Distribution Agreement

In presenting this thesis or dissertation as a partial fulfillment of the requirements for an advanced degree from Emory University, I hereby grant to Emory University and its agents the non-exclusive license to archive, make accessible, and display my thesis or dissertation in whole or in part in all forms of media, now or hereafter known, including display on the world wide web. I understand that I may select some access restrictions as part of the online submission of this thesis or dissertation. I retain all ownership rights to the copyright of the thesis or dissertation. I also retain the right to use in future works (such as articles or books) all or part of this thesis or dissertation.

Signature:

_____ _____

Allyson Mateja Date

Estimation of Persistent Organic Pollutants (POPs) in Fish Collected at Sapelo Island, GA: Statistical Methods Using Left Censored Data

By

Allyson Mateja

Master of Science in Public Health

Department of Biostatistics and Bioinformatics

Lance Waller, Ph.D.

Committee Chair

Robert Lyles, Ph.D.

Committee Member

Estimation of Persistent Organic Pollutants (POPs) in Fish Collected at Sapelo Island, GA: Statistical Methods Using Left Censored Data

By

Allyson Mateja

B.S.

The College of William & Mary

2015

Thesis Committee Chair: Lance Waller, Ph.D.

An abstract of

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Science in Public Health

in Biostatistics

2019

Abstract

Estimation of Persistent Organic Pollutants (POPs) in Fish Collected at Sapelo Island, GA: Statistical Methods Using Left Censored Data

By Allyson Mateja

Background: Persistent organic pollutants (POPs) are resistant to degradation and bioaccumulate in food webs. Although many have been banned, they are still pervasive threats to marine animal health. Dolphins in the Turtle/Brunswick River Estuary (TBRE), a nationally designated Superfund site, were exposed to very high levels of polychlorinated biphenyls (PCBs). Sapelo Island, GA has minimal urbanization, so was used as a control for the TBRE. However, research shows dolphins at Sapelo Island also have elevated PCB levels.

Objective: We were interested in examining whether there is a difference in analyte levels between fish at four different sites at Sapelo Island and in three different species for different chemical classes. The majority of chemical concentration measurements were below a limit of detection (left-censored).

Methods: The mean for each site/species/analyte combination was estimated using a maximum likelihood approach and a two-way ANOVA (model [1]). This model was simplified to assume no interaction between site and species (model [2]), and to a one-way ANOVA to compare only between species (model [3]).

Results: Sea trout have lower summed mean concentrations across all chemical classes compared to mullet and silver perch, with the exception of Aroclor 1268, for which mullet have the lowest concentrations. Mullet and silver perch have similar summed mean concentrations across chemical classes, with mullet higher for DDTs and metals and silver perch higher for pesticides and PCBs. We found statistically significant differences between species for all chemical classes except polybrominated diphenyl ethers (PBDEs) in model [2]. Only pesticides were significantly different across species in model [1]. Aroclor 1268 was significantly different between sites in model [2], after removal of outliers. Otherwise, we found no significant differences between sites.

Conclusion: Our results are limited due to large amounts of left-censored data and large variances of estimates. Even after removing outliers, most analytes did not have maximum likelihood estimates with stable standard errors. The small number of fish and large number of left-censored measurements create statistical challenges for accurate estimation and must be considered when interpreting results. We recommend future studies include larger sample sizes or focus on analytes present at higher concentrations.

Estimation of Persistent Organic Pollutants (POPs) in Fish Collected at Sapelo Island, GA: Statistical Methods Using Left Censored Data

By

Allyson Mateja

B.S.

The College of William & Mary

2015

Thesis Committee Chair: Lance Waller, Ph.D.

A thesis submitted to the Faculty of the

Rollins School of Public Health of Emory University

in partial fulfillment of the requirements for the degree of

Master of Science in Public Health

in Biostatistics

2019

Acknowledgements

I would like to thank the entire Department of Biostatistics and Bioinformatics at the Rollins School of Public Health for providing me with an excellent education and well-rounded knowledge of biostatistics. Specifically, I would like to thank Lance Waller, Ph.D. and Bob Lyles, Ph.D. who served as my thesis advisor and reader, respectively, and provided me with the methodology and guidance necessary to complete this thesis.

I would also like to thank Lisa Hoopes, Ph.D. of the Georgia Aquarium for allowing me the opportunity to work on this project, providing me with the data, and helping extensively with the background and with my understanding of the significance of this work. I am grateful for her patience in allowing me to complete this project on my own time.

I would like to thank my fellow students in the Department. I would not have gotten through the program without your constant help and support. Thank you to everyone for always being willing to go over course topics in the computer lab, and for the opportunity to represent the department in student government.

Finally, I would like to thank my family and friends in all of the other departments at Rollins. It was especially helpful to have the help of others who understood the process and who were encouraging every step of the way. Writing this thesis gave me an appreciation for the challenges of real-world data analysis beyond the classroom and helped me grow as a biostatistician and student.

Table of Contents

1. Introduction 1

2. Background 2

Study site 4

Sample collection and processing 6

Data Description 8

Approaches for handling left-censored data 8

3. Methods 13

Maximum Likelihood Estimation 15

Extreme Observations 25

4. Results 26

DDTs 27

PCBs 39

PBDEs 44

Metals 46

Pesticides 51

5. Discussion 53

Limitations 56

Future Work 58

Conclusions 60

6. References 64

7. Appendix 67

1. Introduction

Persistent organic pollutants (POPs) threaten marine animal health due to bioaccumulation and biomagnification in high-level predators[11]. Bottlenose dolphins (*Tursiops truncatus*) in the Turtle/Brunswick River Estuary (TBRE), a nationally designated Superfund site, have been shown to have been exposed to very high levels of POPs, such as the polychlorinated biphenyl (PCB) congener Aroclor 1268, which was produced at this site[6]. Sapelo Island, GA, about 40 miles north of the TBRE, has minimal urbanization, so was originally used as a control site in studies of exposures in the TBRE[27]. However, recent research shows bottlenose dolphins at Sapelo Island also have elevated PCB levels, leading some to hypothesize that fish off Sapelo Island may similarly have high exposure levels to many chemicals which may transfer through the local food web[6].

The Georgia Aquarium led a study in which three different species of fish were caught at four different sites off the coast of Sapelo Island. One-hundred forty-four different analytes were measured in order to compare concentration levels of various POPs across sites and species. Over 60% of the measured data points are below the limit of detection; an accurate estimate for these analytes are unknown, and the exposure level is between zero and that limit (yielding a high proportion of left-censored data). Current statistical methods to handle left-censored data include substituting one-half the limit of detection or the limit of detection divided by the square root of two for all left censored data points. However, this method is not reliable with such a large amount of left-censored data. Other statistical methods to handle left-censored data include maximum likelihood estimation and multiple imputation. Little work has been done to compare these statistical methods in data sets with such high percentages of concentration measurements below the limit of detection. This project was conducted in the spirit of One Health, an

interdisciplinary effort to address public health issues through collaborations between animal, human, and environmental health[25]. The goal of this thesis was to examine these different statistical methods of comparing analyte concentration levels between sites and species with high levels of left-censored data.

2. Background

Persistent organic pollutants (POPs), including pesticides, polychlorinated biphenyls (PCBs), polybrominated diphenyl ethers (PBDEs), and dichlorodiphenyltrichloroethane (DDTs), are fat soluble compounds that are resistant to chemical and biological degradation, widely distributed, and bioaccumulate in food webs[5]. Although many have been banned or out of production for many years, concentrations are still found at toxicologically relevant levels, and thus are still pervasive threats to marine animal health[11]. For example, DDT was previously used in the United States as a pesticide but banned in this country in 1972; however, it is still used in some African countries to control malaria[3]. Similarly, the manufacture of PCBs, used primarily as coolants and lubricants in electrical equipment, was halted in the United States in 1977 because of evidence that they harmed the environment and human health[2,10]. Pesticides such as mirex, used for fire ant control in the Southeastern U.S., and aldrin/dieldrin have also been out of production in the United States since the 1970s[1,4]. However, use of aldrin and dieldrin for termite control remained persistent until 1987[4]. Finally, PBDEs are lipophilic compounds that are used primarily as flame retardants, as well as in epoxy resins, polyesters, and textiles[8]. All of these compounds can have widespread impacts on environmental, human, and animal health[1,2,3,4].

Oceans often act as sinks for these POPs from wastewater outfalls (sewage treatment plants), atmospheric deposition, and downstream runoff from industrial and urban areas[5]. Higher

levels are typically associated with industrial areas, as POPs can get into the environment via emissions to the air and surface waters from manufacturing plants and release during the life cycle of consumer products[5]. The Turtle/Brunswick River Estuary (TBRE) in Brunswick, GA was designated as a National Priority List (Superfund) site in 1996 due to extensive environmental contamination, specifically from the PCB mixture Aroclor 1268[6]. Aroclor 1268 was the most highly chlorinated Aroclor (the trade name for combinations of various congeners of PCBs) manufactured and was used as a fire retardant in the navy and in diverse industrial applications[10]. Striped mullet (*Mugil cephalus*) and spotted sea trout (*Cynoscion nebulosus*) in the TBRE had PCB concentrations three times higher than levels in fish 100 km north[6]. In addition, bottlenose dolphins (*Tursiops truncatus*) in the TBRE have been shown to have been exposed to some of the highest levels of PCBs near this Superfund site compared to other regions sampled along the Atlantic seaboard[6].

Marine mammals can be used as bio-monitors to indicate the presence and levels of POPs in the coastal environment. For example, bottlenose dolphins can be used to reflect the POP contamination in that area, and to establish geographic trends of environmental contamination because they live in population subgroups that demonstrate high site fidelity for coastal embayments[17]. In addition, they accumulate POPs in their lipid-rich blubber throughout their lives, and biomagnification occurs as they are a top-level marine predator[26]. POPs can be damaging to bottlenose dolphins because research shows such exposures can lead to compromised immune systems, increased disease susceptibility, and negative effects on reproduction, including a delay in the first reproductive event and death of fetuses during gestation[26]. In 1987-1988, there was a mass mortality event of coastal bottlenose dolphins caused

by *Morbillivirus*-induced disease along the mid-Atlantic coast of the US, and contaminant-induced immunosuppression was hypothesized as a contributing factor[14].

Study site

Sapelo Island, GA, located about 40 miles north of the TBRE, is mainly undeveloped and is composed mostly of salt marsh and sand; ninety-seven percent of the island consists of nature preserves owned by the state of Georgia[27]. The only private land on the island is inhabited by a Geechee/Gullah community, descendants of slaves from a large plantation on the island[27]. The Geechee/Gullah community relies on local seafood as a staple in their diet, so any potential human health effects of these POPs is of concern[9]. In humans, current epidemiological and experimental evidence suggests that background exposure to POP mixtures can result in an increased risk of Type II Diabetes (T2D)[19]. In addition, there is some evidence in animal studies that low exposure to POPs can cause obesity, although this relationship has not been shown to be consistently true in humans[19]. Epidemiological studies in humans have shown associations between long-term exposure to POPs and high cholesterol, reproductive impairment, thyroid disorders, and weakened immune systems[9]. In addition, POPs have been shown to compromise liver function in rats, causing fat accumulation, lipid toxicity, and nonalcoholic fatty liver[19]. It is possible that in the future, more strict advisories on seafood consumption may exist due to the mixture of contaminants found in fish[9].

The Sapelo Island National Estuarine Research Reserve (SINERR) is a state-federal partnership between the Georgia Department of Natural Resources and the National Oceanic and Atmospheric Administration[23]. Given that Sapelo Island has minimal urbanization, it was

originally used as a control for the TBRE, and nearby bottlenose dolphins were examined for POP levels. However, Balmer et al. demonstrated that bottlenose dolphins in SINERR also have elevated PCB and Aroclor 1268 levels, a somewhat surprising result, based on the undeveloped nature of Sapelo Island[6]. Since Aroclor 1268 is hydrophobic, water transport is unlikely, and it is likely that sediment or prey from the TBRE is contaminated and is transferred to the Sapelo Island area[6]. This evidence suggests that Aroclor 1268 contamination extends further outside the TBRE than previously documented, although the exact route of transport has yet to be determined. The PCB profiles between Sapelo Island and TBRE are more similar to that observed in fish tissues than for sediment, indicating that fish (and their tissues) are a more likely source of transport than sediment[30]. Still, in order to further document potential transfer from fish to bottlenose dolphins, more research needs to be done to examine the predator/prey associations between the bottlenose dolphins and their lower trophic level prey, such as mullet.

The goal of the present study was to monitor the prey of bottlenose dolphins near Sapelo Island, GA to assess POP levels in order to determine the source of high PCBs (including Aroclor 1268) in the blubber of Georgia animals. Fish were caught off the coast of Sapelo Island at four different sites – Bell Marsh Road (1), Cabretta (2), UGA (3), and the Main House (4) (Figure 1). Three different species of fish were caught and examined at each site – silver perch/yellowtail (*Bairdiella chrysoura*), sea trout, and mullet. Bottlenose dolphins are opportunistic predators and there is large geographic variation in their diet[7]. The three specific fish species above are all prey of the bottlenose dolphin, with mullet being the most commonly consumed[7].

Figure 1: Map of Sapelo Island, GA, with indications of the four sites where fish were caught for this study[24].

We were primarily interested in examining whether there are differences between the four different sites and the three different species for each chemical class (DDTs, PCBs, PBDEs, pesticides, and metals), with a specific focus on the PCB congener Aroclor 1268, which is specific to the Brunswick Superfund site. While site and species differences were both taken into consideration, species differences were of particular importance and interest. It may be difficult to distinguish differences between sites, given the similarity in physical characteristics (including the possibility of mixing of water and species due to both fish and water movement) of the coastal sites (sites 2, 3 and 4).

Sample collection and processing

The Georgia Aquarium was responsible for all methods related to study design and sample collection and processing. For each site/species combination, three composites of fish

were caught by local fisherman on Sapelo Island. Each composite is made up of three to five fish that were combined before concentrations of environmental exposures were measured. Mullet between 15-30 centimeters (total length), spotted sea trout 33 centimeters (total length) or larger, and silver perch between 7-15 centimeters (total length) were caught, wrapped in foil, and placed in plastic bags before being placed on ice. Care was taken to ensure that the buckets that the fish were placed in were free of any contaminants that could interfere with study results. Fish were then sent to the National Oceanic and Atmospheric Administration (NOAA) laboratory in Charleston for homogenization, extraction, and assays to measure POPs. The sampling scheme is shown in Table 1. At site number one, only mullet was caught, and at site number two, only silver perch and mullet were caught as a result of fishermen eating the fish instead of providing them to the research study. This provides direct links between this study and human health. Only sites numbers three and four have data for all three species of fish, leading to nine different site/species combinations.

Table 1: Distribution of samples that were collected for our study. Colored squares indicate samples that have been collected.

SILVER PERCH/YELLOWTAIL														
SITE 1 (Bell Marsh Road)														
Composite 1					Composite 2					Composite 3				
1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
SEA TROUT														
SITE 1 (Bell Marsh Road)														
Composite 1					Composite 2					Composite 3				
1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
MULLET														
SITE 1 (Bell Marsh Road)														
Composite 1					Composite 2					Composite 3				
1	2	3	4	5	1	2	3	4	5	1	2	3	4	5

Site 2 (Cabretta)														
Composite 1					Composite 2					Composite 3				
1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Site 2 (Cabretta)														
Composite 1					Composite 2					Composite 3				
1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Site 2 (Cabretta)														
Composite 1					Composite 2					Composite 3				
1	2	3	4	5	1	2	3	4	5	1	2	3	4	5

Site 3 (UGA)														
Composite 1					Composite 2					Composite 3				
1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Site 3 (UGA)														
Composite 1					Composite 2					Composite 3				
1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Site 3 (UGA)														
Composite 1					Composite 2					Composite 3				
1	2	3	4	5	1	2	3	4	5	1	2	3	4	5

Site 4 (Main House)														
Composite 1					Composite 2					Composite 3				
1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Site 4 (Main House)														
Composite 1					Composite 2					Composite 3				
1	2	3	4	5	1	2	3	4	5	1	2	3	4	5
Site 4 (Main House)														
Composite 1					Composite 2					Composite 3				
1	2	3	4	5	1	2	3	4	5	1	2	3	4	5

Data Description

Data points are available for a total of 27 composites. In each composite, the level of 144 different analytes (chemicals) was measured. The analytes are in the following chemical classes: DDTs, PCBs, PBDEs, pesticides, and metals (see Appendix I for a full listing of all measured analytes). The organics (DDTs, PCBs, PBDEs, and pesticides) were all measured in ng/g, while the inorganics (metals) were measured in μg/g. Of note, aluminum, zinc, and iron had much higher measurements than the other metals, so they were analyzed as a separate group. Congener Arcoclor 1268 was singled out as an analyte of focus given its significance to the Brunswick Superfund site. According to Kucklick et al., Aroclor 1268 is composed of PCB congeners 201, 180+193, 207, 194, 202, 187, 196, 199, 208, 209, and 206[17]. PCBs 196 and 199 were not measured in this study, so they are excluded from the summed Aroclor 1268 concentration here. PCBs 196 and 199 compose 7.1% and 9.1%, respectively, of total PCBs in sediment adjacent to the Aroclor 1268 Brunswick Superfund site[17].

