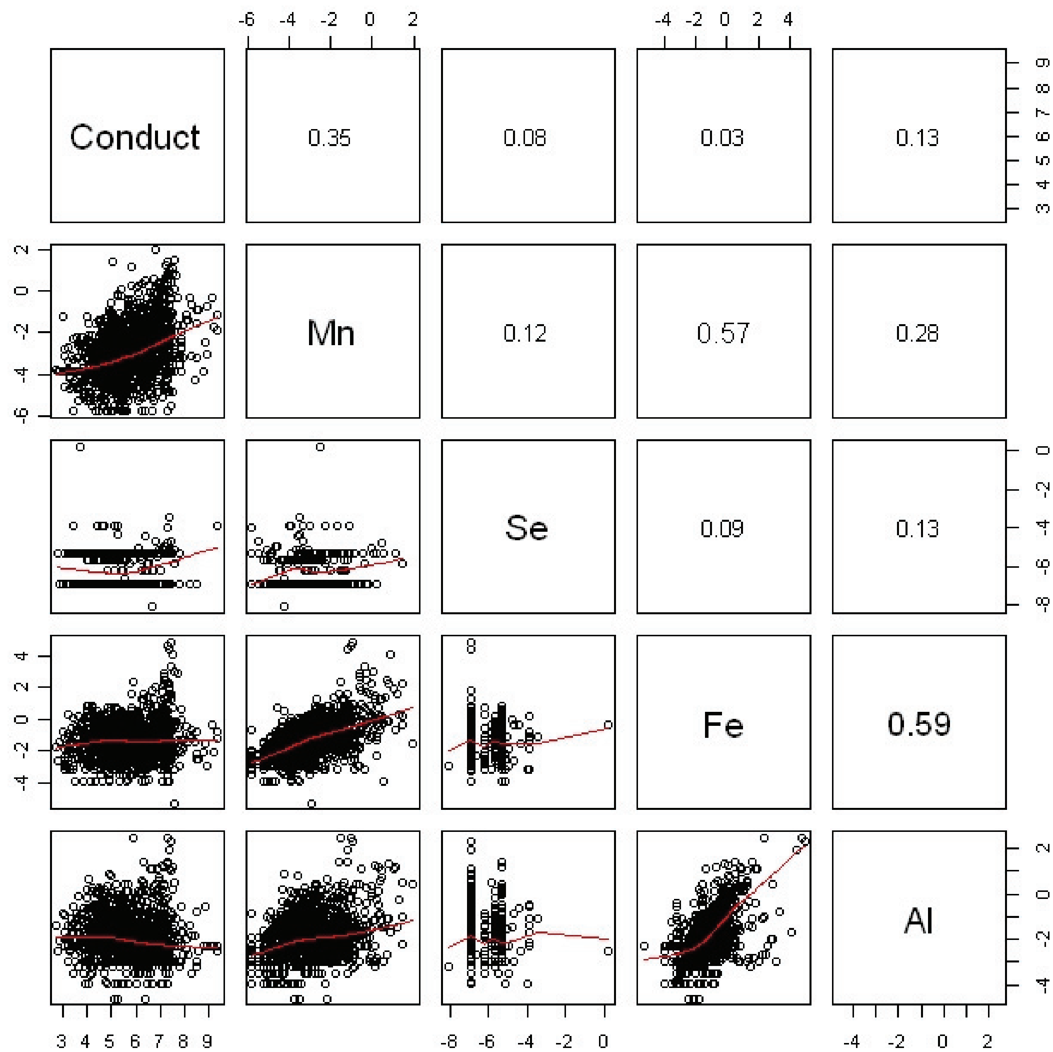

Figure 11d. Total metals. Matrix of scatter plots and absolute Spearman correlation coefficients between conductivity (µS/cm) and total metal concentrations (mg/L) in the streams of Ecoregions 69 and 70 in West Virginia. All variables are logarithm transformed. The smooth lines represent the locally weighted scatter plot smoothing (LOWESS) lines (span = 2/3).

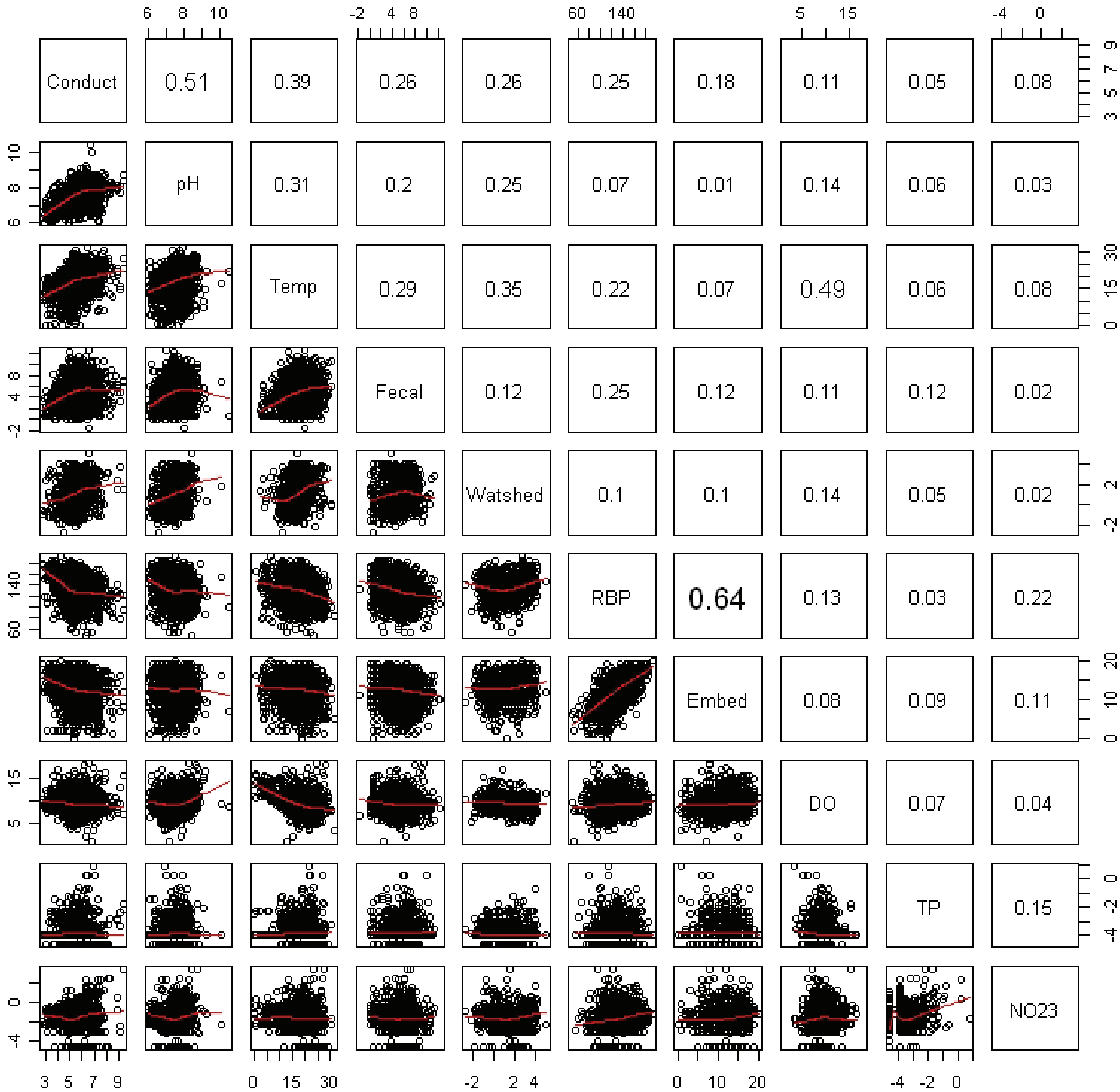

Figure 11e. Other water quality parameters. Matrix of scatter plots and absolute Spearman correlation coefficients between environmental variables in the streams of Ecoregions 69 and 70 in West Virginia. The smooth lines are locally weighted scatter plot smoothing (LOWESS) lines (span = 2/3). Conduct is logarithm transformed specific conductance (μS/cm); Temp is water temperature (°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).

APPENDIX A
CAUSAL ASSESSMENT

ABSTRACT

Because associations in the field are not necessarily causal, this appendix reviews the evidence that salts are a cause of impairment of aquatic macroinvertebrates in streams in Ecoregions 69 and 70 of West Virginia. The goal is to establish that salts are a general cause, not that they cause all impairments, nor that there are no other causes of impairment, nor that they cause the impairment at any particular site. The evidence is organized in terms of six characteristics of causation. The inferential approach is to weigh the body of evidence, as is done in epidemiology. The results are positive; the available evidence indicates that salts, as measured by conductivity, are a common cause of impairment of aquatic macroinvertebrates in the region of concern. The following appendix (B) addresses the potential for other variables to confound the effects of salts.

A.1. INTRODUCTION

To assure that that the association of conductivity with the extirpation of aquatic taxa reflects a causal relationship, we use epidemiological arguments. The most widely accepted epidemiological approach was first used to show that smoking causes cancer in humans (Hill, 1965; U.S. DHEW, 1964). It consists of weighing the available evidence on the basis of causal considerations. As in the case of tobacco smoke, conductivity represents a mixture, and its effects are not necessarily immediately apparent following exposure. Hill's approach for establishing a probable causal relationship has been adapted for ecological applications (Fox, 1991; U.S. EPA, 2000; Suter et al., 2002; Cormier et al., 2010). We rely on the same approach to demonstrate that mixtures of ions that elevate conductivity in streams in the Mountain and Plateau Regions of Central Appalachia are causing local extirpation of species.

The causal characteristics used in this assessment are described in Cormier et al. (2010) (see Table A-1). Each causal characteristic is defined and related to Hill's considerations and to the types of evidence in the *Stressor Identification (SI) Guidance* (U.S. EPA, 2000) and the Causal Analysis/Diagnosis Decision Information System (CADDIS) Web site (http://www.epa.gov/caddis). The SI and CADDIS types of evidence indicate the types of information that are potentially available to demonstrate characteristics of causation from Cormier et al. (2010). Hill's considerations are a mixture of types of evidence, sources of information, and quality of information, but they are included because they are traditional.

A.1.1. Assessment Endpoints

This causal assessment evaluates whether the aqueous salinity, as measured by conductivity, is capable of causing local extirpation of stream biota in an area of the Central

Appalachia including Ecoregions 69 (Central Appalachia) and 70 (Western Alleghany Plateau) (Woods et al., 1996). These regions include parts of the states of Ohio, Pennsylvania, Maryland, West Virginia, Kentucky, and Tennessee. The entities of concern are benthic invertebrates, possibly including rare and threatened species. The effect is local extirpation from streams in their natural range. Depending on the type of evidence, different biological measurement endpoints are used. In particular, the number of ephemeropteran genera is used in many of the quantitative analyses, because most of the sensitive genera are Ephemeroptera and the number of genera is a good summary of the consequences of extirpation. However, the assessment is of general causation in the regions of concern, not for any specific genus or location.

A.1.2. Data Sets

The same data sets used in the derivation of the aquatic life benchmark were used in the causal assessment, particularly the West Virginia Department of Environmental Protection's (WVDEP's) WABase. In addition, evidence was drawn from the literature involving laboratory studies, a data set from U.S. Environmental Protection Agency (U.S. EPA) Region 3 described in Pond et al. (2008a), and geographic information, and related information described in Appendix F.

A.1.3. Weighting

The evidence is weighted using a system of plus (+) for supporting conductivity as a cause, minus (−) for weakening and zero (0) for no effect. (Both neutral evidence and ambiguous evidence have no effect on the inference.) 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

Note that these scores are for particular pieces of evidence, not for causation as a whole. For example, a particular study may convincingly demonstrate that a source exists that is associated with elevated conductivity in the region, so it is scored + + +, but alone it is not convincing evidence that conductivity causes extirpation of biota.

Any relevant evidence receives a single plus, minus, or zero to register the evidence and to indicate a decreased or increased support for a causal relationship (see Table A-2). The strength of evidence is considered next. The strength of a relationship is indicated by the magnitude of a measure of association (e.g., a correlation coefficient) or the number of relationships that display the causal characteristic. After strength, the other possible unit of weight is assigned depending on causal characteristic and on the type of evidence. Additional considerations that may result in a higher score are presented in Table A-3.

A.2. EVIDENCE OF CHARACTERISTICS OF CAUSATION

A.2.1. Co-occurrence

Because causation requires that causal agents interact with unaffected entities; they must co-occur in space and time. Co-occurrence corresponds to Hill's *consistency*, SI's *co-occurrence*, and CADDIS's *co-occurrence in space and time*.

A.2.1.1. *Correlation of Cause and Effect*

In the Watershed Assessment Branch Data Base (WABbase), conductivity and the number of ephemeropteran genera were moderately correlated ($r = -0.63$) (see Figure A-1). This relationship holds even when elevated levels of potential alternative causes (confounders) are removed (see Figure A-2). In the data set created by Pond et al. (2008a), ephemeropteran genera and conductivity were highly correlated ($r = -0.90$).

A.2.1.2. *Contingency Table*

We constructed a contingency table of the presence of Ephemeroptera at sites near background conductivity (<200 µS/cm) and higher conductivities (>1,500 µS/cm) and recorded the ratio of presence or absence of mayflies (see Table A-4). It shows that mayflies co-occur with low conductivity but that all mayfly species are absent from more than ¾ of sites where conductivity is high. This analysis supplements the correlations by emphasizing the difference between high and low conductivity sites with respect to a clear endpoint, the absence of all Ephemeroptera.

We also compared the number of genera at sites with lower conductivities (<200 µS/cm) and higher conductivities (>1,500 µS/cm) with and without the co-occurrence of other parameters that are somewhat correlated with conductivity or are known biological stressors (see Tables B-1, B-2, B-3, and B-4). Whatever the level of the other parameter, when conductivity was low, Ephemeroptera were well represented and occurred less often or not at all at high conductivity. Hence, the potentially confounding agents were not responsible for the observed co-occurrence of conductivity and biological impairments. Other analyses of potential confounders are described in Appendix B on confounding.

A.2.1.3. *Co-occurrence in Paired Watersheds Over Time*

Conductivity is shown to increase after the construction of valley fill coal mining operations, and the number of ephemeropteran genera is low relative to a paired unmined watershed (see Table A-5).

> A.2.1. Summary. In sum, when conductivity is low, the number of genera is high. Even when other stressors are absent, where conductivity is high, the number of genera is low (see Figure A-2). The evidence for co-occurrence of conductivity with biological effects is strong, relevant, consistent, and of high quality and is, therefore, conclusive (see Table A-6).

A.2.2. Preceding Causation

Each causal relationship is a result of a web of preceding cause and effect relationships that begin with sources and include pathways of transport, transformation, and exposure. Evidence of sources of a causal agent increases confidence that the causal event actually occurred and was not a result of a measurement error, chance, or hoax (Bunge, 1979). Although preceding causation was not recognized by Hill, it corresponds to a type of evidence in the U.S. EPA's SI and CADDIS process, *causal pathway*.

A.2.2.1. *Complete Source to Cause Pathway*

Because exposure to aqueous salts does not require transport or transformation (i.e., organisms are directly exposed to salts in water immediately below sources), only evidence of the occurrence of sources is relevant. Potential sources of increased conductivity in the region include surface and underground coal mining, effluent from coal preparation plants and associated slurry impoundments, effluent from coal fly ash impoundments, winter road maintenance, brines from natural gas and coalbed methane drilling operations, treatment of waste water, human and animal waste, scrubbers at coal fired electric plants, and demineralization of crushed rock (Ziegler et al., 2007). The ionic composition of these waters is not uniform (see Table A-7). In particular, bicarbonate and sulfate are the dominant anions in streams at mined and unmined sites, but Marcellus Shale brine is almost entirely chloride salts. However, only four sites were found to have elevated conductivity with high chloride and low sulfate, so shale brines are rarely the sole dominant source of conductivity (see Section 2.3).

A.2.2.2. *Evidence from Literature*

High conductivity leachate has been shown to flow from valley fills created during coal mining operations (Bryant et al., 2002; Merricks et al., 2007). In contrast, conductivity increases only slightly following clear-cutting and burning. Dissolved mineral loading may be increased slightly by harvesting but also declines quickly as vegetation re-establishes (Swank and Douglass, 1977). Golladay (1988) and Arthur et al. (1998) found increases in nitrogen and phosphorus export in logged catchments in the Appalachians but minor differences in calcium, potassium, or sulfate concentrations between logged and undisturbed watersheds. Likens et al. (1970) actually found sulfate concentrations to decrease following clear cutting and experimental suppression of forest growth by herbicides.

A.2.2.3. *Co-occurrence of Sources and Conductivity*

Conductivity increases where surface mining operations occur in a watershed and not in an adjacent unmined watershed (see Table A-5) and are higher overall in mined watersheds with valley fill than in unmined watersheds (see Table A-7).

A.2.2.4. *Characteristic Composition*

Correlation and regression analyses suggest that, in Ecoregions 69 and 70, conductivities above 500 µS/cm contain high levels of the ions of Ca^{2+}, HCO_3^-, Mg^{2+}, and SO_4^{2-} (see Figure 10a−b) which is consistent with large scale surface coal mining and valley fill sources (Pond et al., 2008a). In contrast, the dominant ions of municipal waste water and of Marcellus Shale brine are Na^+ and Cl^-, which rarely dominate conductivity in those regions (see Section 2.3 and Table A-7). Therefore, the causal assessment relates primarily to mixtures of salts typical of alkaline coal mine drainage and associated valley fill discharges.

A.2.2.5. *Correlation of Conductivity with Sources*

Scatter plots of conductivity levels were generated for seven land cover classifications: open water, agriculture, residential, barren, valley fill, abandoned mine lands, and forested (see Appendix F for methods). From 2,151 sites in Ecoregion 69D described in the WVDEP WABbase, 191 <20 km^2 watersheds were found for which there were macroinvertebrate samples identified to the genus level with at least one chemistry sample and TMDL land cover information. Small <20 km^2 subwatersheds were selected to reduce confounding from multiple sources. These subwatersheds drained into the Coal, Upper Kanawha, Gauley, and New Rivers (see Figure F-2. Land use and land cover were generated from publically available databases (see Appendix F). Land use and land cover were arc sine square root transformed to better depict the upper and lower portions of the distribution. Scatter plots and Spearman rank correlations of six land use categories and conductivity are shown in Figure A-3.

Although conductivity typically increases with increasing land use (Herlihy et al., 1998), the densities of agricultural and urban land cover were relatively low, and a clear pattern of increasing conductivity and increasing land use is not seen. At relatively low urban land use, the range of conductivity is highly variable (see Figure A-3). This may be caused by unknown mine drainage, deep mine break-outs, road applications, poor infrastructure condition (e.g., leaking sewers or combined sewers), or other practices. In contrast, there is a clear pattern of increasing conductivity as percent area in valley fill increases, and decreasing conductivity with increasing forest cover (see Figure A-3).

> A.2.2. Summary. In sum, large scale surface mining and associated valley fills constitute a common and long-term source of high conductivity water in this region. The evidence for this source is abundant and of high quality. Hence, the evidence of preceding causation leading to high conductivity is conclusive (see Table A-8).

A.2.3. Interaction and Physiological Mechanisms

Causal agents alter affected entities by interacting with them through a physical mechanism. Evidence that a mechanism of interaction exists for a proposed causal relationship strengthens the argument for that relationship. This characteristic corresponds to Hill's *plausibility*, SI's *mechanism*, and CADDIS's *mechanistically plausible cause*.

A.2.3.1. *Mechanism of Exposure*

Aqueous salts are dissolved ions that are readily available for uptake by aquatic organisms as they pass over their respiratory surfaces.

A.2.3.2. *Mechanism of Effect*

The internal fluids of freshwater organisms are saltier than the water in which they live. As a result, freshwater organisms must use many physical structures and physiological mechanisms to maintain a balance of water content and ionic content. To maintain the balance of ions, they excrete hypotonic urine; possess impermeable scales, cuticles or exoskeletons; and use semipermeable membranes to redistribute ions (Bradley, 2009, Evans, 2008a, b; Wood and Shuttleworth, 2008; Thorp and Covich, 2001). Other methods of absorption include rectal pumping of water in Odonates and drinking by Megalopterans and Coleopterans. Anion, cation, and proton transport include passive, active, uniport, and co-transport (Nelson and Cox, 2005). Many freshwater invertebrates have chloride cells that actively take up chloride and other ions through gills (Komnick, 1977; Bradley, 2009). Members of the orders Ephemeroptera, Plecoptera, and Heteroptera have chloride cells on their body surfaces. Some dipterans and a few Trichopterans have chloride epithelia, and anal papillae are present on other members of these orders.

Numerous specific mechanisms are involved in the toxicity of high-conductivity solutions; one is discussed here for the sake of illustration. The ion regulation system includes antiport anion exchange proteins that co-transport Cl^- against the concentration gradient into the cell simultaneously with HCO_3^- down the concentration gradient and out of the cell (Nelson and Cox, 2005; Bradley 2009). If external HCO_3^- is high, the gradient is not favorable for Cl^- uptake (Avenet and Lingnon, 1985).

Some physiological processes are especially dependent on proper ionic balance. These include nerve conduction, muscle contraction, and secretion. Reproduction, including fertilization, polymerization of egg mass coverings, and embryonic development depend on ionic balance, graphically illustrated by the swelling of fish eggs upon fertilization (Tarin et al., 2000). At the organismal level, effects of aqueous salts on aquatic arthropods include mortality (Kefford et al., 2003, 2005a) and reduced growth, reproduction, and hatching success (Clark et al., 2004a; Hassell et al., 2006; Kefford and Nugegoda, 2005; Kefford et al., 2004, 2006, 2007; Nielsen et al., 2003; Brock et al., 2005). These effects strongly suggest that population density can be reduced over generations of persistent exposure to elevated conductivity (Zalizniak et al., 2007).

A.2.3. Summary. In sum, aquatic organisms are directly exposed to aqueous salts, and the relative amounts and concentration of salts may exceed the capacity of organisms to regulate their internal ionic composition. The importance of osmoregulation and ionic homeostasis has been demonstrated in diverse animal models with results published in the peer-reviewed literature. The evidence is drawn from a long history of physiological investigations (see Table A-9).

A.2.4. Specific Alteration

A specific cause induces a specific effect in particular receptors. This alteration is obscured in many studies by broad definitions of causes and effects, but, when a specific effect of a cause is characterized, it strengthens the evidence for a causal relationship. If the specific effect of a cause occurs with no other causes, it can be diagnostic of that cause. This characteristic corresponds to *specificity* in Hill's considerations and in the SI's types of evidence, and to *symptoms* in CADDIS.

A.2.4.1. *Specificity of Genera*

In a paper focusing on mayflies, principal component analysis sorted mined and residential sites from reference sites primarily on the basis of specific conductance and pH (Pond, 2009). In the same study, a nonmetric multidimensional scaling model strongly associated Ephemerella, Drunella, Cinygmula, Epeorus, and Ameletus with the low conductivity reference sites and no mayflies or Caenis, Baetis, Isonychia, and Stenonema with the higher conductivity sites. The first group has XC_{95} <600 µS/cm, and the second group of tolerant mayflies had XC_{95} >729 µS/cm.

The derivation of 95th percentile extirpation concentration (XC_{95}) values and species sensitivity distributions in this document demonstrated that a characteristic set of genera (primarily Ephemeroptera) were extirpated at relatively low conductivities and others were resistant. The relative sensitivities are consistent with the findings of Pond et al. (2008a) and with analyses of data from Kentucky (see Appendix E). This is not meant to suggest that conductivity is the only possible cause of loss of these genera. Rather, it indicates that the loss of those genera consistently occurs where conductivity is elevated. If a random set of genera were lost, it might suggest that various causes were acting that co-occur with elevated conductivity, but that was not the case.

A.2.4.2. *Specificity of Assemblages*

Using an independent data set collected in West Virginia, nonmetric multidimensional scaling was applied to biological metrics, and sites were sorted into distinct ordination space characterized by low, medium, and highly elevated conductivities associated with surface mines with valley fill (Pond et al., 2008a).

> A.2.4. Summary. In sum, some genera are sensitive to conductivity, and others are not. The evidence for effects specific to high conductivity is reasonably strong, relevant, consistent, and of high quality and is, therefore, supportive (see Table A-10).

A.2.5. Sufficiency

For an effect to occur, sufficiently susceptible entities must experience a sufficient magnitude of exposure. This characteristic corresponds to *biological gradient* in Hill's considerations. In SI and CADDIS, multiple types of evidence may demonstrate sufficiency including *stressor-response in the field, laboratory tests of site media, manipulation of exposure* and *stressor-response from laboratory studies*.

A.2.5.1. *Laboratory Tests of Defined Ion Mixtures*

Mount et al. (1997) tested the acute lethality of several mixtures of salts to two planktonic crustaceans (*Ceriodaphnia dubia* and *Daphnia magna*) and a fish (*Pimephales promelas*). A mixture of K_2SO_4 and $KHCO_3$ salts was the most toxic combination of salts tested in the study. The 48-hour LC_{50} for Ceriodaphnia with K_2SO_4 and $KHCO_3$ corresponds to 438 µS/cm. The 96-hour LC_{50} for *Pimephales promelas* also with K_2SO_4 and $KHCO_3$ corresponds to

1,082 μS/cm. The ion matrix of alkaline mine drainage normally contains little K^+, and instead, Ca^{2+} and Mg^{2+} are the dominant cations. Conductivity measurements below some valley fills were greater than 4,000 μS/cm. This study demonstrates that mixtures of specific salts can be acutely lethal at concentrations corresponding to conductivities measured in the region. The Mount et al. (1997) model has been used to estimate that salt mixtures in some streams below valley fills are sufficient to cause acute lethality in Ceriodaphnia (U.S. EPA, 2009). However, these tests are marginally relevant. The crustaceans are not taxonomically similar to the invertebrate species that are affected, the 48-hour test durations are far shorter than the life-cycle exposures in the field, and the effect (acute lethality) is unlikely to be the cause of population-level effects in the field. Life-cycle effects on local insects are likely to occur at much lower levels of conductivity (U.S. EPA, 2009). However, these tests do indicate that the ion mixture could be toxic to common surrogate laboratory organisms used to evaluate toxicity.

A.2.5.2. *Laboratory Tests of Mine Discharges*

Kennedy et al. (2003, 2004, 2005) tested coal mine discharge waters in Ohio with *Ceriodaphnia dubia* and a mayfly (*Isonychia bicolor*). In 7-day lethality tests, the mayfly was about three times as sensitive as the crustacean. Lowest observed effect concentrations (LOECs) for survival of mayflies (mid to late-instars) at 20°C occurred at 1,562, 966, and 987 μS/cm in three tests. These values bracket the Isonychia XC_{95} of 1,177 μS/cm. Ceriodaphnia tests with simulated effluent containing only major ions showed that the toxicity of this effluent was not due to heavy metals or Se (Kennedy et al., 2005).

Echols et al. (2009) performed 10−14 day toxicity tests of coal processing effluent from Virginia with *Isonychia bicolor*. They obtained LOEC values for survivorship in three tests of 1,508 to 4,101 μS/cm. The lower toxicity of these waters may be due to the dominance of sodium, which has the lowest toxicity of the common cations (Mount et al., 1997). In any case, it is not surprising that these acute lethality tests yield higher conductivity levels than the Isonychia XC_{95} (1,177 μS/cm) which is a result of full life-cycle exposures and effects. In particular, in the final test which yielded the lowest LOEC, survival at the end of the test was approximately 25% and still declining. The effluent contained no detectable toxic trace metals or metalloids except selenium (8.5 μg/L), so the authors stated that the toxicity was likely due to salinity.

A.2.5.3. *Laboratory Tests of Ambient Waters*

Waters from below valley fills in the region of concern were tested by Merricks et al. (2007). *Ceriodaphnia dubia* LC_{50} values in 48-hour tests were established for some but not all waters from Lavender Fork with undiluted concentrations of 2,497−3,050 μS/cm. These tests used relevant mixtures of ions, but the test species, duration and endpoint have low relevance and are likely to underestimate toxicity in the field.

A.2.5.4. *Field Exposure-Response Relationships*

As Hill suggested, a biological gradient in the field suggests that the exposures reach levels that are sufficient to cause effects. Pond et al. (2008a, b) reported that the number of taxa decreases as conductivity increases or as the amount of surface mining and associated valley fills increases. Analyses conducted for this report using the WABbase data sets show that as conductivity increases the total number of genera and the number of ephemeropteran genera decrease at conductivity levels shown to extirpate sensitive genera (see Figure A-1). This analysis shows not only the co-occurrence of elevated conductivity and loss of stream biota but also that there is a regular exposure-response relationship that extends to the lowest observed concentrations (evidence of sufficiency). The same data set was also modeled by partitioning for potential confounding parameters. Streams with higher temperatures (>22°C), low pH (<6), poor habitat (<135) and high fecal coliform (>400 colonies/100 mL) were excluded. The effect of conductivity was still strong (see Figure A-2). Also, the distributions of individual genera show that as conductivity increases the occurrence and capture probability decreases for many genera (see Appendix D).

A.2.5.5. *General Knowledge*

The susceptibility of organisms and communities is a function of genetic, evolutionary, developmental, and physiological legacies. Numerous studies have characterized the disappearance of freshwater taxa with increasing salt concentration (Remane, 1971; Wetzel, 2001). Species native to the Mid-Atlantic Highlands (Ecoregions 69 and 70) have evolved for very low conductivity water and are expected to decline as salinity increases above background. Conductivities below 70 µS/cm were common in forested areas.

A.2.5. Summary. In sum, the available evidence suggests that conductivities in the region of concern reach levels that are sufficient to cause effects on stream communities (see Table A-11). Most laboratory toxicity tests have been conducted with species, effects and exposures that have low relevance and sensitivity to salinity. However, they still show that ambient salinities observed in the regions of concern can cause severe effects. More to the point, saline coal mine effluents from the region are lethal to a mayfly species at conductivities similar to its extirpation concentration. Correlations of conductivity and stream biological metrics confirm that conductivity is strongly associated with gradients of biological response down to the levels where sensitive genera are extirpated. These relationships are strong even when other stressors were present. Finally, general studies of the effects on aquatic organisms of changes in salinity suggest that the observed magnitude of increases in salinity in these regions is sufficient to cause the extirpation of some species, but the studies were not conducted at levels as low as those occurring in this region.

A.2.6. Time Order

Logically, a causal event occurs before an effect is observed. Evidence of time order is provided by changes in the invertebrate assemblages after the introduction of a source that increased conductivity. This characteristic corresponds to *temporality* in Hill's considerations and in the SI types of evidence and to *temporal sequence* in CADDIS.

We could not obtain conductivity and biological survey data for before and after a valley fill or other source of saline effluents began operation. Hence, this characteristic of causation is scored No Evidence (NE).

A.2.7. Evaluation of the Body of Evidence

Conclusions concerning causality are based on the weight of evidence from all types of evidence that support or weaken all of the characteristics of causation. The property of the body of evidence was termed *coherence* by Hill. In SI and CADDIS, it is divided into two considerations: the *consistency* of evidence and the *coherence* of evidence (i.e., the reasonableness of explanations of any inconsistencies in the evidence).

This causal assessment found that the available evidence supports a causal relationship between mixtures of matrix ions in streams of Ecoregions 69 and 70 and biological impairments. That conclusion is based on evidence showing that the relationship of conductivity to the loss of aquatic genera has the characteristics of causation.

1. Co-occurrence—The loss of genera occurs where conductivity is high even when potential confounding causes are low but is rare when conductivity is low (+ + +).

2. Preceding Causation—Sources of conductivity are present and are shown to increase stream conductivity in the region (+ + +).
3. Interaction—Aquatic organisms are directly exposed to dissolved salts. Physiological studies over the last 100 years have documented the many ways that physiological functions of all organisms are affected by excess salt or the combinations of ions for which they do not have physiological capacity or mechanisms to regulate (+).
4. Alteration—Some genera and assemblages are affected at sites with higher conductivity while others are not. These differences are characteristic of high conductivity (+ +).
5. Sufficiency—Increased exposure in both concentration and duration to salt affects invertebrates based on both field and laboratory analyses (+ + +).
6. Time order—Conductivity increases, and local extirpation occurs after mining permits are issued, but before and after data are not available (NE).

Other potential causes of the loss of genera in the region include elevated temperatures associated with loss of shade or increased impervious surfaces, siltation from various land use activities, low pH from atmospheric deposition and abandoned mines, aluminum toxicity from abandoned mines, and nutrient enrichment from various sources. Se toxicity has also been implicated. When these causes are minimized, a relationship between conductivity and mayfly richness is still evident (see Appendix B).

This causal assessment does not attempt to identify the constituents of the mixture that account for the effects. Constituents of the mixture in neutral and somewhat alkaline waters that increase as conductivity increases are all considered as contributing to the local extirpation of genera in the region of concern. The dominant ions include HCO_3^-, SO_4^{2-}, Ca^{2+}, K^+, and Mg^{2+}.

REFERENCES

Arthur, MA; Coltharp, GB; Brown, DL. (1998) Effects of best management practices on forest streamwater quality in eastern Kentucky. J Amer Water Resour Assoc 34(3):481−495.

Avenet, P; Lignon, JM. (1985) Ionic permeabilities of the gill lamina cuticle of the crayfish, Astacus leptodactylus (E). J Physiol 363:377−401.

Bradley, TJ. (2009) Animal Osmoregulation. New York:Oxford University Press; pp. 149−151.

Brock, MA; Nielsen, DL; Crossle, K. (2005) Changes in biotic communities developing from freshwater wetland sediments under experimental salinity and water regimes. Freshwater Biol 50:1376−1390.

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/MTMVFChemistryPart1.pdf.

Bunge, M. (1979) Causality and modern science. Third revised edition. New York: Dover Publications, Inc.

Clark, TM; Flis, BJ; Remold, SK. (2004a) Differences in the effects of salinity on larval growth and developmental programs of a freshwater and a euryhaline mosquito species (Insecta: Diptera, Culicidae). J Exp Biol 207:2289−2295.

Cormier, SM; Suter GW, II; Norton, SB. (2010) Causal characteristics for ecoepidemiology. Hum Ecol Risk Assess 16(1):53-73.

Echols, B; Currie, R; Cherry, DS. (2009) Preliminary results of laboratory toxicity tests with the mayfly, *Isonychia bicolor* (Ephemeroptera: Isonychiidae) for development as a standard test organism for evaluating streams in the Appalachian coalfields of Virginia and West Virginia. Environ Monitor Assess. published online 04 November 2009.

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.

Fox, GA. (1991) Practical causal inference for ecoepidemiologists. J Toxicol Environ Health 33(4):359−374.

Golladay, SW. (1988) The effects of forest disturbance on stream stability. Dissertation. Virginia Polytechnic Institute and State University, Blacksburg, Va.

Hassell, KL; Kefford, BJ; Nugegoda, D. (2006) Sub-lethal and chronic lethal salinity tolerance of three freshwater insects: *Cloeon* sp. and *Centroptilum* sp. (Ephemeroptera: Baetidae) and *Chironomus* sp. (Diptera: Chironomidae). J Exp Biol 209:4024−4032.

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, and Soil Pollution 105:377−386.

Hill, AB. (1965) The environment and disease: Association or causation. Proceed Royal Soc Med 58:295−300.

Kefford, BJ; Papas, PJ; Nugegoda, D. (2003) Relative salinity tolerance of macroinvertebrates from the Barwon River, Victoria, Australia. Mar Freshwater Res 54:755–765.

Kefford, BJ; Dalton, A; Palmer, CG; Nugegoda, D. (2004) The salinity tolerance of eggs and hatchlings of selected aquatic macroinvertebrates in south-east Australia and South Africa. Hydrobiologia 517:179−192.

Kefford, BJ; Nugegoda, D. (2005) No evidence for a critical salinity thresholds for growth and reproduction of the freshwater snail Physa acuta. Environ Pollut 54:755−765.

Kefford, BJ; Zalizniak, L; Nugegoda, D. (2006) Growth of the damselfly Ischnura heterosticta is better in saline water than freshwater. Environ Pollut 141:409–419.

Kefford, BJ; Nugegoda, D; Zalizniak, L; Fields, EF; Hassell, KL. (2007) The salinity tolerance of freshwater macroinvertebrate eggs and hatchlings in comparison to their older life-stages. Aquat Ecol 41:335−348.

Kennedy, AJ; Cherry, DS; Currie, RJ. (2003) Field and laboratory assessment of a coal processing effluent in the Leading Creek Watershed, Meigs County, Ohio. Arch Environ Contam Toxicol 44(3):324−331.

Kennedy, AJ; Cherry, DS; Currie, RJ. (2004) Evaluation of ecologically relevant bioassays for a lotic system impacted by a coal-mine effluent, using *Isonychia*. Environ Monit Assess 95:37−55.

Kennedy, AJ; Cherry, DS; Zipper, CE. (2005) Evaluation of ionic contribution to the toxicity of a coal-mine effluent using *Ceriodaphnia dubia*. Arch Environ Contam Toxicol 49:155−162.

Komnick, H. (1977) Chloride cells and chloride epithelia of aquatic insects. Int Rev Cytol 49:285−328.

Likens, GE; Bormann, FH; Johnson, NM; et al. (1970) Effects of forest cutting and herbicide treatment on nutrient budgets in the Hubbard Brook watershed-ecosystem. Ecol Monogr 40:23−47.

Merricks, TC; Cherry, DS; Zipper, CE; et al. (2007) Coal-mine hollow fill and settling pond influences on headwater streams in southern West Virginia, USA. Environ Monit Assess 129(1−3):359−378.

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.

Nielsen, DL; Brock, M; Crossle, K; Harris, K; Healey, M; Jarosinski, I. (2003) The effects of salinity on aquatic plant germination and zooplankton hatching from two wetlands sediments. Freshwater Biol 48:2,214−2,223.

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. (2008a) Downstream effects of mountaintop coal mining: comparing biological conditions using family- and genus-level macroinvertebrate bioassessment tools. J N Am Benthol Soc 27:717−737.

Pond, GJ; Bailey, JE; Lowman, B. (2008b). West Virginia GLIMPSS (genus-level index of most probable stream status): a benthic macroinvertebrate index of biotic integrity for West Virginia's wadeable streams. West Virginia Department of Environmental Protection, Division of Water and Waste Management, Watershed Assessment Branch. Charleston, WV.

Remane, A. (1971) Ecology of brackish water. In: Remane A; Schlieper, C, eds. Biology of brackish water. New York, NY: John Wiley and Sons.

Suter, GW, II; Traas, T; Posthuma, L. (2002) Issues and practices in the derivation and use of species sensitivity distributions. In: L. Posthuma, L; Suter, GW, II; Traas, T; eds. Species sensitivity distributions in ecotoxicology,. Boca Raton, FL: Lewis Publishers, pp 437−474.

Swank, WT; Douglass, JE. (1977) Nutrient budgets for undisturbed and manipulated hardwood ecosystems in the mountains of North Carolina. In: Correll, DL; ed. Watershed research in eastern North America: A workshop to compare results. Edgewater, MD: Smithsonian Institution Press; pp 343−363.

Tarin, JJ; Cano, C. Tarbin, JJ. (2000) Fertilization in protozoa and metazoan animals: Cellular and molecular aspects. Berlin: Springer-Verlag.

Thorp, JH; Covich, AP; editors. (2001) Ecology and classification of North American freshwater invertebrates. Academic Press, New York.

U.S. DHEW (Department of Health Education and Welfare). (1964) Smoking and Health: Report of the Advisory Committee to the Surgeon General., Washington, D.C. Public Health Service Publication 1103.

U.S. EPA. (Environmental Protection Agency). (2000) Stressor identification guidance document. Washington, DC. EPA/822/B-00/025.

U.S. EPA (Environmental Protection Agency). (2009) 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.

Wetzel, RG. (2001) Limnology. 3rd edition. San Diego, California: Academic Press.

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/600R-96/077.

Zalizniak, LB; Kefford, BJ; Nugegoda, N. (2007) Effects of pH on salinity tolerance of selected freshwater invertebrates. Aquat Ecol 43:135−144.

Ziegler, CR; Suter, GW, II; Kefford, BJ. (2007) Candidate Cause: Ionic Strength. Available online at www.epalgov/caddis (accessed 10/13/09).

Table A-1. Definitions of causal characteristics

Characteristic	Description
Co-occurrence	The cause co-occurs with the unaffected entity in space and time
Preceding causation	Each causal relationship is a result of a larger web of cause and effect relationships
Time order	The cause precedes the effect
Interaction	The cause physically interacts with the entity in a way that induces the effect
Alteration	The entity is changed by the interaction with the cause
Sufficiency	The intensity, frequency, and duration of the cause are adequate, and the entity is susceptible to produce the type and magnitude of the effect

Table A-2. 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	+, 0, −
Strength	Increase score
Other qualities	Increase score

Table A-3. Other considerations used to weight the evidence concerning the influence of potentially confounding variables

Quality of evidence	Alternative outcomes
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
Independence of association	Independent or confounded

Table A-4. Presence of mayflies contingent on stream conductivity (data from WABbase)

	Mayflies present	Mayflies absent	Total
Near background conductivity (<200 μS/cm)	889 96.2%	35 3.8%	924
High conductivity (≥1,500 μS/cm)	28 21.7%	101 78.3%	129
Total	917	136	1,053

Table A-5. Temporal increase of conductivity 2 years after permitting of mining operations

	Never mined Ash Fork			Permit 1994, 1996 Boardtree Branch			Permit 1996; Stillhouse		
	1998	2003	2006/07	1998	2003	2007	1998	2003	2007
μS/cm	44[a]	39[b]	42[b]/39[a]	1,396[a]	3,015[b]	3,390[a]	511[a]	3,200[b]	3,970[a]
% E		27.23	29.21			1.23			0
# E		6	4			2			0
# P		5	6			0			0
# EPT		20	14			5			3
TT		41	24			20			8

[a]Single measurement.
[b]Mean value.