Approaches for handling left-censored data

Preliminary analysis of the data shows that 61.29% of all data points fall below the limit of detection (left-censored). For example, of the six different DDT analytes measured, the majority are reported as less than the limit of detection, with the exception of 4, 4'-DDE, for which all data points were reportable and above the detection limit (Figure 2). In some cases, reportable data are less than the detection limit recorded. This is a result of the limit of detection being a function of the mass of the sample. One hundred-ten PCB analytes were measured, and

there is a large distribution of left censored versus reportable data (Figure 3). The majority of PBDEs were found to be below the limit of detection, with the exception of PBDE 47, for which ten out of 27 total data points were reportable (Figure 4). Some pesticides, such as aldrin, only have data points that are below the detection limit; other pesticides, including cis- and trans-nonachlor yield all reportable data (Figure 5). The majority of metals yield reportable data, including aluminum, iron, and zinc, which had considerably higher concentrations than any other metals. Only a few metals, including silver, uranium, and beryllium, had all data points below the limit of detection (Figure 6).

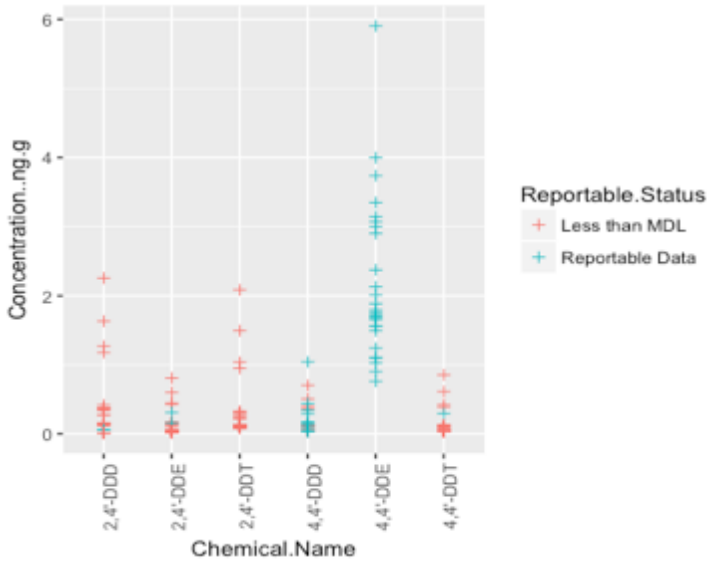


Figure 2: A comparison of data points that are detectable versus those that are below the detection limit for all six DDTs measured.

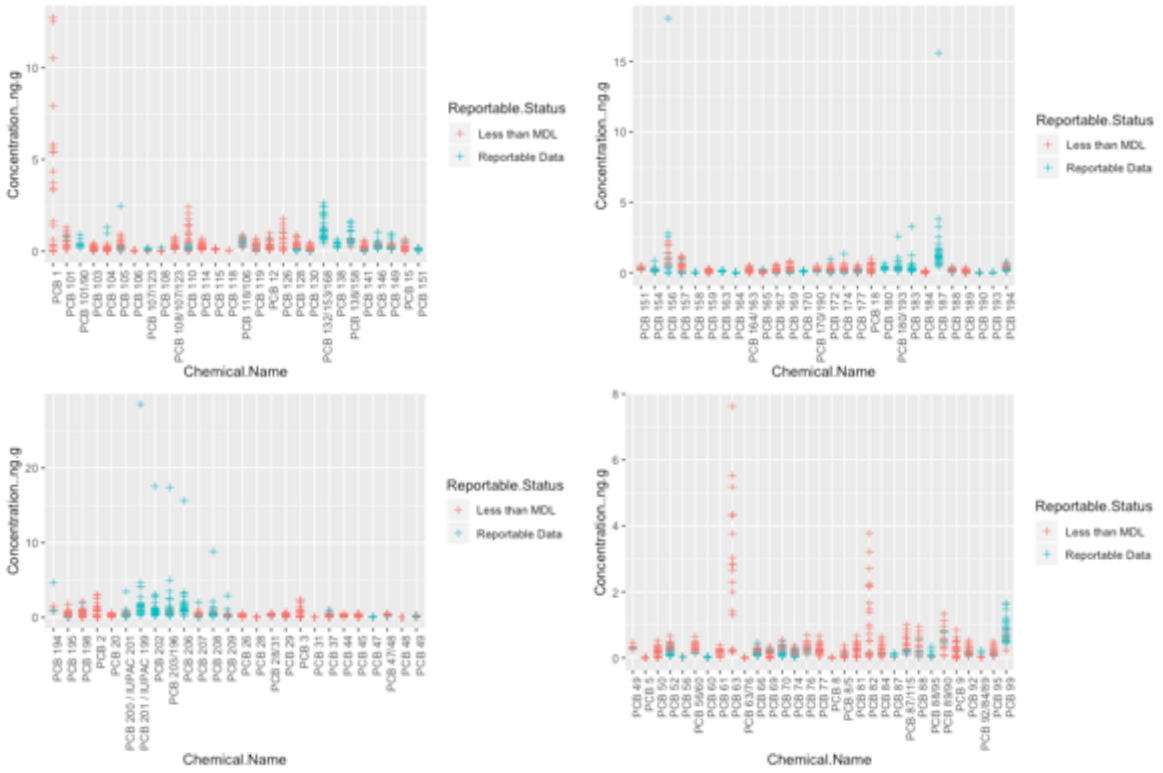


Figure 3: A comparison of data points that are detectable versus those that are below the detection limit for all PCBs measured.

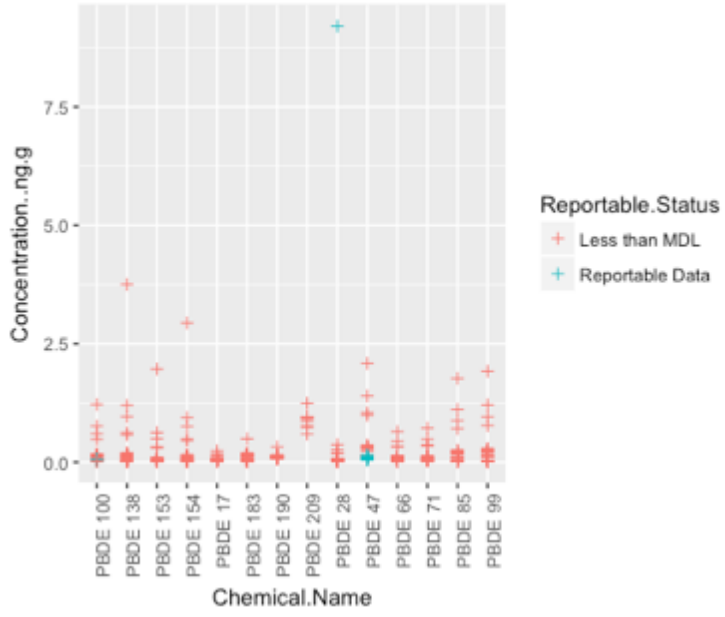


Figure 4: A comparison of data points that are detectable versus those that are below the detection limit for all PBDEs measured.

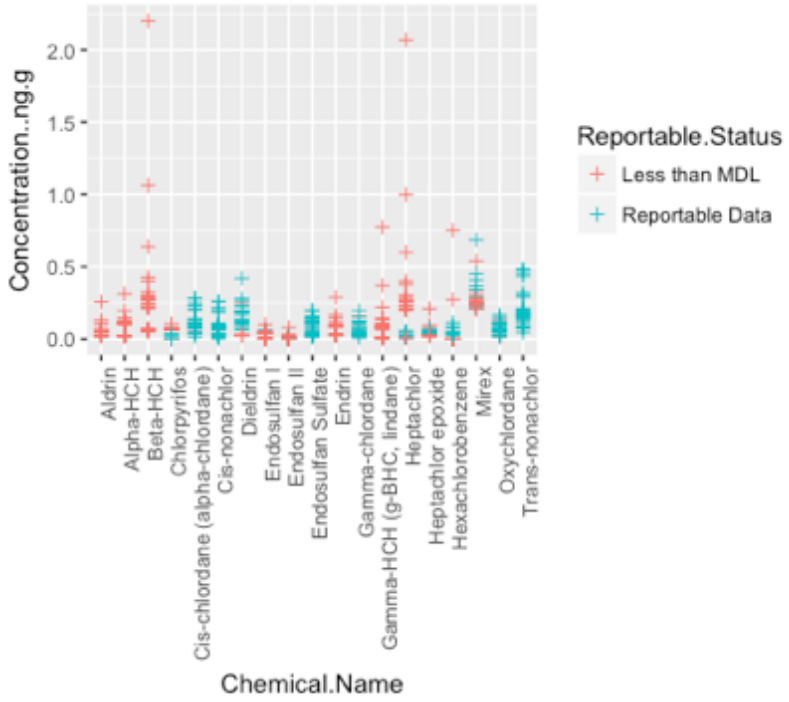


Figure 5: A comparison of data points that are detectable versus those that are below the detection limit for all pesticides measured.

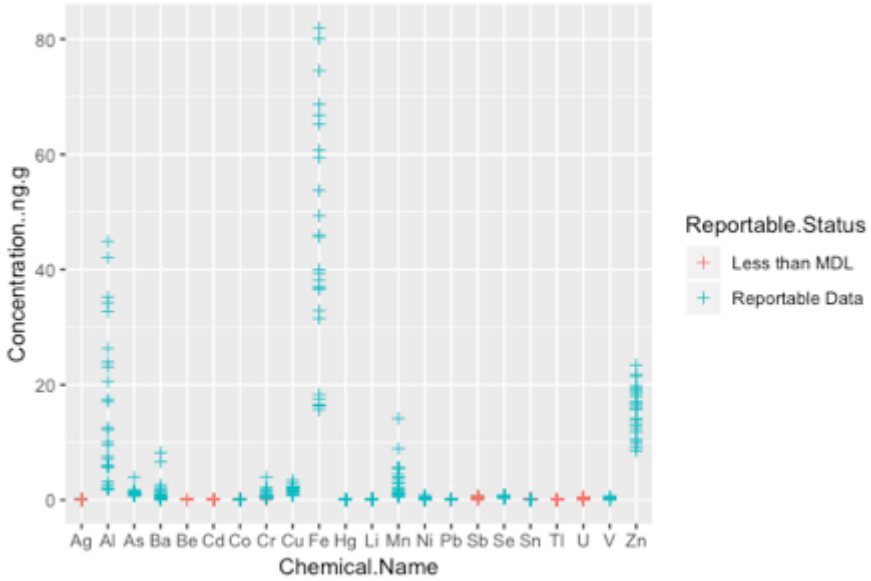


Figure 6: A comparison of data points that are detectable versus those that are below the detection limit for all metals measured. It can be seen that aluminum (Al), iron (Fe), and zinc (Zn) generally have higher concentrations than the other metals that were measured.

With such a large amount of left-censored data, it is difficult to estimate a mean and standard deviation for any given analyte, site, and species combination. For this project, a mean

for each analyte, site, and species combination will be estimated, then each of those means will be summed within each chemical class. The limit of detection is determined as a function of the error involved in measuring the level of each analyte[29]. It is determined for a specific assay by measuring samples with known analyte levels (such as blanks) in series to determine the standard deviation of any background level (noise)[29]. Analyte values that are considered to be below the limit of detection indicates that any level of the analyte that cannot be confidently distinguished from instrument noise; there may or may not be some level of that analyte present. The limit of detection varies from sample to sample because it is calculated as a function of the mass of the sample (as well as the fixed instrument detection limit). Since the mass varies, the amount of sample will affect the calculated detection limit, leading to different limits for each sample. As a result, we see some reportable observations (above the detection limit) from a given analyte that are lower than observations that are below the detection limit for the same analyte in the figures above.

In practice, the most commonly used methods to handle left-censored data involve substitution of either one-half the detection limit, the detection limit divided by the square root of two, or the detection limit itself[16]. This substitution approach is not an ideal method, as it imposes patterns on the data that may not reflect the original pattern of the outcome of interest[16]. Substitution is sometimes employed when relatively few data points are below the limit of detection but would not be an ideal method for our data as over 60% of the data points are left-censored. Maximum likelihood estimation can also be used to estimate the mean and variance of left-censored data, but does require a distributional assumption of the data, typically a log-normal distribution for exposure concentrations[22]. In addition, maximum likelihood estimates often have unstable and inestimable standard errors with large amounts of left-censored data.

Turnbull proposed a non-parametric estimator for the mean of left-censored data, also known as the reverse Kaplan-Meier estimator; that estimates the right-continuous cumulative distribution function using a product-limit estimator[12,13]. All reportable (non-censored) values, x_j, are ordered from smallest to largest; n_j is defined as the number of values (either censored or not) less than or equal to x_j and d_j is the number of non-censored values equal to x_j[12]. For each x, Turnbull's estimator is the product of all $\frac{(n_j - d_j)}{n_j}$ for all $x_j > x$[12]. This results in a step function that "jumps" at each non-censored (reportable) data point[12]. However, this approach does not converge when all of the data points in a given group are below the detection limit, which happens in over half of the combinations of site, species, and analytes in our dataset. Multiple imputation is another technique to handle incomplete data in which each value below the limit of detection is replaced by *m* appropriate values (imputed from the range of the concentration levels) resulting in *m* complete datasets[15]. Each complete (imputed) dataset is analyzed, resulting in *m* mean estimates, which can be combined to create a final estimate[15]. Standard multiple imputation requires the assumption that the data are missing at random and requires some data to be reportable[15]. In general, it is difficult to statistically analyze a dataset where such a large percentage of the data points are below the limit of detection, and it is unclear which of these methods, if any, provide the best estimates of means and standard errors in our setting.

3. Methods

In addition to this dataset having a majority of the data points left-censored, there was also a very small sample size; each combination of site, species, and analyte combination only contained three data points. These issues made it difficult to use some of the existing methods for estimating means with left-censored data, such as Turnbull's estimator, because of situations

when all three of these data points were below the limit of detection. All of the analytes except for 30 had at least one missing data point, and the majority were missing at least half of their measurements (Appendix II). We simply do not have enough data to apply a number of our current statistical methods in a stable manner to the context of our problem. In addition, we do not have complete data to compare across all three species and all four sites. We are missing data from two of the three species (silver perch and sea trout) at site 1, and from one species (sea trout) at site 2. Since we have no data to inform us about the analyte concentrations at these sites, we will only focus on comparing sites 3 and 4 for our analysis.

Estimating the mean of the cumulative concentrations within each chemical class presents two different problems: first, we must estimate the mean of each individual analyte at a given site/species combination. Given the large amount of left-censored data, this estimation is not straightforward. Second, after estimating the mean, we must sum all of the individual analytes within a given class to perform a two-way analysis of variance (ANOVA); in this step, we have already addressed the left-censoring problem, so no further adjustments are necessary. The summing of pollutants within a chemical class are justified because, with the exception of the metals class, all of the analytes within each class are measured on a similar scale. In addition, all analyte measurements within a given chemical class are assumed to be independent of each other. Within the metals, three of them, aluminum, zinc, and iron, are present in much larger concentrations than the other metals. This makes summing all metals together a less reliable exercise than in the other classes, so we will analyze all metals with the exception of aluminum, zinc, and iron, which will be analyzed separately from the others. In total, three groups of metals will be analyzed: metals group 1 (all metals), metals group 2 (aluminum, iron, and zinc only), and metals group 3 (all metals excluding aluminum, iron, and zinc).

An ideal method to estimate the mean level of each analyte at each of the site/species combinations is Turnbull's estimator, a non-parametric method to estimate descriptive statistics involving left-censored data[12, 13]. However, Turnbull's estimator will not converge when *all* of the values are left-censored. This occurs in 725 of the 1320 total site/species/analyte combinations in our data. Only 595/1320 (45.076%) have at least one non-censored data point, and the mean for this site/species/analyte combination can be estimated using a non-parametric approach. Unfortunately, as noted above, we simply do not have enough data to support conducting a separate analysis for each site/species/analyte combination individually.

Maximum Likelihood Estimation

Instead, we use a maximum likelihood estimator to estimate the mean[22]. If there are n measurements, and $m < n$ of those measurements are left-censored at a limit of detection (LOD) c, then the likelihood generally follows the form

$$\prod_{i=1}^{m} \Pr(X < c_i) \prod_{i=m+1}^{n} f(X = x_i)$$

.

In our case, the probability density function for the log-normal distribution is:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}$$

and the cumulative density function is:

$$F(X) = \Pr(X < c) = \Phi\left(\frac{\ln x - u}{\sigma}\right)$$

.

In the above formulas, x represents the measured analyte level, μ represents the mean of a given analyte, σ represents the standard deviation for that analyte, and c represents the limit of detection. The likelihood that we aim to maximize is given by:

$$\prod_{i=1}^{m} \Phi\left(\frac{\ln c_i - u}{\sigma}\right) \prod_{i=m+1}^{n} \frac{1}{x_i \sigma \sqrt{2\pi}} e^{-\frac{(\ln x_i - \mu)^2}{2\sigma^2}}$$

where c=LOD is different for each sample. A limit of this approach is that is requires a distributional assumption, typically the log-normal for exposure data as assumed above, in which the log-transformed concentrations approximate a normal distribution. To see if this was a good fit, we constructed Q-Q plots, substituting in one-half the limit of detection (Figure 7) and the limit of detection itself (Figure 8) for the left censored data points. These plots are similar to each other and show deviations from the normal line for all of the chemical classes. Since the limit of detection varies from sample to sample, using a truncated version of the log-normal is not possible. While the Q-Q plots of the observed data points for each chemical class display tails that tend to verge very far from the log-normal line (Figure 9), this can be misleading because the plots ignore all of the censored data. In addition, because we are working based off of chemical classes, each site/species/analyte grouping likely follows a slightly different model, meaning that the Q-Q plots generally may not represent the grouped data well. We simulated some data from a log-normal distribution with the same mean and variance of each chemical class, and the Q-Q plots still have wide tails (Appendix III).