E = Ephemeroptera; P = Plecoptera; T = Trichoptera; TT = total taxa.

Table A-6. Summary of evidences and scores for co-occurrence

Type of evidence	**Evidence**	**Score**
Correlation of cause and effect	Ephemeroptera were correlated with conductivity in two studies $r = -0.63$ (see Figure A-1) and $r = -0.90$. This is strong quantitative evidence from multiple studies.	+ + +
Contingency table	The contingency table (see Table A-4) provides strong quantitative evidence that high conductivity is strongly associated with severe effects (Ephemeroptera absent at >75% of sites).	+ +
Co-occurrence in paired watersheds over time	24% to 100% difference (see Table A-5) is large and quantitative.	+ +
Overall score	Relevant, strong, consistent.	+ + +

Table A-7. Total cations and anions measured in water originating from surface mined sites with valley fills, unmined sites, or Marcellus Shale brine. Individual ions are presented as a fraction of the total cations or anions. For mined sites, *n* = 13; unmined sites, *n* = 7; Marcellus Shale brine, *n* = 3. Measurements of HCO_3 and NO_3^-N were not available for Marcellus Shale brine sites.

	Mined (Valley Fill)			Unmined			Marcellus Shale Brine		
	Mean	**Median**	**Range**	**Mean**	**Median**	**Range**	**Mean**	**Median**	**Range**
Total Cations (mg/L)	282.4	238.9	72.7–515.2	15.7	15.9	7.0–25.6	23,862.0	21,719.0	8,650.0–41,217.0
Ca	0.48	0.48	0.42–0.55	0.46	0.46	0.37–0.63	0.24	0.23	0.20–0.28
Mg	0.42	0.42	0.28–0.51	0.28	0.27	0.22–0.36	0.02	0.02	0.02–0.02
K	0.04	0.04	0.02–0.05	0.11	0.11	0.06–0.18	0.02	0.01	0.005–0.05
Na	0.06	0.03	0.02–0.25	0.15	0.14	0.06–0.24	0.72	0.70	0.69–0.78
Total Anions (mg/L)	926.8	730.4	228.1–1,734.4	44.7	47.2	21.9–66.5	28,296.1[a]	18,620.8[a]	14,326.3–51,941.3[a]
[b]HCO_3	0.25	0.25	0.06–0.48	0.54	0.57	0.34–0.66	NA	NA	NA
Cl	0.0076	0.0042	0.0032–0.0036	0.07	0.06	0.04–0.11	0.999	0.999	0.998–0.999
NO_3–N	0.0036	0.0031	0.0013–0.011	0.01	0.01	0.002–0.04	NA	NA	NA
SO_4	0.73	0.74	0.51–0.93	0.38	0.35	0.29–0.51	0.0013	0.0011	0.0011–0.0016

[a]Total anions includes only Cl^- and SO_4^{2-}.
[b]HCO_3^- converted from measurement of alkalinity as $CaCO_3$.

NA = not applicable due to lack of data.

Table A-8. Summary of evidences and scores for preceding causation

Type of evidence	Evidence	Score
Complete source-to-cause pathway	Sources are present, and no intermediate steps in the pathway are required.	+
Correlation of conductivity with sources	Figure A-3, $r = 0.61$. This is moderately strong quantitative evidence from the case.	+ +
Evidence from literature	Multiple publications link conductivity to sources in the region and eliminate some other land uses as sources.	+
Co-occurrence of sources and conductivity	When valley fills are present, conductivity is 12- to 90-fold greater than at unmined sites (see Tables A-5 and A-7). This is strong quantitative evidence from the case.	+ +
Characteristic composition	Ambient mixtures of ions have characteristic compositions that can be associated with particular sources. Most sites with elevated conductivities have compositions characteristic of coal mining with valley fill. This is relevant but quantitatively weak evidence.	+
Overall score	Relevant, strong, consistent.	+ + +

Table A-9. Summary of evidences and scores for interaction and physiological mechanism

Type of evidence	Evidence	Score
Mechanism of exposure	Salts readily dissolve in water and interact directly with aquatic organisms.	+
Mechanism of effect	Many mechanistic studies show that osmoregulation and homeostasis of specific ions are sensitive to disruption, particularly in mayflies.	+
Direct evidence	No studies of ionic compensation are available for organisms in the region.	NE
Overall score	Relevant but not case-specific.	+

Table A-10. Summary of evidence and scores for specific alteration

Type of evidence	Evidence	Score
Specificity of genera	Specific genera are consistently sensitive to conductivity. This quantitative evidence is independently confirmed.	+ +
Specificity of assemblage	A model based on specific biology discriminated effects of conductivity associated with mining.	+
Overall score for interaction	Relevant, independently confirmed, and consistent, but only two types of evidence.	+ +

Table A-11. Summary of evidence and scores for sufficiency

Type of evidence	Evidence	Score
Laboratory tests of defined ion mixtures	The tests were high quality, but the species and durations have low relevance for determining the conductivity level at which effects occur, and the effect levels are supportive only if assumptions are made about acute/chronic and intertaxa extrapolations.	0
Laboratory tests of mine discharges	This evidence is relevant in that it comes from nonacid mine effluents in the region and includes an Ephemeropteran; but the ionic mixtures were somewhat different, the effect was lethality and the durations were short. The results for one set of tests matched the XC_{95} for the test genus, but were higher for the other.	+
Laboratory tests of ambient waters	These tests showed acute lethality to an apparently resistant species at high conductivity levels. Its relevance is too low to support or weaken.	0
Field exposure-response relationships for Ephemeroptera	This is strong evidence because it is highly relevant, was obtained independently in two separate data sets, with moderate to strong correlations. It is not convincing in itself because of the potential for confounding, which is treated in Appendix B.	+ +
Field exposure-response relationships for genera	As conductivity increases, genera no longer are observed.	+ +
General knowledge	General knowledge indicates that salinity can cause the loss of species but does not indicate that the salinity levels observed in this case are sufficient.	0
Overall score	The exposure-response relationships in the field, with some support from laboratory studies, provide positive evidence that the conductivity levels observed are sufficient to cause the associated effects.	+ + +

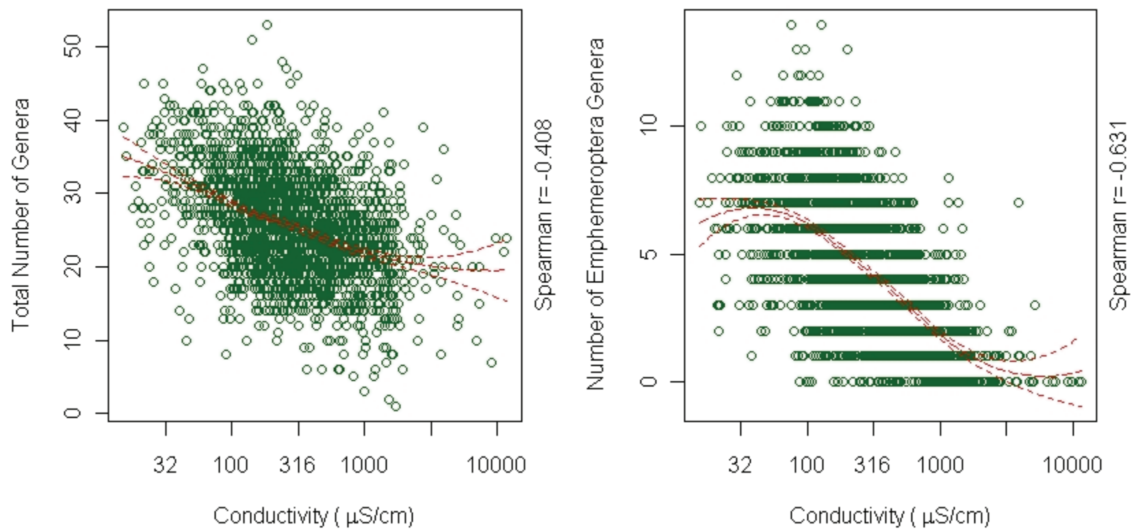

Figure A-1. As conductivity increases, the number of total and ephemeropteran genera decrease.

Data source: WABase

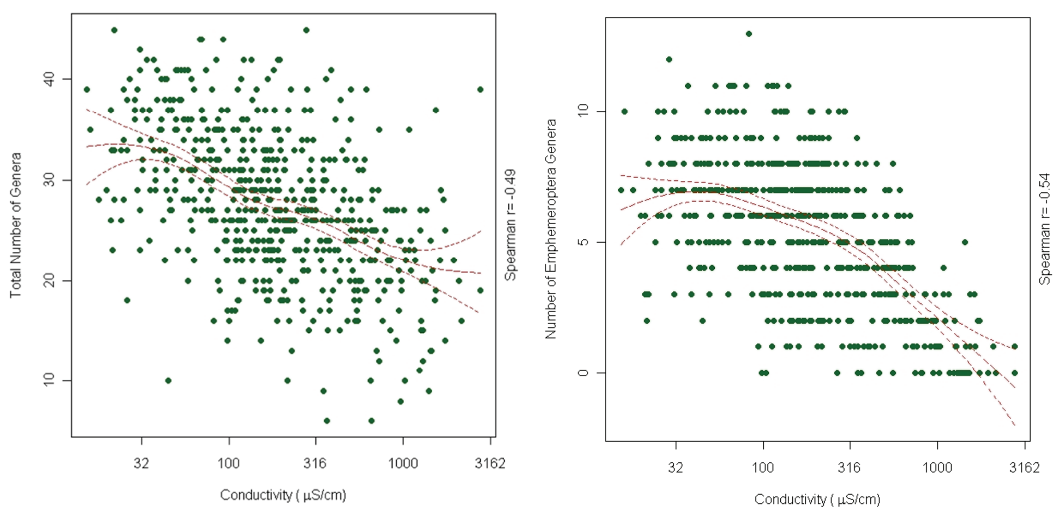

Figure A-2. As conductivity increases, the number of total and ephemeropteran genera decrease even when potentially confounding parameters are minimized. (Excluded: streams with higher temperatures [>22°C], low pH [<6], poor habitat [<135] and high fecal coliform [>400 colonies/100 mL]).

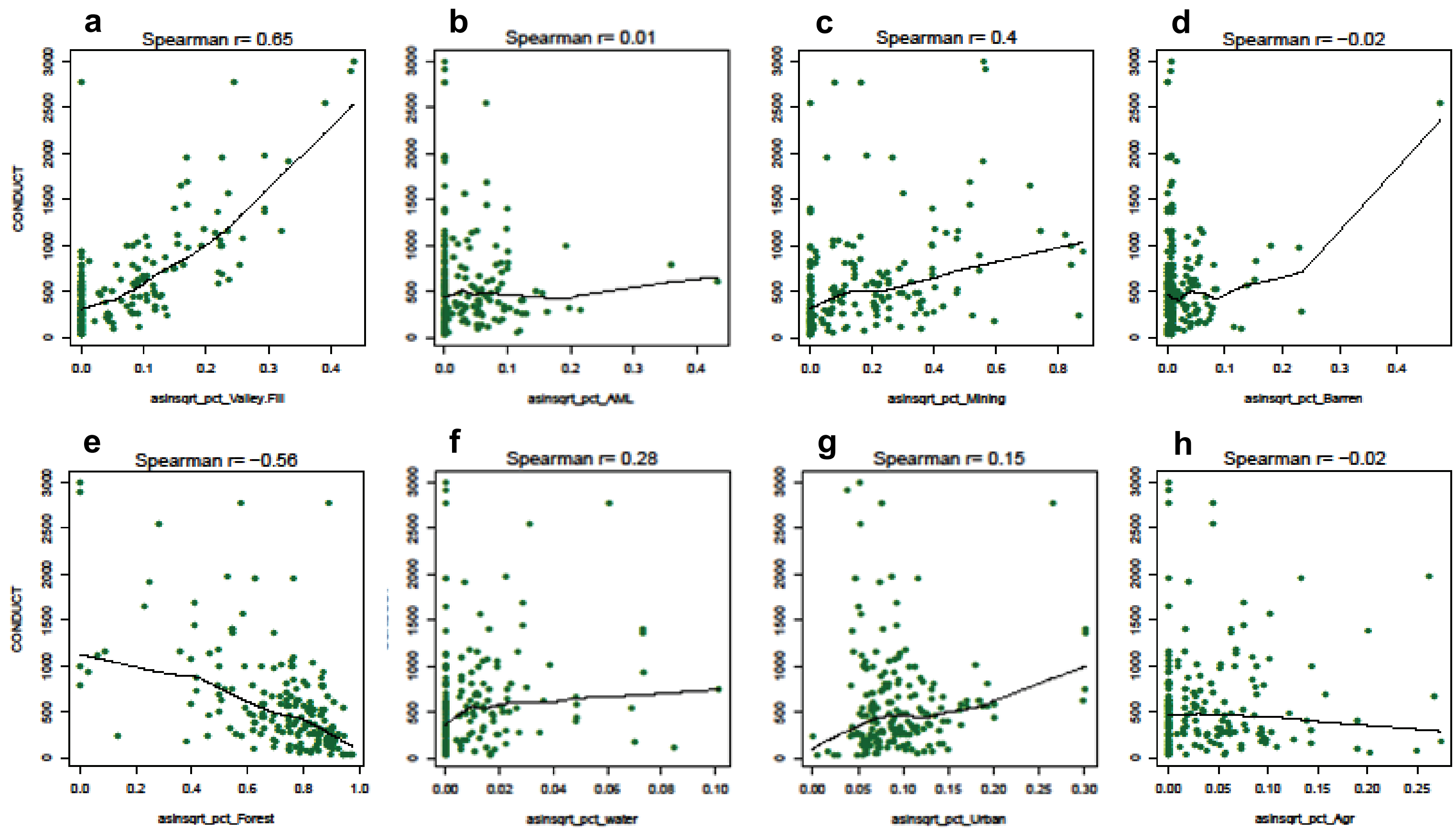

Figure A-3. Conductivity associated with different land uses in 151 watersheds in Ecoregion 69D. There is a clear pattern of increasing conductivity as percent area in valley fill increased, but no pattern with other land use. From left to right, they are (a) mountaintop mining-valley fill, (b) abandoned mine lands, (c) mined, (d) barren, (e) forested, (f) water, (g) urban/residential, and (h) agricultural. Land use and land cover were arc sine square root transformed to better depict the upper and lower portions of the distribution.

APPENDIX B
CONFOUNDING

ABSTRACT

The purpose of this appendix is to determine which, if any, of the variables that may co-occur with conductivity alter our ability to model the relationship between conductivity and occurrence of genera. The point was not to determine whether the confounders are general causes (i.e., can they cause impairments in the region of concern?).

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 significant influence on the causal relationship between salts and macroinvertebrate assemblages. The inference was performed by identifying potential confounders and then determining the occurrence and strength of ten types of evidence for confounding for each of them. 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 influence of Se could not be evaluated 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. We do not argue that these variables have no influence, but their effects are minimal given the streams that would be affected by the aquatic life benchmark.

B.1. INTRODUCTION

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 disturbance. The goals are (1) to define a set of streams in which the effects of elevated conductivity can be identified without significant influence by confounding variables, and (2) to estimate conductivity levels that would protect against the unacceptable effects of salts in those streams (i.e., typical streams receiving high conductivity effluents in the region of concern).

Because of those goals and the nature of the data, it is not appropriate to use multivariate statistics to try to eliminate confounders. Multiple regression methods depend on assumptions of independence, additivity, and normality that are not met. Propensity scores depend on a counterfactual assumption (i.e., it assumes that the confounders can be different than they are). This condition is met in propensity score analyses of epidemiological or econometric studies in which, for example, a cancer patient could be a smoker or not or might live in a city or not. That counterfactual assumption is often not met by ecological data. In particular, the alkaline ions that contribute to elevated conductivity also contribute to raising the pH. Therefore, pH and conductivity are not mechanistically independent, and counterfactual assumptions cannot be applied. Attempts to statistically eliminate the influence of pH would artificially reduce the effects of salinity. However, the epidemiological weight-of-evidence approach used here can make use of the fact that, once we have eliminated acidic sites, the neutral to moderately alkaline pH levels that remain are not toxic to stream organisms.

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 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, 2000; Stewart-Oaten, 1996).

Confounding can also bias a causal model resulting in uncertainty concerning the actual magnitude of the effects. A variety of approaches may be used to determine whether confounders significantly affect the results. 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). We provide a relatively complete list, but we only used Evidence Types 1, 2, 3, 5, 6, and 8.

1. Co-occurrence of confounder and cause: Confounders are correlated with the <u>cause</u> of interest. A low correlation coefficient is evidence against the potential confounder.
2. Co-occurrence of confounder and effect: Potential confounders are correlated with the <u>effect</u> 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 not present, 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 takes the result of the causal assessment as a given and attempts to determine whether any of the known potential confounders interfere with estimating the effects of conductivity to a significant degree. 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 genera is used as a standard metric for the effects of conductivity, which may or may not be confounded. Because the sensitive genera are primarily Ephemeroptera and the endpoint effect is extirpation of 5% of genera, this is an appropriate metric.

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.

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

Any relevant evidence receives a single plus, minus, or zero to register the evidence 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 authors' judgments. After strength, the other possible unit of weight is assigned depending on the type of evidence.

For co-occurrence (Evidence Types 1−4), strength or consistency of the association is the primary consideration. For comparison for any of the potential confounders, the correlation coefficient for conductivity and number of ephemeropteran genera is 0.63, a value in the upper end of the moderate range. Correlations, as measures of co-occurrence, can be scored as in Table B-2.

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

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 confounder is present at high levels, this is supporting evidence for confounding. Note, 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. Therefore, clearly high and low levels of conductivity and the potential confounder are used in 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.

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 (regarding pH, RBP, and fecal coliform) from the correlation of conductivity and Ephemeroptera ($r = 0.63$) in the full data set (see Table B-4).

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.

B.3. WEIGHING

After the individual pieces of evidence had been weighted, the body of evidence for a potential confounder was weighed based on the credibility, diversity, strength, and coherence of the body of evidence (see Table B-6). The body of evidence, rather than a single piece of evidence, was considered to determine how strongly these potential confounders might affect the model.

B.4. POTENTIAL CONFOUNDERS

B.4.1. Habitat Quality

Stream habitat may be modified in reaches that receive high conductivity effluents. Habitat quality was represented by a qualitative index, the RBP derived by the WVDEP, which increases as habitat quality increases.

Although habitat scores were correlated with both conductivity and biological response, which indicates a potential for confounding, low RBP was judged to have little effect on the derivation of the 5th percentile hazardous concentration (HC_{05}) for conductivity (see Table B-7 and B-8).

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 data show no indication of significant confounding associated with fecal coliform counts and effects attributed primarily to organic enrichment (see Tables B-9 and B-10).

B.4.3. Nutrients

Nitrogen and phosphorus may also 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-11).

B.4.4. Deposited Sediment

Sources of salts can be associated with erosion and silt that affect stream organisms. A qualitative measure of embeddedness was evaluated by correlation and by contingency table (see Table B-13). Embeddedness was judged to have little if any effect on the derivation of the HC_{05} for conductivity (see Tables B-12 and B-13).

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 pH is also a major anion moiety that contributes to conductivity. Hence, pH directly reflects a major constituent of conductivity (HCO_3^-) and should not be analyzed as a potential confounder. In addition, salts influence hydrogen ion activity which is measured as pH. In any case, variance in pH was judged to have little effect on the derivation of the HC_{05} for conductivity in waters above pH 7 (see Tables B-14 and B-15).

B.4.6. Low pH

Because low pH from acid mine drainage is known to be an important cause of impairment where it occurs, it was judged a priori to be a potentially important environmental variable. That preconception was supported by the evidence summarized here. 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-16). So, although the benchmark applies to waters with neutral or basic pH, high conductivity appears to also cause effects at low pH.

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. In an analysis of a small data set, Pond et al. (2008a) found that the number of ephemeropteran genera was highly correlated with Se concentration ($r = -0.88$, $n = 20$). In contrast, weak correlations were found in our analysis of the West Virginia data. This result is unreliable, because most of the Se values were detection limits, and many of the detection limits were relatively high, equaling or exceeding the water quality criterion of 5.0 μg/L. In addition, there were too few high Se concentrations in the West Virginia data to perform a contingency table analysis. For these reasons, we did not include a quantitative analysis of potential confounding by Se. The effects of Se 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 lack of stream shading or if saline effluents are heated. Although temperature is moderately correlated with conductivity on an annual basis, the correlation is greatly reduced by seasonal partitioning (see Tables B-17 and B-18). More importantly, elevated temperature does not appear to be associated with the loss of Ephemeroptera.

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-2, B-3, B-4). Conductivity in Spring Branch measured <44−66 μS/cm. Conductivity in Ballard Fork was 464−2,300 μS/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 indicative of 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. Other potential confounders are apparently not responsible for differences between the creeks, because biological quality is not related to habitat quality (embeddedness, total RBP habitat score, and pH). Data are from U.S. EPA Mountain Top Mining studies (Green et al., 2000; Pond et al., 2008a) (see Table B-19).

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. 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-20). WVSCI scores and numbers of EPT taxa were low when conductivity was high regardless of the condition of catchments that provided sources of benthic macroinvertebrates including salt-sensitive genera. Data are from WABbase.

In these two examples, 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.

B.4.10. Catchment Area

Larger streams tend to have more moderate chemical properties than small streams because they receive waters from more sources than small streams, 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 km^2 (or 60 mi^2). 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? We examined these issues by analyzing the influence of stream size (as catchment area) on the effects of conductivity.

Correlation of log conductivity with log catchment area is extremely low ($r = 0.12$). Owing to the large number of sites ($N = 1{,}750$), the regression is statistically significant, but it is almost negligible and accounts for less than 2% of the variability in conductivity. Nearly all reference sites, even those identified as Level II and Level III, had conductivities less than 300 μS/cm.

We categorized streams by catchment area into three groups: small catchments less than 6 mi^2 (15.5 km^2), medium catchments of 6 to 60 mi^2 (155 km^2), and large catchments greater than 60 mi^2. The number of Ephemeroptera (mayfly) taxa declines with increasing conductivity in all streams, independent of classification of catchment area ($r = -0.62$).

We likewise categorized conductivity into three groups by defining low conductivity as <200 μS/cm, and high conductivity as >1,500 μS/cm (see Table B-21). In all three stream size categories, if conductivity was <200 μS/cm, 99% or more of all streams had mayfly populations, but if conductivity was above 1,500 μS/cm, only 50% or fewer streams had mayflies (see Table B-21). Evidence for confounding by catchment area is summarized in Table B-22; the evidence is uniformly negative and we conclude that catchment area has little or no effect on invertebrate response to conductivity.

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

Low pH is a potential confounder, but sites with pH <6 were removed from the data set when calculating the benchmark value. The influence of Se could not be evaluated due to poor data and should be investigated. However, toxic levels of Se appear to be relatively uncommon. Other potential confounders were eliminated from consideration with some confidence. We do not argue that these variables have no influence, but their effects appear to be minimal given the inevitable variability in sites to which the benchmark would be applied.

REFERENCES

Green, J, M Passmore, H Childers. 2000 GREEN, J., M.. A survey of the condition of streams in the primary region of mountaintop mining/valley fill coal mining. Appendix in Mountaintop mining/valley fills in Appalachia. Final programmatic environmental impact statement. Region 3, US Environmental Protection Agency, Philadelphia, Pennsylvania. (Available from: http://www.epa.gov/ region3/mtntop/pdf/Appendices/Appendix%20D%20Aquatic/Streams%20Invertebrate%20Study-%20EPA/ FINAL.pdf)

Hill, AB. (1965) The environment and disease: Association or causation. Proceed Royal Soc Med 58:295−300.

Pearl, J. (2000) Causality: Models, Reasoning, and Inference. Cambridge U. Press, Cambridge, UK.

Pond, GJ; Passmore, ME; Borsuk, FA; L. Reynolds; CJ Rose. (2008) Downstream effects of mountaintop coal mining: comparing biological conditions using family- and genus-level macroinvertebrate bioassessment tools. J N Am Benthol Soc 27:717−737.

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.

U.S. EPA (Environmental Protection Agency). (1976) Quality criteria for water. Washington, DC.

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

Table B-2. Weighting co-occurrence using correlations for Approaches 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	$r > 0.75$	+ +

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

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 deviating by <10% $0.56 < r < 0.69$	− − for a lack of change in effect with removal of confounder
	Coefficients deviating by <20% $0.50 \leq r \leq 0.75$	− for a small change in effect with removal of confounder
	Coefficients deviating by >20% $0.50 > r > 0.75$	+ for a strong increase or decrease in effect with removal of confounder

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

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

Table B-7. Evidence and weight for confounding by habitat quality

Approach	Score	Evidence
1. Correlation of cause and confounder	+	RBP score was moderately correlated with conductivity, ($r = -0.29$, $n = 2{,}344$).
2. Correlation of effect and confounder	+	RBP score was (barely) moderately correlated with the number of ephemeropteran genera ($r = 0.26$, $n = 2{,}193$).
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 µS/cm) even when habitat is poor (<115). However, with high conductivity, Ephemeroptera are present at only 40% of sites with poor habitat and 60% of sites with good 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 ($r = -0.50$, $n = 768$) (see Table B-8).
		The SSD and HC_{05} are very similar when the XC_{95} values were calculated with a year-long data set and 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).
Weight of evidence	−	Somewhat confident, evidence is mixed, but the contingency table gives relatively strong negative evidence (Ephemeroptera occur even when habitat is poor), while RBP explains only 6.7% of the variance in ephemeropteran occurrence.

SSD = species sensitivity distribution.

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

	Conductivity <200 µS/cm	Conductivity ≥1,500 µS/cm
Habitat score <115	155/157 (98.7%)	12/31 (39.7%)
Habitat score ≥140	388/390 (99.5%)	13/22 (59.1%)

Table B-9. Evidence and weights for confounding by organic enrichment

Approach	**Score**	**Evidence**
1. Correlation of cause and confounder	− −	Fecal coliform counts were weakly correlated with conductivity (r = 0.25, n = 2,044).
2. Correlation of effect and confounder	− −	Coliform count was not correlated with the number of ephemeropteran genera (r = −0.14, n = 1,349).
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 (see Table B-10).
6. Removal of confounder	− −	(a) When samples >400 colonies/100 mL were removed from the analysis, the correlation of conductivity with Ephemeroptera was unchanged (r = −0.63, n = 1,671). (b) The SSD and HC_{05} are very similar when the XC_{95} is calculated with a year-long data set and 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).
Weight of evidence	**− −**	Very confident: All negative, some strongly negative. No treatment for confounding.

Table B-10. Number of sites with high and low conductivity with high and low levels of fecal coliform with Ephemeroptera present in streams

	Conductivity <200 µS/cm	Conductivity ≥1,500 µS/cm
Coliform <400 colonies/100 mL	658/662 (99.4%)	36/77 (46.7%)
Coliform >400 colonies/100 mL	233/237 (98.3%)	20/42 (47.6%)

Table B-11. Evidence and weights for confounding by nutrients

Approach	Score	Evidence
1. Correlation of cause and confounder	− −	Conductivity was not correlated with nitrate and nitrite ($r = 0.08$, $n = 1{,}265$) or total phosphorus ($r = 0.05$, $n = 1{,}190$).
2. Correlation of effect and confounder	− −	Ephemeroptera was not correlated with nitrate and nitrite ($r = 0.037$, $n = 1{,}184$) or total phosphorous ($r = 0.001$, $n = 1{,}186$).
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 similar ($r = -0.54$, $n = 999$). When samples with total phosphorus >0.04 mg/L were removed from the analysis, the correlation of conductivity with the number of Ephemeroptera was similar ($r = -0.56$, $n = 999$).
Weight of evidence	− −	Moderately confident: all negative, none strongly negative. No treatment for confounding.

NA = not applicable.

Table B-12. Evidence and weights for confounding by deposited sediment

Approach	Score	Evidence
1. Correlation of cause and confounder	− −	The WABbase embeddedness score is weakly correlated with conductivity ($r = -0.18$, $n = 2{,}202$).
2. Correlation of effect and confounder	− −	The WABbase embeddedness score is weakly correlated with Ephemeroptera ($r = -0.22$, $n = 2{,}198$).
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 (see Table B-13).
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 ($r = -0.61$, $n = 1{,}089$).
Weight of evidence	**− −**	Very confident: all negative, some strongly. No treatment for confounding.

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 µS/cm	**Conductivity >1,500 µS/cm**
Embeddedness score <7	56/58 (96.6%)	7/16 (43.8%)
Embeddedness score >15	208/225 (92.4%)	6/15 (40%)

Table B-14. Evidence and weights for confounding by high pH

Approach	Score	Evidence
1. Correlation of cause and confounder	+	Conductivity was moderately correlated with pH between 7 and 9.0, ($r = 0.45$, $n = 1{,}911$).
2. Correlation of effect and confounder	− −	High pH was weakly correlated with Ephemeroptera ($r = 0.19$, $n = 1{,}907$).
3. Contingency of high level of cause and confounder	0	In a contingency table (see Table B-15), there were too few streams with high pH to provide evidence for or against confounding.
5. Levels of confounder is known to cause effects	−	U.S. EPA (1976) Water Quality Standards indicate that water quality 6.5−9 is protective of freshwater fish.
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 ($r = -0.63$, $n = 1{,}089$) (see Table B-15).
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%).
Weight of evidence	**−**	Reasonably confident: majority negative. No treatment for confounding.

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

	Conductivity <200 µS/cm	Conductivity ≥1,500 µS/cm
pH <4.5	16/19 (84.2%)	0/14 (0%)
pH >8.5	3/3 (100%)	4/8 (50%)

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

Approach	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).
Weight of evidence	+	Potential confounding: majority positive. Correction for confounding may be advisable.

Table B-17. Evidence and weights for confounding by temperature

Approach	Score	Evidence
1. Correlation of cause and confounder	0	Conductivity was moderately correlated with year-round temperature, $r = 0.39$, $n = 2{,}221$. When the correlation was recalculated for the spring and summer index periods, conductivity was less correlated, but still moderately, with the summer ($r = 0.29$, $n = 961$). Spring temperatures, however, were weakly correlated with conductivity ($r = 0.16$, $n = 1{,}199$).
2. Correlation of effect and confounder	−	Temperature was weakly correlated with Ephemeroptera year round ($r = -0.196$, $n = 2{,}363$) and in summer ($r = -0.12$, $n = 961$) and not correlated in spring ($r = -0.04$, $n = 1{,}195$).
3. Contingency of high level of cause and confounder	− − −	Ephemeroptera were present at >98−100% 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-18), which is contrary to expectations, if temperature were contributing to the impairment.
5. Levels of confounder is known to cause effects	−	Temperatures rarely exceeded 20°C and, therefore, are not likely to cause extirpation of genera.
6. Removal of confounder shows it is important	− −	When high temperatures (>22°C) were deleted, the correlation of conductivity and Ephemeroptera was barely changed ($r = -0.61$, $n = 1{,}788$).
Weight of evidence	**− − −**	Moderately confident: all negative, some strongly negative. No treatment for confounding.

Table B-18. 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 (98.7%)	9/27 (33.3%)
Temperature >22°C	78/78 (100%)	24/43 (55.8%)

Table B-19. Comparison of low conductivity Spring Branch with high conductivity Ballard Fork

Stream Name	Date	Embd.	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

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

Table B-20. 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	# EPT Taxa	WVSCI
2003	44.6	Forested	44	148	6	90.72
2004	44.6	Forested	37			-
1998	25.1	Mined	805	155	3	67.62
2003	25.1	Mined	2,087	153	1	58.45
2003	11.9	Mixed Forest and Mine	1,702	157	2	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	3	66.73
2004	0.5	Mixed Forest, Mine, & Urban	836	-	-	-
1998	0	Mixed Forest, Mine, & Urban	590	131	3	65.94

[a]Land use refers to catchment land use of tributaries upstream from the sampled sites in Twentymile Creek.

EPT taxa = Number of Ephemeroptera, Plecoptera, and Trichoptera taxa; WVSCI = West Virginia Stream Condition Index.

Table B-21. Number (and percent) of streams with Ephemeroptera present: small, medium and large streams and low, medium and high conductivity (pH > 6).

	Conductivity < 200 µS/cm	Conductivity > 1,500 µS/cm
Small streams	426/430 (99%)	16/39 (41%)
Medium streams	205/207 (99%)	19/38 (50%)
Large streams	70/70 (100%)	1/2 (50%)

Table B-22. Evidence and weights for confounding by catchment area

Approach	**Score**	**Evidence**
1. Correlation of cause and confounder	− −	Log catchment area was very weakly correlated with log conductivity (*r* = 0.12, *n* = 1,750).
2. Correlation of effect and confounder	− − −	Log catchment area was not correlated with the number of ephemeropteran genera (*r* = 0.05, *n* = 1,750).
3. Contingency of high level of cause and confounder	− − −	In a contingency table (see Table B-21), large catchment area did not change the probability of finding Ephemeroptera at either high or low conductivity.
6. Removal of confounder	− −	Overall correlation of conductivity with Ephemeroptera in all sites was *r* = − 0.62, *n* = 1,750). When correlation was repeated for each of the size classes, the correlations were −0.48 (large streams; *n* = 165), and − 0.64 (small and medium streams; *n* = 942 and 653).
Weight of evidence	**− − −**	Very confident: All negative, some strongly negative. No treatment for confounding.

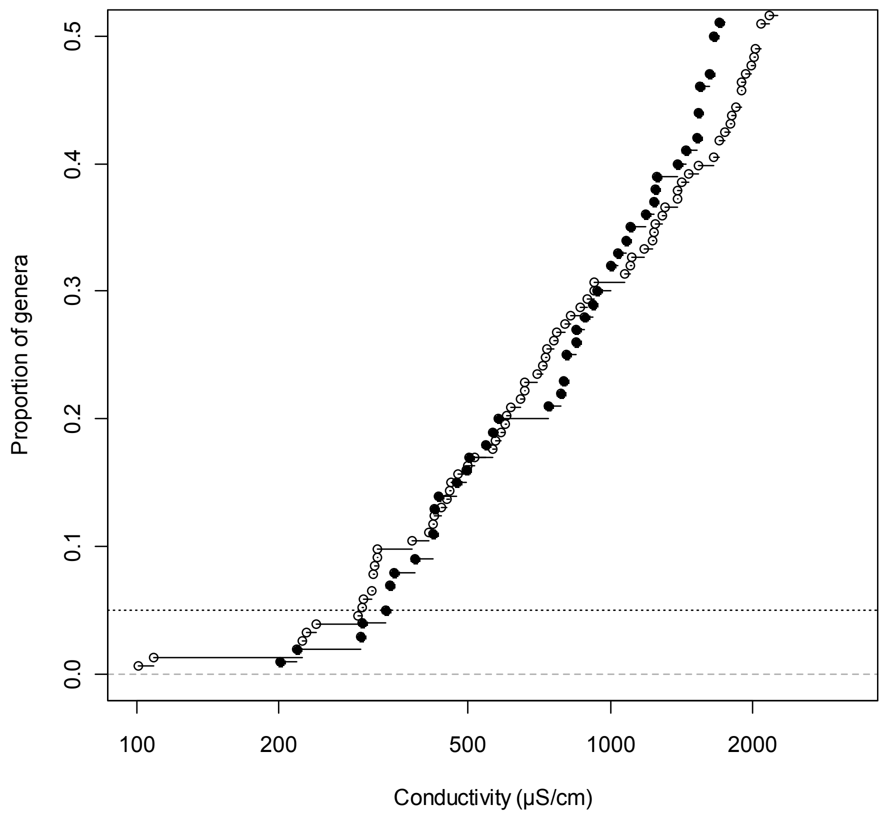

Figure B-1. Species sensitivity distribution for all year, pH >6 and all sites (open circles) and for sites with pH ≥6, Rapid Bioassessment Protocol ≥135 and fecal coliform ≤400 colonies/100 mL (closed circles). Habitat disturbance and organic enrichment have little influence; the HC_{05} for the constrained data set is 300 µS/cm based on 111 genera. The upper and lower confidence bounds on that value are 225 µS/cm and 350 µS/cm, respectively.

Figure B-2. Topographical map of Spring Branch (blue triangle) and Ballard Fork (red triangle) sampling stations.

Figure B-3. 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 as examples.

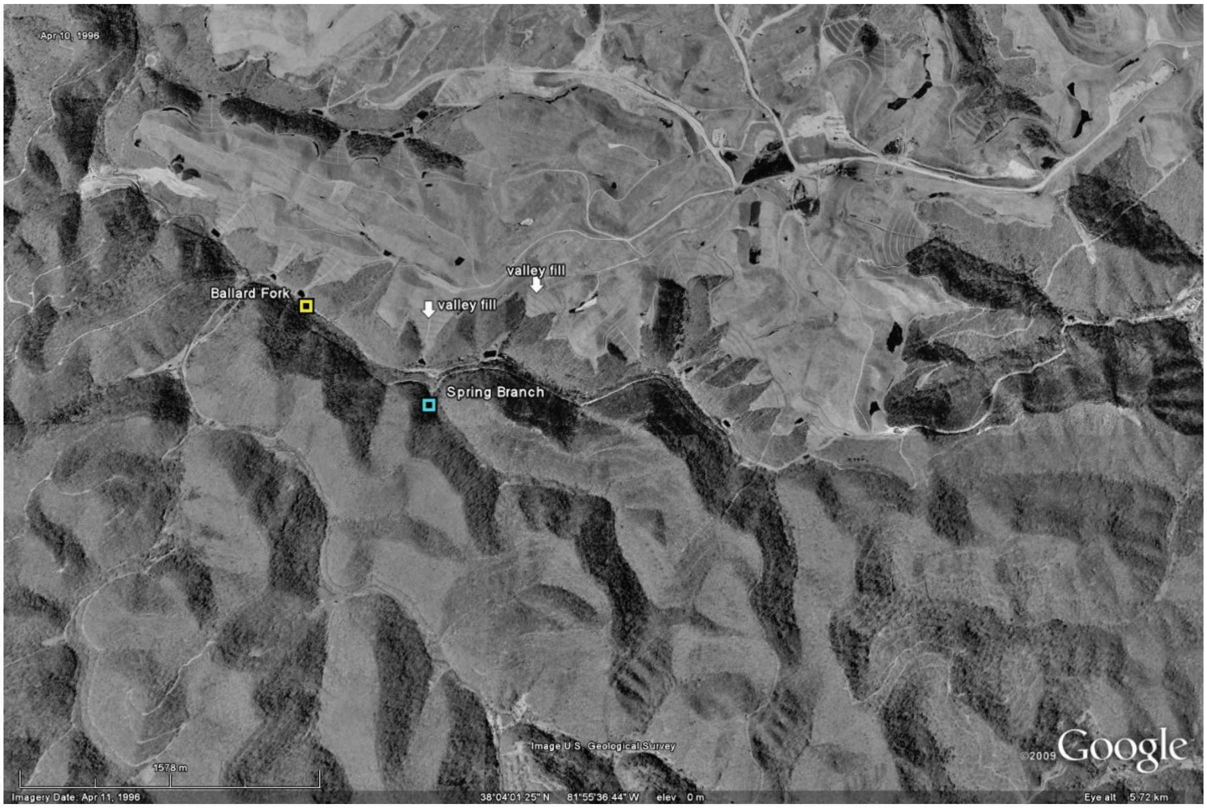

Figure B-4. Same area as Figure 2. Aerial imagery (April 10, 1996) with superimposed sampling locations of spring branch (turquoise square) and Ballard Fork (yellow square). The many upstream valley fills in Ballard Fork are easily seen.