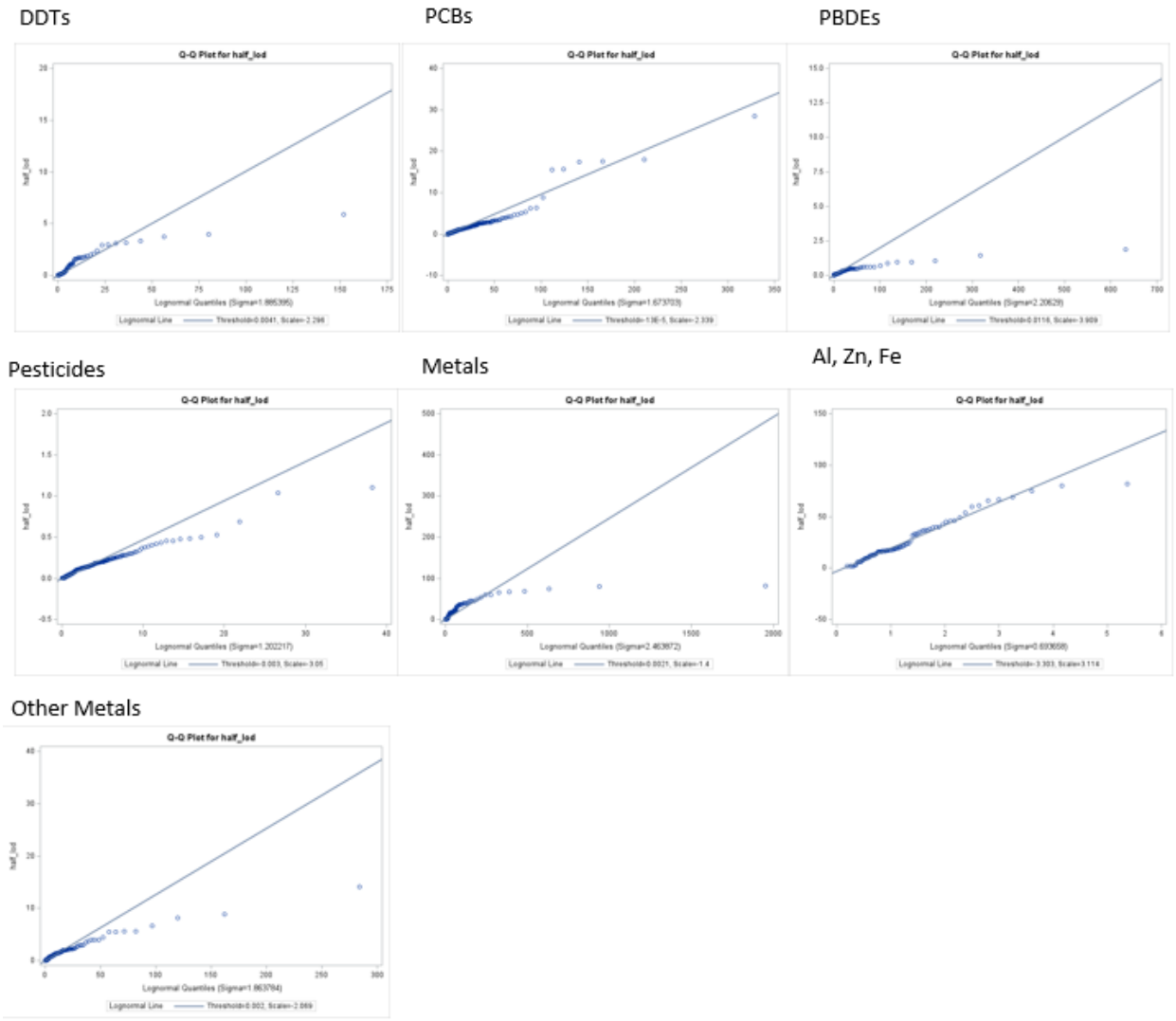


Figure 7: Q-Q plots substituting one half the detection limit for left censored data, plotted against the log-normal line

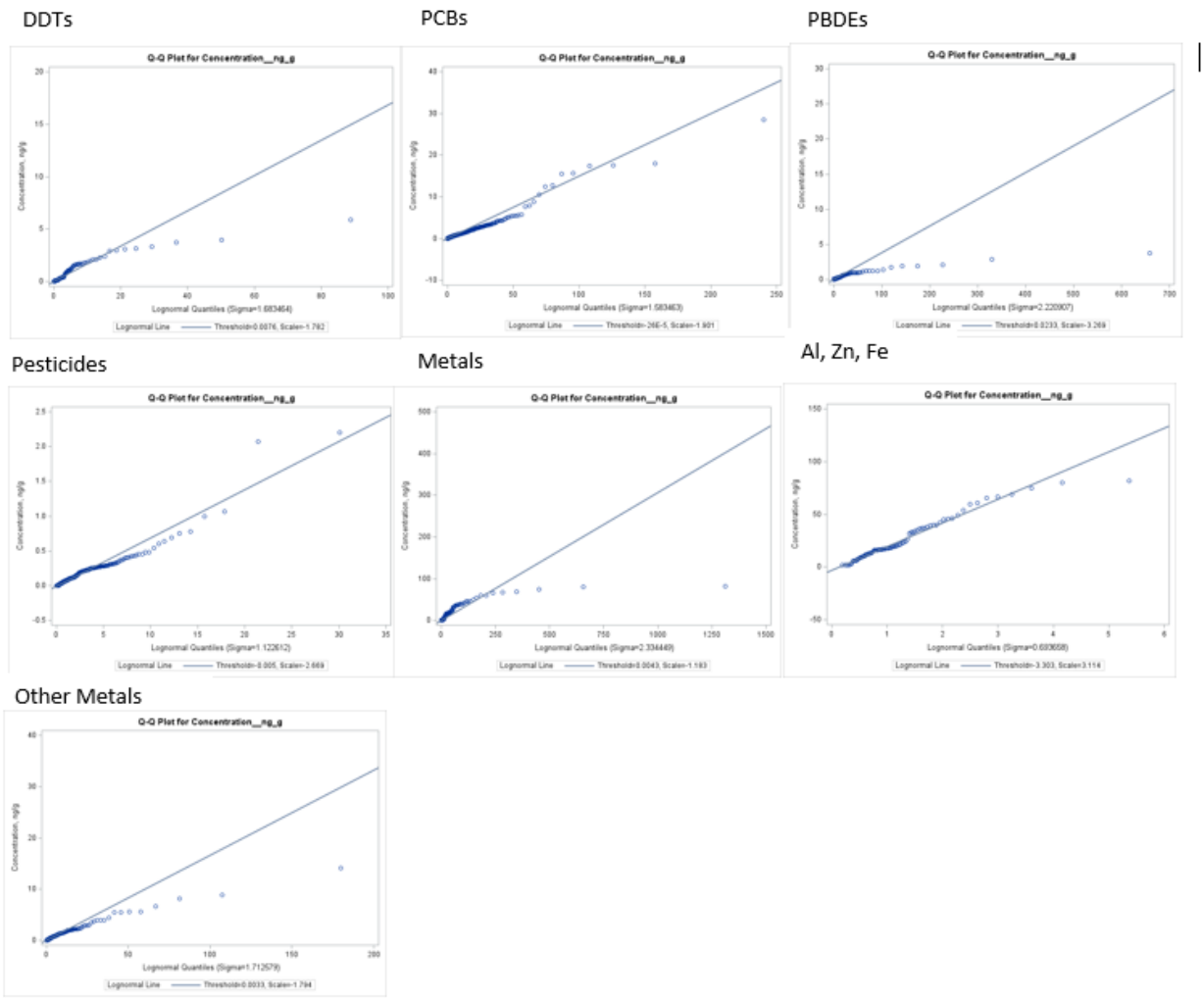


Figure 8: Q-Q plots substituting the detection limit for left censored data, plotted against the log-normal line

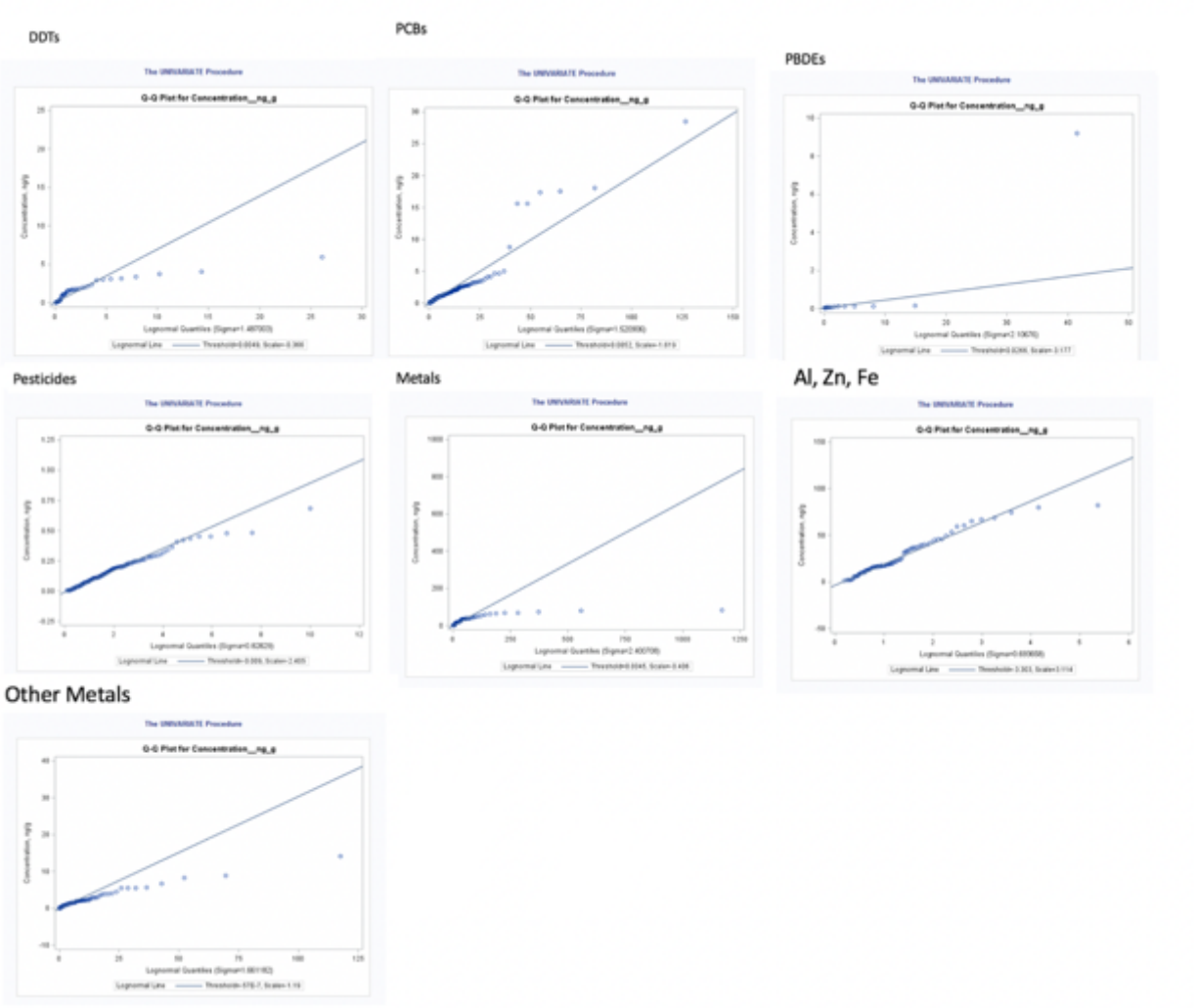


Figure 9: Q-Q plots using the observed data only, plotted against the log-normal line

Then, we attempted to fit a linear regression model using effects coding to account for the site and species:

1] $Y = \mu + \alpha_1 X_1 + \alpha_2 X_2 + \beta_A Z_A + \gamma_{1A} X_1 Z_A + \gamma_{2A} X_2 Z_A + \varepsilon$, where:

Y = concentration

$$X_1 = \begin{cases} 1 \; if \; species = mullet \\ 0 \; if \; species = sea \; trout \\ -1 \; if \; species = silver \; perch \end{cases}$$

$$X_2 = \begin{cases} 1 \; if \; species = sea \; trout \\ 0 \; if \; species = mullet \\ -1 \; if \; species = silver \; perch \end{cases}$$

$$Z_A = \begin{cases} 1 \; if \; site = 3 \\ -1 \; if \; site = 4 \end{cases}$$

In the linear model using effects coding, μ represents the overall mean, which in this case in the overall mean concentration of an analyte across all sites and species. Each α represents the difference between the mean of that species and the overall mean. For example, α_1 is the difference between the mean analyte concentration for mullet and the overall mean. Similarly, α_2 is the difference between the mean analyte concentration for sea trout and the overall mean, and β_A is the difference between the mean analyte concentration for site 3 and the overall mean. The maximum likelihood parameter γ_{1A} is the mean analyte concentration for mullet at site 3, minus the mean analyte concentration for mullet, minus the mean analyte concentration for site 3, plus the overall mean. Likewise, the maximum likelihood parameter γ_{2A} is the mean analyte concentration for sea trout at site 3, minus the mean analyte concentration for sea trout, minus the mean analyte concentration for site 3, plus the overall mean. Finally, ε represents the error term.

We fit this linear regression model for each chemical class, substituting half the detection limit for censored data, and observed the residual plots that result (Appendix IV). The plots suggest a good fit for log-normal data for DDTs, PBDEs, and pesticides, but not for PCBs and metals. The same was true when the detection limit itself was used for censored data (Appendix V). Therefore, using a log-normal distribution and the maximum likelihood approach for DDTs, PBDEs, and pesticides appears to be justified. However, due to limitations of using Q-Q plots with censored data noted above, prior literature for such concentrations assuming a log-normal

distribution[17, 22, 26], and the lack of an alternative distribution, the log-normal was also used for metals and PCBs as a baseline.

We used PROC NLMIXED in SAS® to evaluate the maximum likelihood estimates for the parameters in the linear regression model listed above: μ, α_1, α_2, β_A, γ_{1A}, and γ_{2A} (see Appendix VI for example SAS® code). An estimate for σ^2 is also given by SAS® PROC NLMIXED. We defined the outcome as the log-transformed concentration or detection limit where appropriate. This was done for each analyte individually. From there, the mean concentration for each analyte can be determined for each site and species combination.

To determine the mean concentration for each composite within each analyte, we used the linear regression model, the calculated maximum likelihood estimates for the parameters, and the effects coding (note that these results are still on the log scale):

$E[Y \mid \text{mullet, site 3}] = \gamma_{1A} + \mu + \alpha_1 + \beta_A$

$E[Y \mid \text{sea trout, site 3}] = \gamma_{2A} + \mu + \alpha_2 + \beta_A$

$E[Y \mid \text{silver perch, site 3}] = \mu - \alpha_1 - \alpha_2 + \beta_A - \gamma_{1A} - \gamma_{2A}$

$E[Y \mid \text{mullet, site 4}] = \mu + \alpha_1 - \beta_A - \gamma_{1A}$

$E[Y \mid \text{sea trout, site 4}] = \mu + \alpha_2 - \beta_A - \gamma_{1A}$

$E[Y \mid \text{silver perch, site 4}] = \mu - \alpha_1 - \alpha_2 - \beta_A + \gamma_{1A} + \gamma_{2A}$

After obtaining the mean of each site/species combination for each analyte, we exponentiated them to get the log-concentration back to the original concentration scale. When the outcome is on the log-normal scale, the mean concentration is $e^{(\mu+\frac{\sigma^2}{2})}$. To compare the data between sites and species, all the mean concentrations within each chemical class were summed together[5, 6, 11, 17, 30]. This is valid because the sum of the expected value is the same as the expected value of the sum. As previously mentioned, aluminum, zinc, and iron had consistently

much higher concentrations than the other metals, so while all metals were analyzed together (group 1), separate analyses on just the summed aluminum, zinc, and iron (group 2) concentrations, as well as the summed concentration of all metals excluding aluminum, iron, and zinc (group 3), were performed.

To compute the variance for each site/species/analyte combination, we calculate the variance of the log-normal distribution via: $[e^{\sigma^2} - 1]e^{(2\mu + \sigma^2)}$. This formula gives us the variance of the analyte concentration for that species at that site, rather than the variance of the estimated *mean* analyte concentration. Since the analytes are assumed to be independent of each other, the sum of the variances from each analyte provides the variance for the entire chemical class. To compare the summed concentrations across sites and species, a two-way ANOVA with a randomized block design was used, as there is only one observation per site/species cell[18]. SAS® PROC GLM was used to obtain the relevant p-values for the effect of site and species. Both site and species are treated as fixed effects. Since the variance is not homogenous by cell, the analysis was weighted by the inverse of the variance. The randomized block design has some limitations, namely that it assumes no interaction between site and species and there is no test for this interaction. Tukey's method was used to compare the pairwise mean sums. P-values were taken from the Type III sum of squares table in the output of SAS® PROC GLM.

We are attempting to estimate six parameters in our maximum likelihood model using at most 18 data points (for each analyte, there are three measurements – one from each composite – at each of two sites and for each of three species). In most cases, we do not have enough data for a stable estimate of the standard error of the maximum likelihood estimates, leading to inflated variance estimation of the analytes. There was a very large discrepancy in variance, particularly for metals and PCBs. Stable estimates will only occur when all eighteen measurements for each

analyte have at least one reportable estimate (in other words, the analyte measurement from each composite at each site and each species was above the detection limit). This occurs for only 30 out of the 144 various analytes. Drawing conclusions from our results is difficult because it is not clear that we have enough data to support our conclusions. Therefore, we will also create a maximum likelihood model assuming no interactions between site and species. This is a large assumption but allows us to only estimate four different parameters with our limited data. Therefore, we create model [2]:

[2] $Y = \mu + \alpha_1 X_1 + \alpha_2 X_2 + \beta_A Z_A + \varepsilon$, where Y, X_1, X_2, Z_A, μ, α_1, α_2, β_A, and ε are defined the same as in model [1].

Then, the mean concentration for each site/species combination within each analyte can be determined as above, using the linear regression model, the calculated maximum likelihood estimates for the parameters, and the effects coding (again, note that these results are still on the log scale):

$E[Y \mid \text{mullet, site 3}] = \mu + \alpha_1 + \beta_A$

$E[Y \mid \text{sea trout, site 3}] = \mu + \alpha_2 + \beta_A$

$E[Y \mid \text{silver perch, site 3}] = \mu - \alpha_1 - \alpha_2 + \beta_A$

$E[Y \mid \text{mullet, site 4}] = \mu + \alpha_1 - \beta$

$E[Y \mid \text{sea trout, site 4}] = \mu + \alpha_2 - \beta_A$

$E[Y \mid \text{silver perch, site 4}] = \mu - \alpha_1 - \alpha_2 - \beta_A$

These log-concentration means can be converted to the concentration scale using the same formula, $e^{(\mu+\frac{\sigma^2}{2})}$, as before, and the variance for each can also be determined similarly, using the formula $[e^{\sigma^2} - 1]e^{(2\mu+\sigma^2)}$. A two-way ANOVA using the randomized block design and weighting by the inverse of the variance can be used to determine any differences by site and

species. Tukey's test can be used to test for any significant pairwise differences. This variation on the MLE approach allows us to utilize our limited data to estimate fewer parameters, with the goal of achieving more stable estimates.

Finally, we utilize maximum likelihood to compare exclusively by species. Given that sites 3 and 4 are located relatively close to each other on Sapelo Island and the analyte concentrations are likely similar to each other at those two sites, it is reasonable to combine them. In addition, fish likely move between the two sites. For this method, more data points are used to estimate fewer parameter estimates, which may make our estimates more robust. Therefore, we create model [3]:

[3] $Y = \mu + \alpha_1 X_1 + \alpha_2 X_2 + \varepsilon$, where Y, X_1, X_2, μ, α_1, α_2, and ε are defined the same as in model [1].

Then, the mean log concentration for each site/species combination within each analyte can be determined as above, using the linear regression model, the calculated maximum likelihood estimates for the parameters, and the effects coding (again, note that these results are still on the log scale):

$E[Y \mid \text{mullet}] = \mu + \alpha_1$

$E[Y \mid \text{sea trout}] = \mu + \alpha_2$

$E[Y \mid \text{silver perch}] = \mu - \alpha_1 - \alpha_2$

These log-concentration means can be converted to the concentration scale using the same formula, $e^{(\mu+\frac{\sigma^2}{2})}$, as before, and the variance for each can also be determined similarly, using the formula $[e^{\sigma^2} - 1]e^{(2\mu+\sigma^2)}$. Then, we can compare these three mean summed concentrations to determine if there are any differences between species ignoring the distinction between sites 3 and 4. Unfortunately, we cannot obtain a p-value for these analyses, since we are

comparing only one observation per cell. However, general trends can be examined to determine the relationship between the analyte levels across species.

Extreme Observations

In order to maximize data support, we also examined outliers in our data. An outlier was defined as have a reportable analyte concentration level that was more than three standard deviations away from the mean concentration level (substituting in the limit of detection for values that were left-censored) of that analyte. There is a trade-off here between the quality of our data and the quantity; we want as much data as possible to inform our maximum likelihood estimates and provide stability. However, we also want to ensure that the data we are using to inform our estimates falls inside a standard range of values. Although these measurements are valid and may not be due to measurement error, it is important to ensure that they are not too different from the other measurements. Using this method of identifying outliers, we eliminated three data points from the metals chemical class, one each from manganese (mullet from site 4), chromium (silver perch from site 4), and arsenic (silver perch from site 4); one data point from the pesticides chemical class (mirex), which was a silver perch from site 3; one data point from the DDTs chemical class (4, 4'-DDD), which was a mullet from site 4; no data points from the PBDE chemical class; and 19 data points from the PCBs chemical class, one each of PCB 104, PCB 146, PCB 149, PCB 154, PCB 156, PCB 172, PCB 174, PCB 180/193, PCB 183, PCB187, PCB 194, PCB 200/IUPAC 201, PCB 201/IUPAC 199, PCB 202, PCB 203/196, PCB 206, PCB 207, PCB 208, and PCB 209. All PCB outliers were silver perch, and all except PCB 104 were from site 4. Additionally, all PCB outliers, with the exception of PCB 104, were from a single

composite. We then refit models [1], [2], and [3], to this reduced data set to compare these estimates to those obtained from using the full data. It is interesting to note that the majority of outliers were silver perch from site 4, and there were no outliers that were sea trout.

The maximum likelihood approach is not a perfect solution for our data analysis, because it does not allow us to compare all four sites and all three species in one model. However, it is currently our best option for incorporating all of the data that we have. A different model can be used if we can assume no interaction between site and species or if we are not interested in comparing between sites. The maximum likelihood estimation method has some limitations – it assumes that the log-normal distribution is appropriate for our outcomes, which may not be the case. In particular, it has been shown that the log-normal distribution may not be a good match for PCBs and metals, leading us to question if assuming the log-normal is a good idea. However, because of the small sample size, it is not logical to attempt to customize the distribution any further, and the log-normal assumption provides a set of baseline results for future comparisons with more complicated models. In addition, it is not clear if we have enough data for maximum likelihood to work well, and therefore some sacrifices in our analysis must be made. For example, we lose information when we ignore interactions between site and species and combine the data from sites 3 and 4.

4. Results

In general, and as we might expect, we see an increase in the number of stable estimates as the number of parameters we are estimating decreases. For example, in model [1], 40 analytes lead to stable maximum likelihood estimates. In model [2], there are 57, and in model [3], there are 60. When we remove the outliers, these numbers change to 42 stable estimates for model [1],

58 stable estimates for model [2], and 70 stable estimates for model [3]. Only ten total PBDE measurements are reportable, so there is not enough reportable data in each site/species combination to result in stable estimates for the analytes in this chemical class. Every other chemical class has at least one stable estimate. This shows that estimating fewer parameters, using data without any outliers will increase the stability of our estimates. Although this is interesting and important to note, it still does not lead to stable estimates for the majority of analytes. This leads us to conclude that there is simply not enough data available to fully monitor the prey of bottlenose dolphins near Sapelo Island, GA and assess their POP levels. Results presented below include estimated sums from all analytes, even those that did not have stable estimates.

DDTs

Preliminary graphical analysis, substituting in the limit of detection for left-censored data points, appears to show that the summed DDT concentration is lowest in sea trout and highest in silver perch; in addition, all three species appear to have densities that are right skewed (Figure 10, Appendix VII). A similar distinction for site in not seen; it is not obvious which of the four sites tend to have higher summed concentrations of DDTs (Figure 11, Appendix VIII). These graphs provide a nice visual representation of our data; however, given that they do not take censoring into account, they may not be accurate. In model [1], sea trout appears to have a lower summed DDT concentration at both site 3 and 4 as compared with mullet and silver perch (Table 2). Mullet has the highest summed DDT concentration at both sites, and sea trout has the lowest.

Table 2: Results of model [1], by chemical class. Values are displayed as mean (variance).

	DDTs	
	Site	
Species	3	4
Mullet	2.92 (0.90)	3.90 (1.48)
Sea Trout	1.59 (0.26)	1.24 (0.17)
Silver Perch	2.30 (0.64)	2.28 (0.55)

	PBDEs	
	Site	
Species	3	4
Mullet	0.30 (0.0003)	0.35 (0.0012)
Sea Trout	0.40 (0.0018)	0.45 (0.0024)
Silver Perch	0.19 (0.0001)	0.52 (0.0010)

	PCBs	
	Site	
Species	3	4
Mullet	19.31 (5.27)	17.10 (20.58)
Sea Trout	14.38 (11.93)	16.09 (6.40)
Silver Perch	29.31 (21.27)	52.20 (68.56)

	Aroclor 1268	
	Site	
Species	3	4
Mullet	7.59 (3.42)	4.48 (1.02)
Sea Trout	6.47 (2.43)	9.14 (4.93)
Silver Perch	7.64 (3.75)	26.42 (41.15)

	Metals Group 1 (All)	
	Site	
Species	3	4
Mullet	77.97 (52.92)	142.55 (221.06)
Sea Trout	35.39 (8.55)	33.72 (7.53)
Silver Perch	77.18 (45.56)	72.11 (36.46)

	Metals Group 3	
	Site	
Species	3	4
Mullet	8.13 (1.50)	18.73 (13.85)
Sea Trout	4.44 (0.33)	4.84 (0.38)
Silver Perch	7.84 (1.04)	9.04 (1.25)

	Metals Group 2 (Al, Zn, Fe)	
	Site	
Species	3	4
Mullet	24.71 (51.42)	49.67 (207.22)
Sea Trout	13.48 (8.22)	12.71 (7.15)
Silver Perch	26.84 (44.53)	25.13 (35.22)

	Pesticides	
	Site	
Species	3	4
Mullet	1.24 (0.02)	1.43 (0.04)
Sea Trout	0.48 (0.004)	0.53 (0.006)
Silver Perch	1.67 (0.04)	1.82 (0.04)

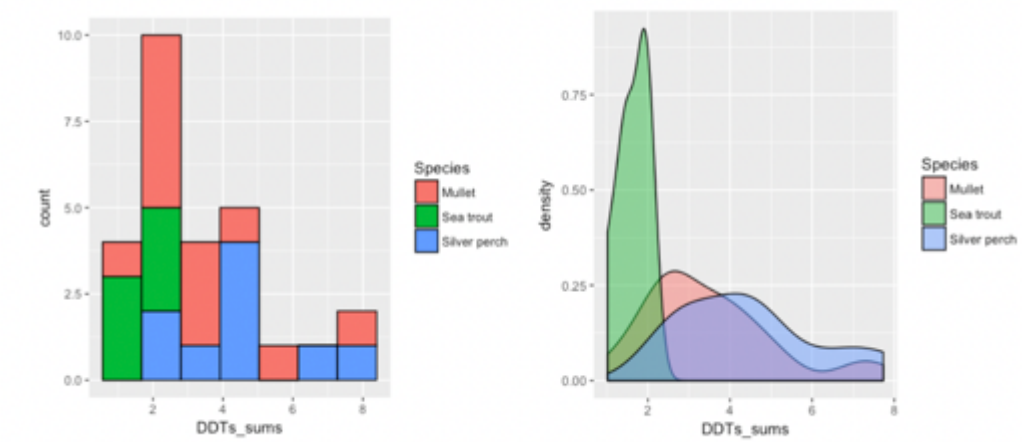


Figure 10: Histogram and density of summed DDT concentrations after substituting the limit of detection for left censored data points, separated by species.

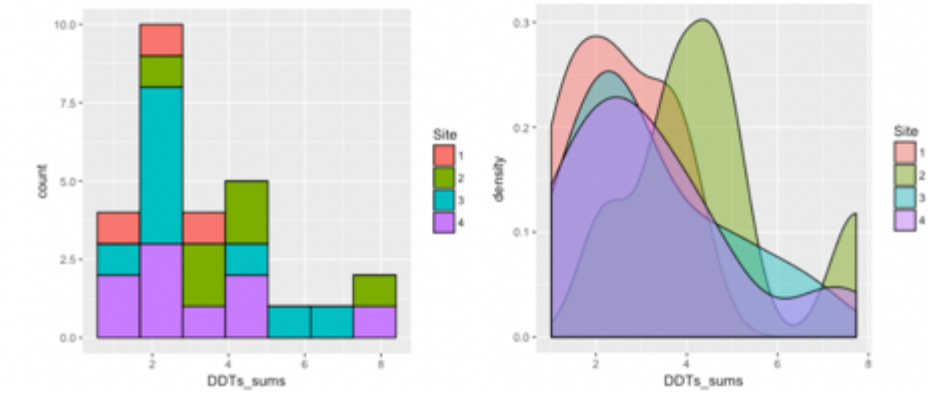


Figure 11: Histogram and density of summed DDT concentrations after substituting the limit of detection for left censored data points, separated by site.

This same pattern is seen in model [2] (Table 3); in model [3], mullet has the highest summed DDT concentration overall, followed by silver perch, then sea trout (Table 4). When one outlier was removed, the same general trends are seen for all three models; mullet has the highest summed DDT concentration overall, followed by silver perch, then sea trout for both sites and when sites are combined (Tables 6, 7, and 8). The variance for DDTs in all three models is not inflated and is reasonable. However, in model [1], the difference in summed mean DDT concentrations was not significantly different for either site or species (Table 5. Model 1: site, p-value = 0.7369, species, p-value = 0.094). In model [2], the difference in summed mean

DDT concentrations was significantly different between species but not site (Table 5. Model 2: site, p-value = 0.094, species, p-value = 0.0103). By Tukey's test, all three pairwise comparisons of species were significantly different from each other (Table 5. Sea trout vs. mullet p-value = 0.011, sea trout vs. silver perch p-value = 0.0267, silver perch vs. mullet p-value = 0.05). When one outlier was removed, the same pattern of significance was seen. In model [1] without outliers, there is no significant difference by site or species (site, p-value = 0.3637, species, p-value = 0.0783). However, in model [2], there is a significant difference across species (site, p-value = 0.7245, species, p-value = 0.0054), with Tukey's test indicating that there is a pairwise difference between all three species (Table 9. Sea trout vs. mullet p-value = 0.006, sea trout vs. silver perch p-value = 0.0135, silver perch vs. mullet p-value = 0.0323).

Table 3: Results of model [2], by chemical class. Values are displayed as mean (variance).

	DDTs	
	Site	
Species	3	4
Mullet	3.4 (1.25)	3.62 (1.49)
Sea Trout	1.48 (0.24)	1.61 (0.33)
Silver Perch	2.58 (0.67)	2.41 (0.65)

	PBDEs	
	Site	
Species	3	4
Mullet	0.4 (0.00081)	0.37 (0.001)
Sea Trout	0.48 (0.0033)	0.34 (0.0013)
Silver Perch	0.42 (0.00053)	0.38 (0.0006)

	PCBs	
	Site	
Species	3	4
Mullet	16.75 (4.25)	21.17 (25.62)
Sea Trout	14.62 (14.62)	17 (11)
Silver Perch	43.8 (40.95)	38.37 (52.3)

	Aroclor 1268	
	Site	
Species	3	4
Mullet	5.23 (2.45)	7.21 (4.49)
Sea Trout	7 (4.8)	9.62 (8.82)
Silver Perch	13.37 (16.53)	18.56 (30.62)

	Metals Group 1 (All)	
	Site	
Species	3	4
Mullet	99.03 (190.75)	114.26 (255.23)
Sea Trout	32.73 (13.73)	37.2 (17.32)
Silver Perch	70.74 (73.78)	80.44 (93.98)

	Metals Group 3	
	Site	
Species	3	4
Mullet	10.75 (5.18)	14.52 (10.41)
Sea Trout	4.15 (0.38)	5.31 (0.67)
Silver Perch	7.59 (1.18)	9.73 (2.24)

	Metals Group 2 (Al, Zn, Fe)	
	Site	
Species	3	4
Mullet	88.28 (185.57)	99.74 (244.83)
Sea Trout	28.58 (13.38)	31.89 (16.65)
Silver Perch	63.14 (72.6)	70.71 (91.74)

	Pesticides	
	Site	
Species	3	4
Mullet	1.41 (0.039)	1.32 (0.037)
Sea Trout	0.59 (0.0066)	0.55 (0.0067)
Silver Perch	1.81 (0.056)	1.72 (0.053)

Table 4: Results of model [3], by chemical class. Values are displayed as mean (variance).

	DDTs	**PBDEs**	**Pesticides**	**Metals Group 1 (All)**	**Metals Group 2 (Al, Zn, Fe)**	**Metals Group 3**	**PCBs**	**Aroclor 1268**
Mullet	3.52 (1.37)	0.43 (0.001)	1.35 (0.038)	106.66 (237.44)	94.01 (228.75)	12.65 (8.69)	63.39 (51.65)	6.27 (3.61)
Sea Trout	1.54 (0.27)	0.26 (0.0007)	0.56 (0.007)	34.97 (16.72)	30.24 (16.13)	4.73 (0.58)	18.3 (26.31)	8.3 (7.05)
Silver Perch	2.51 (0.68)	0.43 (0.0006)	1.79 (0.055)	75.56 (89.93)	66.93 (87.98)	8.63 (1.95)	79.99 (61.58)	16.01 (24.59)

Table 5: P-values from model [1] and [2]. Bolded cells indicate significant differences (p-value < 0.05).