APPENDIX C
EXTIRPATION CONCENTRATION VALUES FOR INVERTEBRATES

	Genus	Both			Spring		Summer	
		XC_{95}	*N*	Ref.	XC_{95}	*N*	XC_{95}	*N*
1	Ablabesmyia	>11,646	162	5	3,162	56	11,646	106
2	Acentrella	1,289	748	31	1,289	422	769	326
3	Acroneuria	2,630	480	60	1,649	138	2,320	342
4	Alloperla	228	96	15	319	82		
5	Ameletus	599	192	30	388	189		
6	Amphinemura	805	561	42	1,468	556		
7	Antocha	>6,468	538	18	3,725	162	6,468	376
8	Argia	9,790	75	NA			9,790	71
9	Asellus	925	33	2				
10	Atherix	>11,646	156	3			11,646	149
11	Atrichopogon	2,257	42	3			2,257	40
12	Attenella	574	34	1				
13	Baetis	1,383	1,509	72	1,383	642	1,494	867
14	Baetisca	918	47	NA			646	32
15	Bezzia	381	62	2	563	39	11,227	127
16	Bezzia/Palpomyia	4,713	306	26	3,725	179		
17	Boyeria	>7,340	173	5	1,468	52	7,340	121
18	Brillia	1,746	91	6	1,083	51	2,768	40
19	Caecidotea	>4,713	137	1	1,083	51	4,713	62
20	Caenis	3,884	541	8	1,175	168	3,884	373
21	Calopteryx	3,489	52	NA			3,489	45
22	Cambarus	1,228	464	44	1,649	289	1,340	175
23	Cardiocladius	>2,257	185	2	1,530	86	2,257	99
24	Centroptilum	1,075	90	6	1,175	59	1,075	31
25	Ceratopsyche	>6,468	885	29	3,314	232	6,468	653
26	Chaetocladius	>5,057	182	4	1,650	76	5,057	106
27	Chelifera	>3,341	152	9	1,650	64	3,341	88
28	Cheumatopsyche	>9,180	1,612	57	2,493	562	9,180	1,050
29	Chimarra	>3,972	490	11	1,175	100	3,972	390
30	Chironomus	>11,646	105	1	5,120	48	11,646	57
31	Chrysops	>11,646	76	1			11,646	51

	Genus	Both			Spring		Summer	
		XC_{95}	*N*	Ref.	XC_{95}	*N*	XC_{95}	*N*
32	Cinygmula	224	81	17	347	80		
33	Cladotanytarsus	>11,646	103	5	3,162	57	11,646	46
34	Clinocera	>4,713	60	6	1,573	33		
35	Conchapelopia	518	135	7	1,175	120		
36	Corbicula	9,790	184	NA	1,175	51	9,790	133
37	Cordulegaster	1,468	42	3				
38	Corydalus	>11,227	311	1	1,117	49	11,227	262
39	Corynoneura	2,006	149	5	1,650	82	2,768	67
40	Crangonyx	2,169	105	7	796	65	2,169	40
41	Cricotopus	>11,227	605	24	3,725	274	11,227	331
42	Cricotopus/Orthocladius	>6,468	1,054	13	3,725	493	7,340	561
43	Cryptochironomus	>3,489	287	3	3,162	129	3,489	158
44	Dasyhelea	>3,341	66	3			3,341	51
45	Demicryptochironomus	322	81	6	322	67		
46	Diamesa	>4,713	457	14	3,725	294	5,057	163
47	Dicranota	>7,010	327	43	1,649	160	7,010	167
48	Dicrotendipes	>11,646	192	1	3,314	79	11,646	113
49	Dineutus	9,790	46	NA				
50	Diphetor	648	134	17	653	88	701	46
51	Diplectrona	2,523	594	60	3,725	233	2,523	361
52	Diploperla	318	99	2	357	96		
53	Dixa	722	68	16	1,650	34	794	34
54	Dolophilodes	864	339	46	1,323	145	618	194
55	Drunella	294	172	18	660	153		
56	Dubiraphia	>7,370	141	3	3,162	33	7,370	108
57	Eccoptura	462	65	6	518	43		
58	Ectopria	1,386	311	32	1,175	121	1,570	190
59	Epeorus	316	359	57	500	280	247	79
60	Ephemera	736	138	20	1,175	45	627	93
61	Ephemerella	302	362	38	434	347		
62	Eukiefferiella	1,930	501	28	1,649	271	1,979	230
63	Eurylophella	476	173	19	280	98	554	75

	Genus	Both			Spring		Summer	
		XC95	*N*	Ref.	XC95	*N*	XC95	*N*
64	Ferrissia	4,884	91	NA			4,884	80
65	Fossaria	5,057	30	NA				
66	Gammarus	>4,713	215	10	1,800	67	4,713	148
67	Glossosoma	1,650	154	7	1,650	38	925	116
68	Haploperla	423	235	27	497	182	603	53
69	Heleniella	1,700	62	7			2,768	35
70	Helichus	>11,646	328	18	1,650	164	11,646	164
71	Hemerodromia	>9,790	607	8	3,725	109	9,790	498
72	Heptagenia	313	68	3	269	55		
73	Hexatoma	>9,790	818	65	1,059	393	9,790	425
74	Hydroporus	822	32	1	810	30		
75	Hydropsyche	>7,010	981	21	3,725	234	7,010	747
76	Hydroptila	>11,646	278	4	3,162	51	11,227	227
77	Isonychia	1,177	712	16	1,175	234	1,068	478
78	Isoperla	459	485	39	694	428	704	57
79	Krenopelopia	2,320	61	2			2,320	36
80	Lanthus	2,087	66	7	1,175	34	1,702	32
81	Larsia	2,630	96	3	1,289	76		
82	Lepidostoma	109	88	12	796	74		
83	Leptophlebia	224	85	8	805	70		
84	Leucrocuta	425	219	29	1,175	158	418	61
85	Leuctra	2,087	1,170	84	1,175	665	2,257	505
86	Limnophila	2,768	49	10	322	33		
87	Limnophyes	>5,120	88	1	5,120	31	1,979	57
88	Limonia	>5,057	62	1			5,057	57
89	Lirceus	1,303	70	6	517	51		
90	Maccaffertium	1,111	174	13	680	82	1,177	92
91	Macronychus	1,890	39	4				
92	Microcylloepus	3,341	94	2			3,341	82
93	Micropsectra	>6,468	220	25	5,120	107	6,468	113
94	Microtendipes	>3,489	507	34	1,383	198	3,489	309
95	Microvelia	2,523	46	3			2,523	32

	Genus	Both			Spring		Summer	
		XC_{95}	*N*	Ref.	XC_{95}	*N*	XC_{95}	*N*
96	Nanocladius	1,485	50	NA				
97	Natarsia	1,842	54	1			1,842	36
98	Neophylax	323	122	36	578	116		
99	Nigronia	>9,790	726	37	3,162	204	7,340	522
100	Nilotanypus	2,630	112	3	731	49	2,630	63
101	Nixe	316	77	3	357	73		
102	Ochrotrichia	2,791	32	1				
103	Optioservus	9,790	1,429	65	1,890	500	7,370	929
104	Orconectes	3,162	205	2	3,162	56	1,978	149
105	Orthocladius	3,341	272	10	805	117	3,341	155
106	Oulimnius	2,791	219	27	1,650	73	2,440	146
107	Pagastia	1,800	46	2				
108	Palpomyia	1,870	40	NA				
109	Parachaetocladius	1,239	151	27	509	33	1,205	118
110	Paragnetina	2,087	39	3				
111	Parakiefferiella	1,700	75	2	1,006	39	1,896	36
112	Paraleptophlebia	439	432	46	496	295	488	137
113	Parametriocnemus	>4,713	1,450	72	2,493	687	4,713	763
114	Paraphaenocladius	>6,468	71	2			6,468	46
115	Paratanytarsus	>3,489	108	2	3,314	48	3,489	60
116	Paratendipes	11,227	78	NA	1,800	34	11,227	44
117	Peltoperla	659	124	12	1,650	73	745	51
118	Perlesta	3,314	314	8	3,162	289		
119	Phaenopsectra	2,332	89	2			2,332	64
120	Physella	>9,790	143	1	1,276	61	9,790	82
121	Pisidium	1,795	34	2				
122	Plauditus	927	286	12	847	209	2,257	77
123	Polycentropus	4,713	357	41	1,443	154	2,768	203
124	Polypedilum	>4,884	1,604	75	1,890	686	5,057	918
125	Potthastia	1,896	60	1	1,534	33		
126	Procloeon	701	78	3	347	31	524	47
127	Promoresia	589	78	5			467	53

	Genus	Both			Spring		Summer	
		XC_{95}	*N*	Ref.	XC_{95}	*N*	XC_{95}	*N*
128	Prosimulium	565	89	21	808	82		
129	Prostoma	2,553	41	NA				
130	Psephenus	>9,790	853	36	1,479	329	2,553	39
131	Pseudochironomus	>11,646	31	2			7,370	524
132	Pseudolimnophila	1,418	130	11	731	78	1,740	52
133	Psychomyia	1,106	38	3			1,106	33
134	Pteronarcys	660	105	25	499	41	907	64
135	Pycnopsyche	299	40	10	502	31		
136	Remenus	101	35	3	183	35		
137	Rhagovelia	2,030	51	3				
138	Rheocricotopus	3,489	556	11	1,346	225	3,489	331
139	Rheopelopia	1,247	125	4	1,534	37	1,014	88
140	Rheotanytarsus	>3,489	938	28	1,346	252	3,489	686
141	Rhyacophila	1,890	379	58	1,650	225	5,057	154
142	Serratella	500	46	2				
143	Sialis	>11,227	261	3	3,725	52	11,227	209
144	Simulium	>6,468	1,084	26	1,800	408	6,468	676
145	Sphaerium	>9,790	39	NA				
146	Stempellina	617	33	8				
147	Stempellinella	892	304	26	562	120	1,075	184
148	Stenacron	769	249	15	316	105	850	144
149	Stenelmis	>9,790	1,217	27	3,162	539	9,790	678
150	Stenochironomus	1,613	40	NA				
151	Stenonema	729	905	61	875	331	687	574
152	Stictochironomus	3,162	39	NA				
153	Stylogomphus	>6,468	117	1			6,468	99
154	Sublettea	2,440	179	2	929	43	2,087	136
155	Sweltsa	761	294	43	1,650	126	777	168
156	Tabanus	>9,790	60	1	3,162	35		
157	Tallaperla	452	87	16			302	63
158	Tanytarsus	9,180	1,194	64	1,650	481	9,180	713
159	Thienemanniella	>9,790	389	9	2,493	203	9,790	186

	Genus	Both			Spring		Summer	
		XC_{95}	*N*	Ref.	XC_{95}	*N*	XC_{95}	*N*
160	Thienemannimyia	>6,468	1,326	56	1,649	536	6,468	790
161	Tipula	1,979	590	36	1,649	320	2,169	270
162	Tokunagaia	1,070	43	NA			1,070	35
163	Tribelos	2,257	45	NA				
164	Tricorythodes	2,006	44	NA			2,006	43
165	Tvetenia	>2,768	727	40	1,649	370	5,057	357
166	Utaperla	240	47	2			198	32
167	Wormaldia	1,533	73	8	796	31	1,746	42
168	Yugus	603	72	12	796	48		
169	Zavrelia	413	81	6	347	60		
170	Zavrelimyia	>2,768	240	11	834	112	4,884	128

Empty cells indicates fewer than 30 occurrences during that season.

XC_{95} = 95th percentile extirpation concentration reported as µS/cm; NA = not applicable because it never occurs at WVDEP reference locations; Both = March through October; Spring = Sampled March through June; Summer = July through October; Ref. = number of times the genus was observed at a reference site.

APPENDIX D
GRAPHS OF OBSERVATION PROBABILITIES AND CUMULATIVE DISTRIBUTION FUNCTIONS FOR EACH GENUS

The purpose of Appendix D is to help the reader visualize the changes in the occurrence of each genus as conductivity increases. Figure D-1 contains general additive models of the relationship between capture probability of the genus and conductivity, ordered from the lowest to the highest XC_{95} value. 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 line fitted to the probabilities is for visualization. The vertical red line indicates the XC_{95}. Note that different genera respond differently to increasing salinity. For example, 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.

Figure D-2 contains the weighted cumulative distribution function (CDF) and the associated 95^{th} percentile extirpation concentration value arranged in alphabetical order by genus. Each point shows the weighted proportion of samples with each genus present at (F(x)) the conductivity less than the indicated conductivity value (μS/cm). The conductivity at the 95^{th} percentile is the XC_{95} (arrow). The CDF was calculated from observations from March through October (all year; black connected points) from March through June (spring; green connected points), and from July through October (summer; red connected points). As there were fewer than 30 observations such as for Drunella between July and October, no CDF was developed for the summer index period. In a CDF, genera that are affected by increasing conductivity (e.g., Drunella) show a steep slope and asymptote well below the measured range of 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 perceptible asymptote.

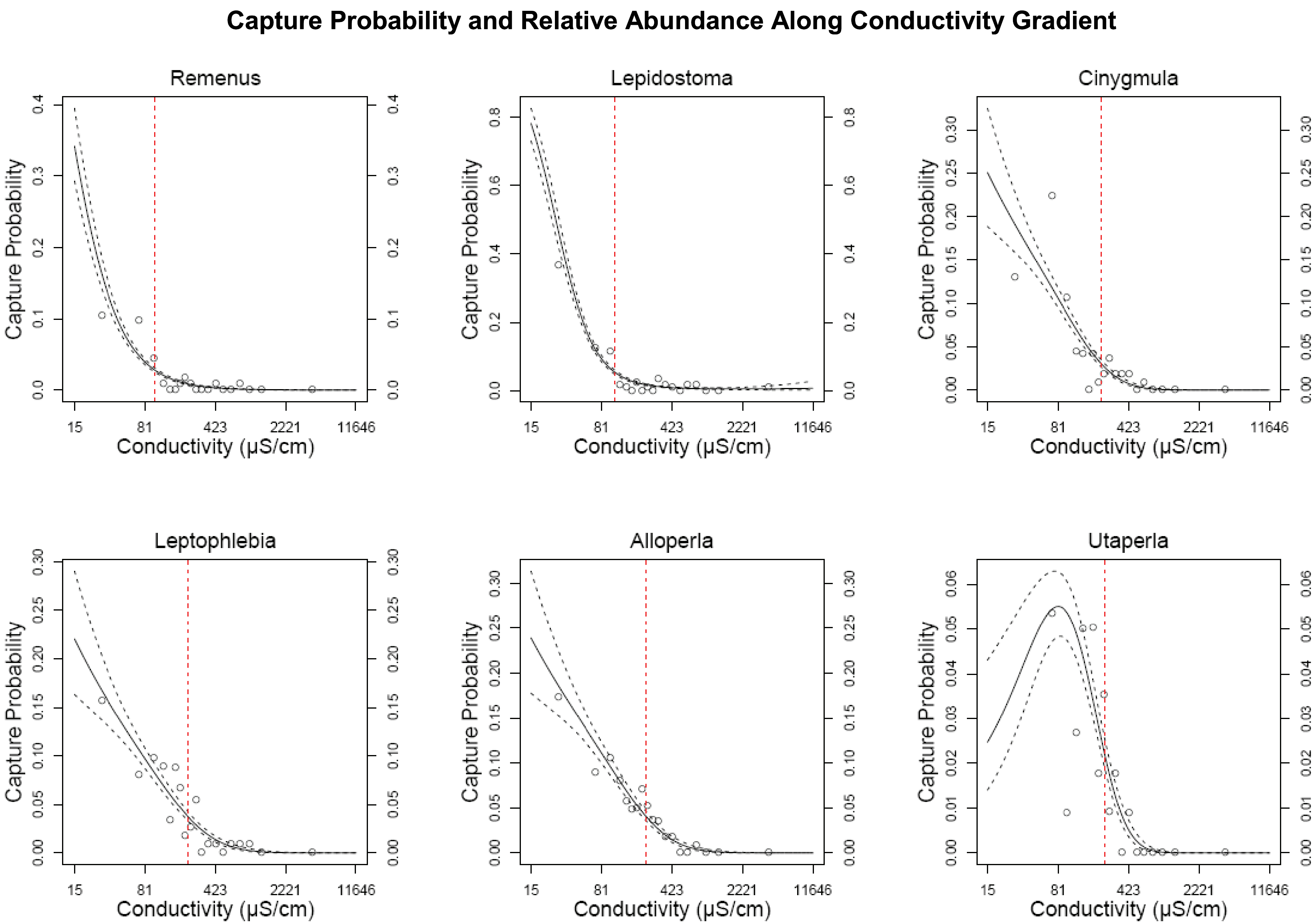

Figure D-1. Observation probabilities for each genus. Sites with pH <6 excluded.

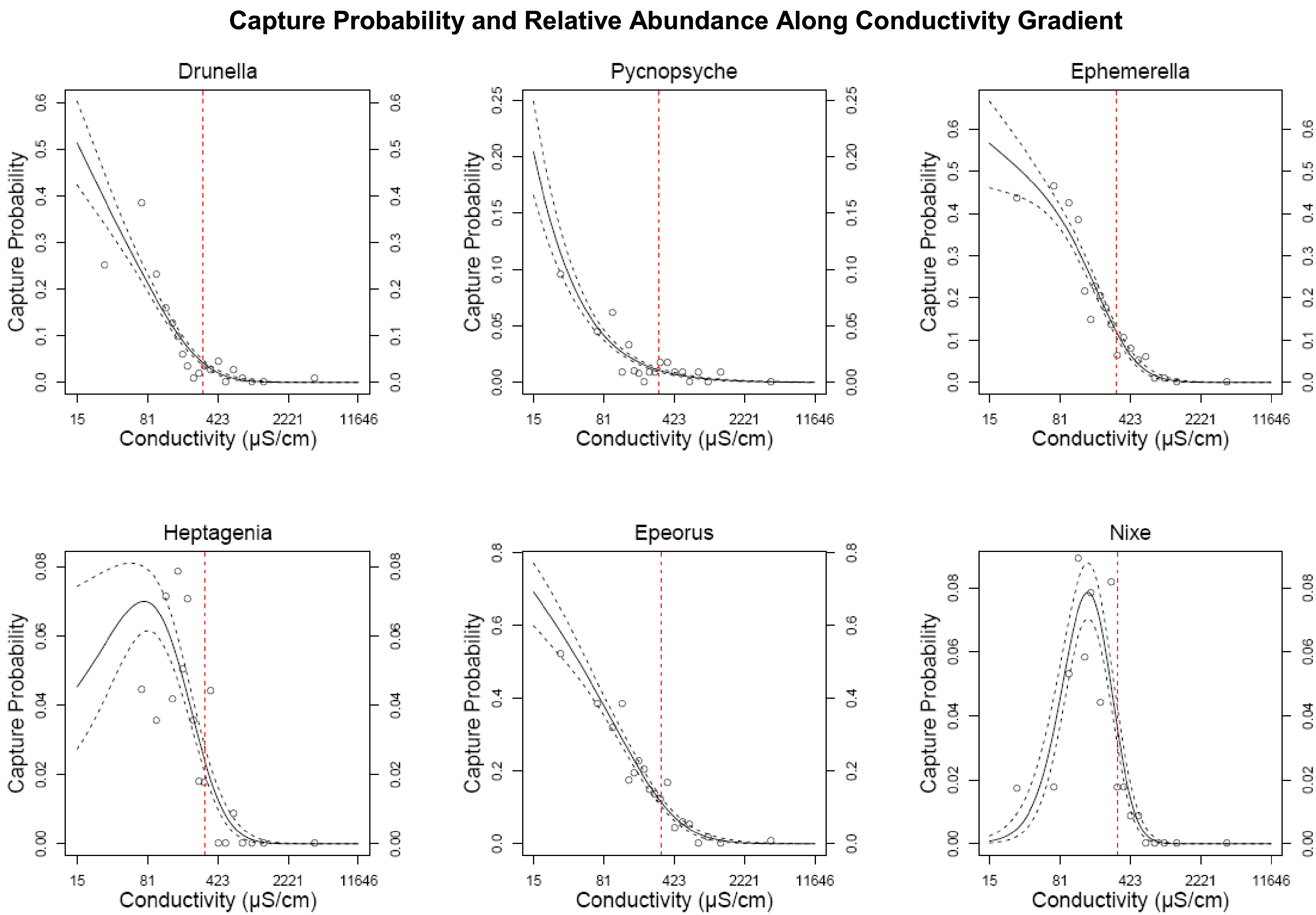

Figure D-1. Observation probabilities for each genus (continued).

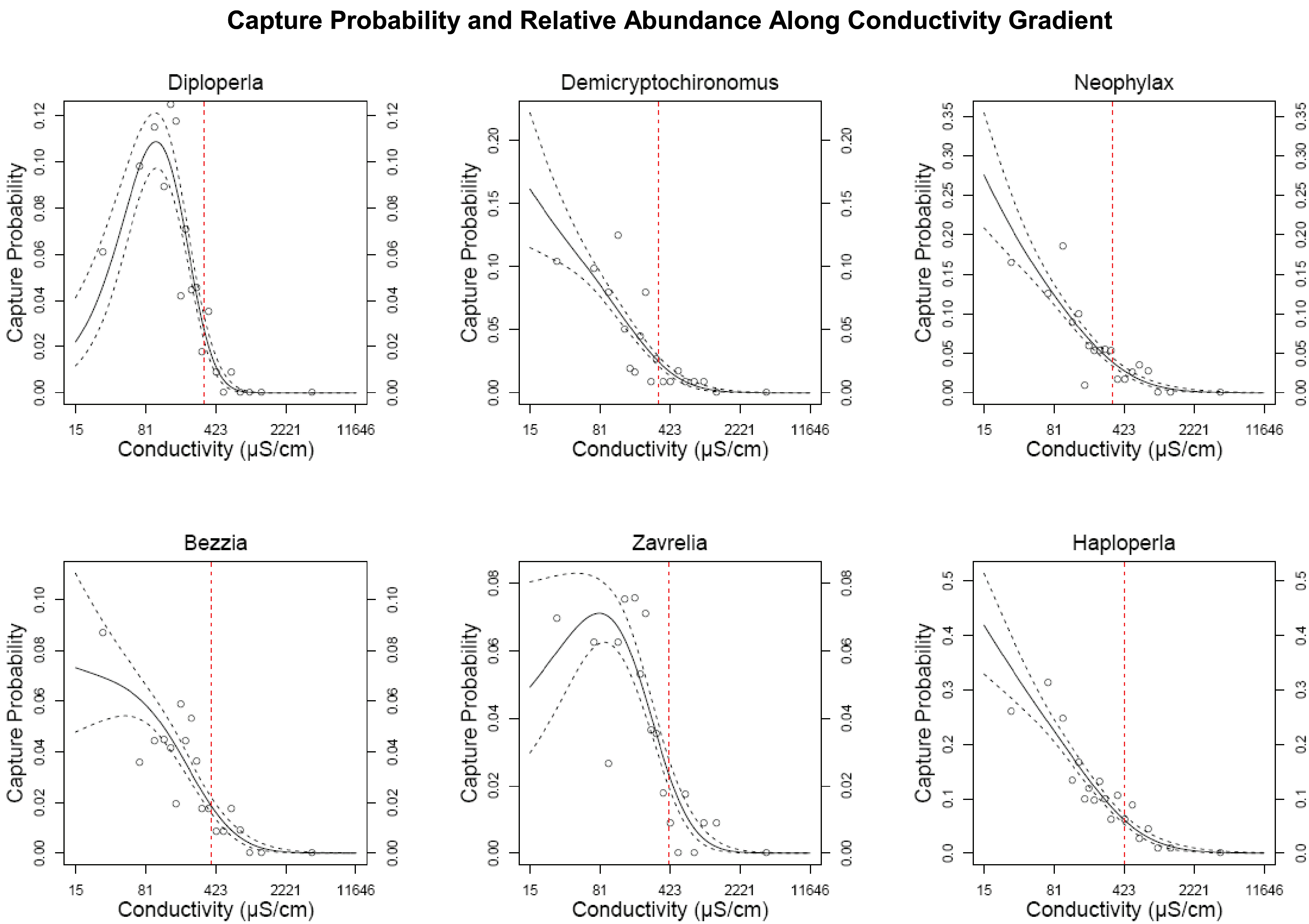

Figure D-1. Observation probabilities for each genus (continued).

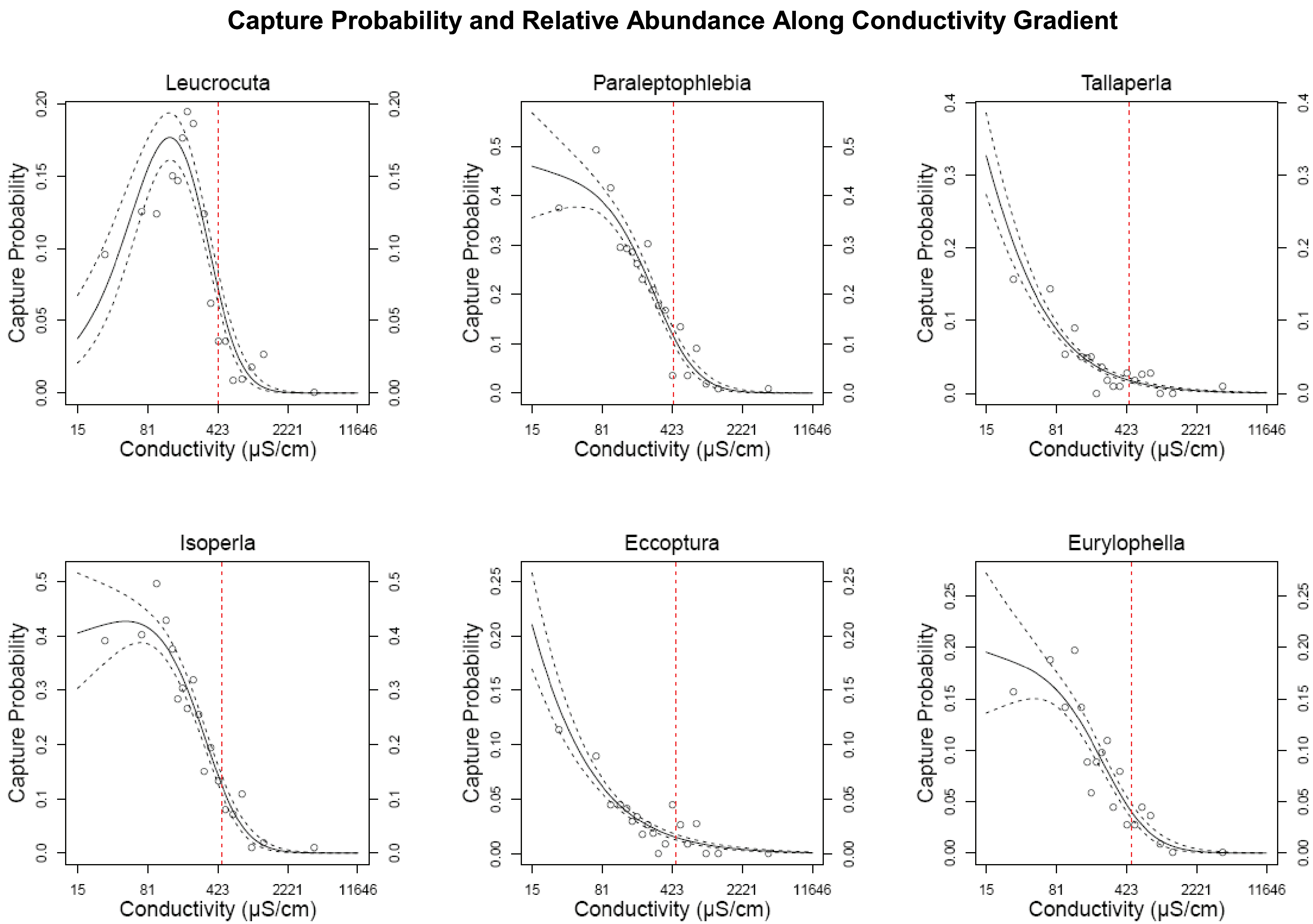

Figure D-1. Observation probabilities for each genus (continued).

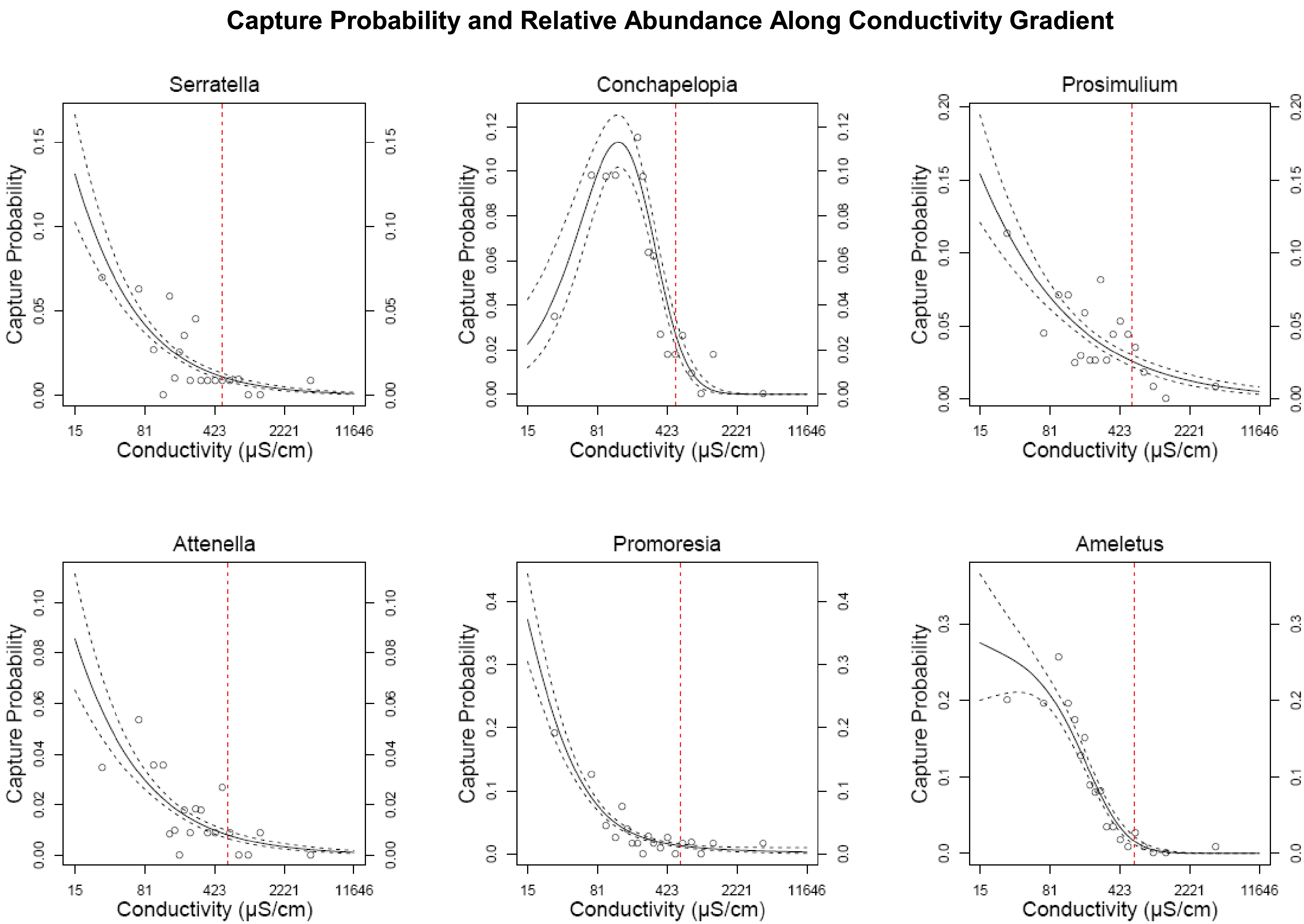

Figure D-1. Observation probabilities for each genus (continued).

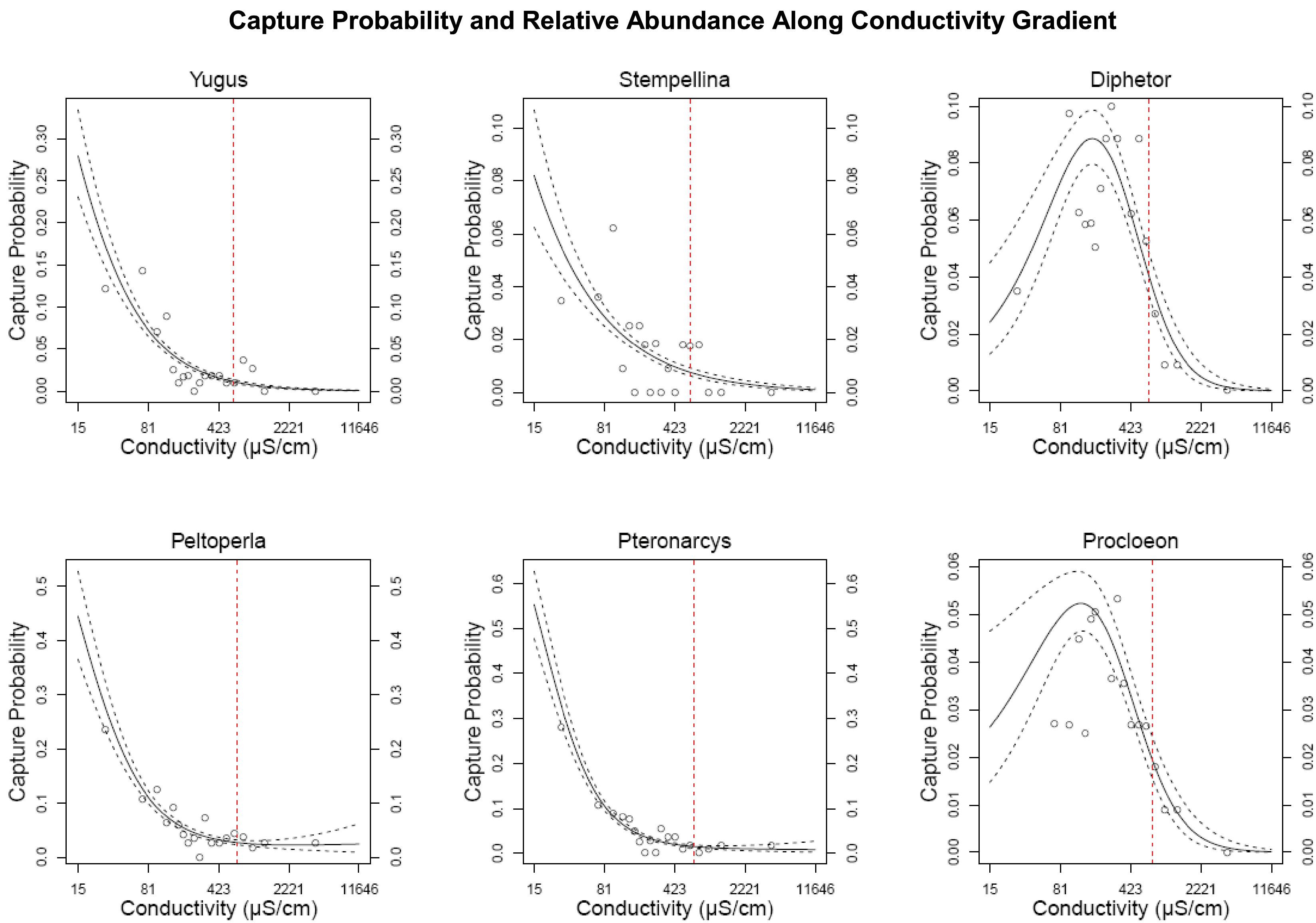

Figure D-1. Observation probabilities for each genus (continued).

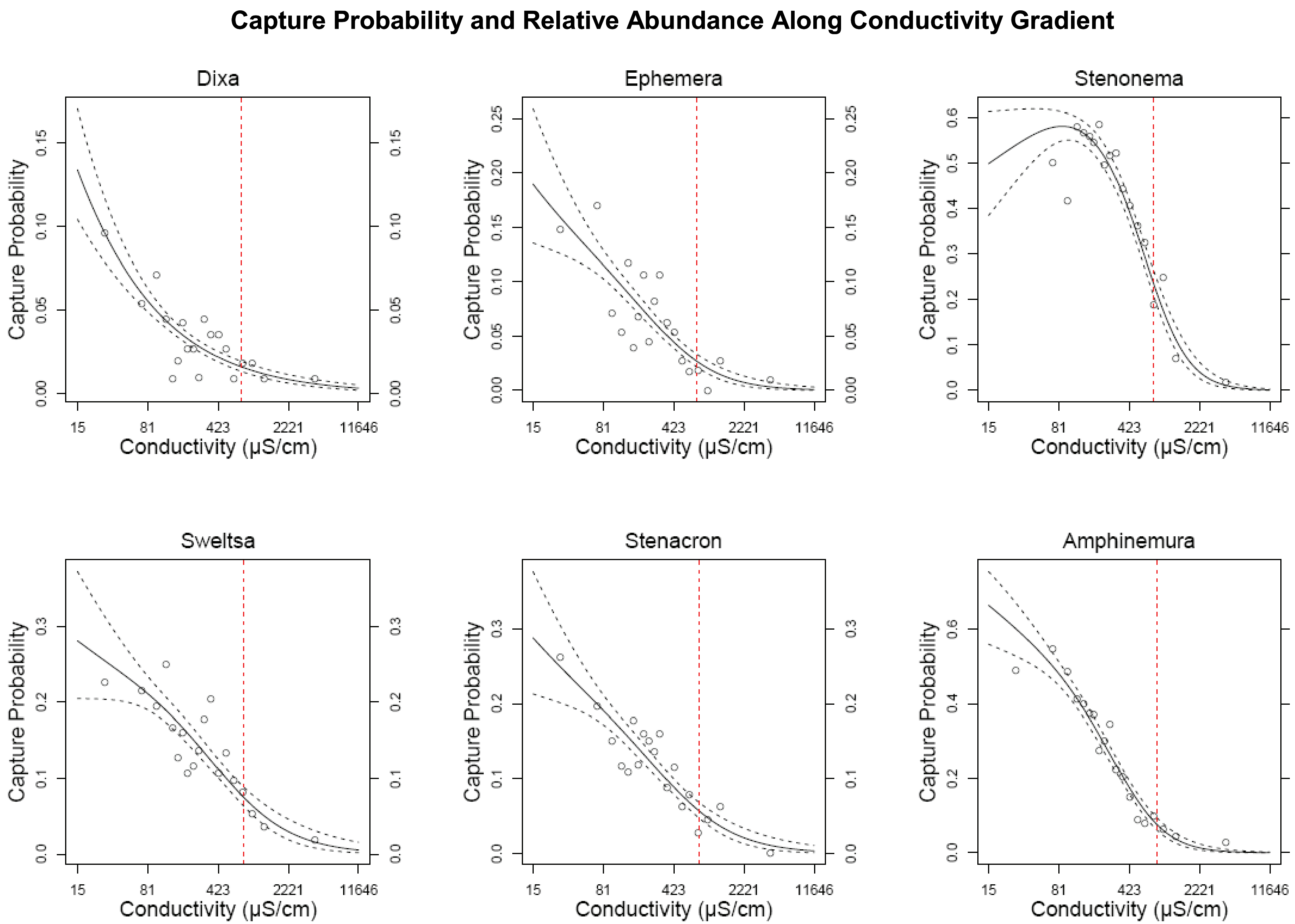

Figure D-1. Observation probabilities for each genus (continued).

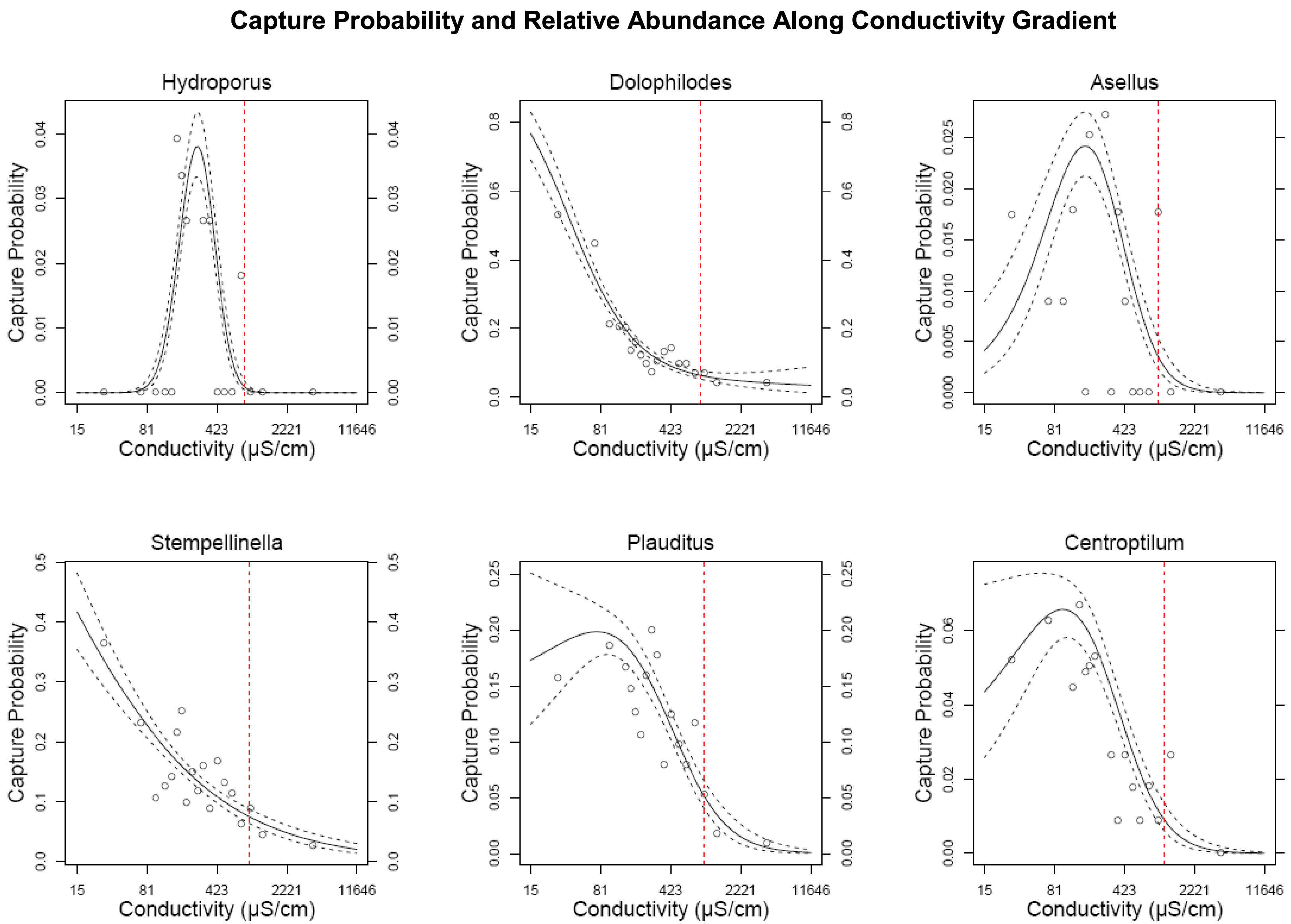

Figure D-1. Observation probabilities for each genus (continued).

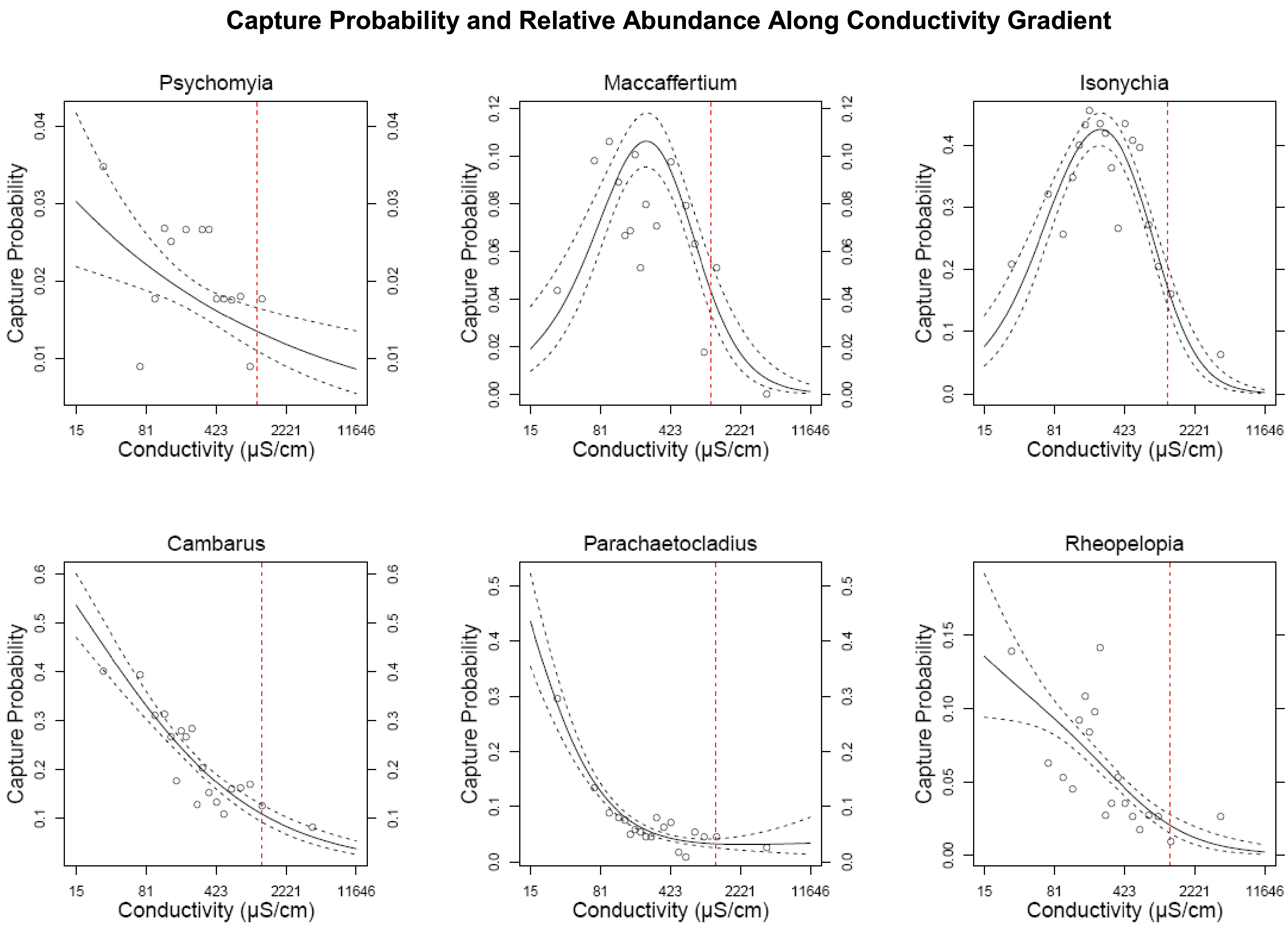

Figure D-1. Observation probabilities for each genus (continued).

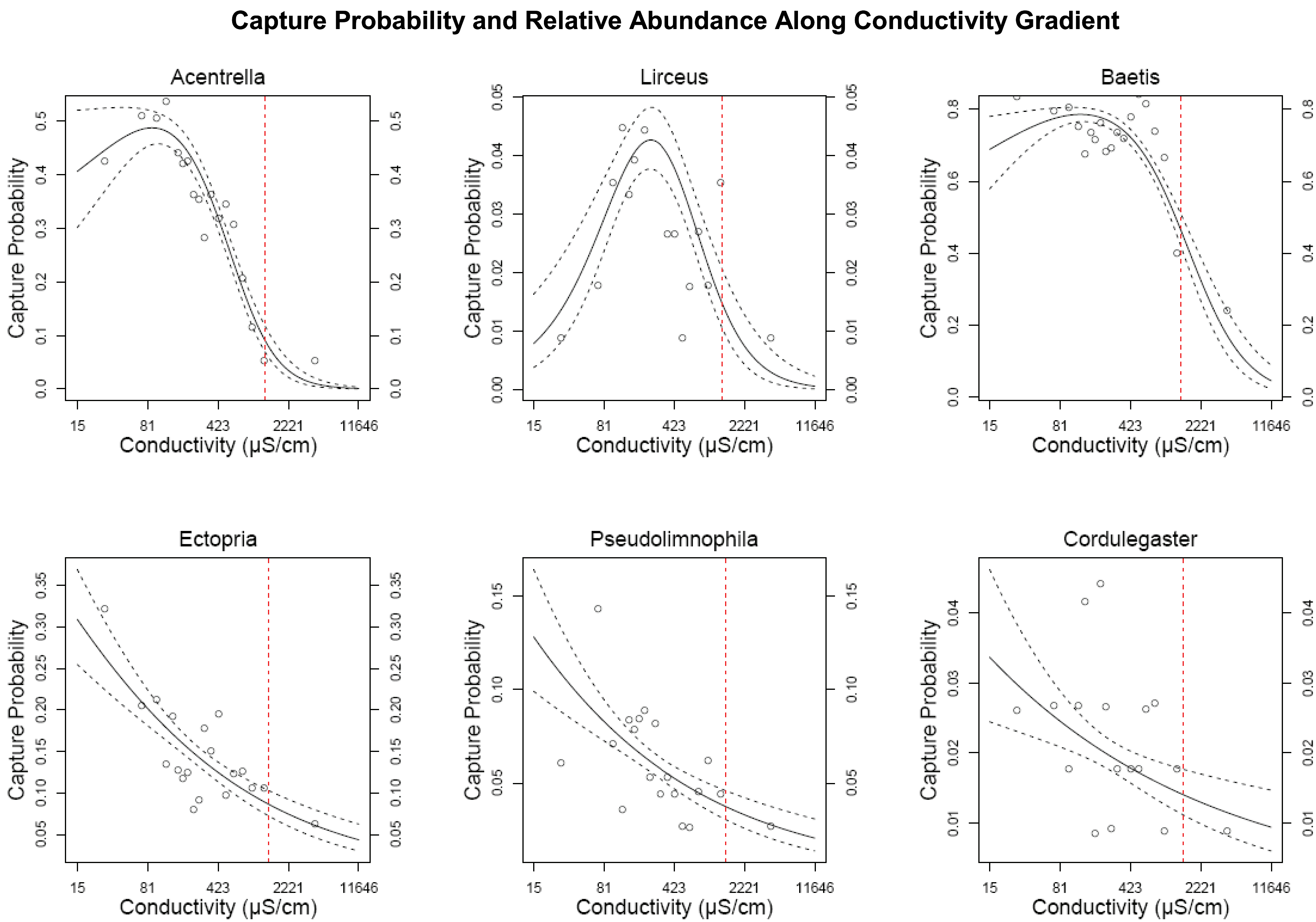

Figure D-1. Observation probabilities for each genus (continued).

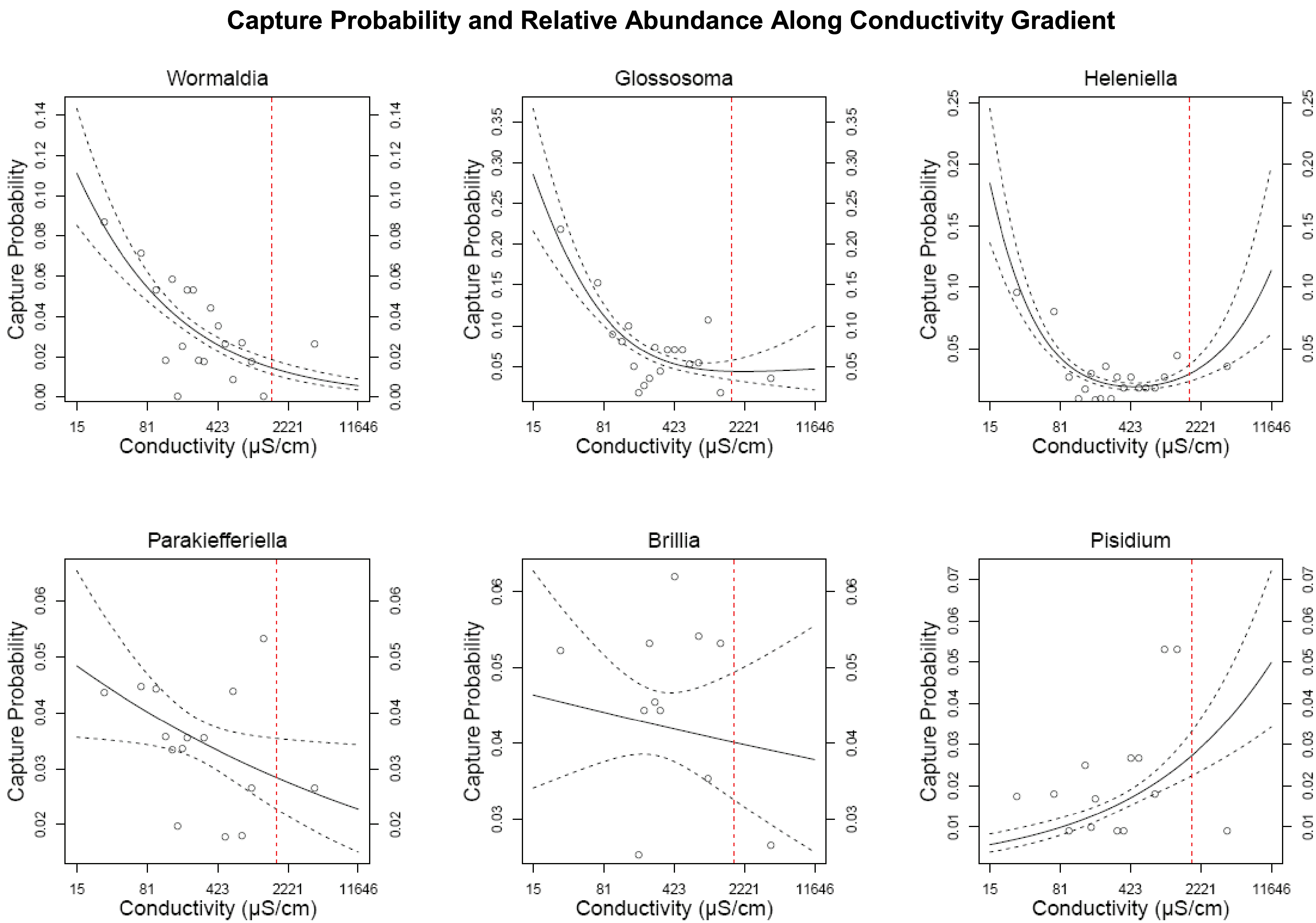

Figure D-1. Observation probabilities for each genus (continued).

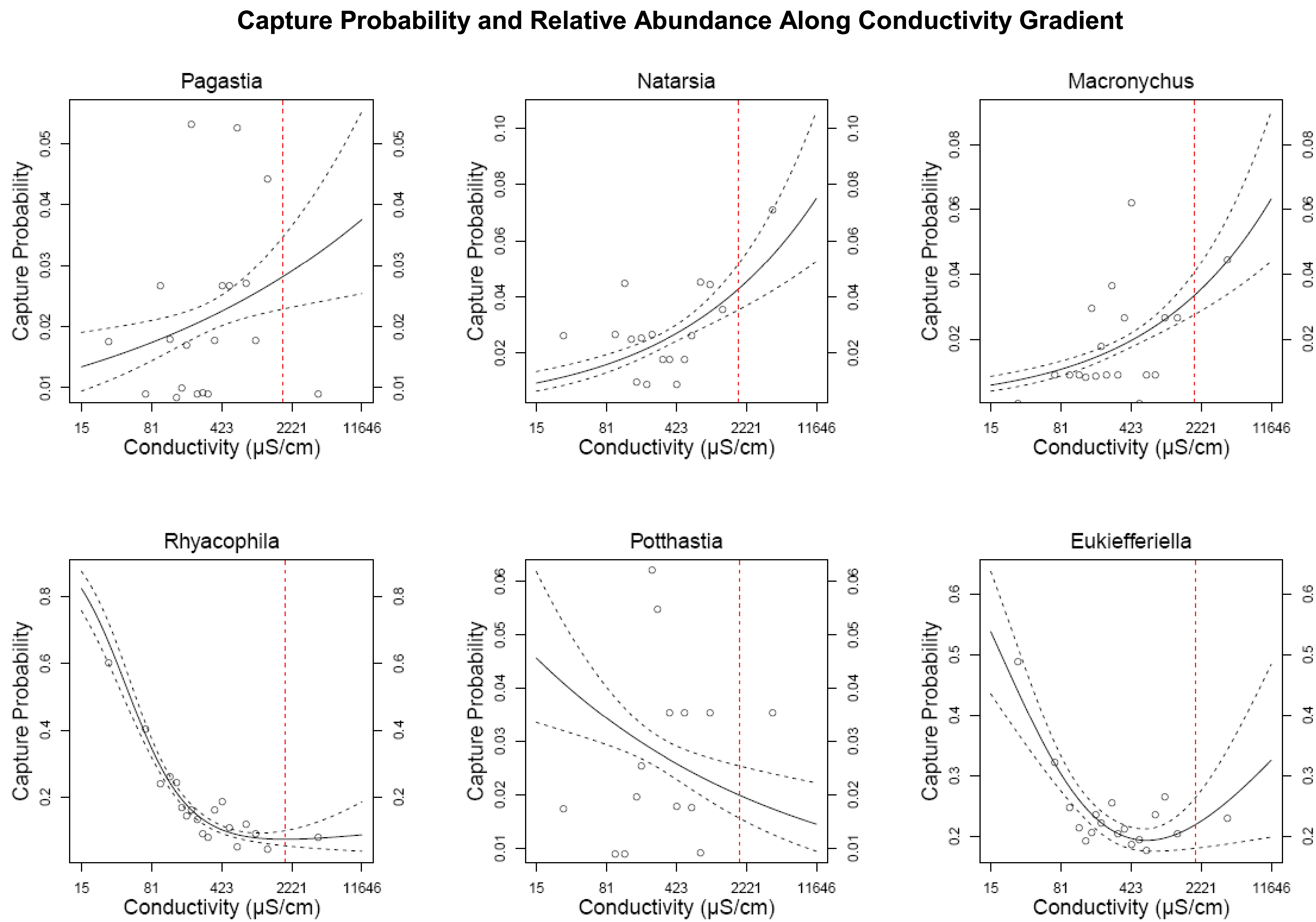

Figure D-1. Observation probabilities for each genus (continued).

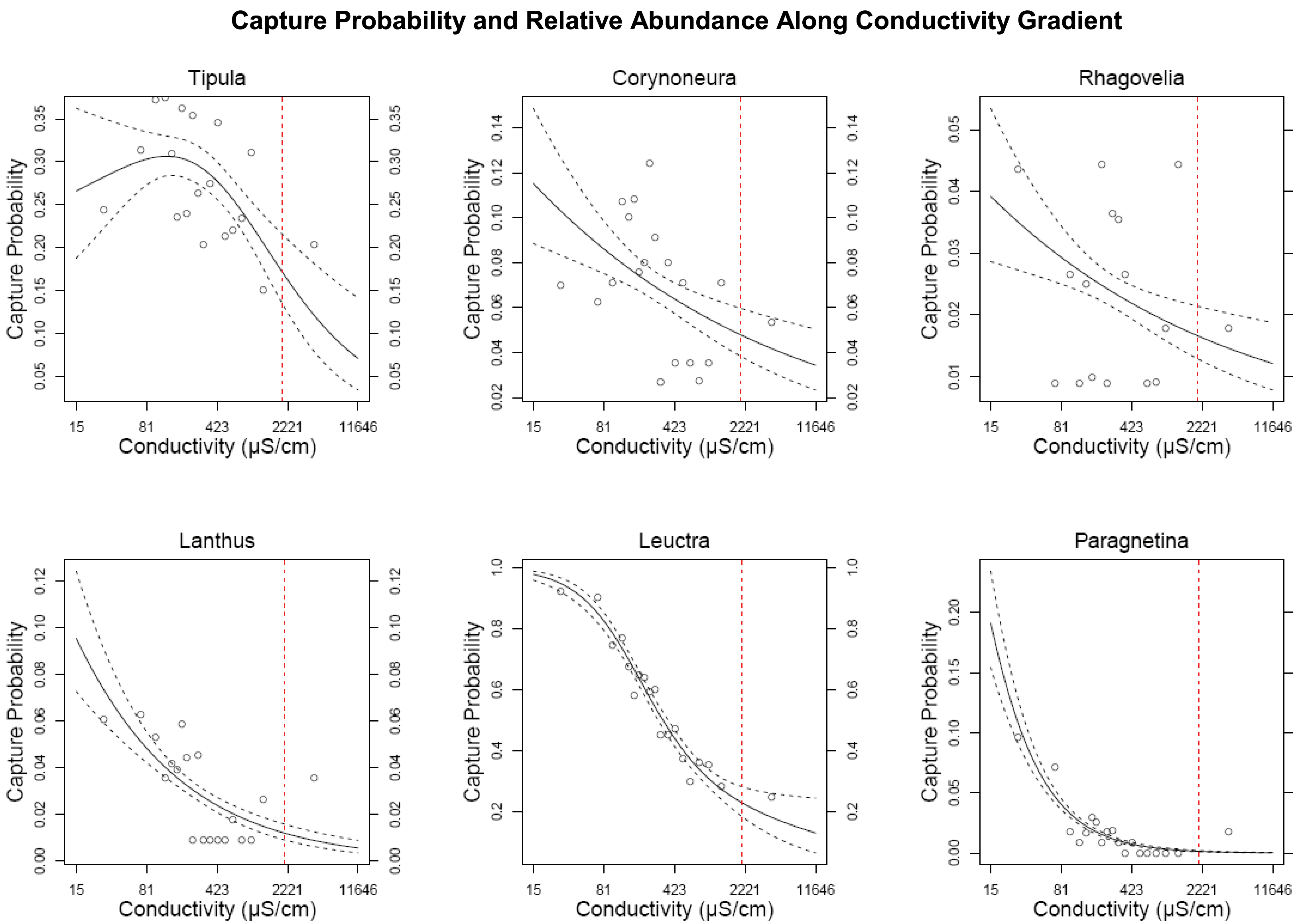

Figure D-1. Observation probabilities for each genus (continued).

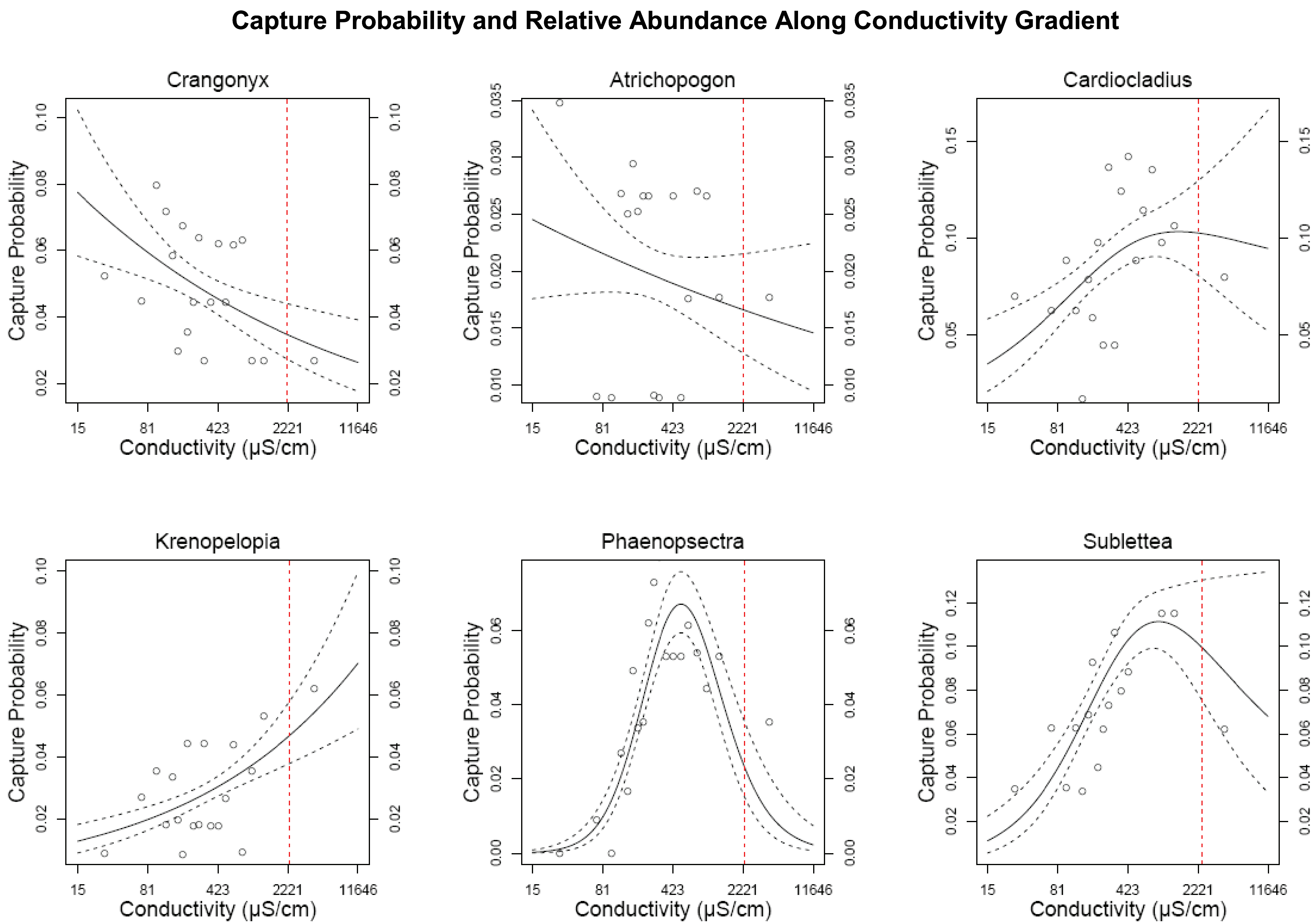

Figure D-1. Observation probabilities for each genus (continued).

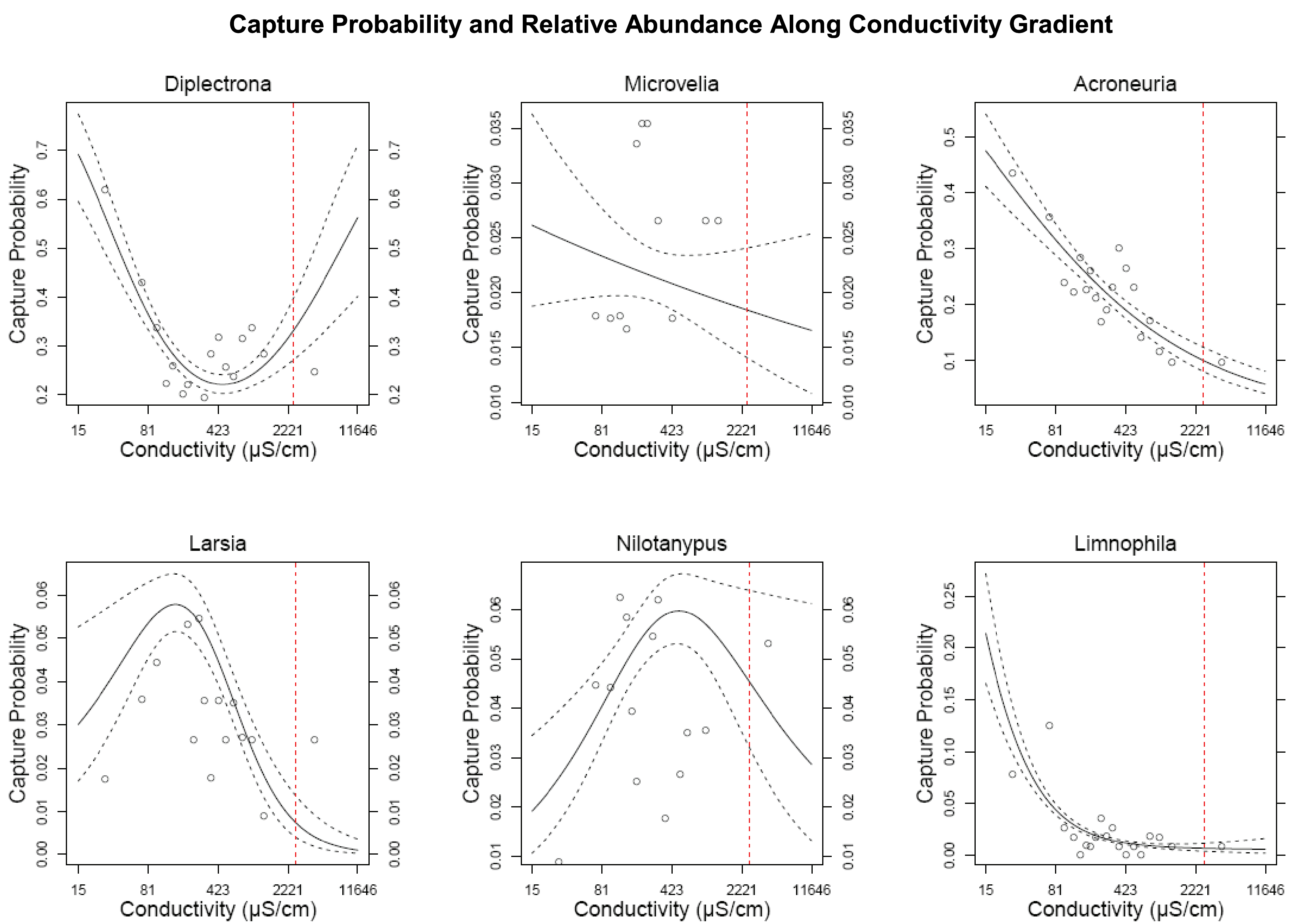

Figure D-1. Observation probabilities for each genus (continued).

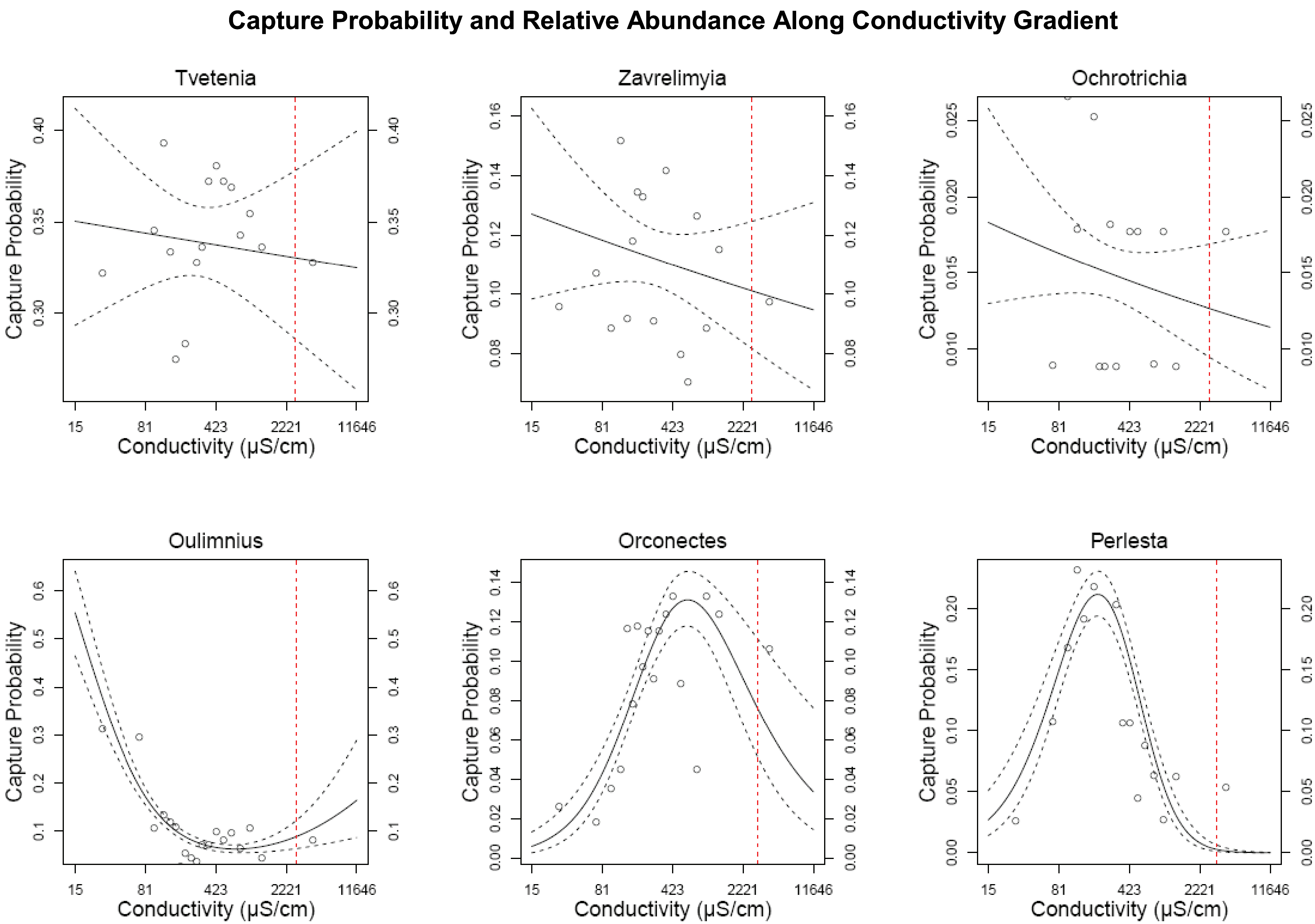

Figure D-1. Observation probabilities for each genus (continued).

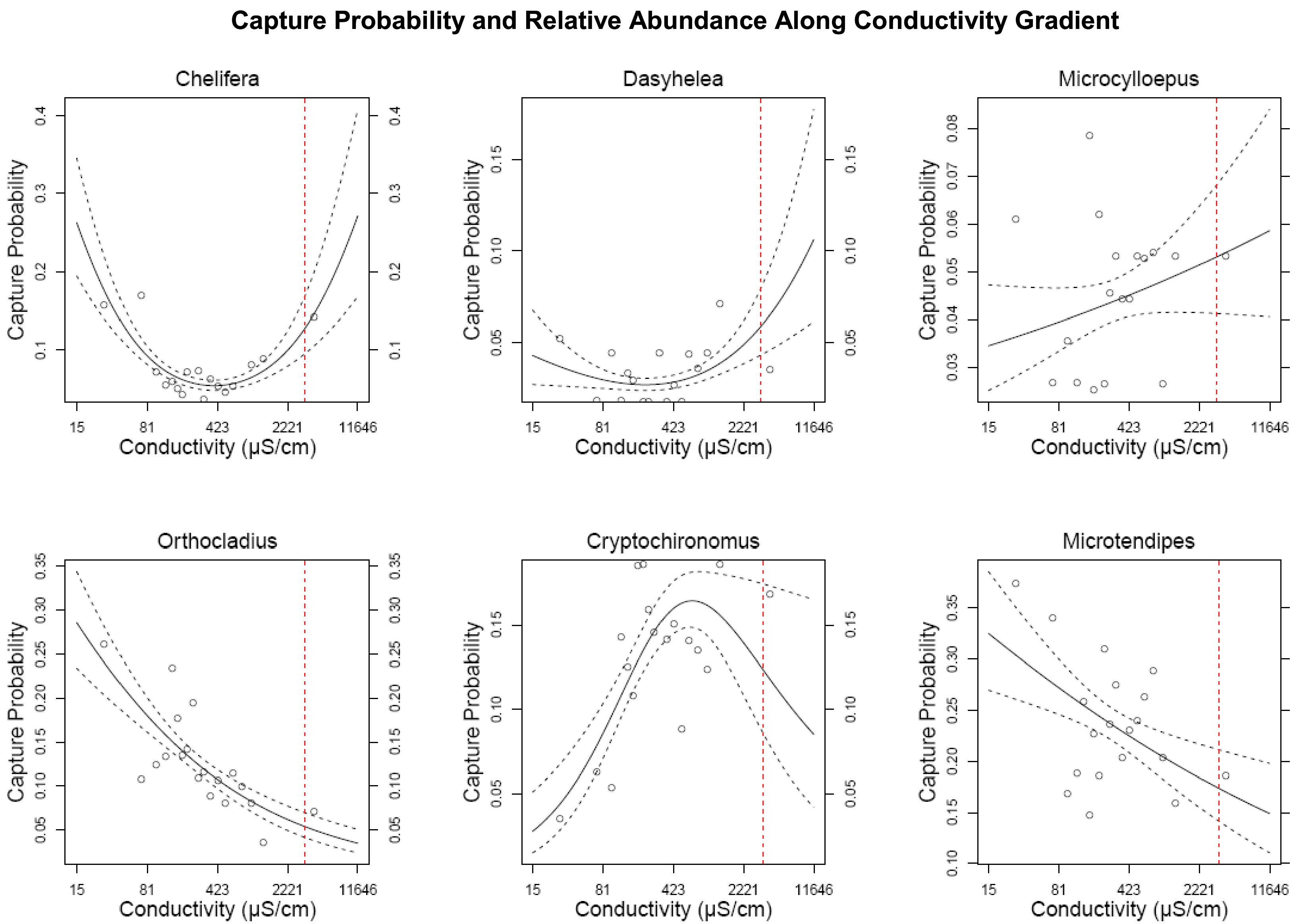

Figure D-1. Observation probabilities for each genus (continued).

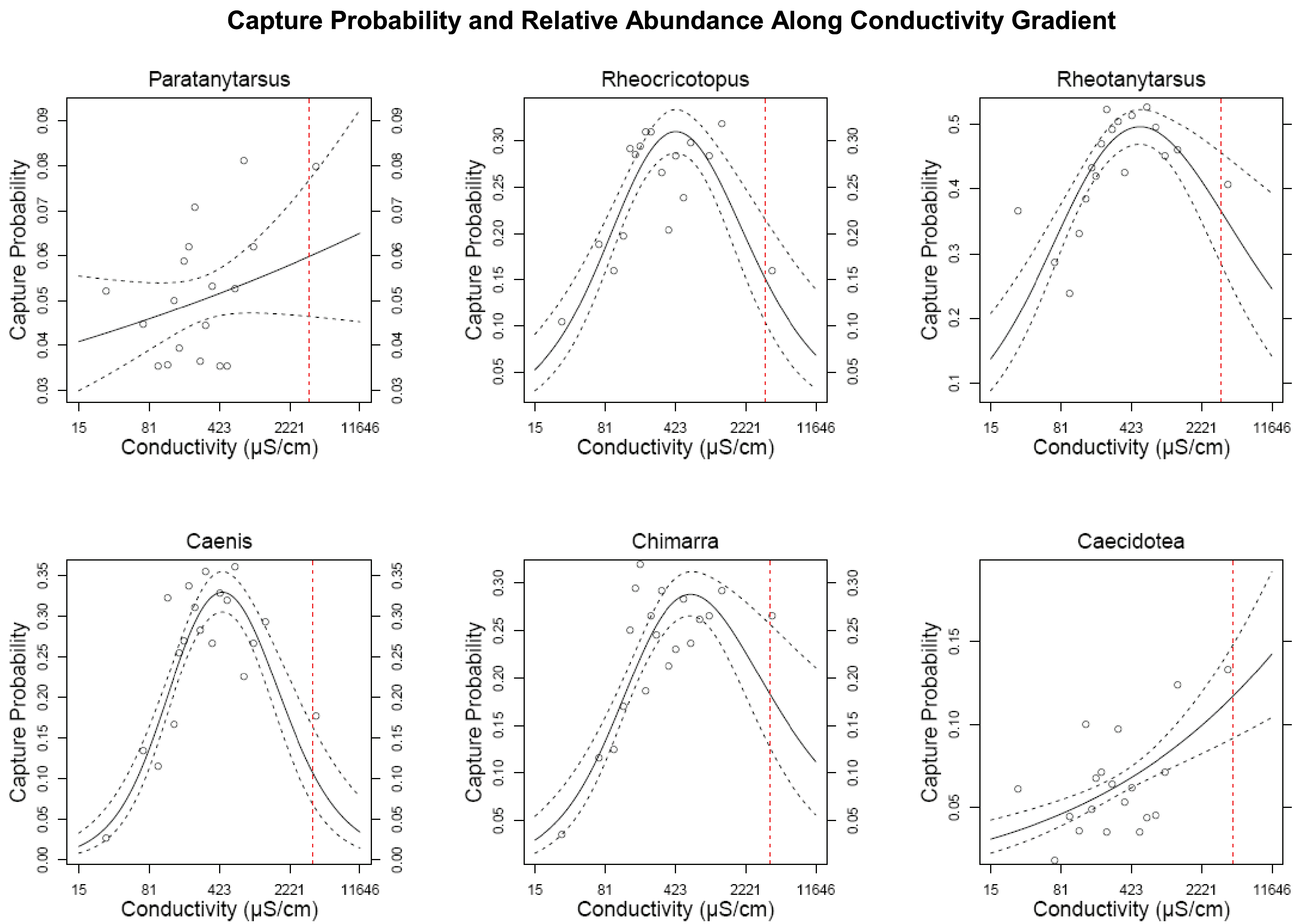

Figure D-1. Observation probabilities for each genus (continued).

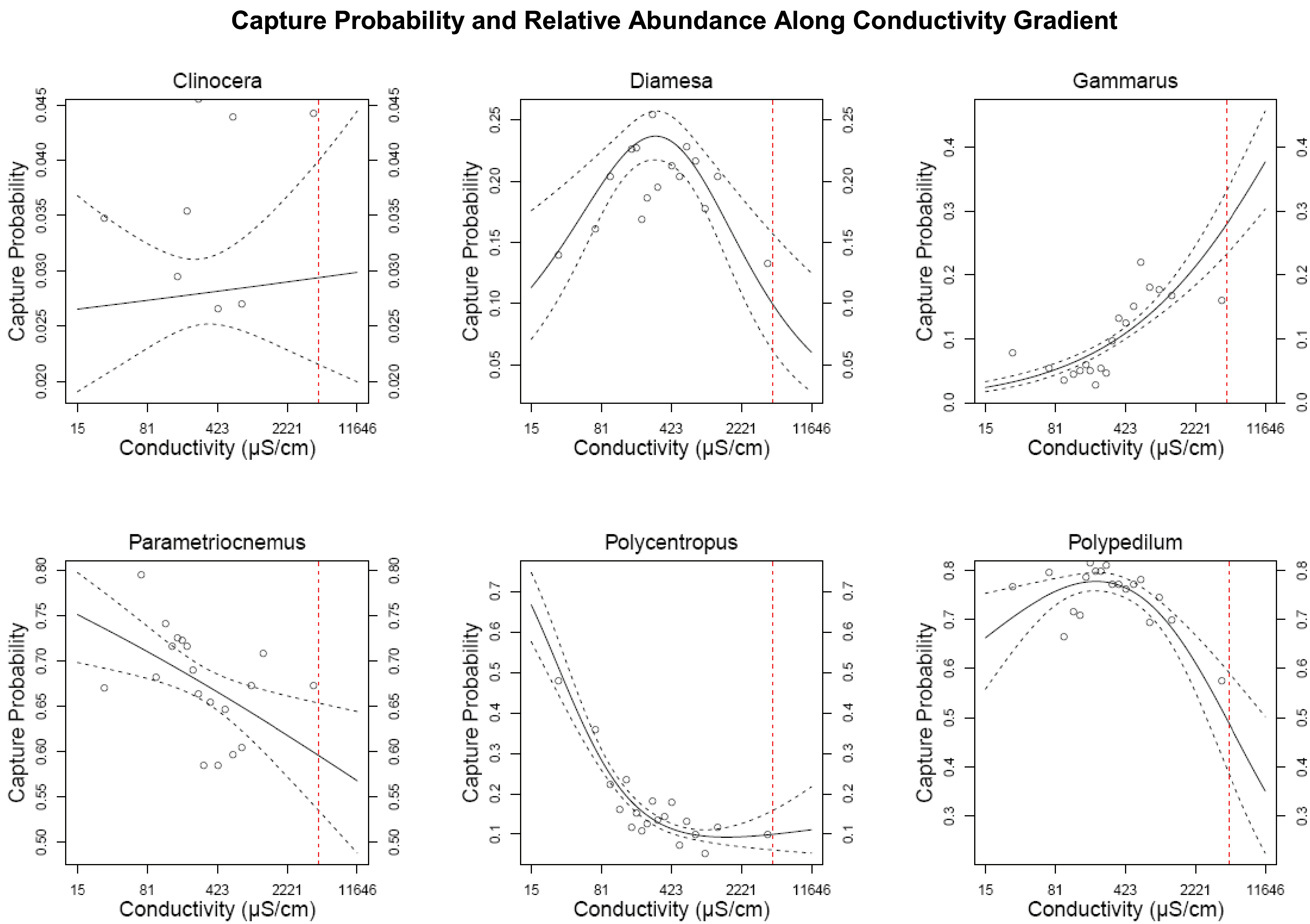

Figure D-1. Observation probabilities for each genus (continued).