	DDTs	
	Site	Species
Model 1	0.3637	0.0783
Model 2	0.7245	**0.0054**

	Tukey	Model 2	
	Mullet	**Sea Trout**	**Silver Perch**
Mullet		0.006	0.0323
Sea Trout	**0.006**		0.0135
Silver Perch	**0.0323**	**0.0135**	

	Metals Group 1 (All)	
	Site	Species
Model 1	0.7581	0.1174
Model 2	0.1201	**0.0045**

	Tukey	Model 2	
	Mullet	**Sea Trout**	**Silver Perch**
Mullet		0.0066	0.0437
Sea Trout	**0.0066**		0.0087
Silver Perch	**0.0437**	**0.0087**	

	Metals Group 2 (Al, Zn, Fe)	
	Site	Species
Model 1	0.9438	0.1749
Model 2	0.1126	**0.0038**

	Tukey	Model 2	
	Mullet	**Sea Trout**	**Silver Perch**
Mullet		0.0056	0.0387
Sea Trout	**0.006**		0.0071
Silver Perch	**0.039**	**0.0071**	

	Metals Group 3	
	Site	Species
Model 1	0.891	0.1748
Model 2	0.0975	**0.0146**

	Tukey	Model 2	
	Mullet	**Sea Trout**	**Silver Perch**
Mullet		0.022	0.0872
Sea Trout	**0.022**		0.0256
Silver Perch	0.0872	**0.0256**	

	PCBs	
	Site	Species
Model 1	0.5568	0.0757
Model 2	0.3554	**0.0375**

	Tukey	Model 2	
	Mullet	**Sea Trout**	**Silver Perch**
Mullet		0.6857	0.0495
Sea Trout	0.6857		0.0376
Silver Perch	**0.0495**	**0.0376**	

	Aroclor 1268	
	Site	Species
Model 1	0.6068	0.5325
Model 2	**0.0404**	**0.0009**

	Tukey	Model 2	
	Mullet	**Sea Trout**	**Silver Perch**
Mullet		0.0031	0.0009
Sea Trout	**0.0031**		0.0037
Silver Perch	**0.0009**	**0.0037**	

	Pesticides	
	Site	Species
Model 1	0.477	**0.009**
Model 2	0.6895	**0.009**

	Tukey	Model 1	
	Mullet	**Sea Trout**	**Silver Perch**
Mullet		0.02	0.1644
Sea Trout	**0.02**		0.0118
Silver Perch	0.164	**0.0118**	

	Tukey	Model 2	
	Mullet	**Sea Trout**	**Silver Perch**
Mullet		0.0019	0.0191
Sea Trout	**0.0019**		0.0011
Silver Perch	**0.0191**	**0.0011**	

	PBDEs	
	Site	Species
Model 1	0.1907	0.5844
Model 2	0.179	0.8557

Table 6: Results of model [1], with outliers removed, by chemical class. Values are displayed as mean (variance).

	DDTs	
	Site	
Species	3	4
Mullet	2.92 (0.90)	3.74 (1.44)
Sea Trout	1.59 (0.25)	1.24 (0.17)
Silver Perch	2.31 (0.64)	2.29 (0.55)

	PBDEs	
	Site	
Species	3	4
Mullet	0.30 (0.0003)	0.35 (0.0012)
Sea Trout	0.40 (0.0018)	0.45 (0.0024)
Silver Perch	0.19 (0.0001)	0.52 (0.0010)

	PCBs	
	Site	
Species	3	4
Mullet	18.56 (2.13)	16.55 (19.63)
Sea Trout	13.68 (9.95)	15.09 (2.19)
Silver Perch	25.87 (10.59)	33.82 (13.93)

	Aroclor 1268	
	Site	
Species	3	4
Mullet	7.06 (1.13)	4.11 (0.32)
Sea Trout	5.94 (0.78)	8.39 (1.59)
Silver Perch	7.15 (1.13)	11.72 (2.85)

	Metals Group 1 (All)	
	Site	
Species	3	4
Mullet	77.89 (52.76)	141.26 (217.84)
Sea Trout	35.33 (8.45)	33.65 (7.37)
Silver Perch	77.09 (45.37)	70.83 (35.95)

	Metals Group 3	
	Site	
Species	3	3
Mullet	8.05 (1.34)	8.05 (1.34)
Sea Trout	4.38 (0.24)	4.38 (0.24)
Silver Perch	7.75 (0.84)	7.75 (0.84)

	Metals Group 2 (Al, Zn, Fe)	
	Site	
Species	3	4
Mullet	24.71 (51.42)	49.67 (207.22)
Sea Trout	13.48 (8.22)	12.71 (7.15)
Silver Perch	26.84 (44.53)	25.13 (35.22)

	Pesticides	
	Site	
Species	3	4
Mullet	1.24 (0.02)	1.44 (0.04)
Sea Trout	0.51 (0.004)	0.54 (0.003)
Silver Perch	1.57 (0.03)	1.81 (0.04)

Table 7: Results of model [2], with outliers removed, by chemical class. Values are displayed as mean (variance).

	DDTs	
	Site	
Species	3	4
Mullet	3.37 (1.22)	3.46 (1.29)
Sea Trout	1.48 (0.23)	1.54 (0.25)
Silver Perch	2.58 (0.67)	2.41 (0.65)

	PBDEs	
	Site	
Species	3	4
Mullet	0.4 (0.00081)	0.37 (0.001)
Sea Trout	0.48 (0.0033)	0.34 (0.0013)
Silver Perch	0.42 (0.00053)	0.38 (0.0006)

	PCBs	
	Site	
Species	3	4
Mullet	16.97 (2.12)	18.74 (20.61)
Sea Trout	14.75 (11.12)	14.14 (2.71)
Silver Perch	34.87 (14.33)	28.73 (13.38)

	Aroclor 1268	
	Site	
Species	3	4
Mullet	5.35 (0.95)	5.58 (1.0)
Sea Trout	7.09 (1.87)	7.46 (1.97)
Silver Perch	9.15 (2.65)	9.6 (2.8)

	Metals Group 1 (All)	
	Site	
Species	3	4
Mullet	98.69 (189.46)	112.98 (251.86)
Sea Trout	32.82 (13.65)	36.92 (17.06)
Silver Perch	70.44 (73.52)	79.44 (93.15)

	Metals Group 3	
	Site	
Species	3	4
Mullet	10.4 (3.89)	13.24 (7.03)
Sea Trout	4.23 (0.27)	5.04 (0.41)
Silver Perch	7.3 (0.92)	8.73 (1.41)

	Metals Group 2 (Al, Zn, Fe)	
	Site	
Species	3	4
Mullet	88.28 (185.57)	99.74 (244.83)
Sea Trout	28.58 (13.38)	31.89 (16.65)
Silver Perch	63.14 (72.6)	70.71 (91.74)

	Pesticides	
	Site	
Species	3	4
Mullet	1.4 (0.037)	1.33 (0.035)
Sea Trout	0.59 (0.005)	0.58 (0.0051)
Silver Perch	1.74 (0.048)	1.68 (0.044)

Table 8: Results of model [3], with outliers removed, by chemical class. Values are displayed as mean (variance).

	DDTs	**PBDEs**	**Pesticides**	**Metals Group 1 (All)**	**Metals Group 2 (Al, Zn, Fe)**	**Metals Group 3**	**PCBs**	**Aroclor 1268**
Mullet	3.41 (1.25)	0.43 (0.001)	1.35 (0.036)	105.79 (234.65)	94.01 (228.75)	11.77 (5.9)	62.25 (47.93)	5.5 (0.98)
Sea Trout	1.51 (0.24)	0.26 (0.0007)	0.58 (0.005)	34.88 (16.5)	30.24 (16.13)	4.64 (0.37)	16.95 (20.13)	7.27 (1.92)
Silver Perch	2.51 (0.68)	0.43 (0.0006)	1.74 (0.047)	74.88 (89.25)	66.93 (87.98)	7.95 (1.27)	69.89 (31.29)	9.38 (2.72)

Table 9: P-values from model [1] and [2] using data with outliers removed. Bolded cells indicate significant differences (p-value < 0.05).

	DDTs	
	Site	Species
Model 1	0.3637	0.0783
Model 2	0.7245	**0.0054**

	Tukey	Model 2	
	Mullet	**Sea Trout**	**Silver Perch**
Mullet		0.006	0.0323
Sea Trout	**0.006**		0.0135
Silver Perch	**0.0323**	**0.0135**	

	Metals Group 1 (All)	
	Site	Species
Model 1	0.7581	0.1174
Model 2	0.1201	**0.0045**

	Tukey	Model 2	
	Mullet	**Sea Trout**	**Silver Perch**
Mullet		0.0066	0.0437
Sea Trout	**0.0066**		0.0087
Silver Perch	**0.0437**	**0.0087**	

	Metals Group 3	
	Site	Species
Model 1	0.891	0.1748
Model 2	0.0975	**0.0146**

	Tukey	Model 2	
	Mullet	**Sea Trout**	**Silver Perch**
Mullet		0.022	0.0872
Sea Trout	**0.022**		0.0256
Silver Perch	0.0872	**0.0256**	

	PCBs	
	Site	Species
Model 1	0.5568	0.0757
Model 2	0.3554	**0.0375**

	Tukey	Model 2	
	Mullet	**Sea Trout**	**Silver Perch**
Mullet		0.6857	0.0495
Sea Trout	0.6857		0.0376
Silver Perch	**0.0495**	**0.0376**	

	Aroclor 1268	
	Site	Species
Model 1	0.6068	0.5325
Model 2	**0.0404**	**0.0009**

	Tukey	Model 2	
	Mullet	**Sea Trout**	**Silver Perch**
Mullet		0.0031	0.0009
Sea Trout	**0.0031**		0.0037
Silver Perch	**0.0009**	**0.0037**	

	Pesticides	
	Site	Species
Model 1	0.477	**0.009**
Model 2	0.6895	**0.009**

	Tukey	Model 1	
	Mullet	**Sea Trout**	**Silver Perch**
Mullet		0.02	0.1644
Sea Trout	**0.02**		0.0118
Silver Perch	0.164	**0.0118**	

	Tukey	Model 2	
	Mullet	**Sea Trout**	**Silver Perch**
Mullet		0.0019	0.0191
Sea Trout	**0.0019**		0.0011
Silver Perch	**0.0191**	**0.0011**	

PCBs

Similar to DDTs, graphical analysis shows that sea trout have lower summed concentrations of PCBs and silver perch have higher summed PCB concentrations, after substituting the limit of detection for the left-censored data points (Figure 12, Appendix VII). Silver perch appear to have higher summed Aroclor 1268 concentrations than the other two species, and mullet appears to have slightly lower summed Aroclor 1268 concentrations than sea trout (Figure 12, Appendix VIII). Again, as with DDTs, there is not a clear distinction between

sites for summed PCB concentrations or summed Aroclor 1268 concentrations (Figure 13). In model [1], sea trout has the lowest summed concentration of PCBs at both sites, and mullet similar summed concentrations (Table 2). In particular, summed PCB concentrations for silver perch at site 4 is extremely high as compared to the other two species and is much higher than silver perch at site 3 (although silver perch at site 3 is also much higher than mullet and sea trout). However, it is important to note that the variance for summed PCB concentrations in mullet and silver perch at site 4 is extremely large. When examining only Aroclor 1268 (a subset of the PCBs) using model [1], mullet, sea trout, and silver perch at site 3 all have similar summed concentrations, while that for silver perch at site 4 is much higher. Again, this may be due to the extremely high variance of silver perch at site 4. Mullet at site 4 has the lowest summed concentration, which is a departure from previous patterns, as typically sea trout is the lowest. However, none of the summed concentrations were significantly different from one another by model [1] (Table 5. PCBs: site, p-value = 0.6726, species, p-value = 0.2273, Aroclor 1268: site, p-value = 0.961, species, p-value = 0.7521).

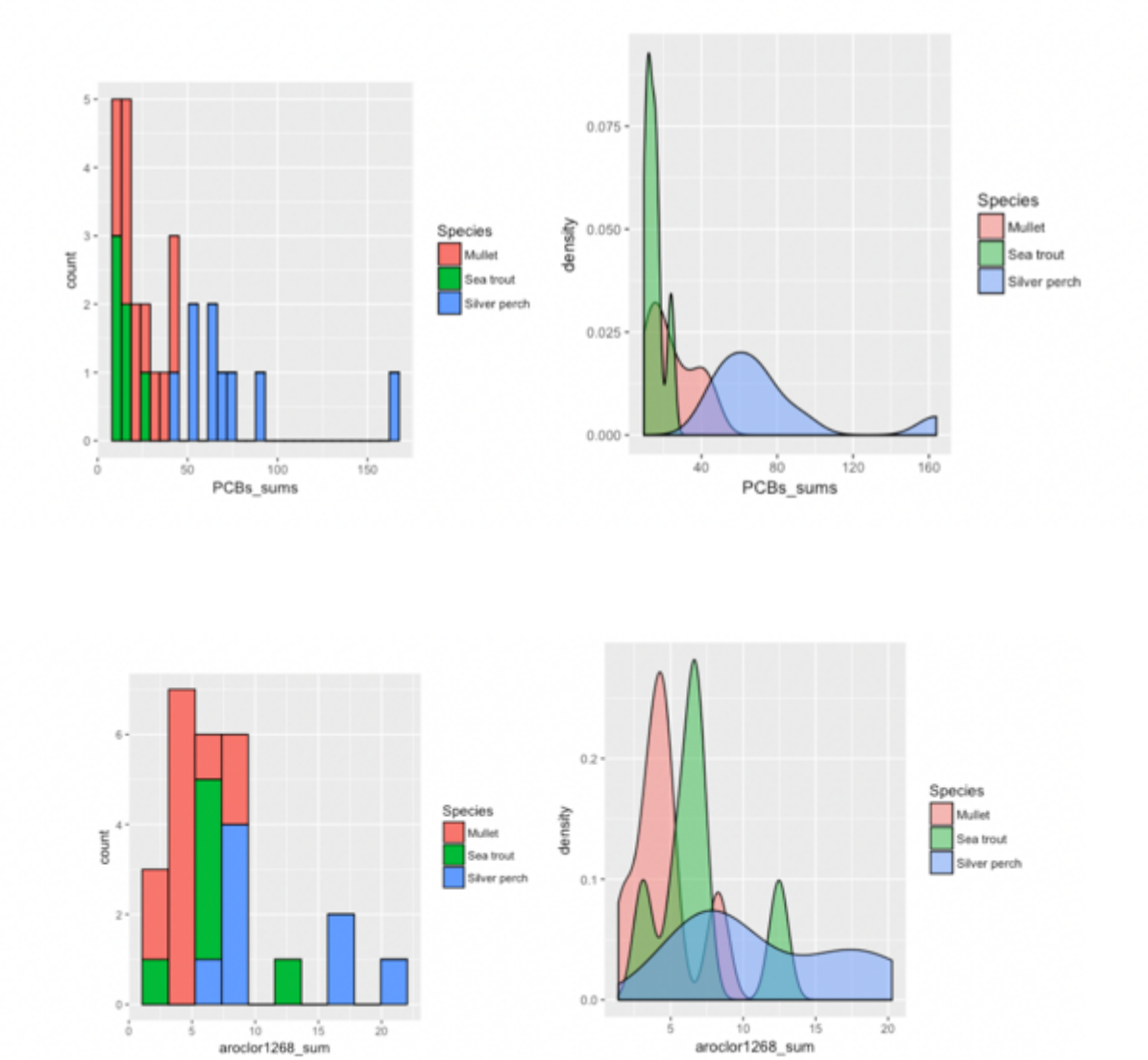


Figure 12: Histogram and density of summed PCBs (top panel) and summed Aroclor 1268 concentration (bottom panel) after substituting the limit of detection for left censored data points, separated by species.

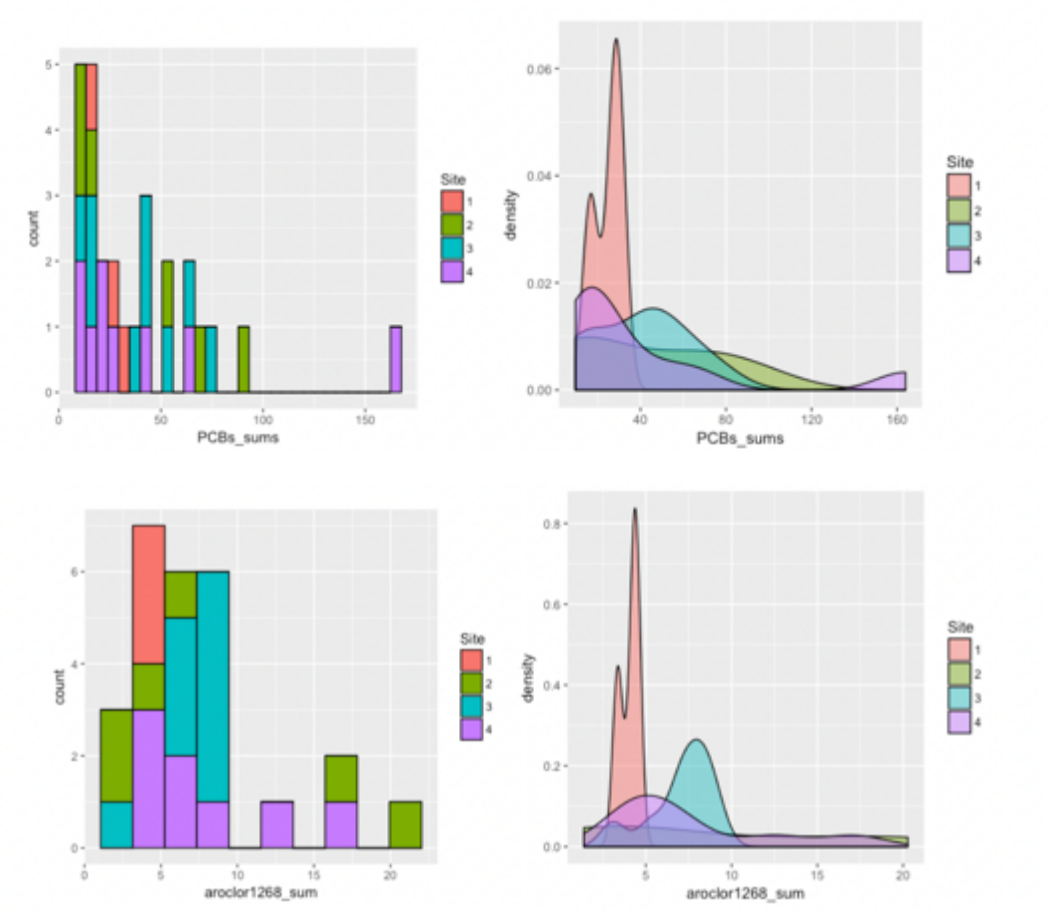


Figure 13: Histogram and density of summed PCB (top panel) and summed Aroclor 1268 concentration (bottom panel) after substituting the limit of detection for left censored data points, separated by site.