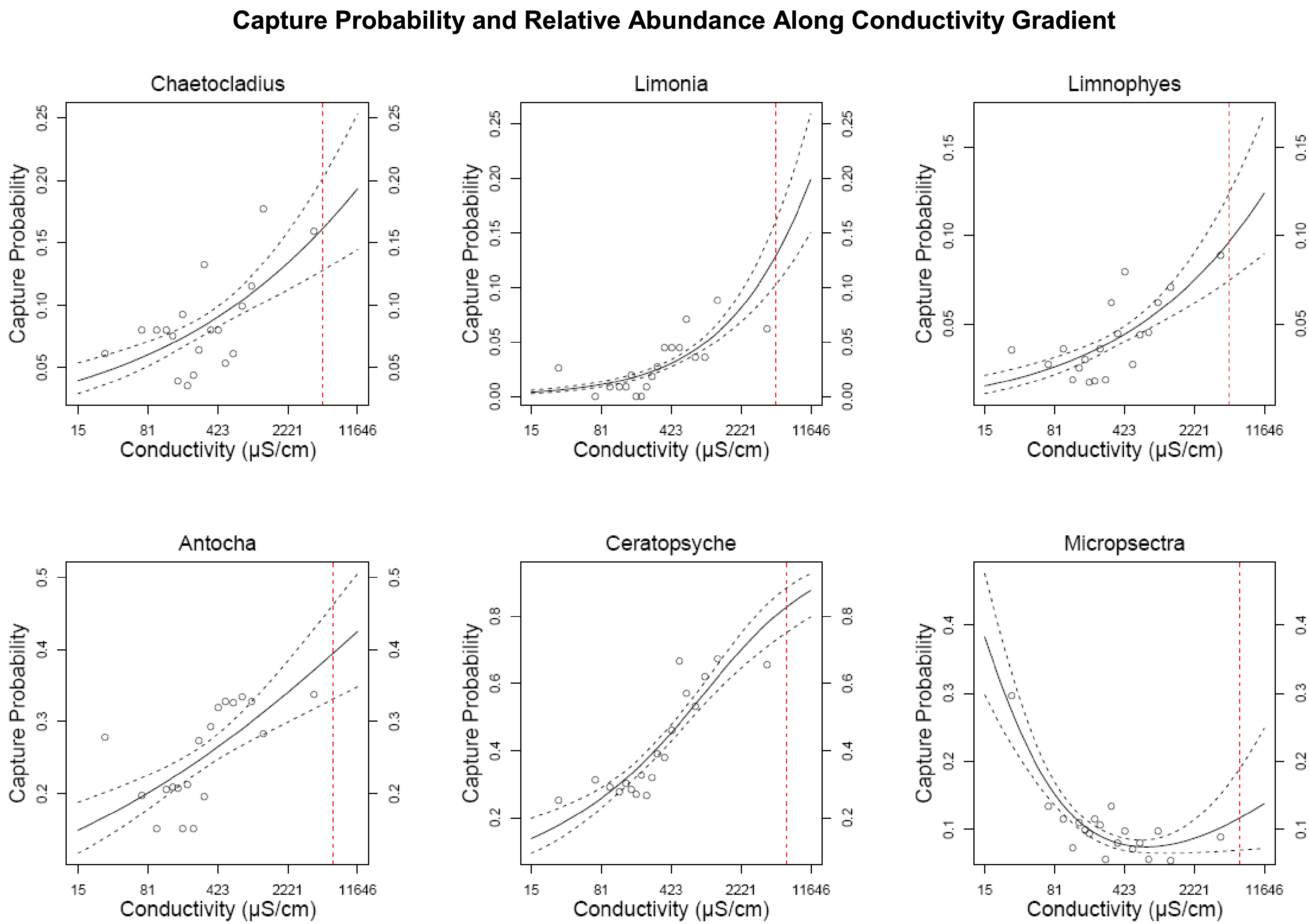

Figure D-1. Observation probabilities for each genus (continued).

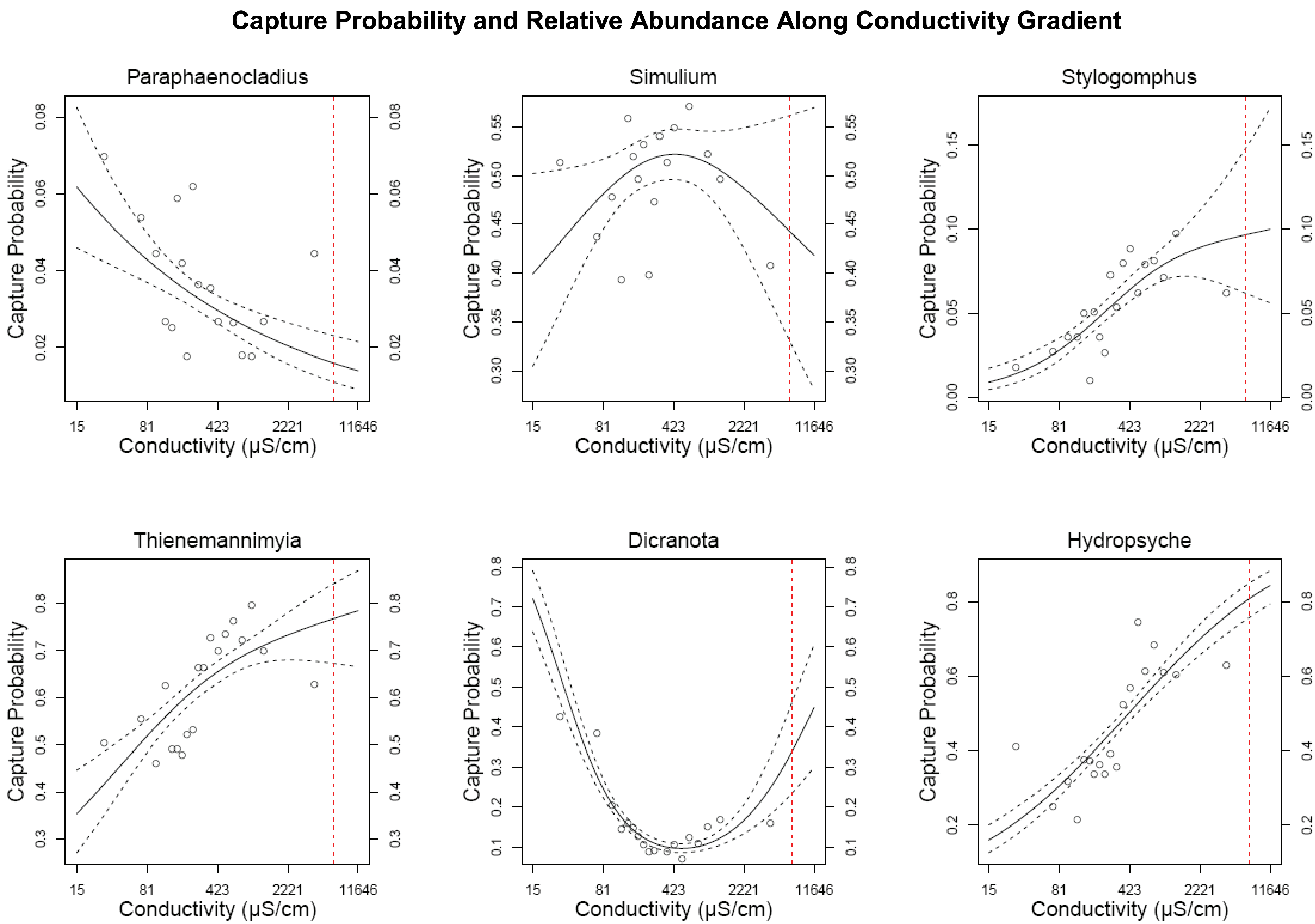

Figure D-1. Observation probabilities for each genus (continued).

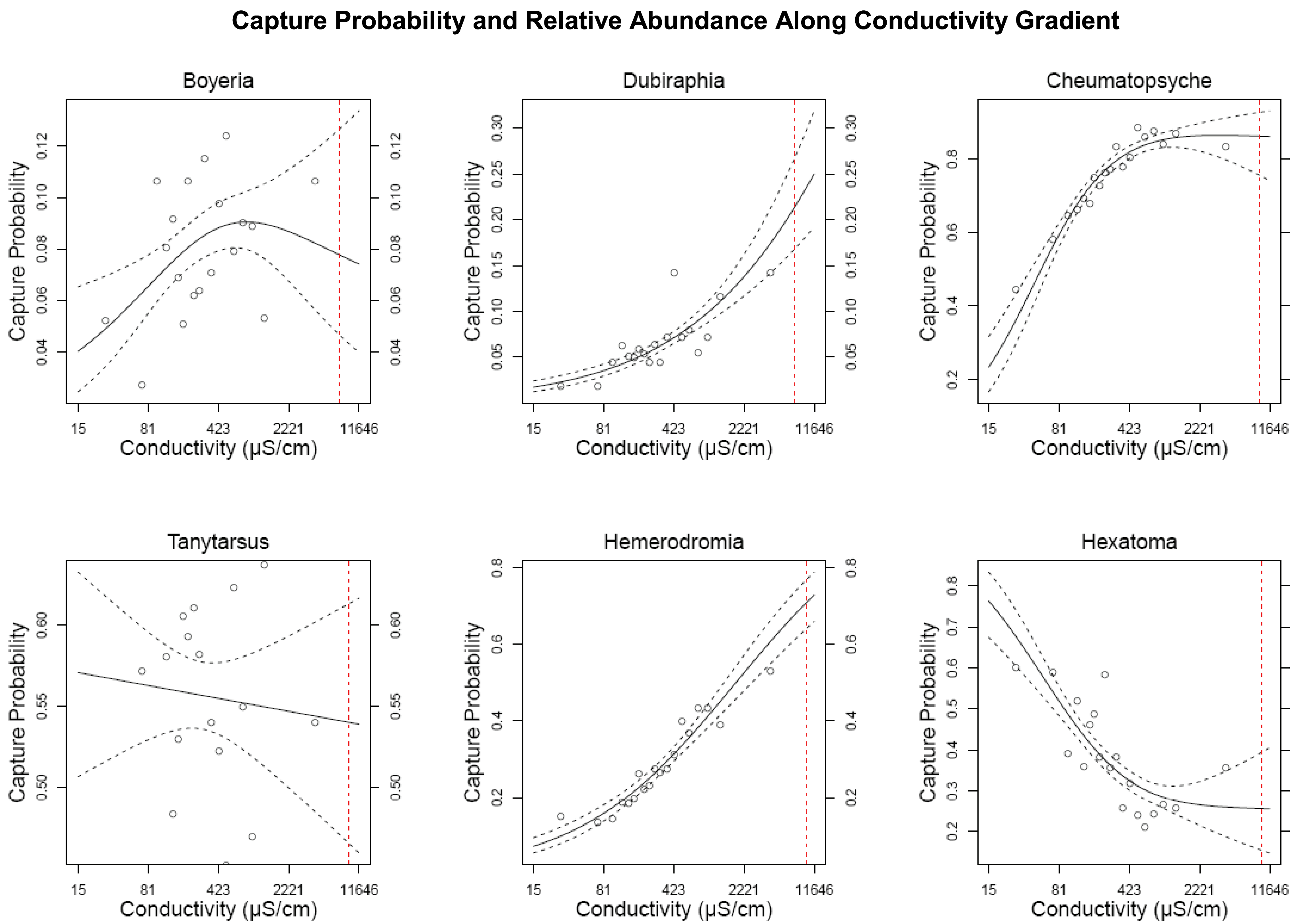

Figure D-1. Observation probabilities for each genus (continued).

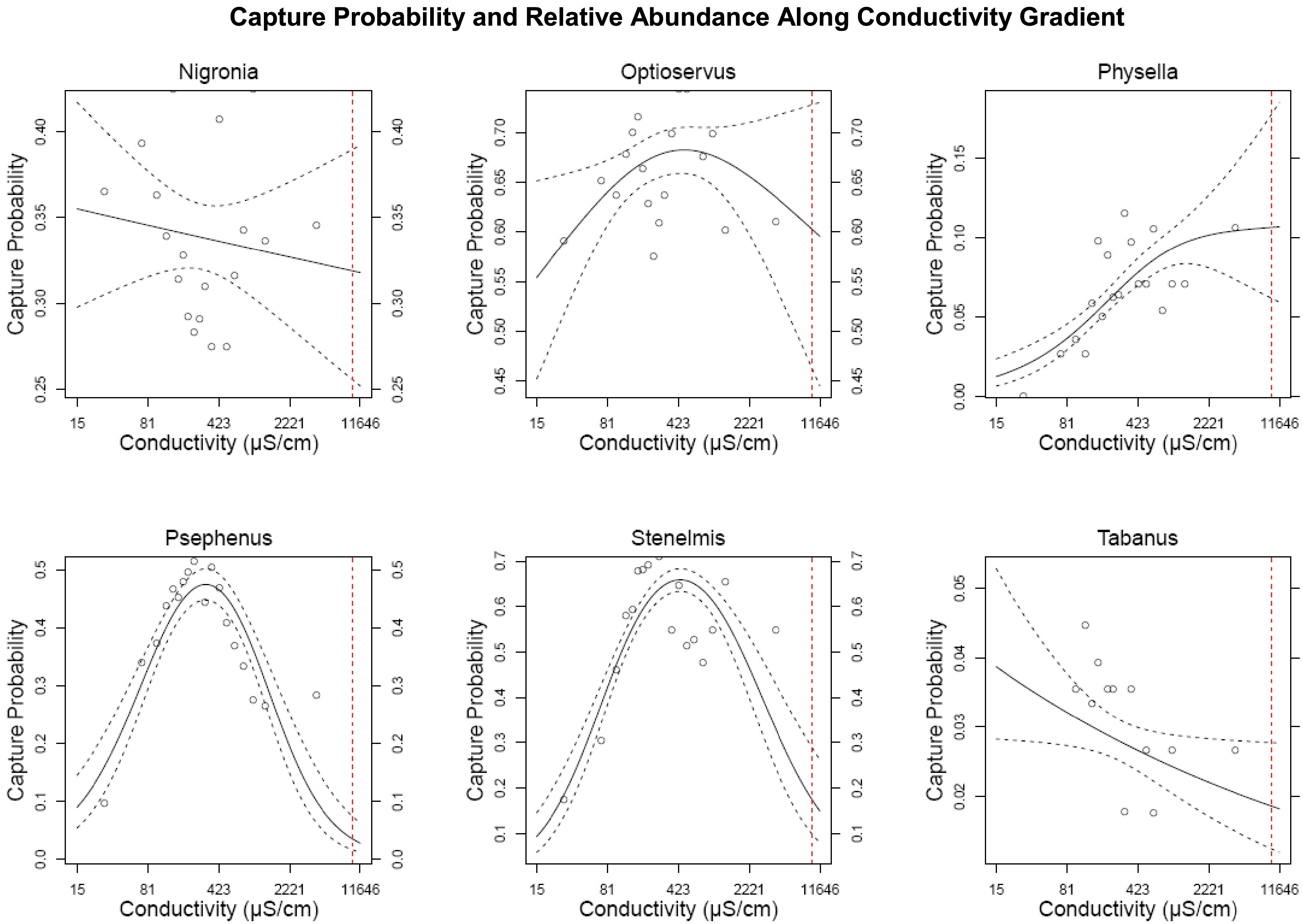

Figure D-1. Observation probabilities for each genus (continued).

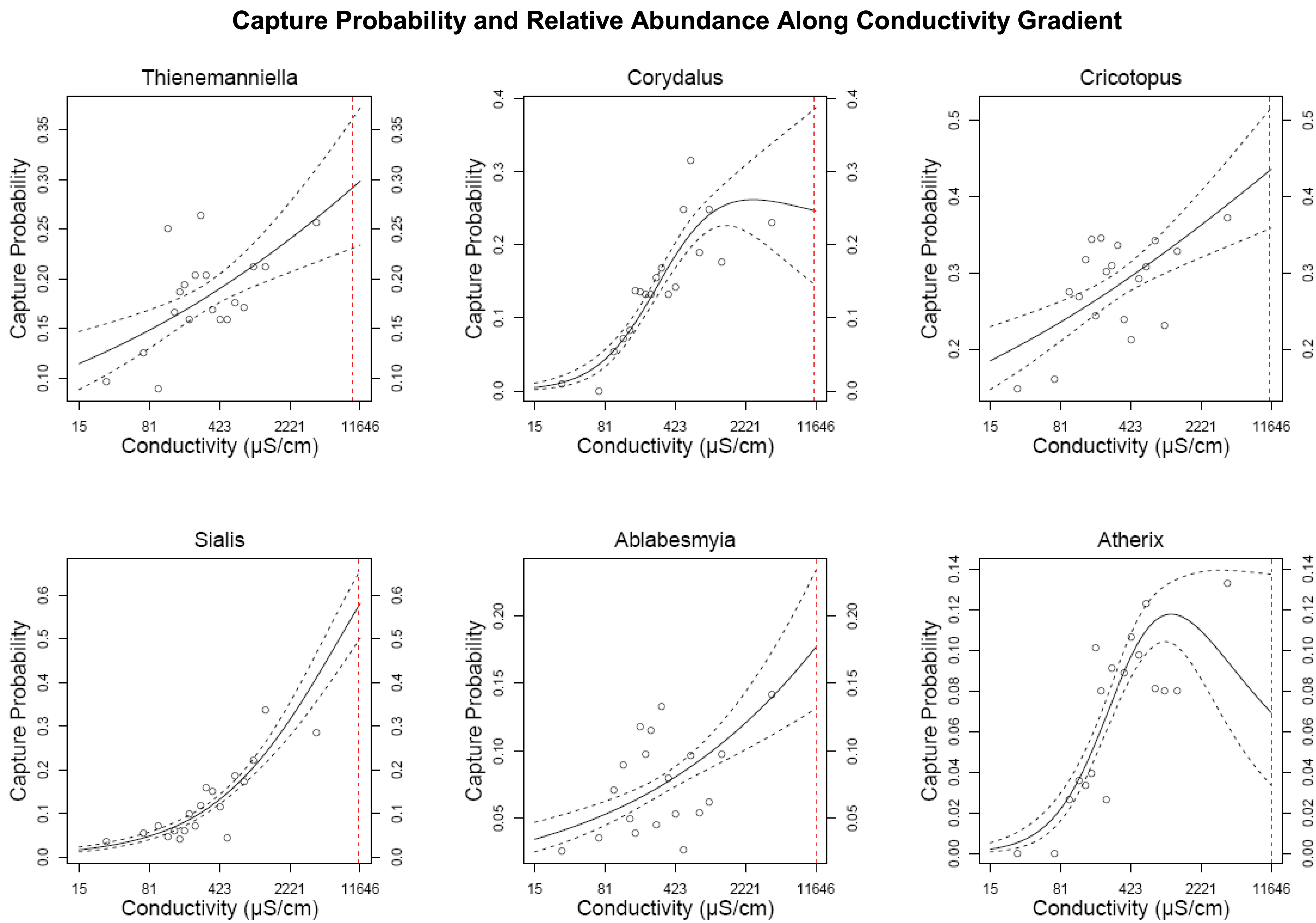

Figure D-1. Observation probabilities for each genus (continued).

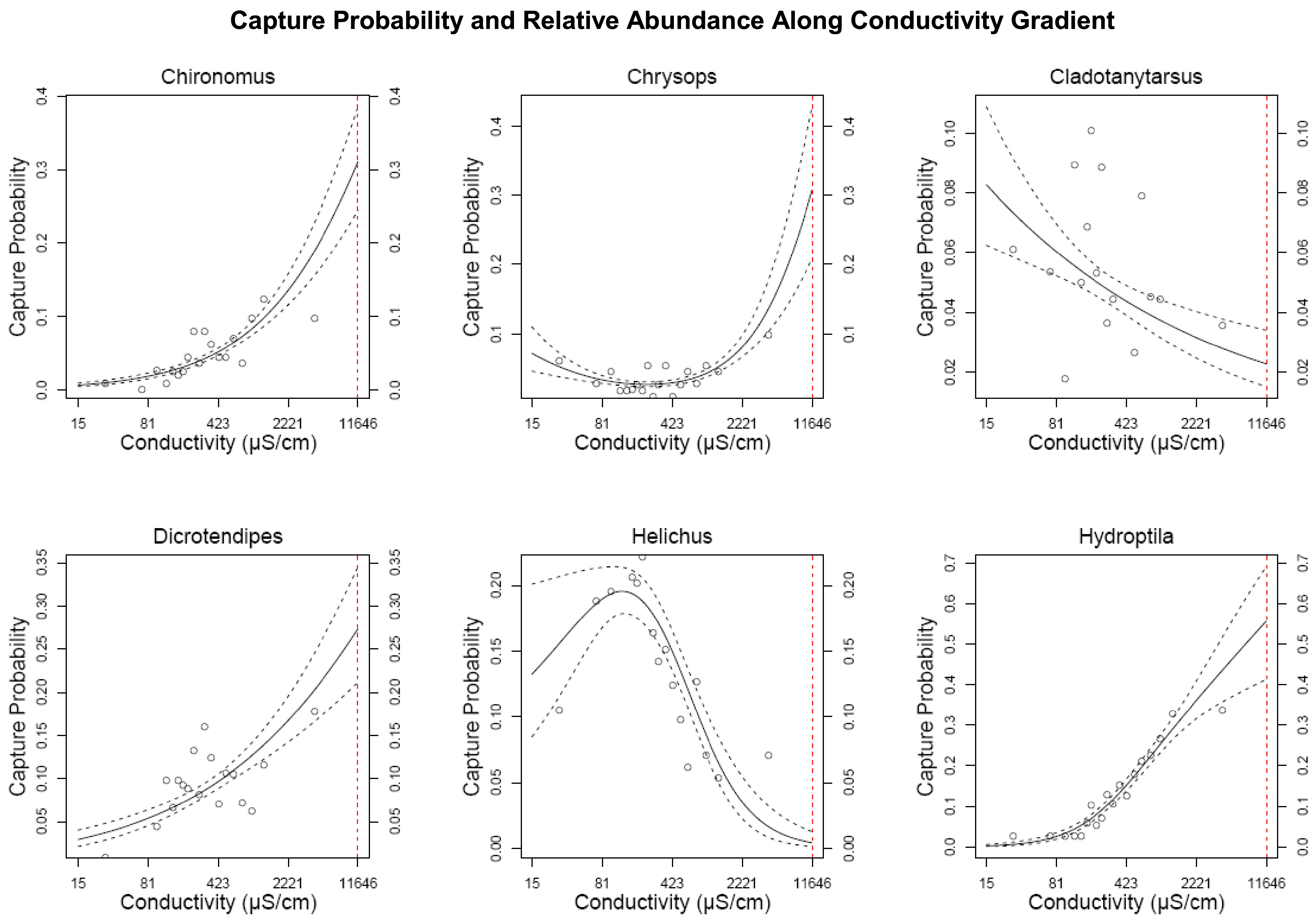

Figure D-1. Observation probabilities for each genus (continued).

Capture Probability and Relative Abundance Along Conductivity Gradient

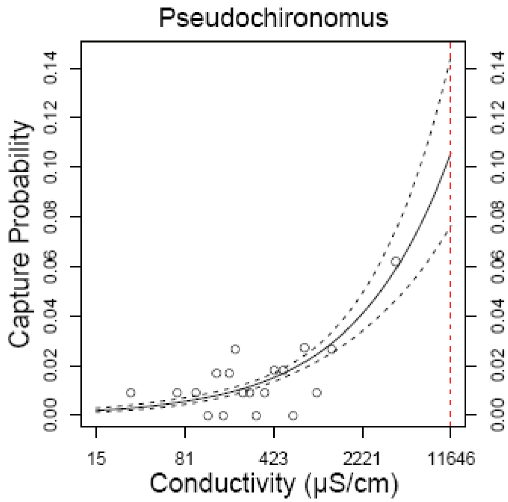

Figure D-1. Observation probabilities for each genus (continued).

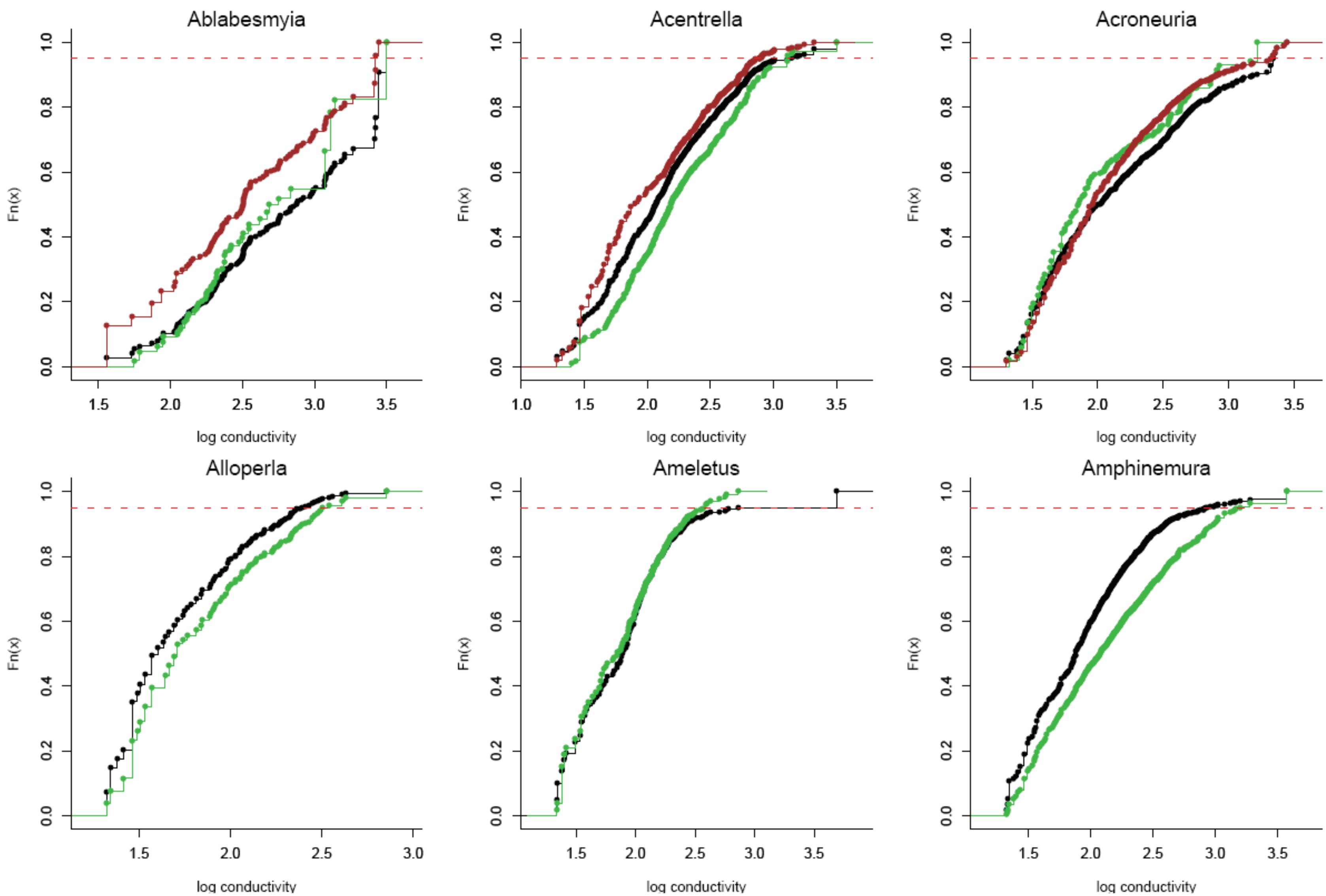

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus. Sites with pH <6 excluded.

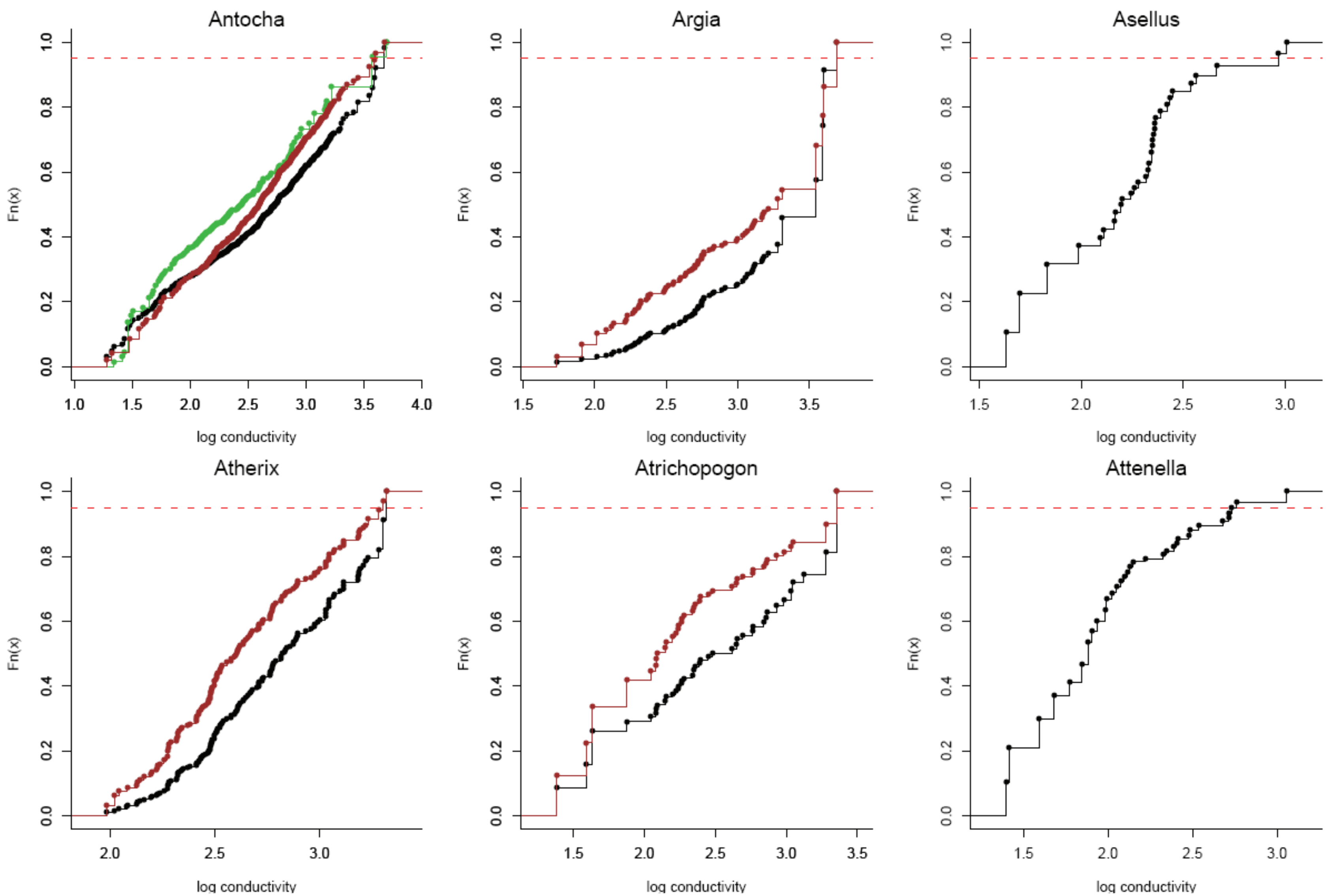

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

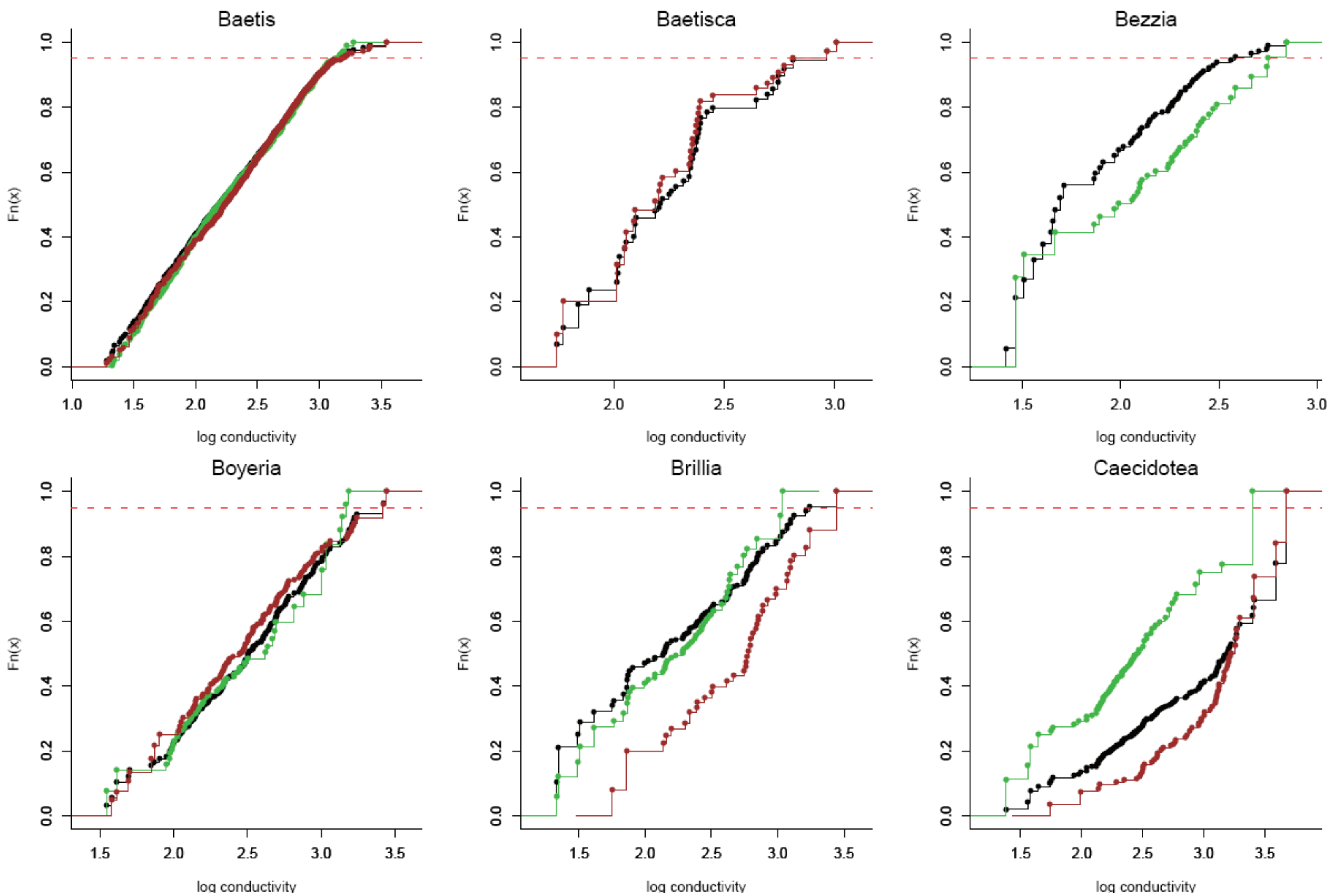

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

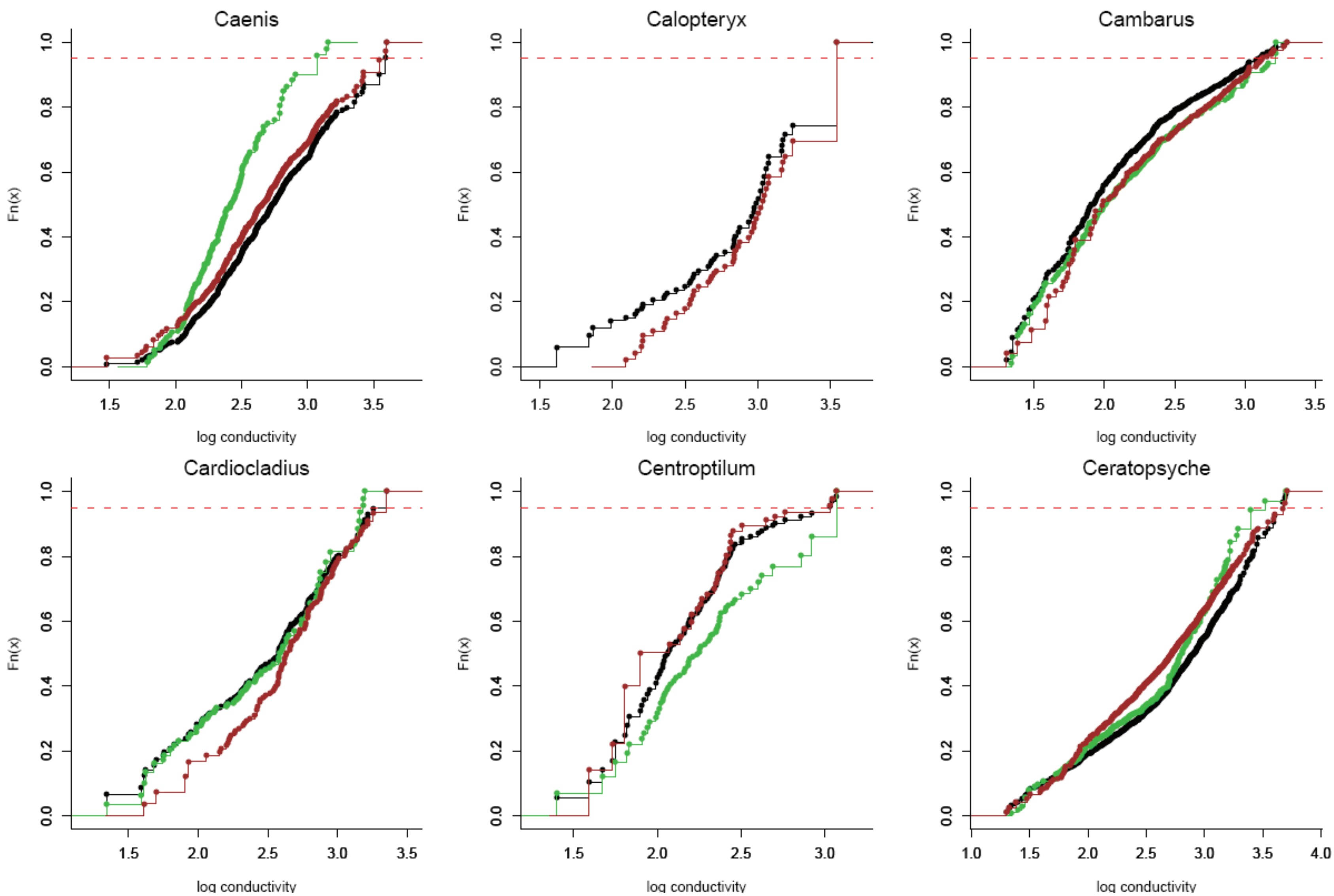

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

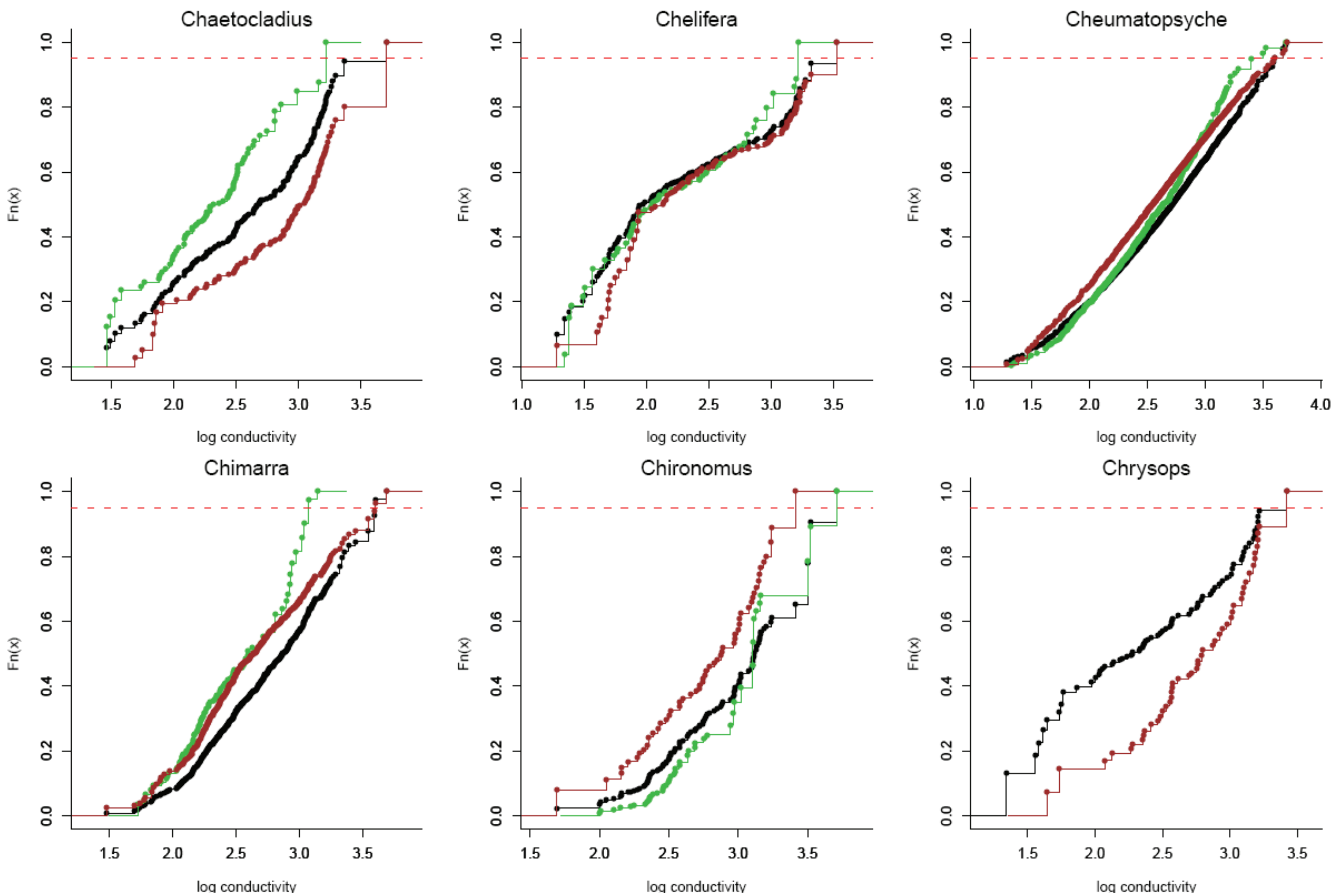

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

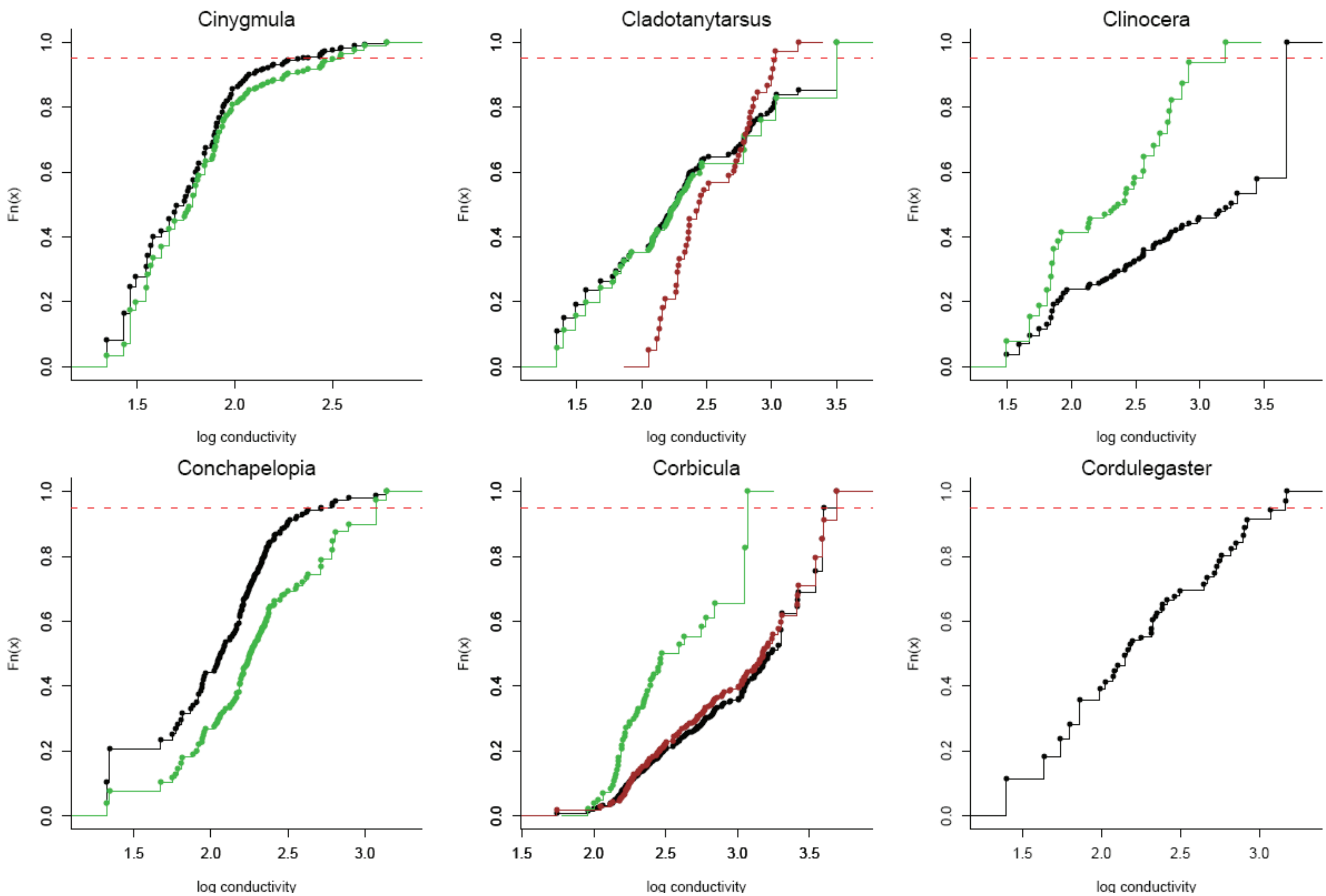

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

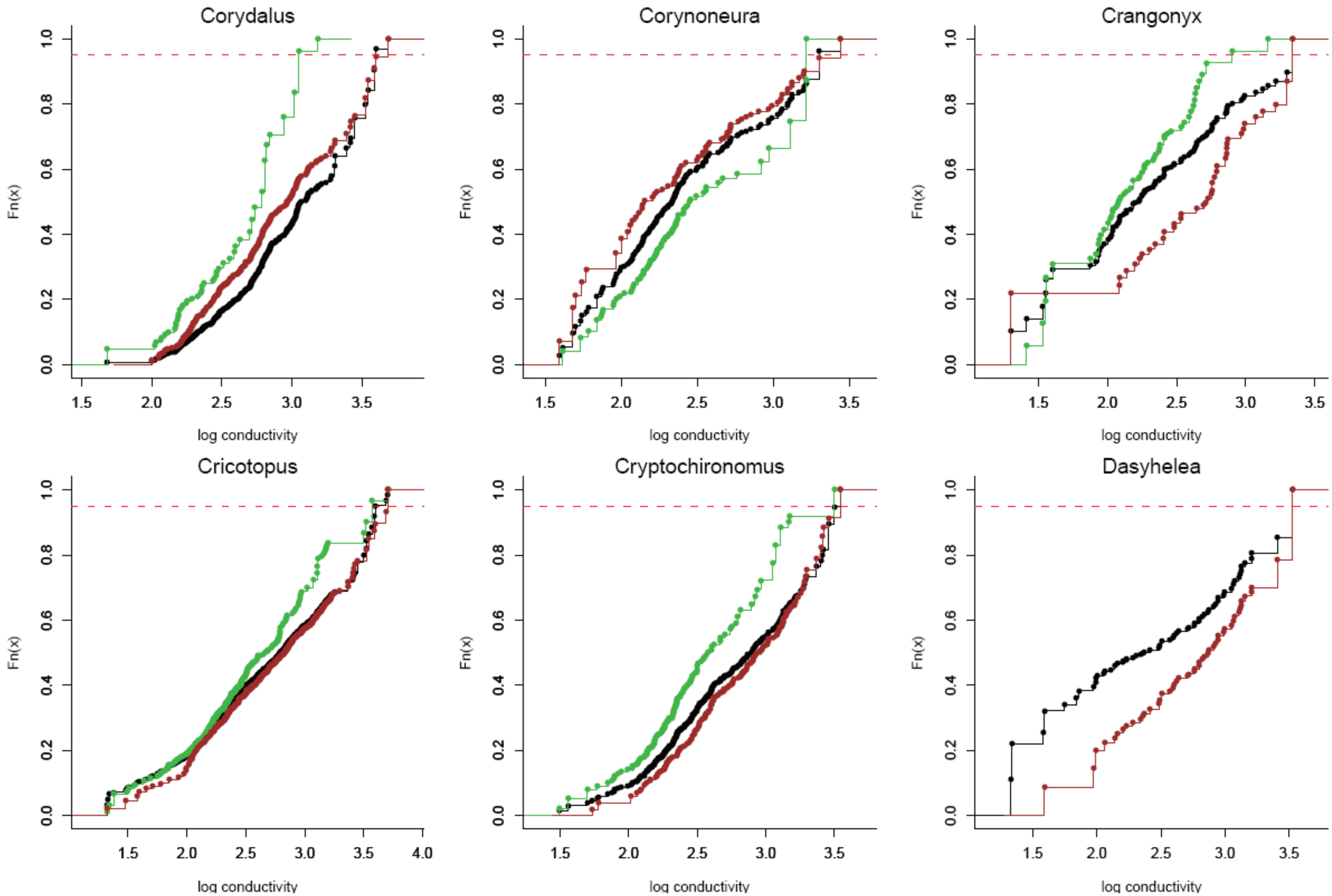

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

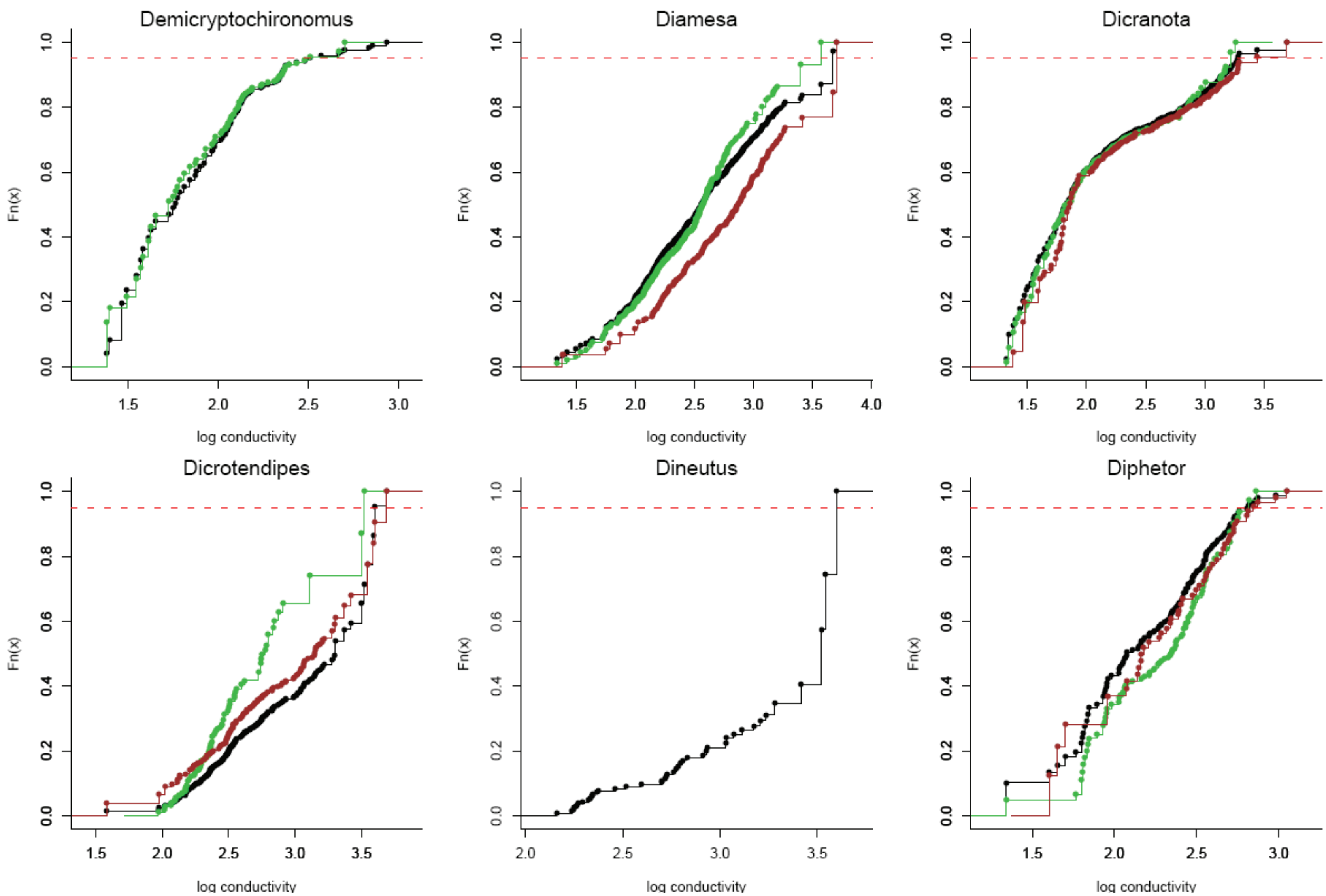

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

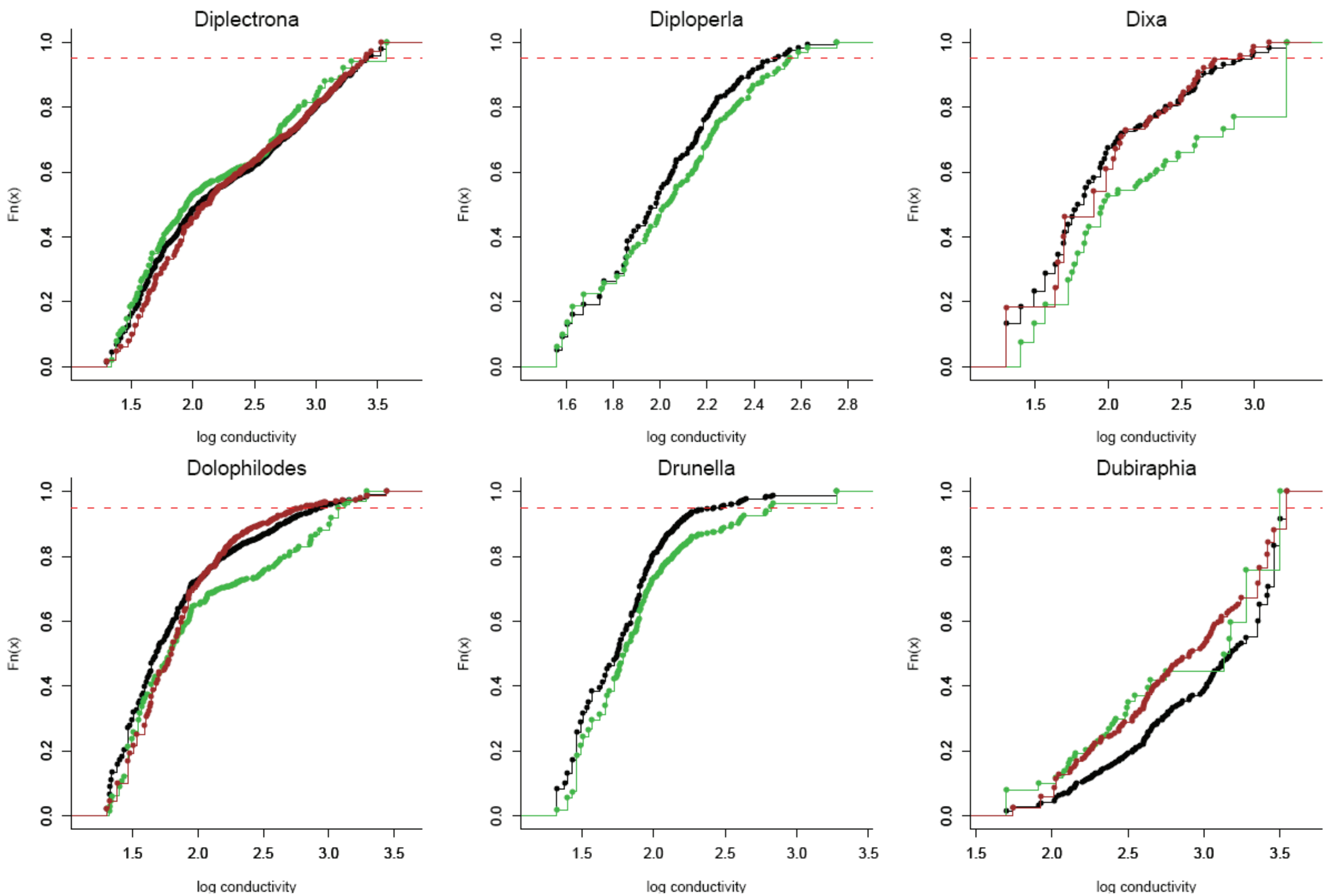

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

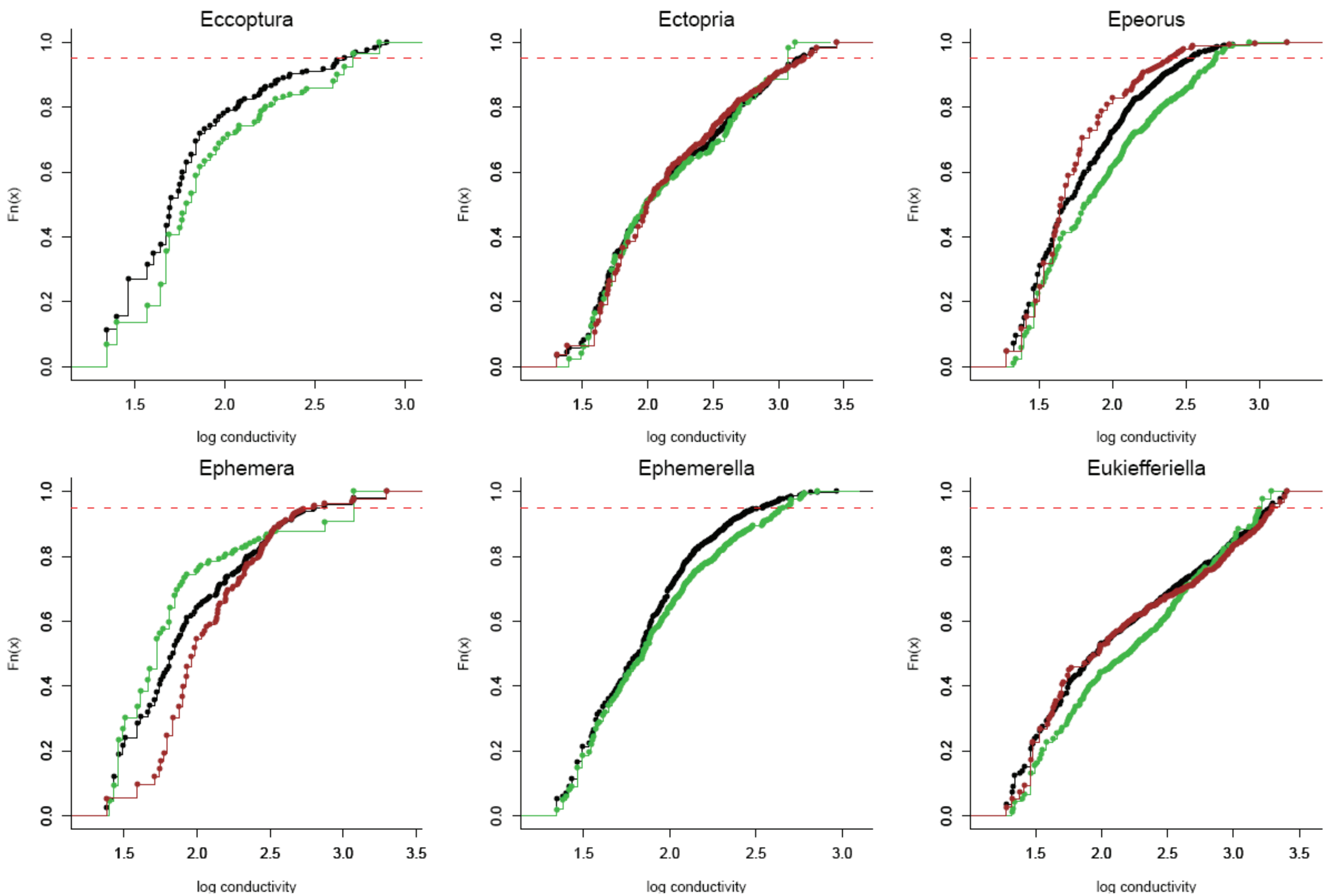

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

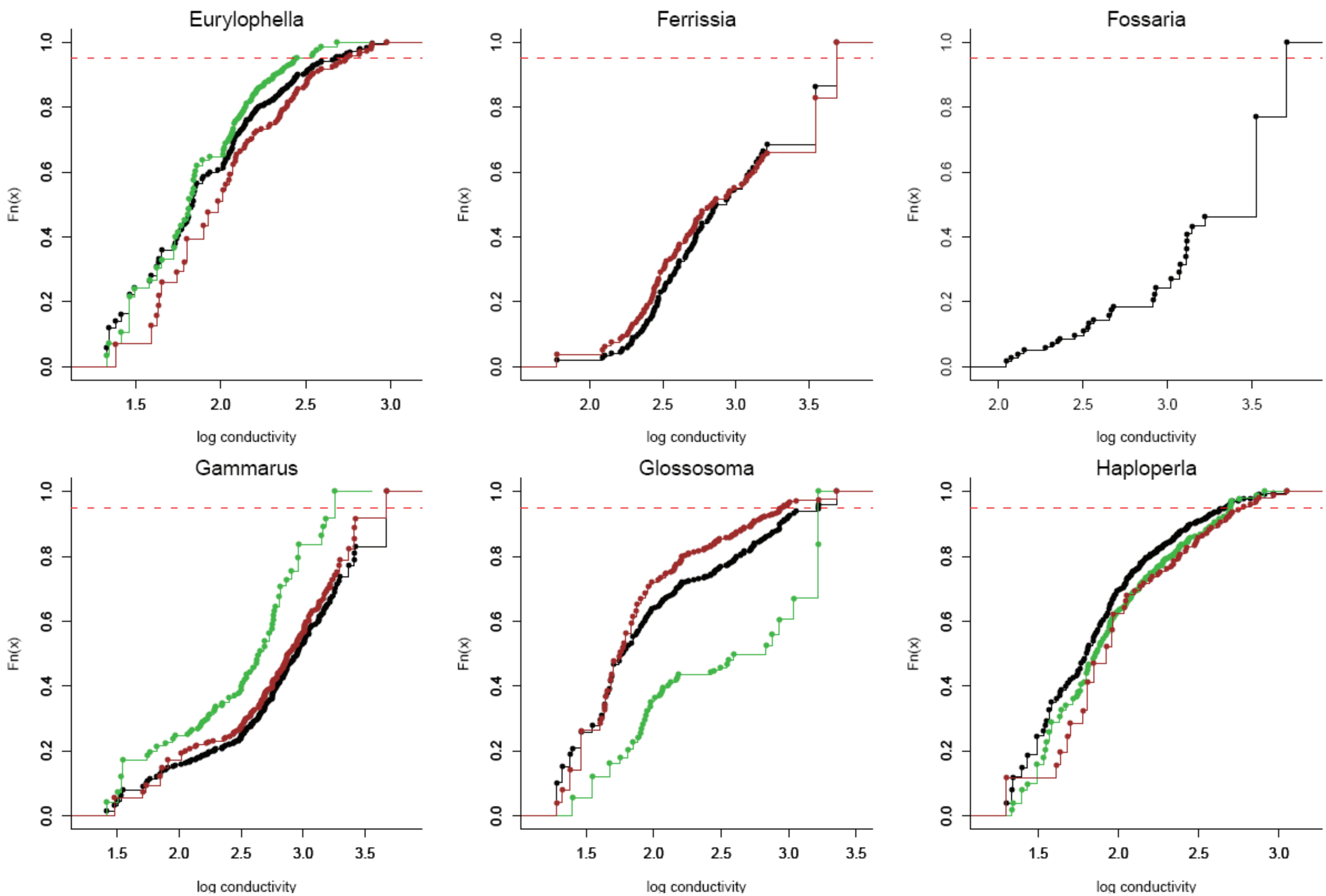

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

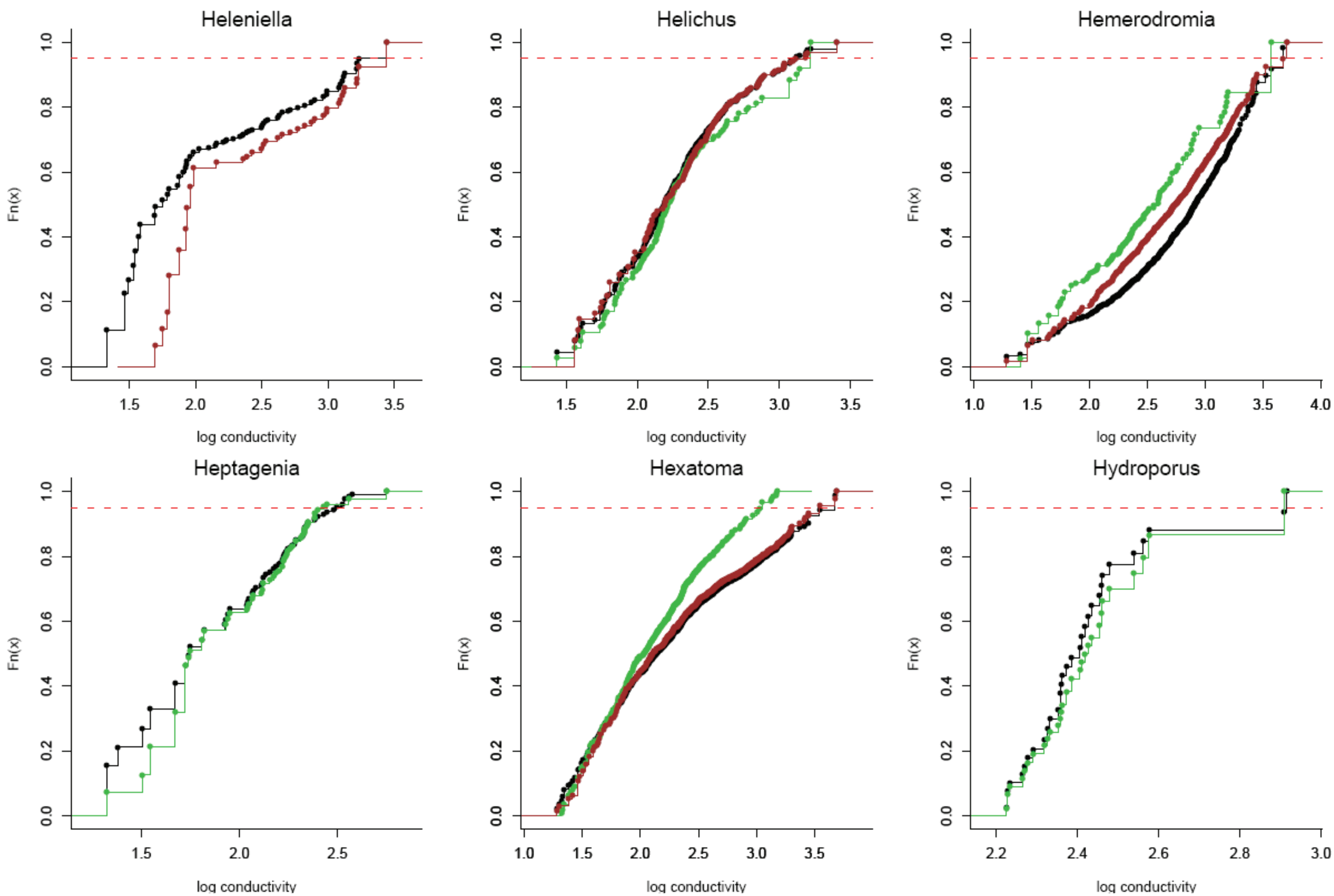

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

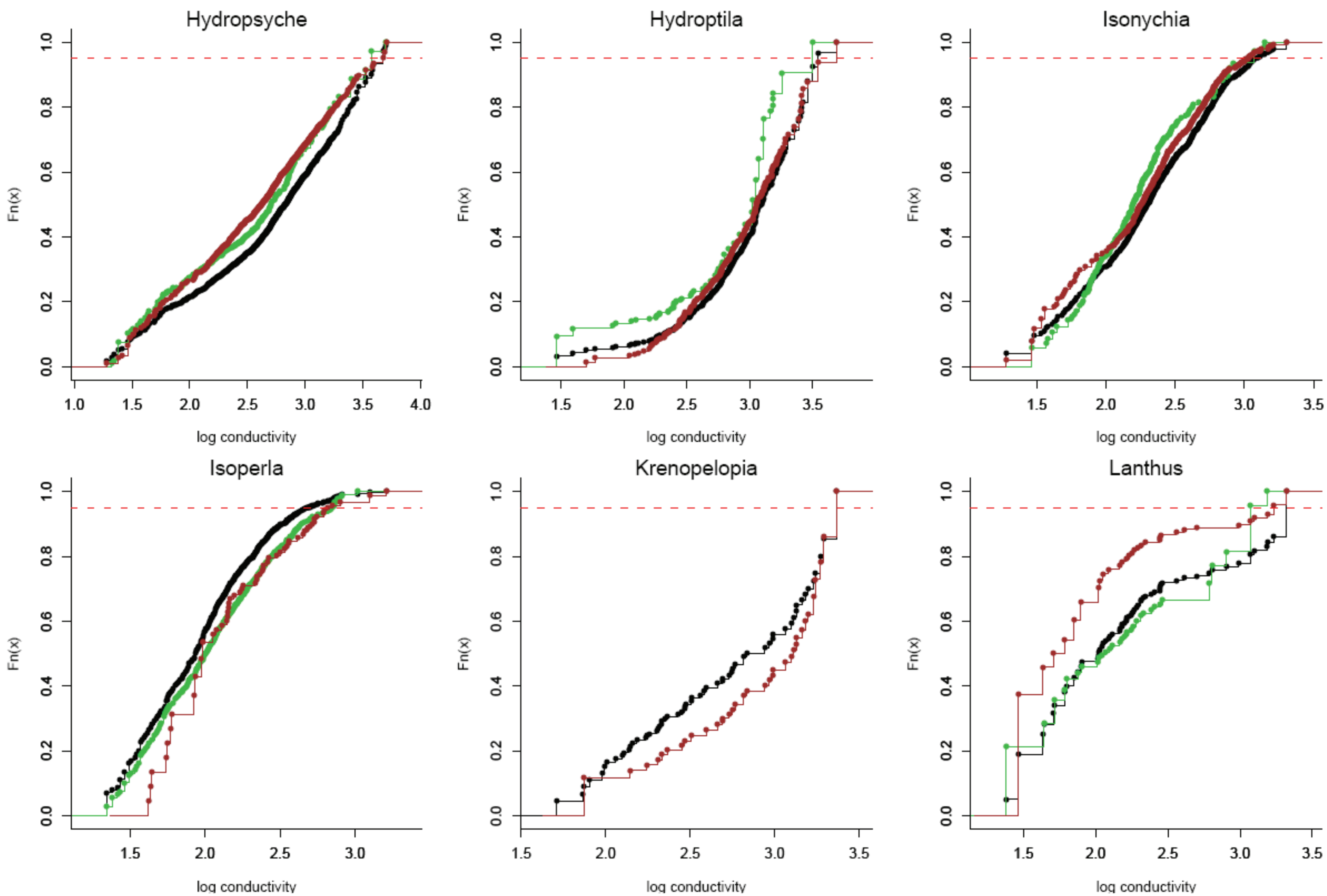

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

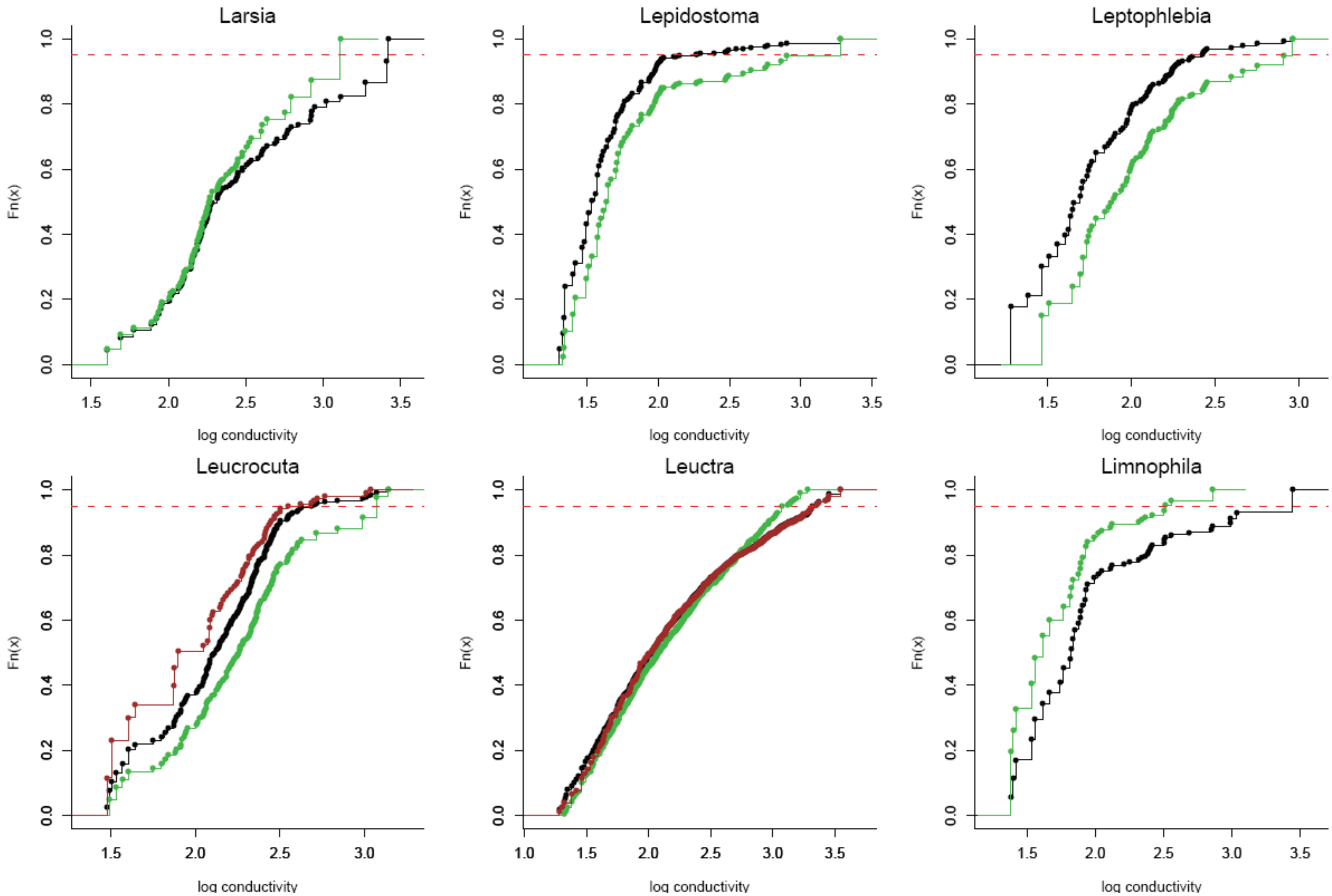

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

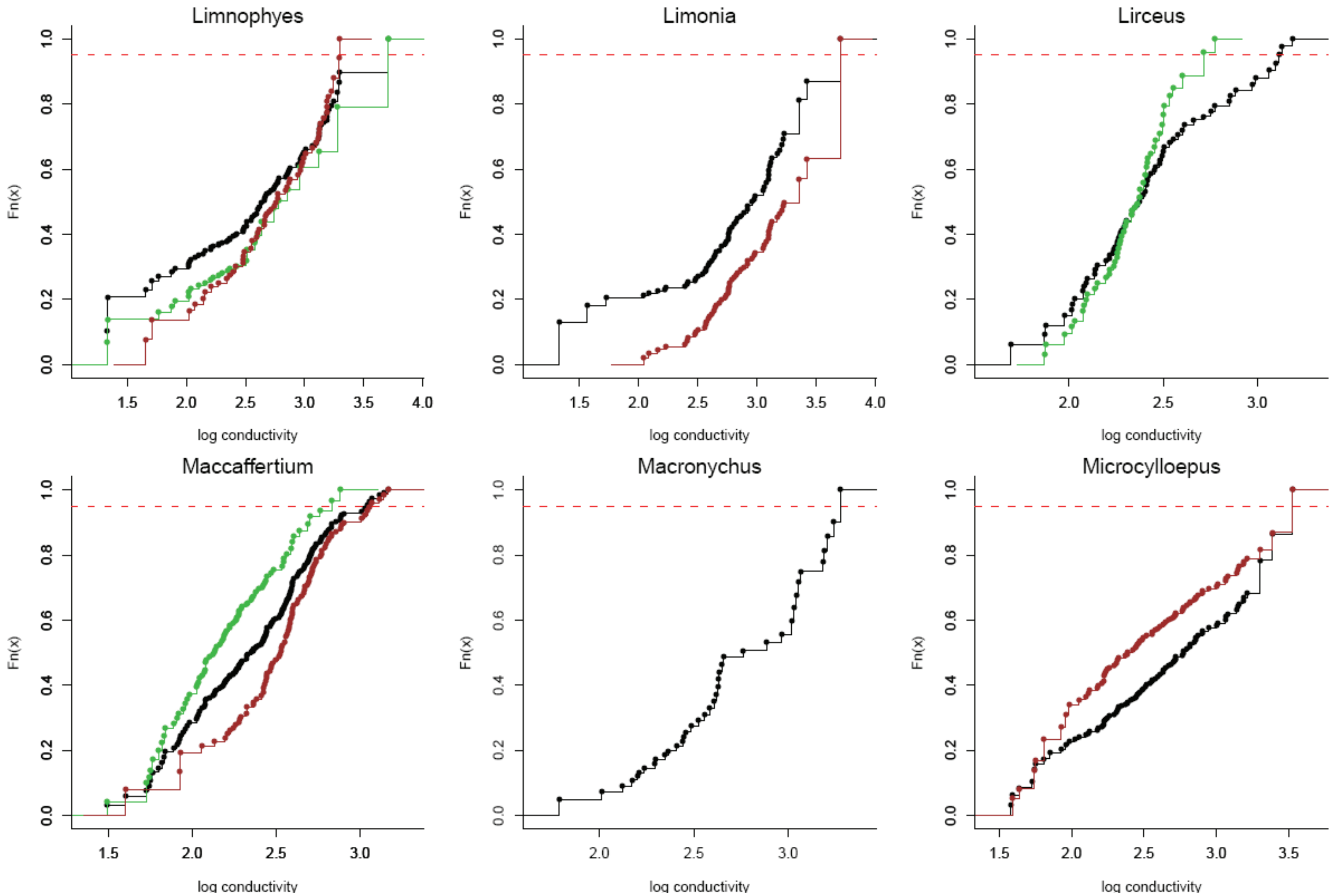

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