In model [2], mullet and sea trout have similar summed concentration of PCBs at site 3, while silver perch is much higher. Similarly, silver perch has a much higher summed PCB concentration at site 4, while those for mullet and sea trout are lower, with sea trout having the lowest (Table 3). Aroclor 1268 follows the same pattern. Silver perch in general has a very large variance at both sites. This is not unreasonable given the number of PCBs that are being summed together, and the large spread in concentration values seen for some analytes. In model [2], we see a significant difference in both summed PCBs and summed Aroclor 1268 concentration across species, but not site, although a site difference is marginally significant for Aroclor 1268

(Table 5. PCBs: site, p-value = 0.3289, species, p-value = 0.0321, Aroclor 1268: site, p-value = 0.0578, species, p-value = 0.0246). Tukey's test found that for both summed PCBs and summed Aroclor 1268 concentrations, there is a significant difference between mullet and silver perch (PCBs: p-value = 0.035, Aroclor 1268: p-value = 0.023), and silver perch and sea trout (PCBs: p-value = 0.0311, Aroclor 1268: p-value = 0.041), but not sea trout and silver perch (PCBs: p-value = 0.614, Aroclor 1268: p-value = 0.164). In model [3], we can see that sea trout has a much lower summed concentration of PCBs than the other two species (Table 4). Mullet and silver perch both have fairly high summed PCB concentrations and large variances, with silver perch again having the largest. For Aroclor 1268 specifically, mullet has the lowest summed concentration, with sea trout being not much larger. Silver perch has the highest summed concentration of Aroclor 1268 (Table 4, Figure 12).

The removal of outliers in the PCB chemical class caused the biggest change in the mean summed concentrations and variances as compared with the other chemical classes. In model [1], we again see that sea trout has the lowest summed concentration of PCBs, followed by mullet, then silver perch at both sites (Table 6). However, mullet and sea trout are similar, and the variances are much less inflated. For Aroclor 1268 specifically, sea trout has the lowest concentration at site 3, but mullet has the lowest concentration at site 4. Mullet and sea trout concentrations are very similar at site 3. As before, there are not significant differences between site or species using model [1] with outliers removed for either PCBs or Aroclor 1268 (Table 9. PCBs: site p-value = 0.5568, species p-value = 0.0757, Aroclor 1268: site p-value = 0.6068, species p -value = 0.5325). In model [2] with outliers removed, we see a similar pattern with the summed PCB concentration. Sea trout has the lowest, followed by mullet, then silver perch, and the concentrations are similar at both sites (Table 6).

However, when we look at Aroclor 1268, we see mullet with the lowest concentration, followed by sea trout and silver perch. Summed PCBs are significantly different by species but not by site, according to model [2] (Table 9. Site p-value = 0.3554, species p-value = 0.0375). Tukey's test showed that there is a difference between silver perch and mullet (p-value = 0.0495), silver perch and sea trout (p-value = 0.0376), but not sea trout and mullet (p-value = 0.6857), just as with model [2] using the full data. Interestingly, Aroclor 1268 showed a significant difference by both site and species (site p-value = 0.0404, species p-value = 0.0009). This is the only time we observed a significant difference by site. By inspection of the mean values (Table 7), the sites do not look to be significantly different. Tukey's test showed a significant pairwise difference between all three fish species (Table 9. Sea trout vs. mullet p-value = 0.0031, silver perch vs. mullet p-value = 0.0009, silver perch vs. sea trout p-value = 0.0037). Finally, model [3] with no outliers shows sea trout with the lowest summed concentration of PCBs, and mullet and silver perch much higher and similar (Table 8). However, according to model [3], mullet has the lowest summed concentration of Aroclor 1268, followed by sea trout and silver perch.

PBDEs

Graphical display of summed concentrations of PBDEs also show that sea trout appears to have the lowest summed concentrations and silver perch has the highest, after substituting the limit of detection for the left-censored data points (Figure 14, Appendix VII). Again, there is not a clear distinction by site for summed PBDE concentrations (Figure 15, Appendix VIII). Similar to DDTs, PBDEs do not appear to differ significantly across sites and species. In model [1],

silver perch at site 3 has the highest summed concentration of PBDEs, while silver perch at site 4 has the highest (Table 2). Summed concentration of PBDEs is similar between sites 3 and 4 in sea trout and mullet. However, summed concentrations of PBDEs are relatively small across all sites and species, and in model [1], the difference in summed mean PBDEs was not significantly different for either site or species (Table 5. Model 1: site, p-value = 0.1907, species, p-value = 0.5844). In model [2], summed concentrations of PBDEs are similar across all three species at both sites, and there is no significant different between site or species (Table 3. Model 2: site, p-value = 0.179, species, p-value = 0.8557). In model [3], sea trout have a lower summed concentration of PBDEs than the other two species, which have the same summed PBDEs (Table 4). Variance estimates for PBDEs are very small and are similar across site and species combinations. No outliers were removed from the PBDE chemical class.

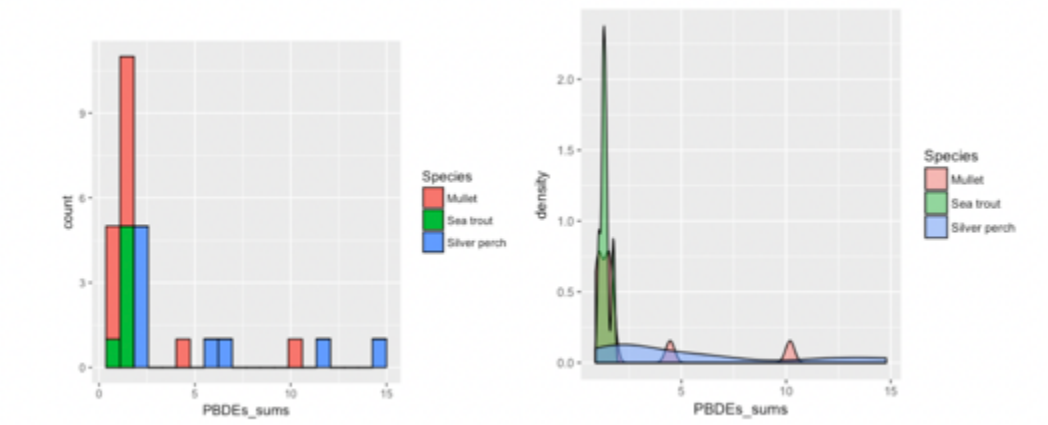


Figure 14: Histogram and density of summed PBDE concentrations after substituting the limit of detection for left censored data points, separated by species.

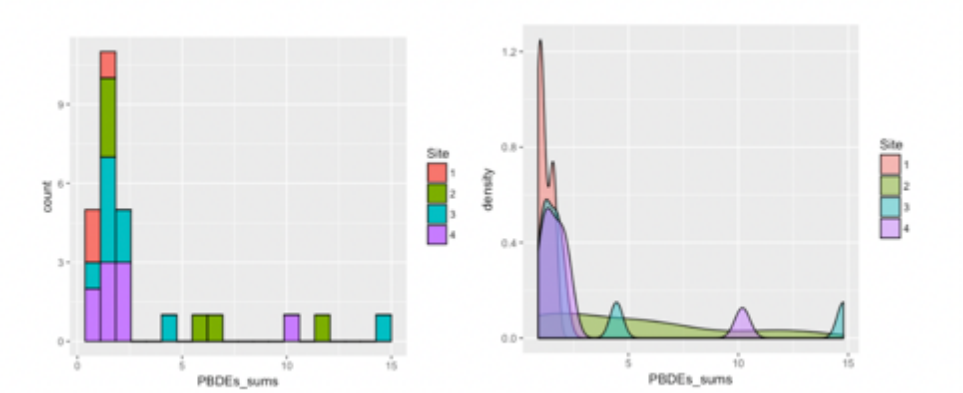


Figure 15: Histogram and density of summed PBDE concentrations after substituting the limit of detection for left censored data points, separated by site.

Metals

To reiterate, metals were analyzed as three separate groups: metals group 1 (includes all metals), metals group 2 (aluminum, iron, and zinc only), and metals group 3 (all metals excluding aluminum, iron, and zinc). Across metals group 1, we see a similar pattern as with the other chemical classes; sea trout has the lowest summed concentrations and silver perch has the highest, but no clear distinction is seen by site after substituting the limit of detection for the left-censored data points (Figure 16, Appendix VII and Figure 17, Appendix VIII). In model [1], sea trout has a much lower summed concentration of all metals at both sites as compared to mullet and silver perch (Table 2). Silver perch has a similar summed concentration of all metals between sites 3 and 4, while mullet at site 4 is much higher than site 3 (which is similar to the summed concentrations in silver perch). However, both mullet and silver perch tend to have large variances at both sites, with the variance for mullet at site 4 being extremely large.

When looking at metals group 2, for which concentrations were higher than for metals group 3, mullet and silver perch have much higher summed concentrations than sea trout in model [1]. Mullet at site 4 again have the highest value, while silver perch at both sites are

comparable to mullet at site 3. Sea trout have similar concentrations at both sites. When examining the summed concentration for metals group 3, mullet at site [4] is again the highest in model [1], although not as drastically different as with metals groups 1 and 2. Sea trout has the lowest concentration and is similar at both site 3 and 4. However, none of these differences were statistically significant (Table 5. Metals Group 1: site, p-value = 0.9376, species, p-value = 0.1167, Metals Group 2: site, p-value = 0.9438, species, p-value = 0.1749, Metals Group 3: site, p-value = 0.5497, species, p-value = 0.1822).

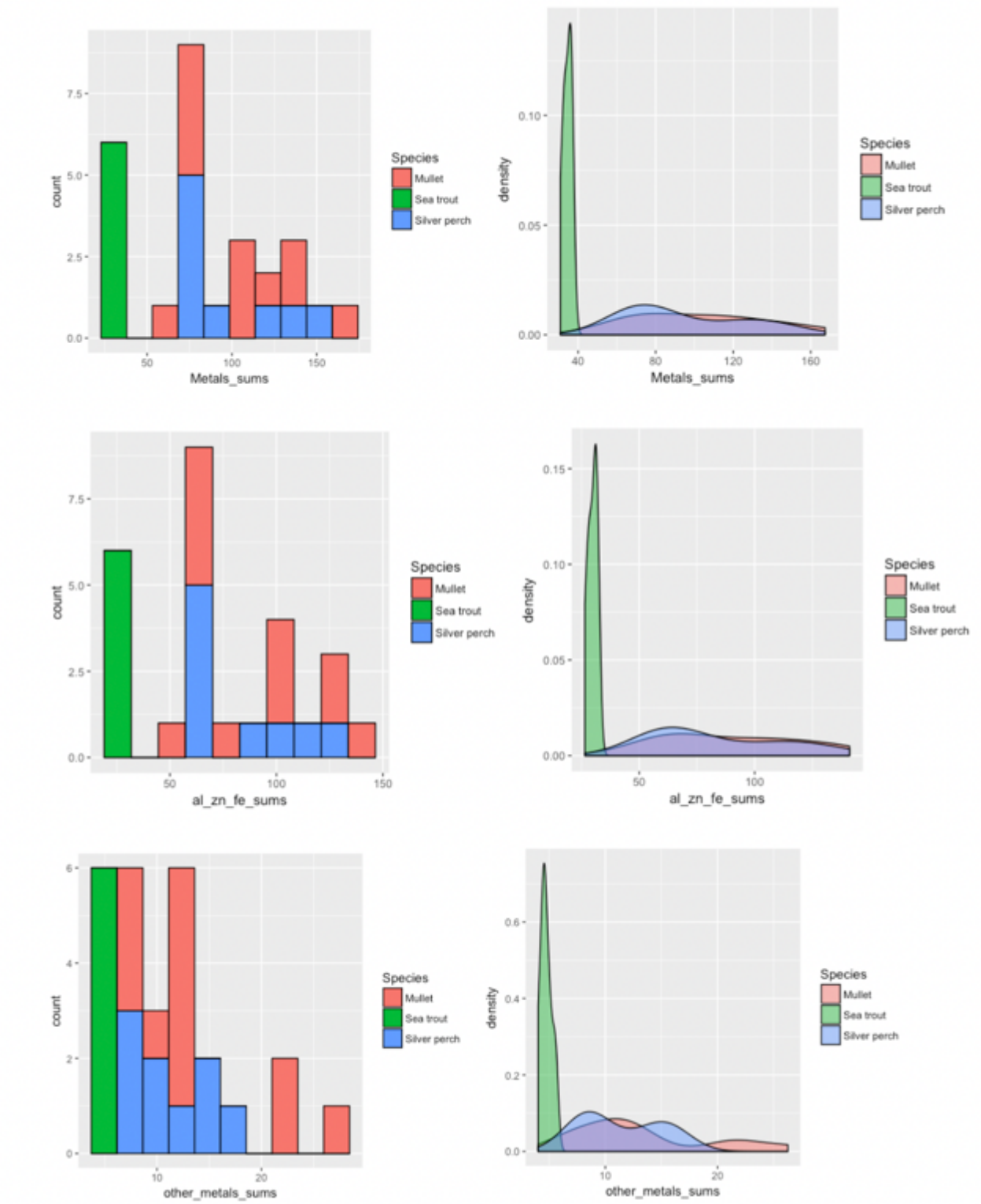


Figure 16: Histogram and density of summed metals (metals group 1) concentrations (top panel), summed aluminum, zinc, and iron (metals group 2) concentrations (middle panel), and summed all other metals (metals group 3) concentrations (bottom panel) after substituting the limit of detection for left censored data points, separated by species.

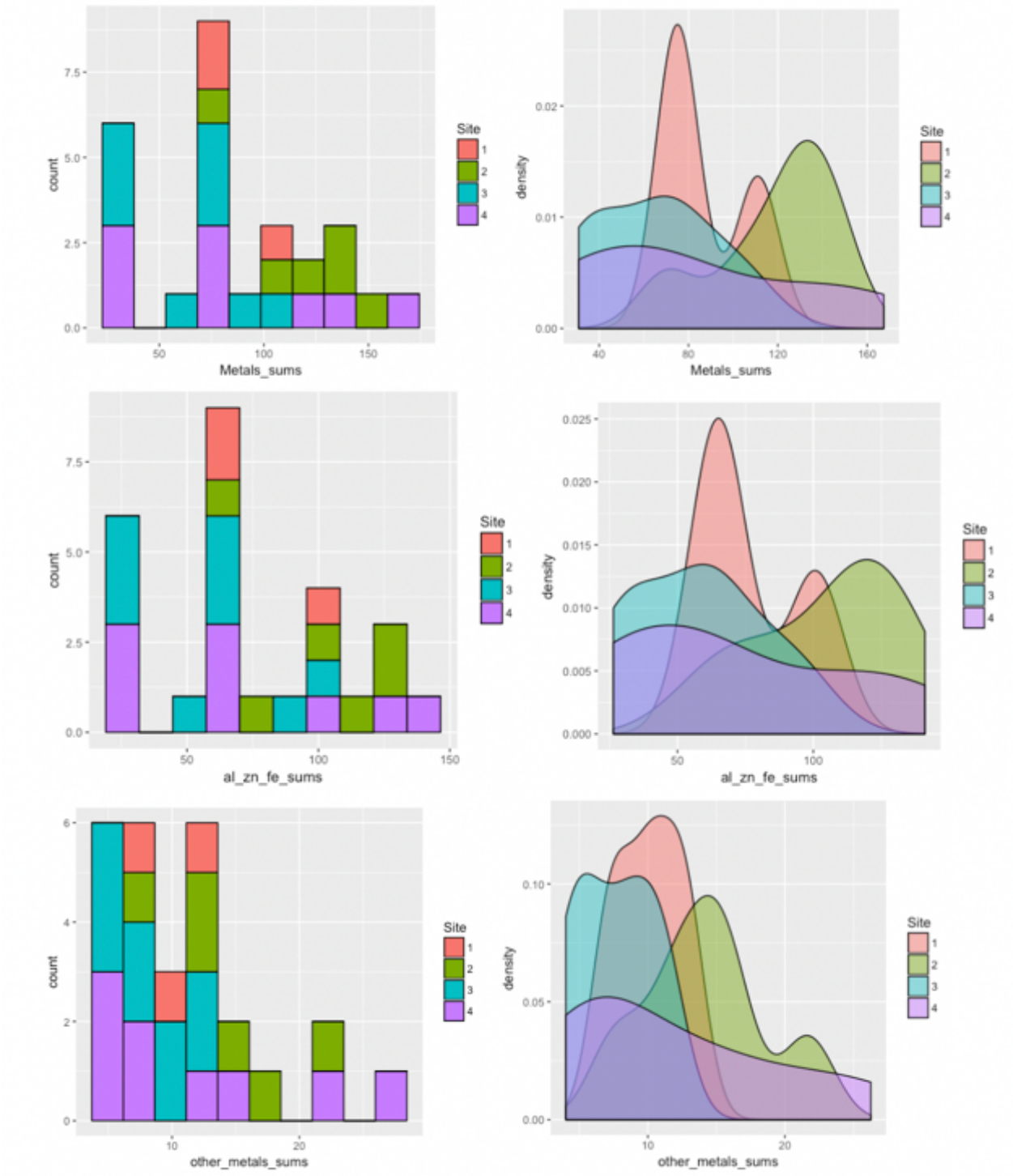


Figure 17: Histogram and density of summed metals (metals group 1) concentrations (top panel), summed aluminum, zinc, and iron (metals group 2) concentrations (middle panel), and summed all other metals (metals group 3) concentrations (bottom panel) after substituting the limit of detection for left censored data points, separated by site.