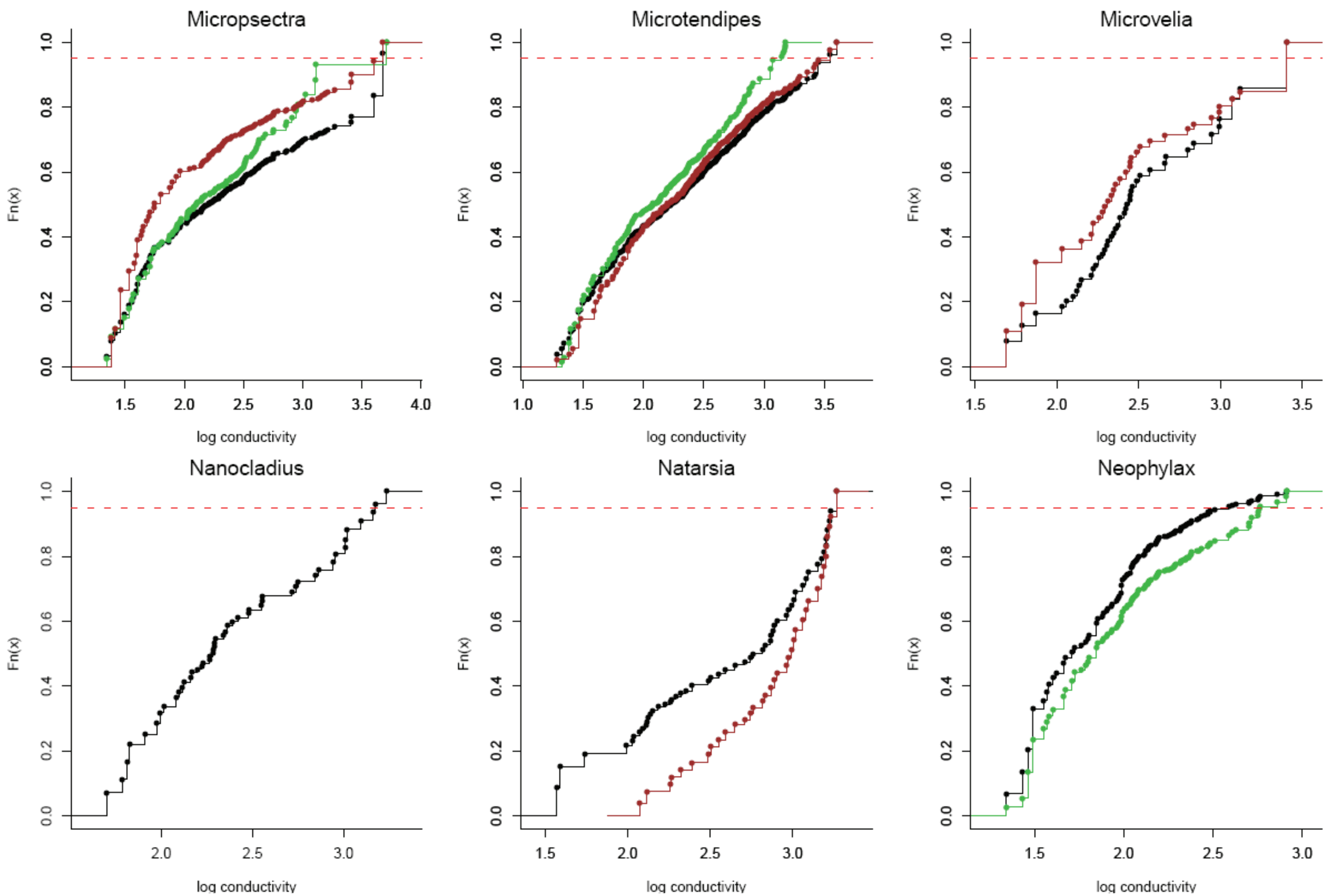

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

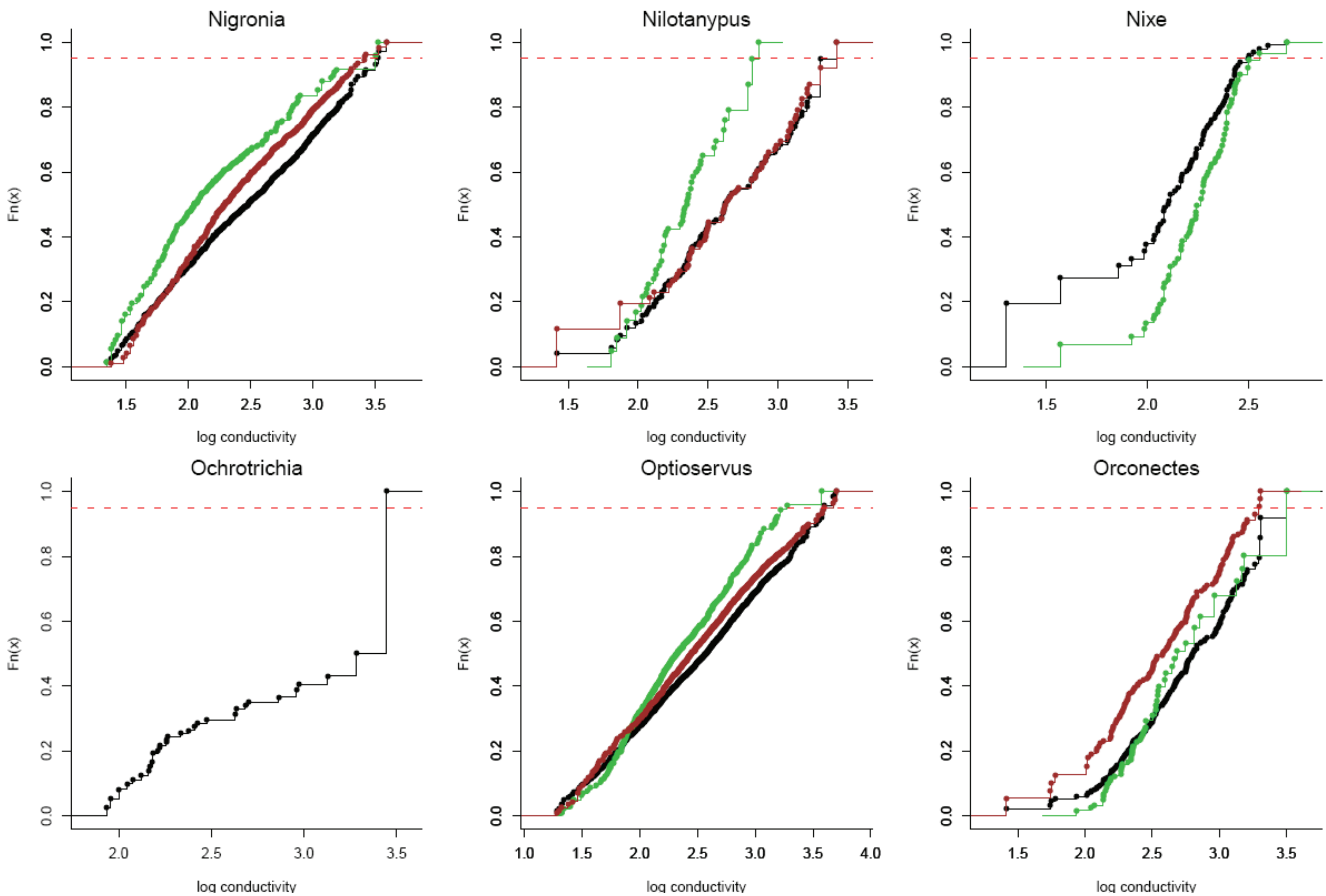

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

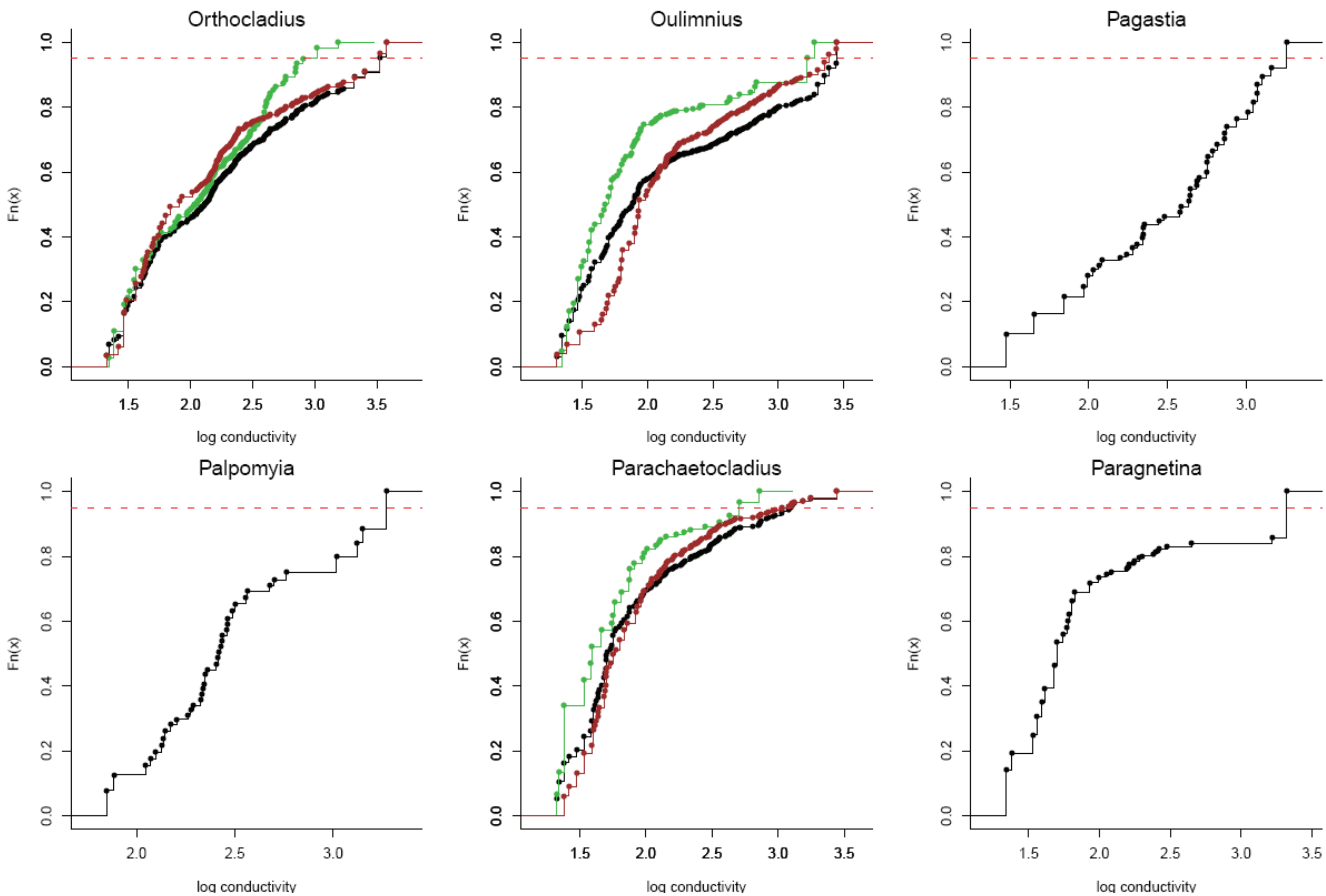

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

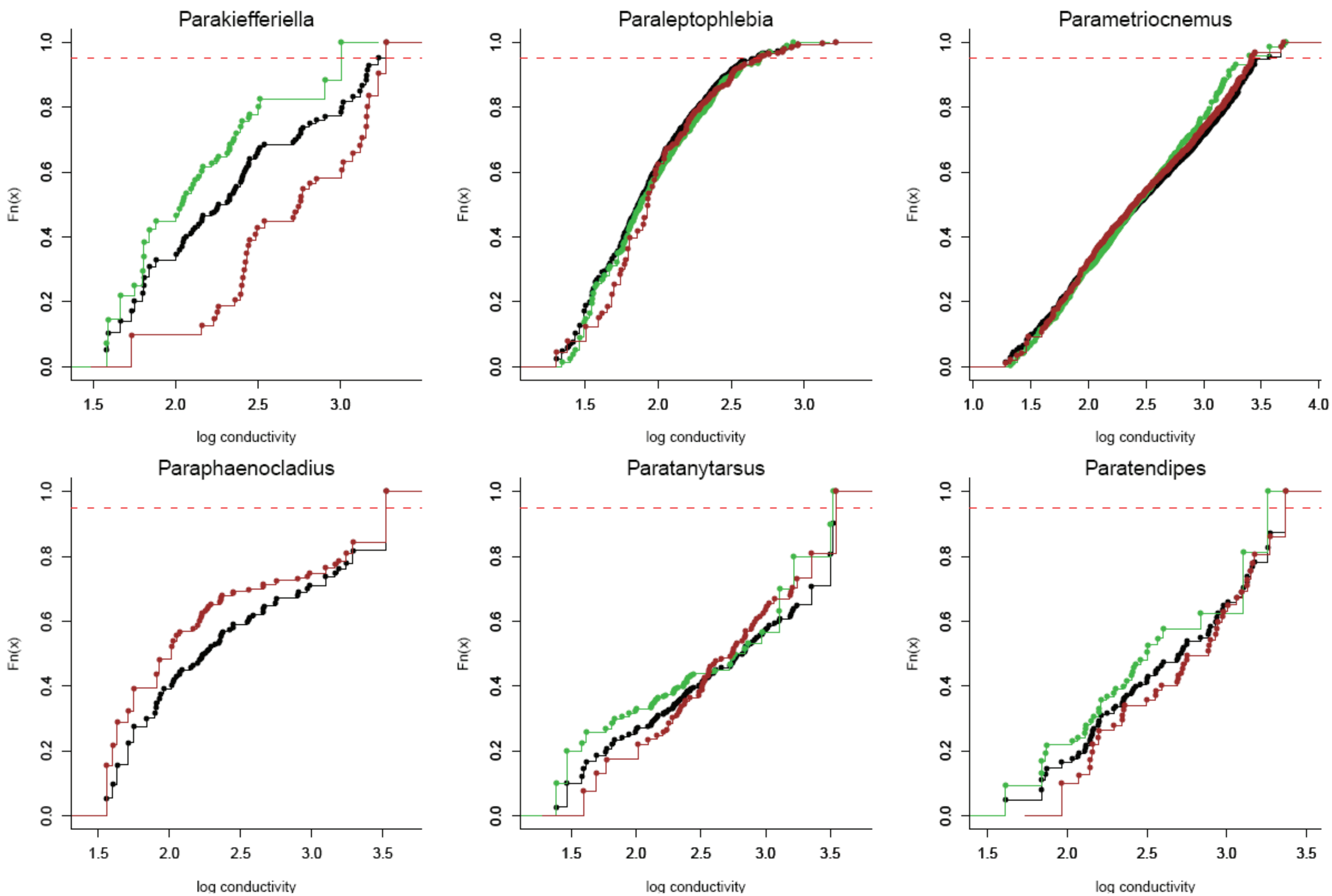

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

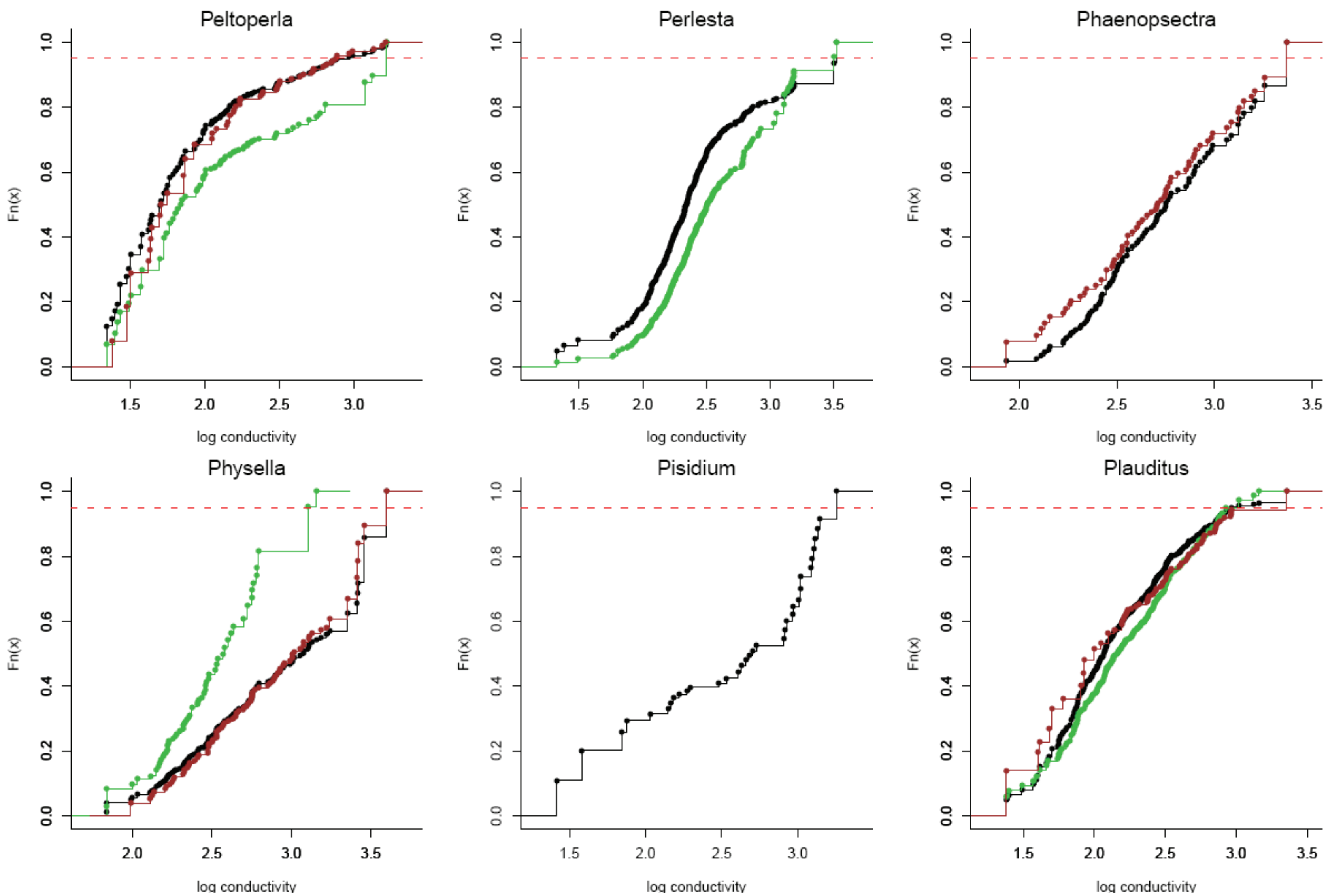

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

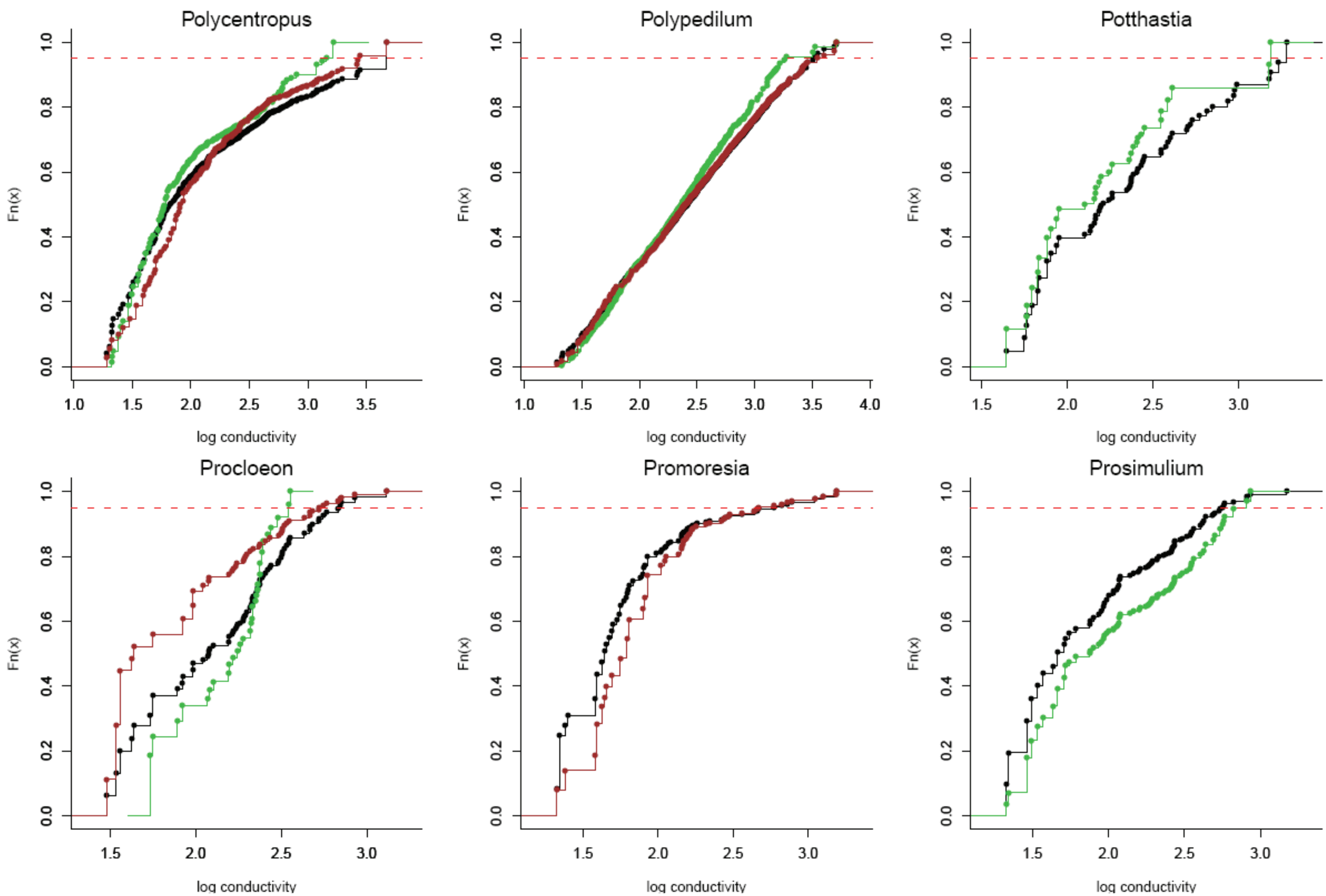

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

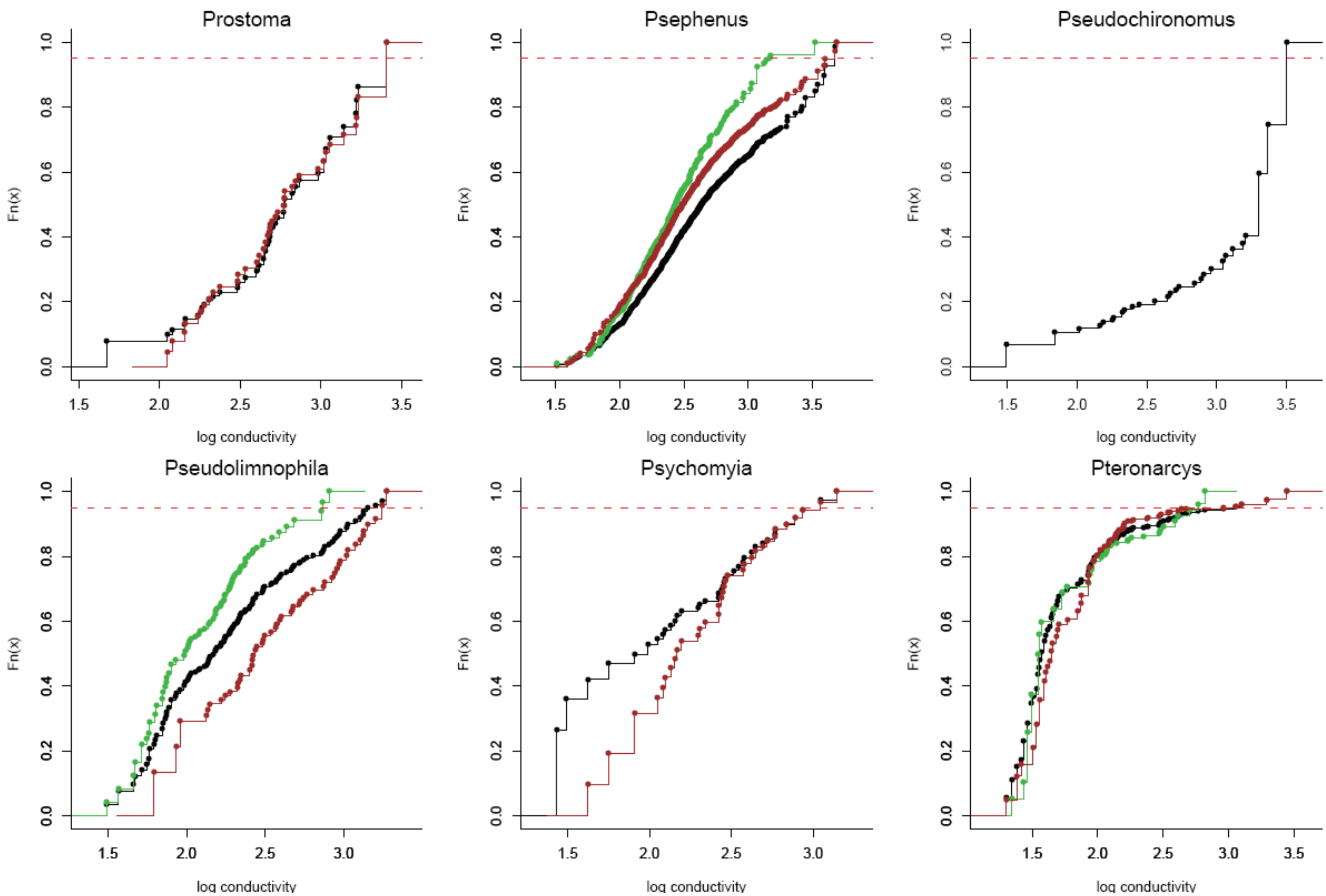

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

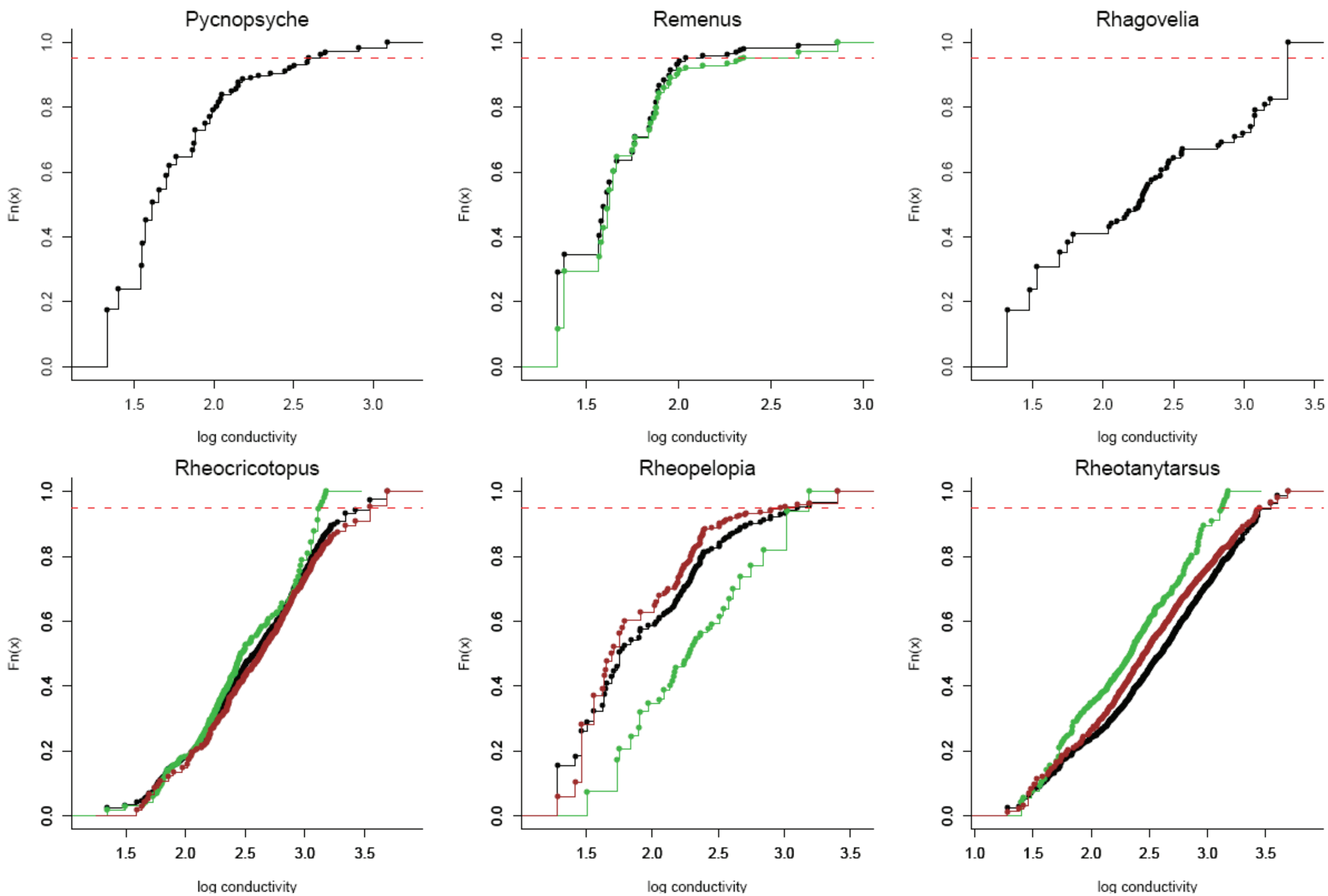

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

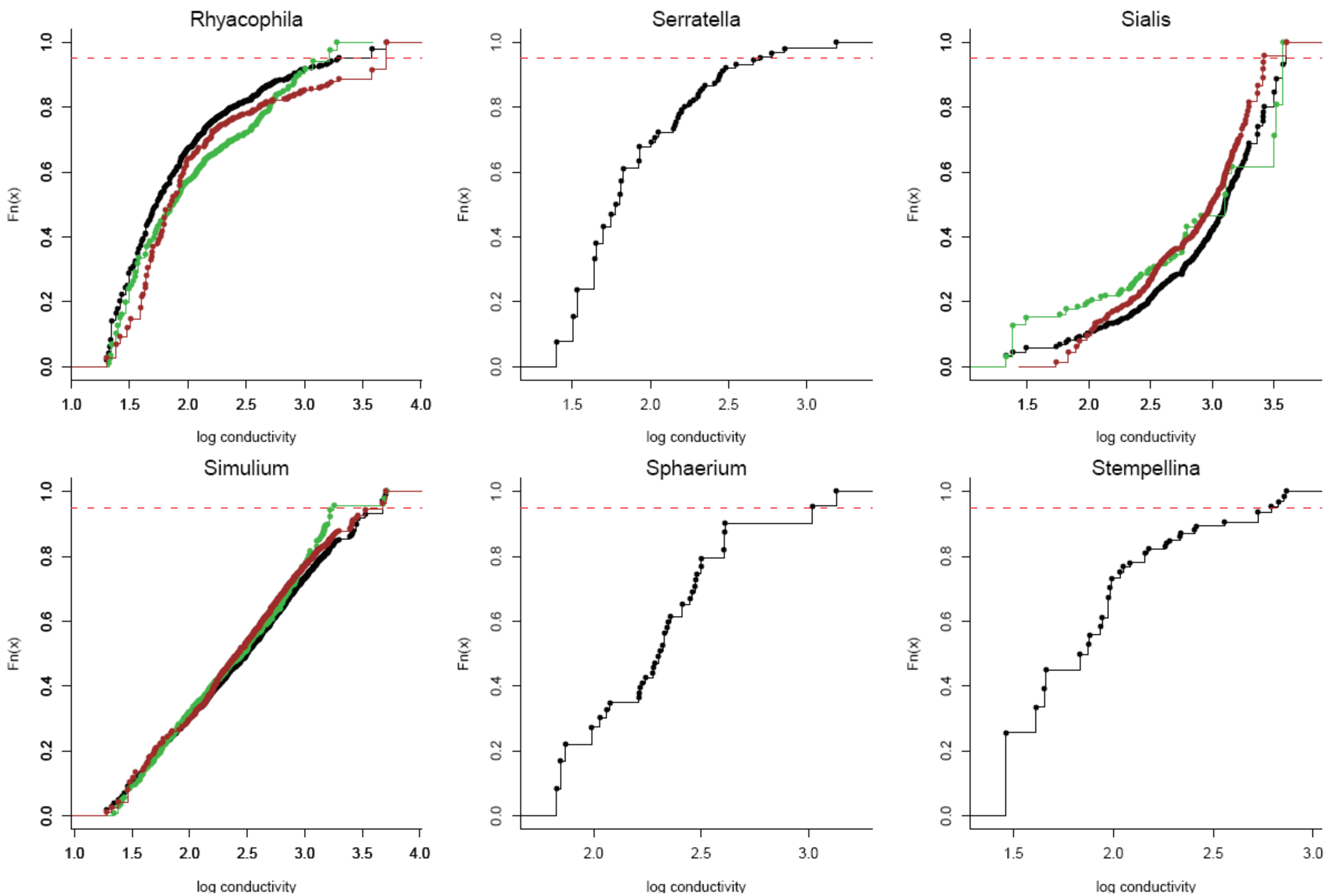

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

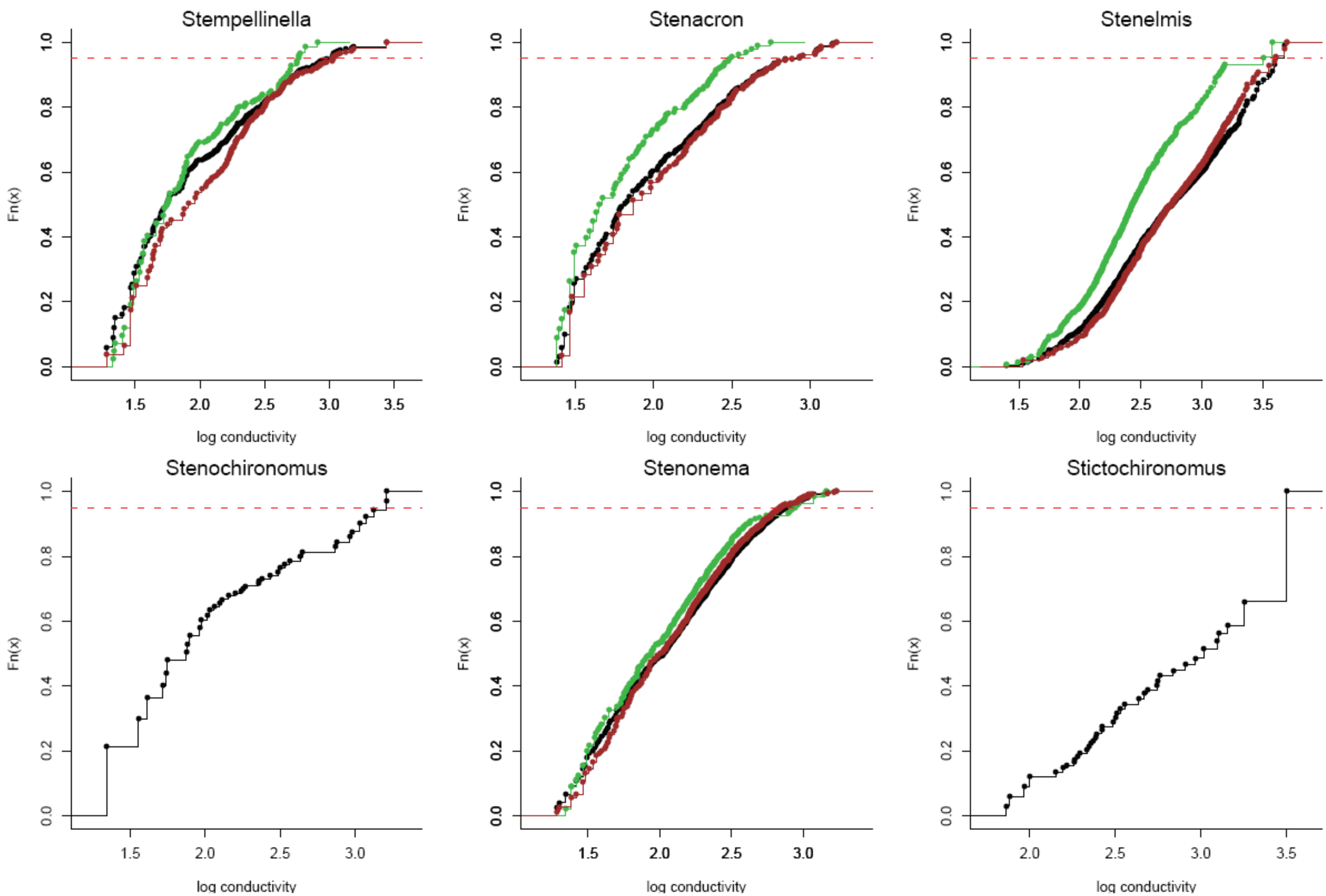

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

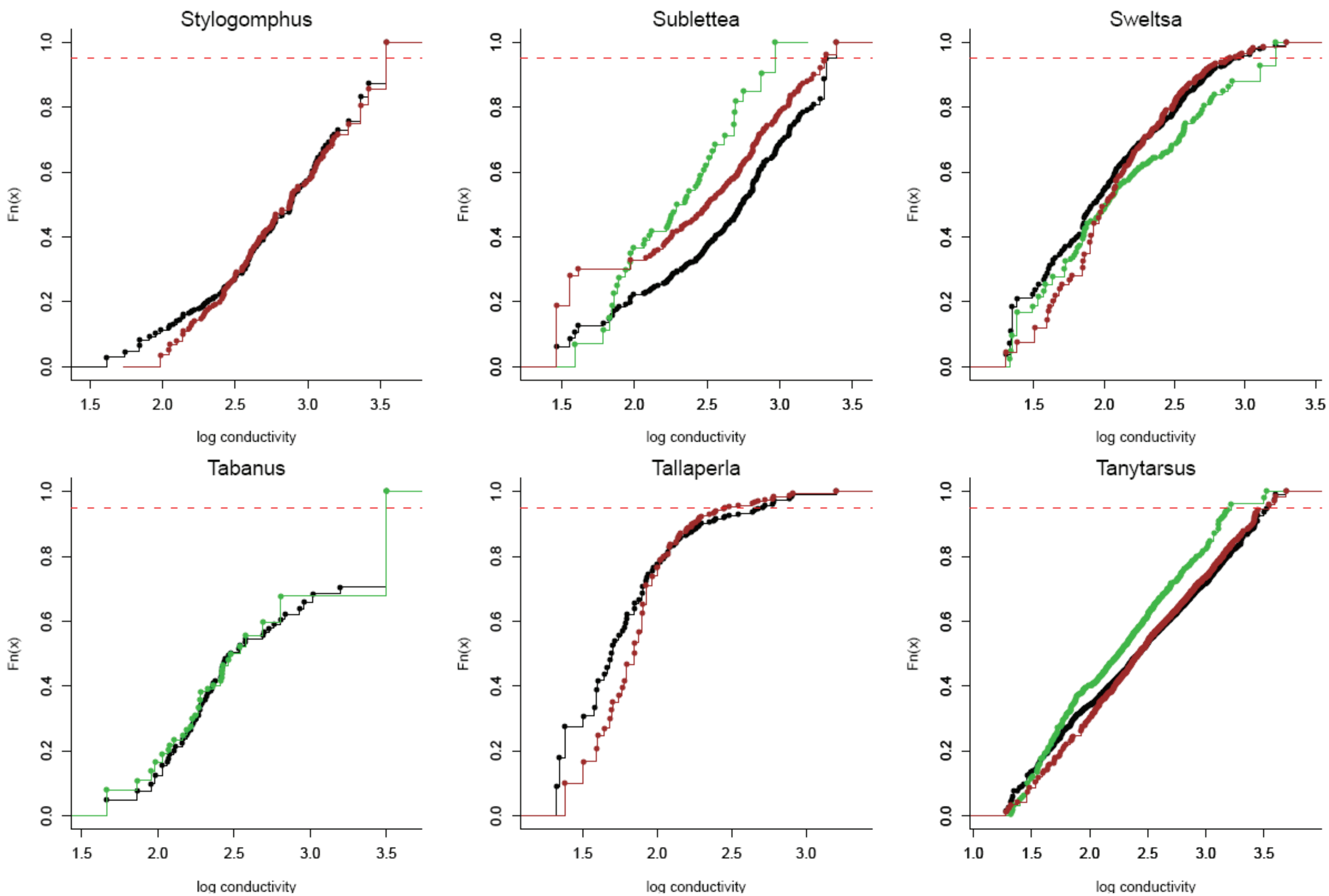

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

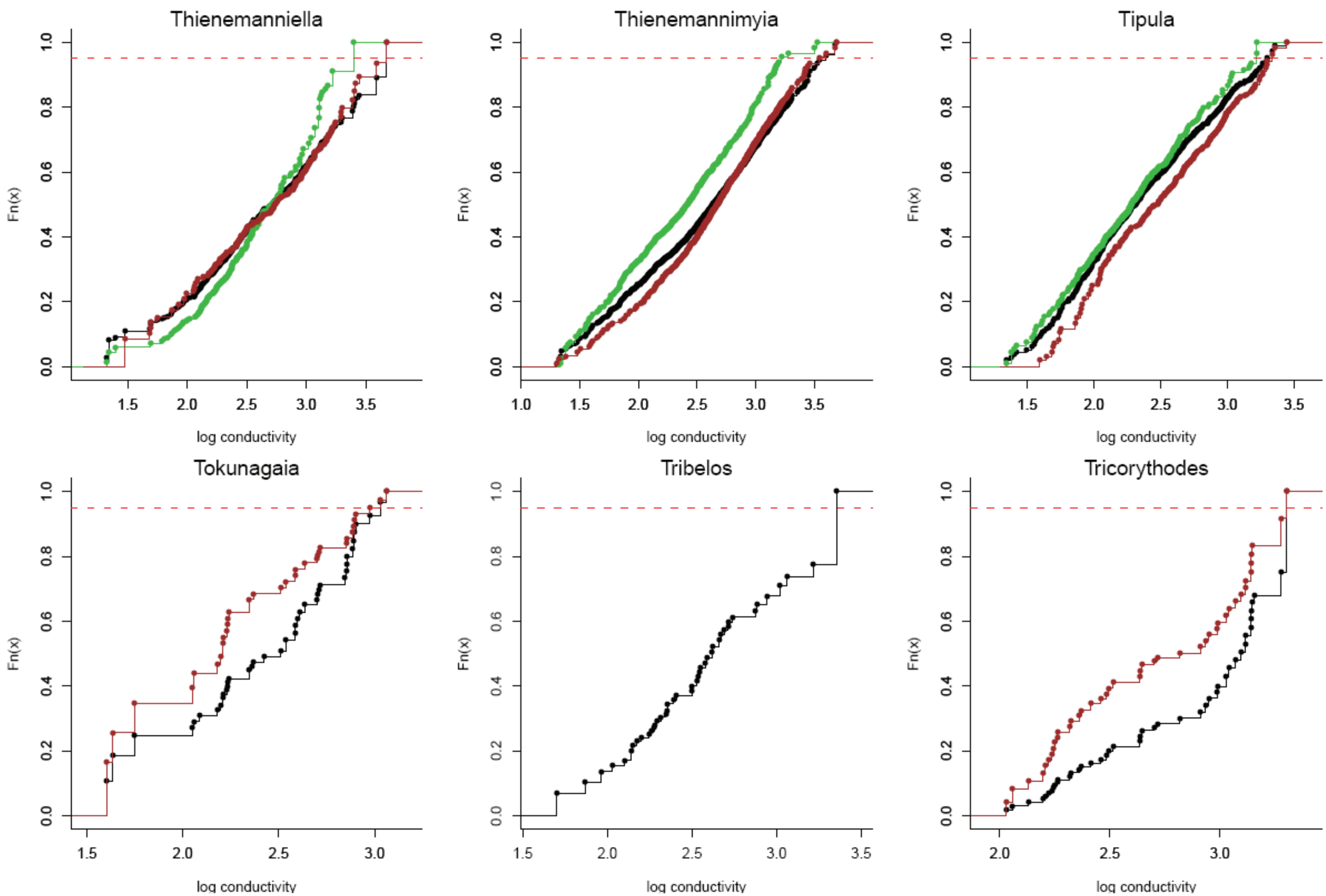

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

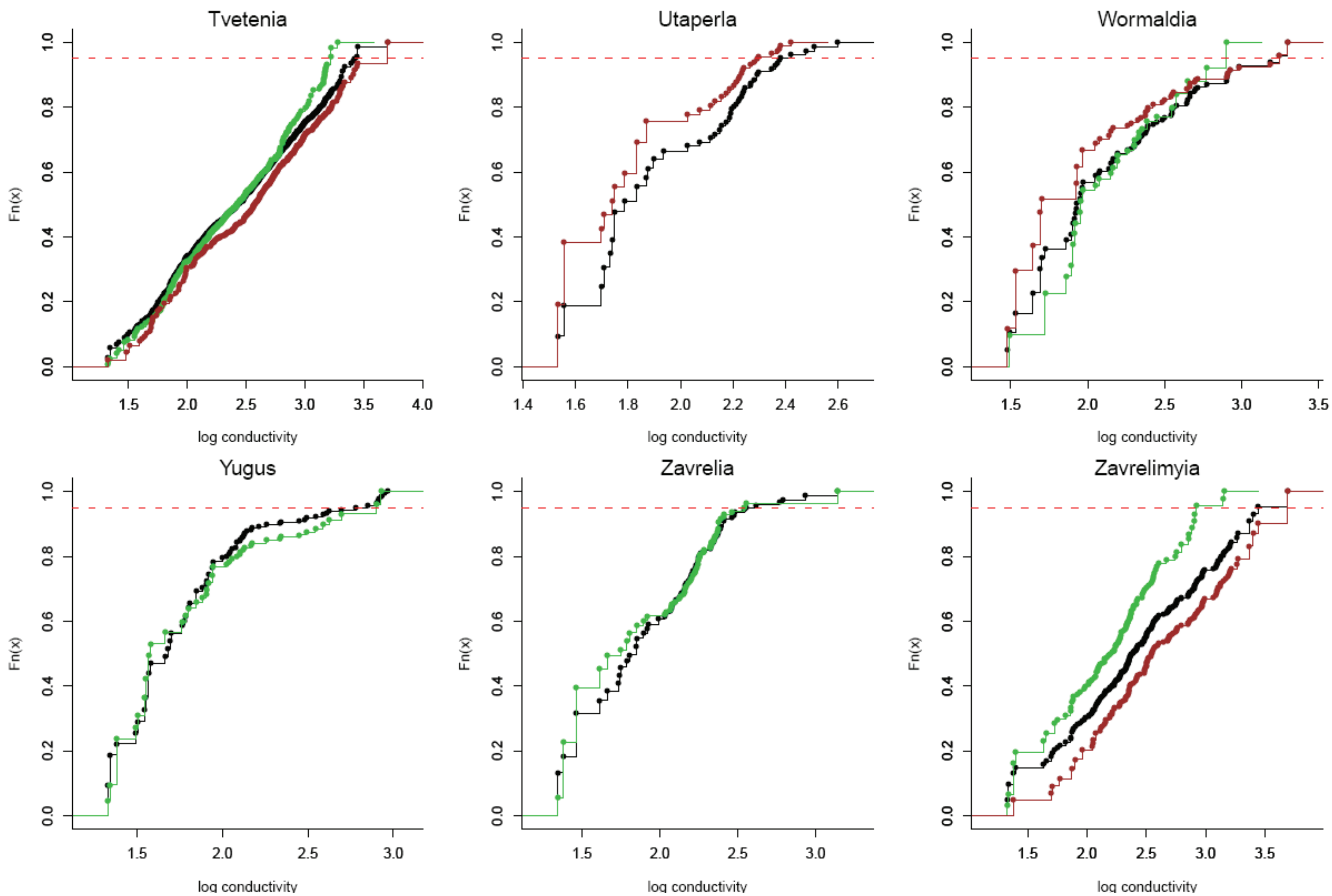

Figure D-2. Cumulative distribution functions of observation probabilities weighted by sampling frequency for each genus (continued).