According to model [2], metals group 1 has the lowest summed concentration in sea trout, and these concentrations are similar at both sites (Table 3). Silver perch have the next lowest concentration, and mullet have the highest. Metals group 1 has a very large variance, driven particularly by the variances of aluminum, iron, and zinc. This same pattern of concentration levels is seen for both metals group 2 and metals group 3. For both groups 2 and 3, there does not seem to be a large difference by site. This is confirmed by our two-way ANOVA results; there is a significant difference between species for all three metals groups, but not by site (Table 5. Metals Group 1: site, p-value = 0.1141, species, p-value = 0.005, Metals Group 2: site, p-value = 0.1126, species, p-value = 0.0038, Metals Group 3: site, p-value = 0.0914, species, p-value = 0.0211). For metals groups 1 and 2, this difference is significant for mullet vs. sea trout (Metals Group 1 p-value = 0.007, Metals group 2 p-value = 0.006), mullet vs. silver perch (Metals Group 1 p-value = 0.048, Metals group 2 p-value = 0.039), and silver perch vs. sea trout (Metals Group 1 p-value = 0.0095, Metals group 2 p-value = 0.0071); however, for metals group 3, the difference was only significant for mullet and sea trout (p-value = 0.033), and sea trout and silver perch (p-value = 0.035), but not silver perch and mullet (p-value = 0.137) (Table 5). Model [3] shows no significant differences in the pattern as seen previously; for all three groups, sea trout has the lowest summed concentration, followed by silver perch, then sea trout (Table 4). This is a deviation from other chemical classes, for which silver perch tends to have the highest concentrations.

Following the removal of some outliers (none of which were aluminum, iron, or zinc analytes), we see similar summed mean concentrations in metals group 1 as before using model [1]. Sea trout still have the lowest summed concentration of all metals, and their concentration is similar across both sites (Table 6). Silver perch have the next highest summed concentrations

(which again is similar across sites), while mullet have the highest concentration, which is much higher at site 4 than site 3. This same pattern is seen for metals group 3. Using model [1], there is not a significant difference between site or species looking either at all metals group 1 or 3 (Table 9. Metals Group 1: site, p-value = 0.7581, species, p-value = 0.1174, Metals Group 3: site, p-value = 0.891, species, p-value = 0.1748). Model [2] without outliers shows the same pattern as with model [1], except there is not as large of a difference between summed concentrations for mullet at the two sites (Table 7). However, there is a significant difference between species in this case, but not for site (Table 9. Metals Group 1: site, p-value = 0.1201, species, p-value = 0.0045, Metals Group 3: site, p-value = 0.0975, species, p-value = 0.0146). As with model [1], for metals group 1 this difference is significant for mullet vs. sea trout (p-value = 0.0066), mullet vs. silver perch (p-value = 0.0437), and silver perch vs. sea trout (p-value = 0.0087); however, for metals group 3, the difference was only significant for mullet and sea trout (p-value = 0.022), and sea trout and silver perch (p-value = 0.0256), but not silver perch and mullet (p-value = 0.0872) (Table 9). Finally, model [3] with no outliers shows the same ordering of species, although the variance for mullet is extremely large (Table 8).

Pesticides

Pesticides also graphically seem to display a pattern in terms of summed concentrations by species, with sea trout having the lowest summed concentration and silver perch having the highest, after substituting the limit of detection for the left-censored data points (Figure 18, Appendix VII). Again, no clear distinction is seen by site (Figure 19, Appendix VIII). In model [1], sea trout have a lower summed concentration of pesticides at both sites 3 and 4, while mullet

and silver perch are similar across both sites (Table 2). There was a significant difference between species in model 1 (Table 5, p-value = 0.004), and according to Tukey's test, that difference was between mullet and sea trout (p-value = 0.0087) and silver perch and sea trout (p-value = 0.0052), although the difference between mullet and silver perch was not significant (p-value = 0.0617). Site was not significantly different in model [1] (p-value = 0.187).

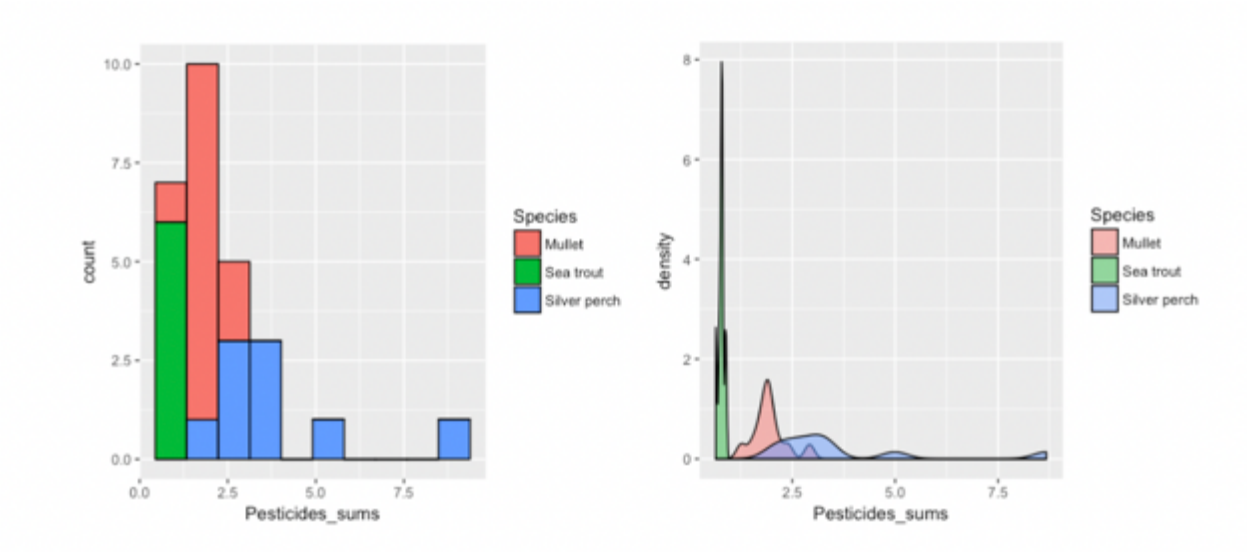


Figure 18: Histogram and density of summed pesticide concentrations after substituting the limit of detection for left censored data points, separated by species.

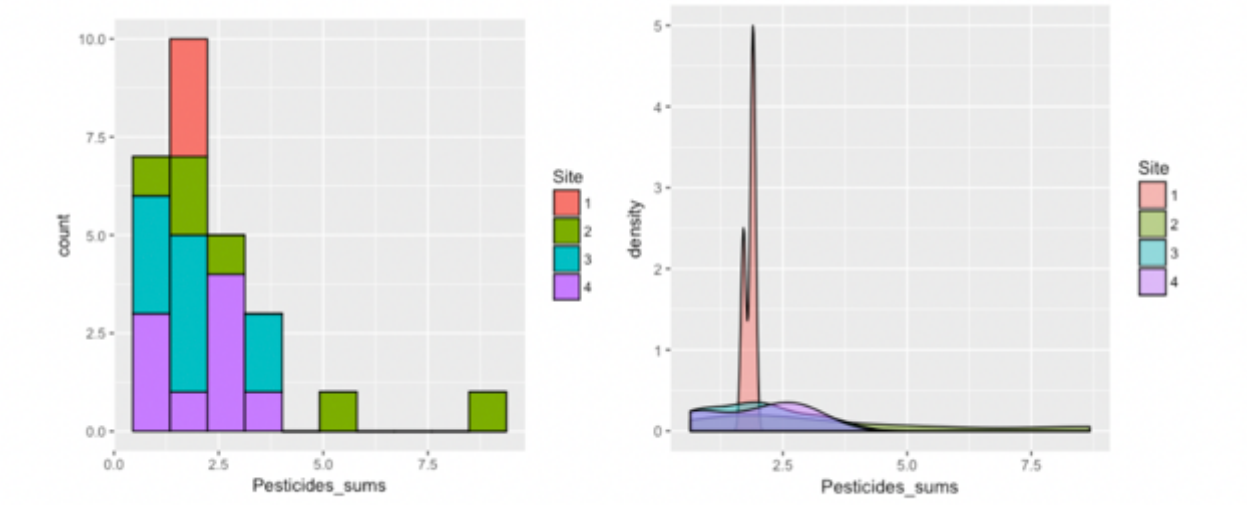


Figure 19: Histogram and density of summed pesticide concentrations after substituting the limit of detection for left censored data points, separated by site.

In model [2], all three fish species have similar summed pesticide concentrations across the two sites (Table 3). However, concentrations in sea trout were much lower than in the other two species, with those in mullet additionally being lower than those in silver perch. In model [2], concentrations by site are not significantly different (Table 5, p-value=0.1158), but are by species (p-value=0.0006). Tukey's test revealed all three pairwise comparisons of species were significantly different from each other (sea trout vs. mullet p-value = 0.001, sea trout vs. silver perch p-value = 0.0008, silver perch vs. mullet p-value = 0.011). In model [3], sea trout had lower summed concentrations of pesticides than mullet and silver perch, which are similar (Table 4). When one outlier was removed, similar patterns are seen for all three models (Table 6, 7, 8). The variance of pesticides was not very large or variable across sites and species. The same significance trends were seen in pesticides when one outlier was removed. In both model [1] and [2], there was a significant different between species, but not site (Table 9. Model 1: site p-value = 0.477, species p-value = 0.009. Model 2: site p-value = 0.6895, species p-value = 0.009). In model [1], there were significant pairwise differences between mullet and sea trout (p-value = 0.02), and sea trout and silver perch (p-value = 0.0118), but not silver perch and mullet (p-value = 0.164). In model [2], there were pairwise differences between all three species of fish (sea trout vs. mullet p-value = 0.0019, sea trout vs. silver perch p-value = 0.0011, silver perch vs. mullet p-value = 0.0191).

5. Discussion

It is promising that we see an increase in the number of stable estimates as the number of parameters that we estimate decreases. However, we are still not able to get enough reliable

estimates from the available sample sizes. We do not get enough stable estimates from all chemical classes to get an accurate summed concentration; for example, PBDEs are never stable in any of the three models because there is simply too much data below the limit of detection. It is challenging to pool such a limited amount of data together for each analyte, especially when some analytes have a wide range of concentration values across sites and species. We are seeing extremely high variances for some of the analytes, which seems to be caused by having some high concentrations and some lower, but no outliers; this, combined with having some measurements below the detection limit, results in large deviations from a mean value. We are currently at the limit of which standard analyses will work; at present, there is not statistical methodology that can reliably deal with so much missing data.

In general, we tend to see stable estimates when either all of the data points in a given site/species/analyte combination are reportable, or only one or two are below the limit of detection. In addition, we notice that stable estimates are present when those few values that are below the limit of detection are relatively high and/or close to the reportable estimates; greater stability is achieved when there is a smaller range of concentration values, either for reportable data or detection limits. Even if there are no outliers, but there is a wide range of values especially with extremely small concentration measurements, estimates are more unstable.

Although we do not have a lot of reportable data above the detection limit, we tend to trust the results from the models without any outliers more than those that include outliers. This is because the range of concentration values is smaller, and we see analytes that had unstable estimates become stable when outliers are removed. In addition, although we cannot make any statistical conclusions from model [3], it did result in the largest number of mean analyte estimates with reasonable and stable standard errors. Therefore, it can still be used to visually

compare summed concentrations across species. However, given that it only includes information on species, loses the site data, and can't be used for statistical tests, we don't recommend its use. We recommend the use of model [2] in future studies. Model [2] loses some biological information (the potential interaction between site and species), but still allows us to compare across both sites and species. Model [2] resulted in a moderate number of stable estimates, which could potentially increase even more in number with additional data.

In general, we feel comfortable comparing the general trends across site and species, but the statistical results from the two-way ANOVA should be examined with caution due to the high variance and low stability of the estimates. It does appear as though sea trout has lower chemical concentrations than the other species, and there do not appear to be any significant differences between the chemical concentrations at the two different sites. The current results can be used to inform the design of future studies; our baseline results suggest the need for larger sample sizes, i.e., more fishermen will need to be employed to catch more fish. Although this will increase the cost of a future study, it will lead to better statistical accuracy in results overall.

This project helps fill in the picture of pollutant levels in the environment near Sapelo Island. We know that bottlenose dolphins located off the coast of Sapelo Island have high levels of the PCB congener Aroclor 1268, despite being miles away from the Superfund site where this compound was produced[6]. Since bottlenose dolphins' blubber allows for bioaccumulation of environmental exposures, this study was done to examine the exposure levels in bottlenose dolphins' prey. The results of this study can help to explain why bottlenose dolphins at Sapelo also have high exposure levels similar to that of bottlenose dolphins near the Brunswick Superfund site.

According to Balmer et al., the chemical class with the highest summed mean concentrations in bottlenose dolphins off the coast of Sapelo Island was PCBs[6]. Summed mean DDT concentrations in bottlenose dolphins were much smaller, summed mean PBDE concentrations were the lowest, and summed mean pesticides were lower than DDTs but higher than PBDEs[6]. Kucklick et al. found the similar trends in the blubber of bottlenose dolphins caught off the coast of Sapelo Island[17]. This trend of high summed PCB concentrations, smaller summed DDT concentrations, even lower summed pesticide concentrations, and the lowest summed PBDE concentrations is consistent in all three species of fish and across the two sites in this study.

Limitations

As noted above, there were a number of analytic challenges in attempting to compare mean differences in sites to species. While all of the mixed models converged, standard error estimates were over-inflated and often unstable. Reasonable standard error estimates for the maximum likelihood parameter estimates were only present in the models for which no left censoring or very little left censoring was present. Using the model without the interaction terms did provide more estimates that were stable, since we are estimating less parameters, with the downside of losing biological information. This assumption is probably reasonable, given that graphical representation of the densities of the summed concentrations for each chemical class (after substituting the limit of detection for the left-censored data), do not appear to display any clear trends for species by site or vice versa (Appendix IX). Additional estimates became stable after removing the site component and only focusing on species. This shows that estimating less

parameters does improve stability, even with limited data. However, we do not have enough reportable data points to reasonably estimate the means of most analytes, even when only estimating three maximum likelihood parameters.

In order to provide additional data to hopefully produce more reliable estimates, we could combine all of the data in a chemical class and fit the same maximum likelihood models. Instead of fitting the model for each analyte, it would be fit by chemical class. This removes many of the intricacies of the data but allows for more data to be used to estimate the mean at each site for each species. However, estimates for most chemical classes using this method were still unstable with large variances, indicating that we may be past the limit of an acceptable amount of left-censored data.

There is a question of whether the independence assumption we used to sum the variance of each analyte within a chemical class is valid. It is possible that analytes within a chemical class are related to each other; perhaps they were both used in the manufacture of a given product, and we therefore expect their concentrations to be related. Although previous literature has consistently summed the concentration measurements, it is possible that this assumption is not valid and should be considered further and evaluated in future research.

In addition, we were not able to compare mean summed concentrations across all sites and species due to the missing data at sites 1 and 2. We do not have information about analyte concentrations from silver perch and sea trout at site 1 and sea trout at site 2. Without any information to inform about the possible values of the missing data there, we are not able to impute or simulate any representative data. This limited our analyses to only sites 3 and 4, which had complete data.

Ideally, we would be able to come up with a mean summed concentration of each chemical class by site and species. However, there are so much data below the detection limit that our model does not have enough flexibility to provide stable estimates necessary to answer our research question. In our highly censored setting, the maximum likelihood estimates have an identifiability problem, and the results may not be accurate due to all of the instability problems we are observing.

Future Work

In the future, we would like to adjust the Q-Q plots created to adjust for the censored data. This has been done previously for right-censored data by using the empirical distribution[28]. It is possible that this technique could be extended to left-censored data as well in order to provide further justification for using the log-normal distribution to obtain our maximum likelihood estimates. Although this would be interesting to examine, it was not explored in this thesis simply because we do not have an alternative distribution to use for this data. This shows another limitation of our analysis; we are limited in the potential distributions we can work with to obtain estimates of our maximum likelihood parameters.

Future work could examine addressing our limit of detection challenges by utilizing a multiple imputation approach to impute the values that are below the limit of detection. Each concentration below the limit of detection could be replaced by 10 appropriate values (imputed from the range of the concentration levels below the limit of detection) resulting in 10 complete datasets[15]. Each complete (imputed) dataset would be analyzed, resulting in 10 mean estimates, which can be combined to create a final estimate[15]. The imputations involved are drawn

conditional on knowing the underlying distribution of the data; in our case, we assume that the data lies between 0 and the known limit of detection. Since each imputed concentration is an individual draw and not a mean, the estimate is valid for a wide range[20]. It is important to note that in for imputed values, there are two variance components, within- and between-imputation[20], that together provide an estimate of the overall variance of our estimates. This technique is preferred over single imputation (such as substituting one half of the LOD), which does not provide a variance estimate.