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

The method for developing the aquatic life benchmark for conductivity was validated by developing XC_{95} and 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 WV 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 319 μS/cm for the full Kentucky data set, which is very close to the West Virginia result.

E.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 Figure E-1). Although the data set is smaller than the West Virginia data set, it was judged to be large enough for validation of the method. 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. The three ecoregions were judged to be similar within the state of Kentucky in terms of water quality, resident biota, and sources of conductivity. Confidence in the quality of reference sites was relatively high owing to the extensively forested areas of the region. Background conductivity was estimated from a probability sample from the U.S. EPA Wadeable Stream Assessment (U.S. EPA, 2006) at the 25^{th} percentile using the Spatial Survey Design Package (sp. survey R package) (Stevens and Olsen, 2004). Background conductivity at the 25^{th} percentile was 63 μS/cm for the Southern Appalachians, which includes Ecoregions 68, 69, and 70. When this value is compared to the 25^{th} percentile from a probabilistic subset of the WV data set, it was similar to the 72 μS/cm value for Ecoregion 69, but much lower than the 153 μS/cm value for Ecoregion 70.

E.2. DATA SOURCES

All data used in this study were taken from the Kentucky Division of Water, Water Quality Branch database (KY EDAS). Chemical, physical, or biological samples were collected from 274 distinct locations during February to October from 1998−2004 (see Table E-1). Like 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 for streams of interest as reference or impaired sites or for other reasons. Most sites have been sampled once during February to September. Quality assurance and standard procedures are described by KDOW (2008). 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.

E.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 WV data used that was used to derive the aquatic life benchmark for conductivity. A sample was excluded from calculations if (1) it lacked a conductivity measurement, (2) the organisms were not identified to the genus level, or (3) the pH was low. 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 sites with high chloride and low sulfate were identified or removed from the Kentucky data set. We evaluated the effects of spring benthic invertebrate emergence, seasonal differences in temperature and conductivities by partitioning the data set into spring (February−June) and summer (July−October) subsets. Eighty-one percent of the 95 genera used to develop the SSD for Kentucky also occurred in the WV SSD. This indicates that sensitive genera still exist in both states. Genera from both states were judged to be similarly susceptible to the effects of conductivity after exploratory analysis. Conductivity ranged from 16 to 2,390 μS/cm for the Kentucky data set and 15 to 11,646 μS/cm for the WV data set.

In the Kentucky database, 365 benthic invertebrate genera were identified. XC_{95} values were not calculated for genera that occurred at <30 sampling sites and therefore, these genera were not used to generate the SSD. Genera that did not occur at reference sites in West Virginia were excluded from the SSD. Of the 365 genera collected, 95 occurred in at least 30 sampling locations in Ecoregions 68, 69, and 70 (see Table E-4). Of the genera occurring in 30 or more samples, all genera occurred in all three ecoregions.

KDOW samples benthic macroinvertebrates using methods similar to WVDEP (KDOW, 2008). KDOW collects 4–0.5 m^2 kick samples in riffle/run habitat and composites them to yield a 1 m^2 sample. KDOW also supplements collections with multi-habitat qualitative sampling. However, for consistency, these qualitative sampling data were not used in model construction, only the riffle/run samples. Another 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. KDOW follows similar field and laboratory quality assurance methods as WVDEP.

E.4. CONCLUSIONS

Despite the differences in method and in location, the HC_{05} was similar: 319 µS/cm for Kentucky compared to 297 µS/cm for West Virginia (see Figures E-2 and E-3, Table E-2). The 95% confidence bounds for the Kentucky data set value are 180 µS/cm and 439 µS/cm which overlap with the West Virginia data set confidence bounds of 225 µS/cm and 305 µS/c. Genera that exhibited a decreasing occurrence with increasing conductivity were among those with the lowest XC_{95} values in both States. Table E-3 shows the 10 lowest XC_{95} values for both West Virginia and Kentucky samples. The 5th percentile occurs near genus 7 for West Virginia samples and genus 5 for Kentucky samples. Table E-4 lists the genera used to construct the SSD from the Kentucky sample and their XC_{95} values.

Based on the similar results, we judged the 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.

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.

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.

U.S. EPA (U.S. Environmental Protection Agency). (2006) Wadeable streams assessment: a collaborative survey of the Nation's STREAMS. Office of Research and Development, Office of Water, Washington, DC. EPA/841/B-06/002. December. Available online at http://www.epa.gov/owow/streamsurvey/ (accessed 12/20/2009).

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.

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

Region	Month												Total
	1	2	3	4	5	6	7	8	9	10	11	12	
68				10		6	18	2	3				39
69		7	14	44	16	16	42	18					182
70			9	21	2	17	21			10			70
													291

Table E-2. HC_{05} values for Kentucky and West Virginia

Kentucky	HC_{05}		West Virginia	HC_{05}
February−October	319 μS/cm		March−October	297 μS/cm
February−June	397 μS/cm		March−June	322 μS/cm
July−October	641 μS/cm		July−October	479 μS/cm

Table E-3. Comparison of the sensitive genera and XC_{95} values

WV			KY		
Rank	Genus	XC_{95}	Rank	Genus	XC_{95}
1	Remenus	101	1	Lepidostoma	132
2	Lepidostoma	109	2	Cinygmula	161
3	Cinygmula	224	3	Wormaldia	161
4	Leptophlebia	224	4	Dolophilodes	317
5	Alloperla	228	5	Drunella	320
6	Utaperla	240	6	Epeorus	324
7	Drunella	294	7	Neophylax	324
8	Pycnopsyche	299	8	Oulimnius	378
9	Ephemerella	302	9	Paraleptophlebia	400
10	Heptagenia	313	10	Ephemerella	467

Table E-4. Extirpation concentration and sample size from Kentucky data set. Highlighted genera are not found at WV reference sites. For a genus with a decreasing capture probability with increasing conductivity, the XC_{95} is reported directly. If the genus did not occur at the higher conductivity levels, it is reported as a greater than or equal to value (≥). Genera with increasing capture probability with increasing conductivity are reported as greater than (>).

	Genus	Both		Spring		Summer	
		XC_{95}	N	XC_{95}	N	XC_{95}	N
1	Ablabesmyia	>1,410	43			879	30
2	Acentrella	≥618	98	762	67	626	31
3	Acroneuria	≥703	105	926	56	703	49
4	Ameletus	≥507	69	762	68		
5	Amphinemura	≥1,287	107	1,980	105		
6	Ancyronyx	≥841	30				
7	Antocha	≥958	49				
8	Argia	>1,410	51			950	34
9	Atherix	>1,650	61			2,000	48
10	Baetis	>1,410	170	1,410	79	1,176	91
11	Boyeria	≥1,410	92	1,410	36	879	56
12	Caenis	>1,410	85	1,410	35	2,340	50
13	Calopteryx	>1,980	35				
14	Cambarus	≥1,132	157	1,163	95	841	62
15	Ceratopsyche	>1,580	102	1,980	34	1,203	68
16	Cheumatopsyche	>1,630	230	1,980	106		
17	Chimarra	>2,000	90	2,260	34	2,000	56
18	Chironomus	>2,340	31				
19	Cinygmula	161	39	183	39		
20	Corbicula	>1,863	84			1,863	60
21	Corydalus	>1,650	121	1,480	33	2,000	88
22	Cricotopus	>2,000	98	1,980	44	2,340	54
23	Diamesa	>1,980	54	1,980	43		

Table E-4. Extirpation concentration and sample size from Kentucky data set (continued)

	Genus	Both		Spring		Summer	
		XC_{95}	N	XC_{95}	N	XC_{95}	N
24	Dineutus	879	45			841	33
25	Diplectro	≥958	102	1,228	76		
26	Diploperla	≥702	35	1,157	34		
27	Dolophilodes	317	31				
28	Drunella	320	37	762	35		
29	Dubiraphia	>1,650	86			1,650	65
30	Eccoptura	≥1,228	31				
31	Eclipidrilus	>1,322	92	1,271	47	934	45
32	Ectopria	≥561	66	505	40		
33	Elimia	≥879	33				
34	Ellagma	894	31				
35	Epeorus	324	65	324	59		
36	Ephemera	561	42				
37	Ephemerella	467	70	485	67		
38	Eukiefferiella	>1,650	54	2,260	38		
39	Eurylophella	≥505	84	526	67		
40	Gomphus	≥1,047	36				
41	Haploperla	485	37	485	32		
42	Helichus	≥1,050	147	1,520	80	1,132	67
43	Hemerodromia	>2,000	123	2,260	50	2,000	73
44	Hexatoma	≥1,069	105	762	62	2,000	43
45	Hydropsyche	>1,650	160	1,980	66	2,000	94
46	Hydroptila	>1,863	58			1,863	41
47	Isonychia	≥1,580	132	894	38	1,863	94
48	Isoperla	≥1,176	80	1,157	77		
49	Lanthus	1,520	34				

Table E-4. Extirpation concentration and sample size from Kentucky data set (continued)

	Genus	Both		Spring		Summer	
		XC_{95}	N	XC_{95}	N	XC_{95}	N
50	Lepidostoma	132	30				
51	Leucrocuta	≥703	45				
52	Leuctra	≥1,113	131	889	91	2,000	40
53	Lirceus	958	35				
54	Macronychus	≥1,650	54			1,863	41
55	Microtendipes	≥675	58			805	32
56	Natarsia	>1,630	45				
57	Neophylax	324	73	431	67		
58	Nigronia	≥1,203	153	1,113	72	1,203	81
59	Oecetis	≥2,000	31				
60	Optioservus	>1,560	178	1,410	70		
61	Orconectes	>1,302	115	1,287	41	1,277	74
62	Orthocladius	>1,480	49	1,480	39		
63	Oulimnius	378	31				
64	Paraleptophlebia	400	76	762	54		
65	Parametriocnemus	>1,630	184	1,980	111	2,000	73
66	Peltoperla	≥1,520	37				
67	Perlesta	≥1,287	51	1,520	33		
68	Physella	≥2,260	52				
69	Plauditus	≥703	55				
70	Polycentropus	570	82	635	43	570	39
71	Polypedilum	≥1,247	157	1,271	69	950	88
72	Procloeon	≥802	42			768	34
73	Prosimulium	≥958	53	401	46		
74	Psephenus	≥738	111	762	59	879	52
75	Pseudocloeon	855	36				

Table E-4. Extirpation concentration and sample size from Kentucky data set (continued)

	Genus	Both		Spring		Summer	
		XC_{95}	N	XC_{95}	N	XC_{95}	N
76	Pseudolimnophila	≥1,050	39				
77	Pycnopsyche	≥802	64	889	45		
78	Rheocricotopus	≥1,163	51				
79	Rheotanytarsus	>1,580	115	1,113	36	1,863	79
80	Rhyacophila	≥565	94	820	74		
81	Sialis	>1,287	63	1,980	32	891	31
82	Simulium	>1,580	179	1,410	92	1,863	87
83	Stenacron	879	90	762	37	879	53
84	Stenelmis	≥1,520	168	1,480	81	1,863	87
85	Stenochironomus	≥802	35				
86	Stenonema	≥993	178	658	68		
87	Stylogomphus	>1,863	89	1,480	40	1,863	49
88	Sweltsa	507	55	435	39		
89	Tanytarsus	≥1,287	118	1,163	54	1,863	64
90	Thienemannimyia	≥1,630	139	1,410	71	2,000	68
91	Tipula	>,630	150	1,980	106	2,340	44
92	Triaenodes	841	31				
93	Tricorythodes	>,000	48			2,340	44
94	Tvetenia	>,203	46				
95	Wormaldia	161	38	1,980	34		

XC_{95} = 95th percentile extirpation concentration reported as μS/cm; NA = not applicable because it never occurs at WVDEP reference locations; Both = February through October; Spring = Sampled February through June; Summer = July through October; Empty cells indicates fewer than 30 occurrences during that season.

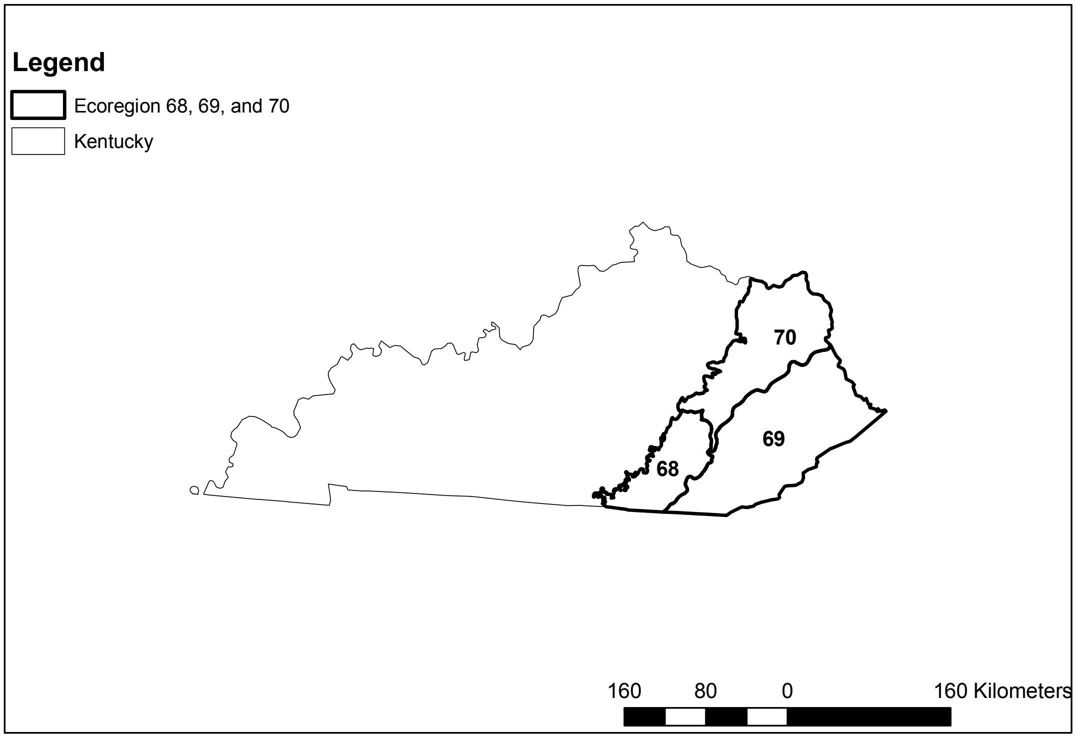

Figure E-1. Location of Southern Appalachia (68), Central Appalachia (69), and Allegheny Plateau (70).

Data source: State outlines from the U.S. EPA Base Map Shapefile Omernik Level III Ecoregions from National Atlas (NationalAtlas.gov), projection: NAD 1983 UTM 17 N. Map made December 21, 2009, by M. McManus.

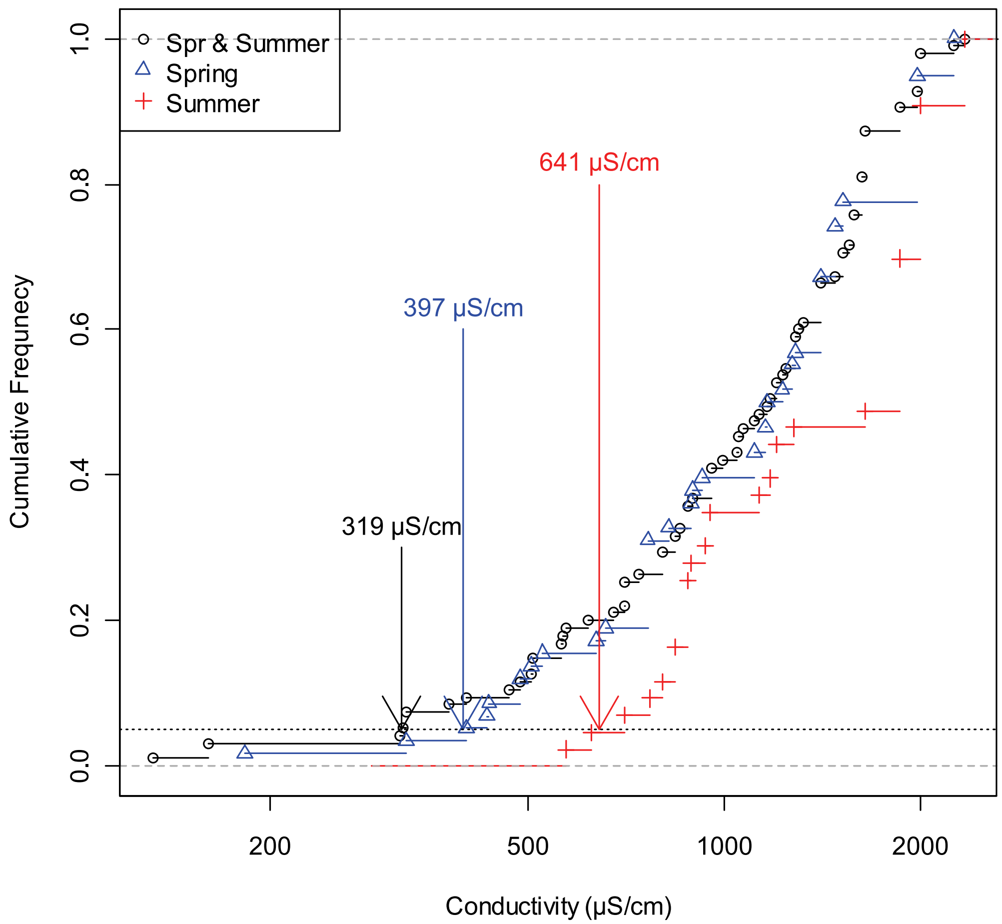

Figure E-2. The species sensitivity distributions for all year (February through October [black circles], February through June [blue triangles], and July through October [red crosses +]). Ninety-five genera are included in SSD using the all year data set. The HC_{05} is the conductivity at the intercept of the CDF with the horizontal line at the 5^{th} percentile. For all year, it is 319 µS/cm with 95% confidence bounds at 180 µS/cm and 439 µS/cm.

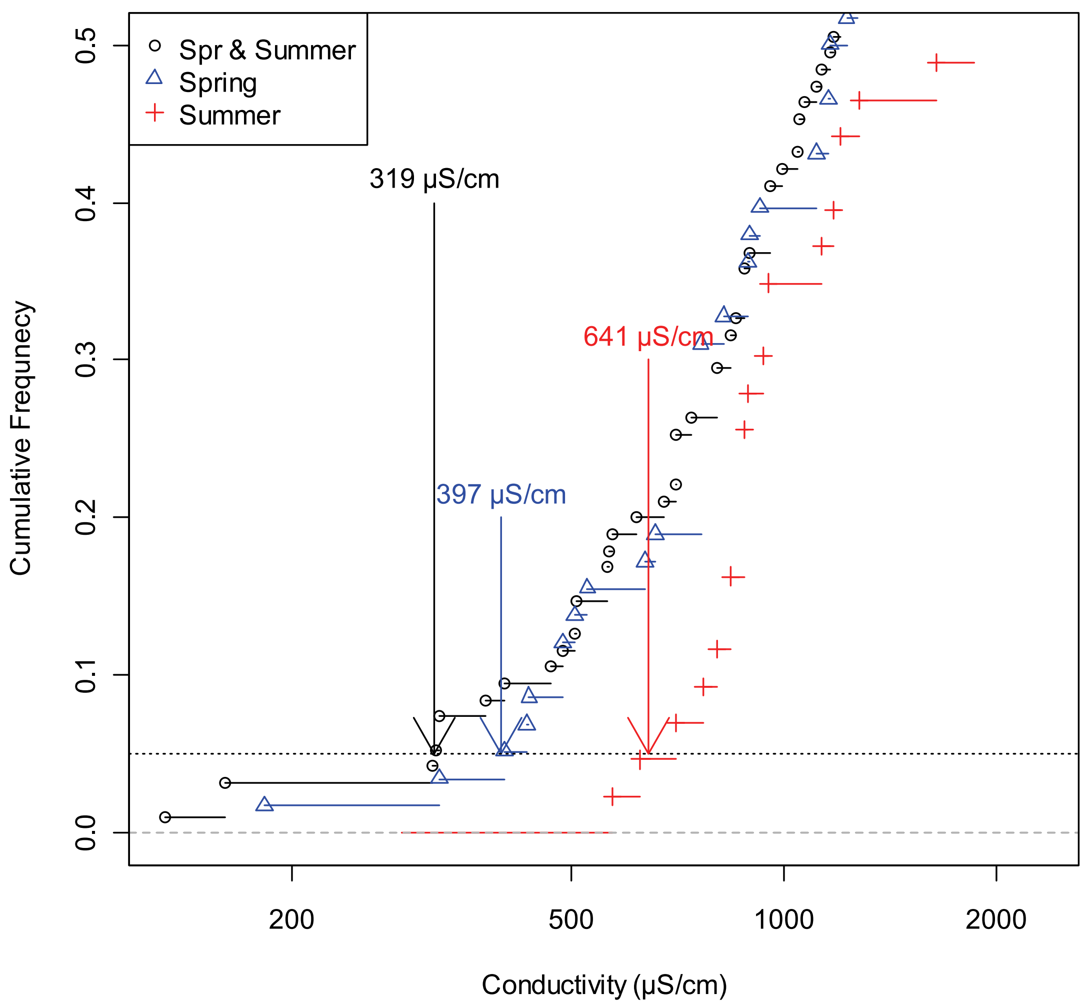

Figure E-3. Species sensitivity distribution for all year (black circles), spring (blue triangles), and summer (red crosses +). Only the lower half genera are shown to better discriminate the points in the left side of the distribution.

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

F.1. OVERVIEW

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. This information was used as evidence of preceding causation in the causal assessment described in Appendix A of this report. A strategy used in this analysis was to limit the watersheds to <20 km^2 to minimize the number of land use and cover types within a single watershed, thereby providing a clearer signal. However, because the region has a long history of mining, persistent effects of mining were potentially present even when there was no current record of past or present mining activity.

The final data set consisted of 191 small watersheds for which macroinvertebrate samples were identified to genus, water chemistry was available from at least one sampling effort, subwatershed area was ≤20 km^2, and detailed land cover information was also available. The data set of 191 sites was drawn from 2,151 sites in Ecoregion 69D described in the West Virginia Department of Environmental Protection's (WVDEP) Watershed Assessment Branch Data Base (WABbase). These 191 tributary watersheds were from the Coal, Upper Kanawha, Gauley, and New Rivers. From each watershed, scatter plots for several parameters were generated for eight land cover classifications: open water, agriculture, urban/residential, barren, valley fill, mining, abandoned mine lands, and forested lands.

Although conductivity typically increases with increasing land use (Herlihy et al., 1998), 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. At relatively low urban land use, the range of conductivity is highly variable. This may be caused by unknown mine drainage, deep mine break-outs, road applications, poor infrastructure condition (e.g., leaking sewers or combined sewers), or other practices. In contrast, there is a clear pattern of increasing conductivity as percent area in valley fill increases and decreasing conductivity with increasing forest cover. Pairs of land use and water quality parameters with moderately strong correlation coefficients ($r \geq |0.50|$) are listed in Table F-1. All other pairs exhibit $r < |0.50|$ except a few with spurious points or composed of only 2 points from which no evaluations could be made. Biological effects measured as the West Virginia Stream Condition Index (WVSCI) score or the genus level index of most probable stream status (GLIMPSS) score were weakly correlated with percent forest cover and percent valley fill with $r > |0.30|$.

F.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 West Virginia data, the West Virginia GIS Technical Center (http://wvgis.wvu.edu/data/data.php), maintains publicly available shapefiles. WVDEP also maintains a publicly available repository of statewide GIS data sets (http://gis.wvdep.org/). All relevant GIS metadata are available for the data housed at each repository site. All GIS coverages used in this U.S. EPA study are in universal transverse mercator (UTM) 1983 Zone 17 and the units are in meters. Table F-2 describes some of the publicly available GIS shapefiles that were used as the total daily maximum load (TMDL) land use base files and the beginning point for determining the 191 stations selected for the analyses described in Section F.3, and as the beginning point for the 191 stations land use analysis described in Section F.4.

F.3. METHODS

The analysis for Appendix A proceeded in two steps; (1) selection of the 191 stations and (2) land use analysis of the 191 stations. Section F.3 describes the selection process for selecting the 191 sample stations, while Section F.4 describes the detailed land use evaluation for each of the 191 stations. Figure F-1 depicts the Ecoregion 69D in relation to the West Virginia State boundary and the 8-digit watershed boundaries, while Figure F-2 shows the locations of the 191 stations within Ecoregion 69D.

191 Stations Selection Process within Ecoregion 69D with TMDL Land Use

- All WVDEP WAB stations located within Ecoregion 69D were selected. This generated 2,151 stations.
- The next station reduction occurred by selecting only 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 must accompany the macroinvertebrate sample from the same station location. This narrowed the available stations to 825.
- To obtain the TMDL associated land use, stations located within the Coal, Upper Kanawha, Gauley, and New River TMDL watersheds were selected. This narrowed the selection to 382 stations.
- Stations were eliminated if detailed land use was not created during the TMDL process. This eliminated 38 stations for a total of 344 stations.
- Next, a station was eliminated if it was located on an undelineated tributary stream that was contained within a larger main stem subwatershed. Failure to remove these would

generate an overestimation of land use from the entire upstream contributing land use and not simply the land use from the tributary where the sample was actually located. This eliminated an additional 33 stations for a total of 311 stations.

- Lastly, the EPA workgroup limited the upstream contributing land use to stations with a total watershed drainage area ≤20 km^2 (4,942.08 acres). The total remaining stations in TMDL watersheds within Ecoregion 69D after this last reduction was 191 stations (see Figure F-2), and the data from these stations were assembled from 1997 to 2007, with the majority of samples having been collected from 2001 to 2006. EPA workgroup consisted of scientists from EPA Office of Research and Development and Region 3, contracted scientists from Tetra Tech Inc., and scientists from the WVDEP.

F.4. LANDUSE MANIPULATIONS

To create the land use for the 191 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 NLCD) 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 F-3 contains the WVDEP TMDL 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.

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 191 sampling stations. The land use representation used in TMDL development 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 191 WAB sampling stations used seven categories consolidated from the TMDL land use (see Table F-3), then included the addition of the valley fill area.

F.5. CORRELATIONS WITH IN STREAM BIOLOGICAL AND WATER QUALITY PARAMETERS

Spearman rank correlations of eight land use categories with conductivity and ion concentration were calculated (see Table F-1). Individual scatter plots and associated correlation coefficients for conductivity can be found in Appendix A (see Figure A-3). Land use and land cover were arc sine square root transformed to better depict the upper and lower portions of the distribution.

REFERENCES

Herlihy, A.T.; Stoddard, J.L.; Johnson, CB. (1998) The relationship between stream chemistry and watershed land cover data in the mid-Atlantic region, U.S. Water, Air, and Soil Pollution 105:377-386.

Shank, M. (2004) Advanced integration of geospatial technologies in mining and reclamation conference, December 7-9, 2004, Atlanta, GA.

Table F-1. Pairs of land use and water quality parameters with correlations coefficients >0.5 in the land use data set

Land use	Water quality parameter	*r*
Percent Forest	Conductivity	−0.56
	Alkalinity	−0.51
	Hardness	−0.65
	Sulfate	−0.54
	Calcium-total	−0.64
	Magnesium-total	−0.58
Percent Mined Area	Hardness	0.56
	Calcium-total	0.51
	Magnesium-total	0.58
Percent Valley Fill Area	Conductivity	0.65
	Alkalinity	0.50
	Sulfate	0.64
	Calcium-total	0.66
	Magnesium-total	0.66

Table F-2. 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
WVDEP GIS data sets	General WVDEP's GIS data repository location	http://gis.wvdep.org/
Base Land use/land cover		
GAP	GAP land use	http://wvgis.wvu.edu/data/dataset.php?ID=62
NLCD 2001	NLCD land use	http://wvgis.wvu.edu/data/dataset.php?ID=269
Other files		
Watershed Boundary Datasets	USGS 8-digit Hydrologic Unit Code boundaries	http://wvgis.wvu.edu/data/dataset.php?ID=123
NHD Streams	National Hydrography Dataset Streams	http://wvgis.wvu.edu/data/dataset.php?ID=235
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
OMR Mining NPDES Permits and Outlets	WVDEP Office of Mining and Reclamation NPDES permit and outlet coverages	http://gis.wvdep.org/data/omr.html
Mining related Fills, Southern West Virginia	WVDEP valley fills coverage from 2003	http://gis.wvdep.org/data/omr.html
Mining Permit Boundaries	WVDEP Mining permit boundaries	http://wvgis.wvu.edu/data/dataset.php?ID=149
Roads_Paved	2000 TIGER/Line GIS and WV_Roads shapefiles	http://wvgis.wvu.edu/data/data.php
Roads_Unpaved	2000 TIGER/Line GIS shapefile and digitized from aerial photographs and topographic maps	http://wvgis.wvu.edu/data/data.php

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

Table F-3. Detailed WV TMDL Land Use Category Derivation and Land Use Derivation used in Appendix A. Base land use categories 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 F-3. 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

[a]Valley fill land use was not part of the base TMDL land use and was specifically incorporated into the detailed land use analysis during the EPA ion study for the 191 stations. See Table F-2 for the source file.
LU = land use; WQ = water quality.

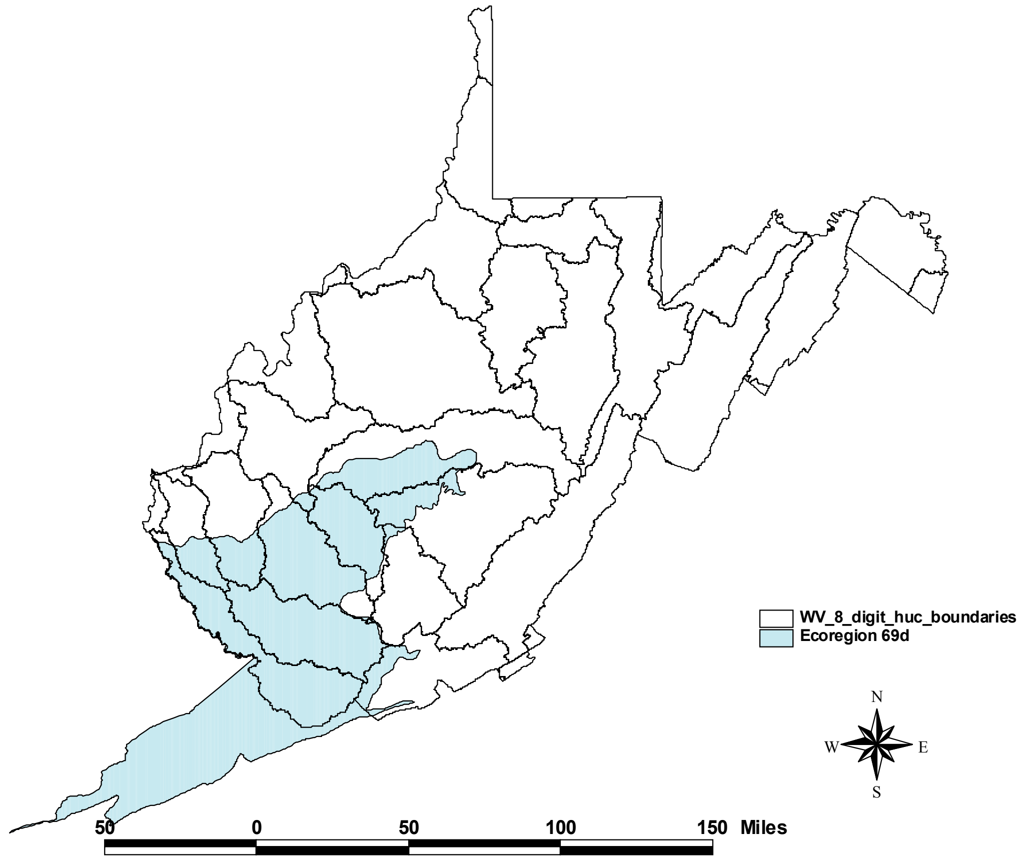

Figure F-1. West Virginia 8-digit Hydrologic Unit Code boundaries, Ecoregion 69D.

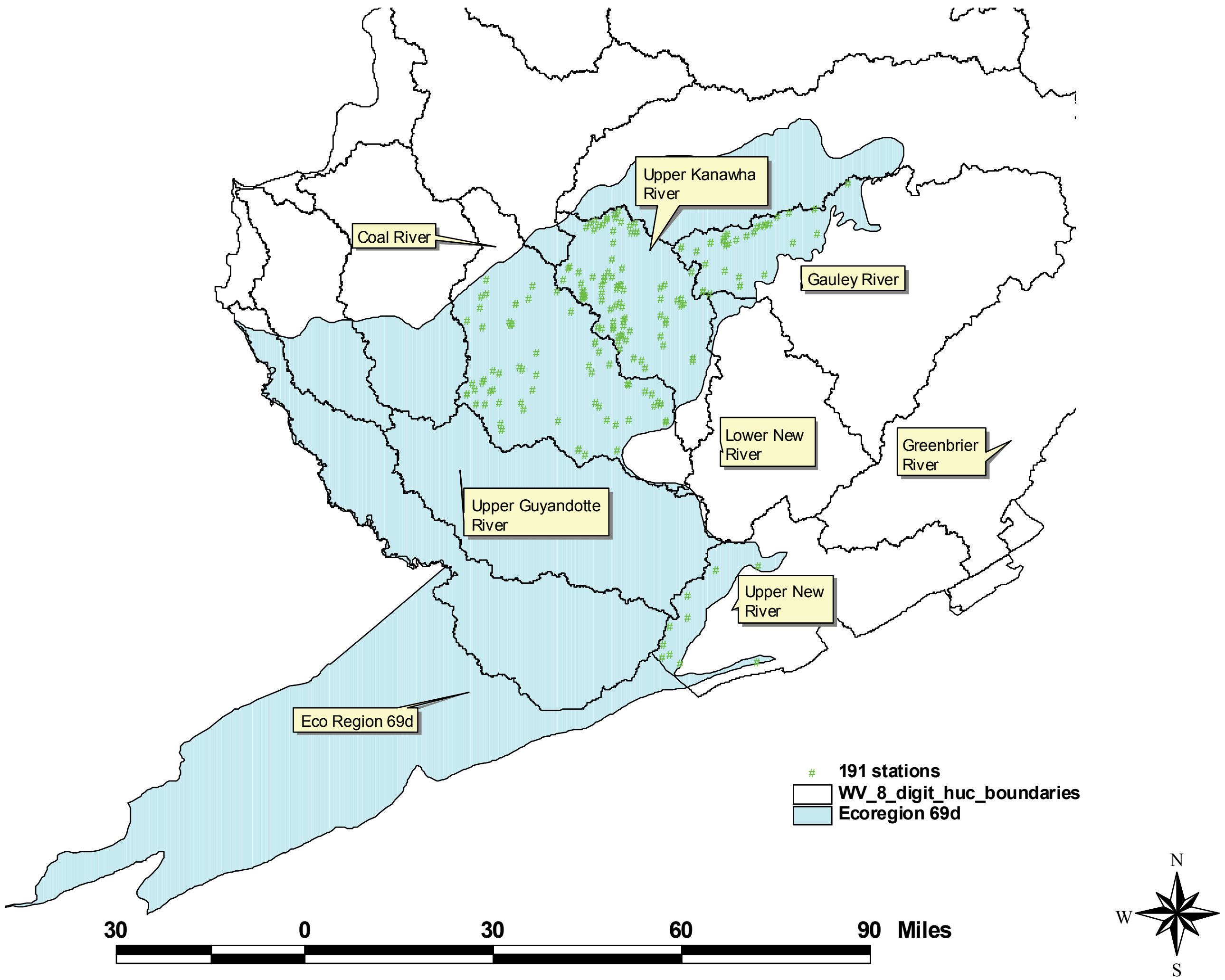

Figure F-2. One hundred ninety-one station locations used in the detailed land use analysis.