Multiple imputation relies on a distributional assumption for the values that are below the limit of detection. It uses the values in the reportable range to inform about the data points below the detection limit. However, as mentioned previously, less than half of the analytes have enough reportable data to inform accurately about the missing data. Therefore, we will not do multiple imputation in this thesis, as we likely need much more data in order for it to perform well. We cannot impute values below the detection limit if we do not know anything about the values above the limit. Previous literature has shown that maximum likelihood estimation and multiple imputation perform similarly and provide similar results[21]. Therefore, it is likely we will come to the same conclusions with both methods. Future work could also explore using a simulation to see the maximum amount of missingness that is tolerated to still produce uninflated variances. Finally, we would like to find a way to incorporate the data from sites 1 and 2, despite the fact that they did not have any data for some of the fish species.

It would be helpful for future work to include a larger sample size with more composites in each site/species combination. We need much larger numbers of fish to provide the most reliable estimates. In addition, focusing only on exposures that are either present at higher levels in general or can be accurately detected at low levels would allow for more stable estimates.

Missing data and data below detection limits are common in environmental exposure studies. It is especially challenging when such a large number of analytes are being measured, and some are on differing scales than others. This work can help to inform about the level of missingness that can be tolerated, as well as the expected results from using maximum likelihood estimation in the presence of left-censored data. The estimation of fewer parameters results in more stable estimates, and this further improves with the removal of outliers. Even having a few reportable data points can result in stable estimates when only a few parameters are estimated. Future studies can focus on analytes that can be reported even in just a few samples if there are specific analytes of interest; however, the purpose of this exploratory study was to gain a general overview of POP levels in the prey of bottlenose dolphins by examining many different POPs.

Conclusions

Overall, we found there to be a significant difference in pesticide concentrations between species, but not site, using both models. Given that this occurred in both models, and pesticides consistently had a majority of analytes with stable standard errors, we believe these results are robust. In model [1], mean summed pesticide concentration was different only between mullet and sea trout and silver perch and sea trout; however, in model [2], it was different between all three species. In addition, PBDE concentrations were not significantly different between sites or species in either model. There were no stable estimates by analytes in the PBDE chemical class, due to such a large amount of the data being below the limit of detection. In general, we found that sea trout tend to have the lowest concentration across all chemical classes, except for Aroclor 1268, for which mullet has the lowest concentration. This is true for all models, both with and without outliers. Mullet and silver perch generally tend to have similar summed mean

concentration levels across the chemical classes, with mullet higher for DDTs and metals and silver perch higher for pesticides and PCBs.

We found that there was a significant difference between species for almost all chemical classes when using model [2] when no interaction between site and species is present in the model. We do not have enough data presently to test if this is a fair assumption and these interaction terms can reasonably be removed from the model. For DDTs and metals, all three species are significantly different from each other both with and without outliers present in the data. Aluminum, iron, and zinc as a distinct group of metals also yield significantly different concentrations between all three species in model [2], while the concentrations of other metals are only significantly different between mullet and sea trout and silver perch and sea trout, again both with and without outliers present. PCB concentrations are only significantly different between mullet and silver perch and between sea trout and silver perch. Aroclor 1268 as a subgroup of PCBs shows significant differences in concentrations between mullet and silver perch and between sea trout and silver perch in model [2] when outliers are present in the data but shows significant differences between all three species when outliers are removed. In addition, Aroclor 1268 shows a significant difference in concentration between site using model [2] when outliers are removed, the only time this happens in the data. However, in looking at the actual mean summed concentrations of Aroclor 1268 at both sites, they do not appear to be that different in magnitude, so this result should be interpreted with caution.

Given that there are no significant differences in concentrations between sites, and the values obtained from models [1] and [2] appear similar at sites 3 and 4, it is justifiable to create model [3]. In model [3], we can only examine trends across species in each chemical class. Results tend to be similar in the model with and without outliers, although the mean summed

concentrations are lower in the model without outliers. We see very large variances, particularly in the metals and PCB classes, although these also improve when outliers are removed. It is important to note, however, that site not being significantly different could be either because there truly is no difference in the contaminant levels between sites 3 and 4, or because we do not have enough power to detect any differences as a result of our small sample size and large amount of missing data.

There was a very large discrepancy in variances of concentrations, particularly for metals and PCBs. Therefore, although there are means that may look significantly different, the variance is so large for some of the larger summed concentrations that it makes sense that the differences are not in fact statistically significant. For example, aluminum, zinc, and iron have high variance which drives the overall high metals variance. However, these estimates of variance (not standard deviation) are valid given the data and the overall large means of these specific metals. All of the analytes with large variances make sense from the data and are a result of either that analyte having a larger mean overall or having a wider spread of data. In addition, with so many PCB analytes measured, it follows that they would generally have larger variances.

After summing together the mean summed concentrations of all five chemical classes, we find that sea trout is the least contaminated species overall. Mullet is the most contaminated species overall. This result is not consistent across chemical classes, as we find mullet has higher summed mean concentrations DDTs and metals and silver perch have higher summed mean concentrations for pesticides and PCBs. The summed mean concentrations for PBDEs are similar between silver perch and mullet, and it is not different which species has a higher summed mean concentration depending on the model and site. Bottlenose dolphins that consume greater amounts of mullet and silver perch as compared to sea trout are more likely to be exposed to

POPs that will then bioaccumulate in their blubber and result in adverse health effects. We know from previous research that bottlenose dolphins most commonly consume mullet[7]. Given that it was the most contaminated in our study, this could mean that the chemical exposure levels in the mullet are driving the high POP levels seen in bottlenose dolphins. Similarly, humans that consume more mullet and silver perch are more likely to be exposed to POPs. Interestingly, mullet was caught at all four sites used in the study and was not kept by the fisherman to eat. Silver perch was only missing at one site, compared with sea trout that was missing at two. Since sea trout was the least contaminated species overall, this is good as it shows humans may already be consuming less contaminated fish.

In closing, we note that the fact that so many of the analyte measures are below the limit of detection is good news from the perspective of pollutant exposures and environmental health. We want these exposure levels to be minimal. This means that fish have generally low exposure levels to the environmental toxicants, meaning that the humans consuming these fish are also hopefully exposed at low levels. We can question whether these fish are really contaminated to a point that would impact their health and human health. However, it is important to note here that the majority of the outliers removed from the analysis were silver perch at site 4. These fish may actually be contaminated. Given such low levels of exposure, it is possible that the impacts of these pollutants are minimal. However, this presents a statistical challenge for obtaining accurate and reliable estimates of concentrations within and between species and locations.

6. References

1. (ATSDR), Agency for Toxic Substances and Disease Registry. *Toxicological Profile for Mirex and Chlordecone*. Atlanta, GA: U.S. Department of Health and Human Services, Public Health Service, 1995. Print.

2. (ATSDR), Agency for Toxic Substances and Disease Registry *Toxicological Profile for Polychlorinated Biphenyls (PCBs)*. Atlanta, GA: U.S. Department of Health and Human Services, Public Health Service, 2000. Print.

3. (ATSDR), Agency for Toxic Substances and Disease Registry *Toxicological Profile for DDT, DDE, DDD*. Atlanta, GA: U.S. Department of Health and Human Services, Public Health Service, 2002. Print.

4. (ATSDR), Agency for Toxic Substances and Disease Registry *Toxicological Profile for Aldrin/Dieldrin*. Atlanta, GA: U.S. Department of Health and Human Services, Public Health Service, 2002. Print.

5. Adams, Jeffrey, et al. "The Relationship between Land Use and Emerging and Legacy Contaminants in an Apex Predator, the Bottlenose Dolphin (*Tursiops truncatus*), from Two Adjacent Estuarine Watersheds." *Environmental Research* 135 (2014): 346-53. Print.

6. Balmer, Brian C., et al. "Relationship between Persistent Organic Pollutants (POPs) and Ranging Patterns in Common Bottlenose Dolphins (*Tursiops truncatus*) from Coastal Georgia, USA." *Science of The Total Environment* 409.11 (2011): 2094-101. Print.

7. Barros, Nélio B., and Randall S. Wells. "Prey and Feeding Patterns of Resident Bottlenose Dolphins (*Tursiops truncatus*) in Sarasota Bay, Florida." *Journal of Mammalogy* 79.3 (1998): 1045-59. Print.

8. de Wit, Cynthia A. "An Overview of Brominated Flame Retardants in the Environment." *Chemosphere* 46.5 (2002): 583-624. Print.

9. Fair, Patricia A., et al. "Perfluoroalkyl Substances (PFASS) in Edible Fish Species from Charleston Harbor and Tributaries, South Carolina, United States : Exposure and Risk Assessment." *Environmental Research* 171 (2019): 266-77. Print.

10. Folland, William R., et al. "Growth and Reproductive Effects from Dietary Exposure to Aroclor 1268 in Mink (*Neovison vison*), a Surrogate Model for Marine Mammals." *Environmental Toxicology and Chemistry* 35.3 (2016): 604-18. Print.

11. García-Álvarez, Natalia, et al. "Assessment of the Levels of Polycyclic Aromatic Hydrocarbons and Organochlorine Contaminants in Bottlenose Dolphins (*Tursiops truncatus*) from the Eastern Atlantic Ocean." *Marine Environmental Research* 100 (2014): 48-56. Print.

12. Gillespie, Brenda W., et al. "Estimating Population Distributions When Some Data Are Below a Limit of Detection by Using a Reverse Kaplan-Meier Estimator." 21.4 (2010): S64-S70. Print.

13. Giolo, Suely. *Turnbull's Nonparametric Estimator for Interval-Censored Data*. 2015. Print.

14. Hansen, Larry J., et al. "Geographic Variation in Polychlorinated Biphenyl and Organochlorine Pesticide Concentrations in the Blubber of Bottlenose Dolphins from the Us Atlantic Coast." *Science of The Total Environment* 319.1 (2004): 147-72. Print.

15. Harel, Ofer, Neil Perkins, and Enrique F. Schisterman. "The Use of Multiple Imputation for Data Subject to Limits of Detection." *Sri Lankan journal of applied statistics* 5.4 (2014): 227-46. Print.

16. Helsel, Dennis. "Much Ado About Next to Nothing: Incorporating Nondetects in Science." *The Annals of Occupational Hygiene* 54.3 (2010): 257-62. Print.

17. Kucklick, John, et al. "Bottlenose Dolphins as Indicators of Persistent Organic Pollutants in the Western North Atlantic Ocean and Northern Gulf of Mexico." *Environmental Science & Technology* 45.10 (2011): 4270-77. Print.

18. Kutner, Michael H., et al. *Applied Linear Statistical Models*. 5 ed. New York, NY: McGraw-Hill/Irwin, 2005. Print.

19. Lee, Duk-Hee, et al. "Chlorinated Persistent Organic Pollutants, Obesity, and Type 2 Diabetes." *Endocrine reviews* 35.4 (2014): 557-601. Print.

20. Little, Roderick J.A., and Donald B. Rubin. *Statistical Analysis with Missing Data*. 2 ed. Hoboken, N.J.: John Wiley & Sons, Inc., 2002. Print.

21. Lyles, Robert H., Dongjie Fan, and Rutt Chuachoowong. "Correlation Coefficient Estimation Involving a Left Censored Laboratory Assay Variable." *Statistics in Medicine* 20.19 (2001): 2921-33. Print.

22. Lynn, Henry S. "Maximum Likelihood Inference for Left-Censored HIV RNA Data." *Statistics in Medicine* 20.1 (2001): 33-45. Print.

23. Sapelo Island National Estuarine Research Reserve. "The History of Sapelo." 2019. Web. January 18, 2019.

24. Sapelo Island National Estuarine Research Reserve. "Nature Trails." 2019. Web.

25. Ryu, Sukhyun, et al. "One Health Perspectives on Emerging Public Health Threats." *J Prev Med Public Health* 50.6 (2017): 411-14. Print.

26. Schwacke, Lori H., et al. "Probabilistic Risk Assessment of Reproductive Effects of Polychlorinated Biphenyls on Bottlenose Dolphins (*Tursiops truncatus*) from the Southeast United States Coast." 21.12 (2002): 2752-64. Print.

27. Severson, Kim. "Taxes Threaten an Island Culture in Georgia." *The New York Times* 2012: A16. Print.

28. Waller, Lance A., and Bruce W. Turnbull. "Probability Plotting with Censored Data." *The American Statistician* 46.1 (1992): 5-12. Print.

29. Whitcomb, Brian W., and Enrique F. Schisterman. "Assays with Lower Detection Limits: Implications for Epidemiological Investigations." *Paediatric and Perinatal Epidemiology* 22.6 (2008): 597-602. Print.

30. Wirth, E. F., et al. "Distribution and Sources of PCBs (Aroclor 1268) in the Sapelo Island National Estuarine Research Reserve." *Environmental Monitoring and Assessment* 186.12 (2014): 8717-26. Print.

7. Appendix

Appendix I. List of all analytes measured.

Chemical Name	Chemical Class
2,4'-DDD	DDTs
2,4'-DDE	DDTs
2,4'-DDT	DDTs
4,4'-DDD	DDTs
4,4'-DDE	DDTs
4,4'-DDT	DDTs
Silver (Ag)	Metals
Aluminum (Al)	Metals
Arsenic (As)	Metals
Barium (Ba)	Metals
Beryllium (Be)	Metals
Cadmium (Cd)	Metals
Cobalt (Co)	Metals
Chromium (Cr)	Metals
Copper (Cu)	Metals
Iron (Fe)	Metals
Mercury (Hg)	Metals
Lithium (Li)	Metals
Manganese (Mn)	Metals
Nickel (Ni)	Metals
Lead (Pb)	Metals
Antimony (Sb)	Metals
Selenium (Se)	Metals
Tin (Sn)	Metals
Thallium (Tl)	Metals
Uranium (U)	Metals
Vanadium (V)	Metals
Zinc (Zn)	Metals
PBDE 100	PBDEs
PBDE 138	PBDEs
PBDE 153	PBDEs
PBDE 154	PBDEs
PBDE 17	PBDEs
PBDE 183	PBDEs

Chemical Name	Chemical Class
PBDE 190	PBDEs
PBDE 28	PBDEs
PBDE 47	PBDEs
PBDE 66	PBDEs
PBDE 71	PBDEs
PBDE 85	PBDEs
PBDE 99	PBDEs
PCB 1	PCBs
PCB 101	PCBs
PCB 103	PCBs
PCB 104	PCBs
PCB 105	PCBs
PCB 108/107/123	PCBs
PCB 110	PCBs
PCB 114	PCBs
PCB 118/106	PCBs
PCB 119	PCBs
PCB 12	PCBs
PCB 126	PCBs
PCB 128	PCBs
PCB 130	PCBs
PCB 132/153/168	PCBs
PCB 138/158	PCBs
PCB 141	PCBs
PCB 146	PCBs
PCB 149	PCBs
PCB 15	PCBs
PCB 151	PCBs
PCB 154	PCBs
PCB 156	PCBs
PCB 157	PCBs
PCB 159	PCBs
PCB 164/163	PCBs
PCB 165	PCBs
PCB 167	PCBs
PCB 169	PCBs
PCB 170/190	PCBs
PCB 172	PCBs

Chemical Name	Chemical Class
PCB 174	PCBs
PCB 177	PCBs
PCB 18	PCBs
PCB 180/193	PCBs
PCB 183	PCBs
PCB 184	PCBs
PCB 187	PCBs
PCB 188	PCBs
PCB 189	PCBs
PCB 194	PCBs
PCB 195	PCBs
PCB 198	PCBs
PCB 2	PCBs
PCB 20	PCBs
PCB 200 / IUPAC 201	PCBs
PCB 201 / IUPAC 199	PCBs
PCB 202	PCBs
PCB 203/196	PCBs
PCB 206	PCBs
PCB 207	PCBs
PCB 208	PCBs
PCB 209	PCBs
PCB 26	PCBs
PCB 28/31	PCBs
PCB 29	PCBs
PCB 3	PCBs
PCB 37	PCBs
PCB 44	PCBs
PCB 45	PCBs
PCB 47/48	PCBs
PCB 49	PCBs
PCB 50	PCBs
PCB 52	PCBs
PCB 56/60	PCBs
PCB 61	PCBs
PCB 63	PCBs
PCB 66	PCBs
PCB 69	PCBs

Chemical Name	Chemical Class
PCB 70	PCBs
PCB 74	PCBs
PCB 76	PCBs
PCB 77	PCBs
PCB 8/5	PCBs
PCB 81	PCBs
PCB 82	PCBs
PCB 84	PCBs
PCB 87/115	PCBs
PCB 88	PCBs
PCB 89/90	PCBs
PCB 9	PCBs
PCB 92	PCBs
PCB 95	PCBs
PCB 99	PCBs
Aldrin	Pesticides
Alpha-HCH	Pesticides
Beta-HCH	Pesticides
Chlorpyrifos	Pesticides
Cis-chlordane (alpha-chlordane)	Pesticides
Cis-nonachlor	Pesticides
Dieldrin	Pesticides
Endosulfan I	Pesticides
Endosulfan II	Pesticides
Endosulfan Sulfate	Pesticides
Endrin	Pesticides
Gamma-chlordane	Pesticides
Gamma-HCH (g-BHC, lindane)	Pesticides
Heptachlor	Pesticides
Heptachlor epoxide	Pesticides
Hexachlorobenzene	Pesticides
Mirex	Pesticides
Oxychlordane	Pesticides
Trans-nonachlor	Pesticides
PBDE 209	PBDEs
PCB 101/90	PCBs
PCB 106	PCBs
PCB 107/123	PCBs

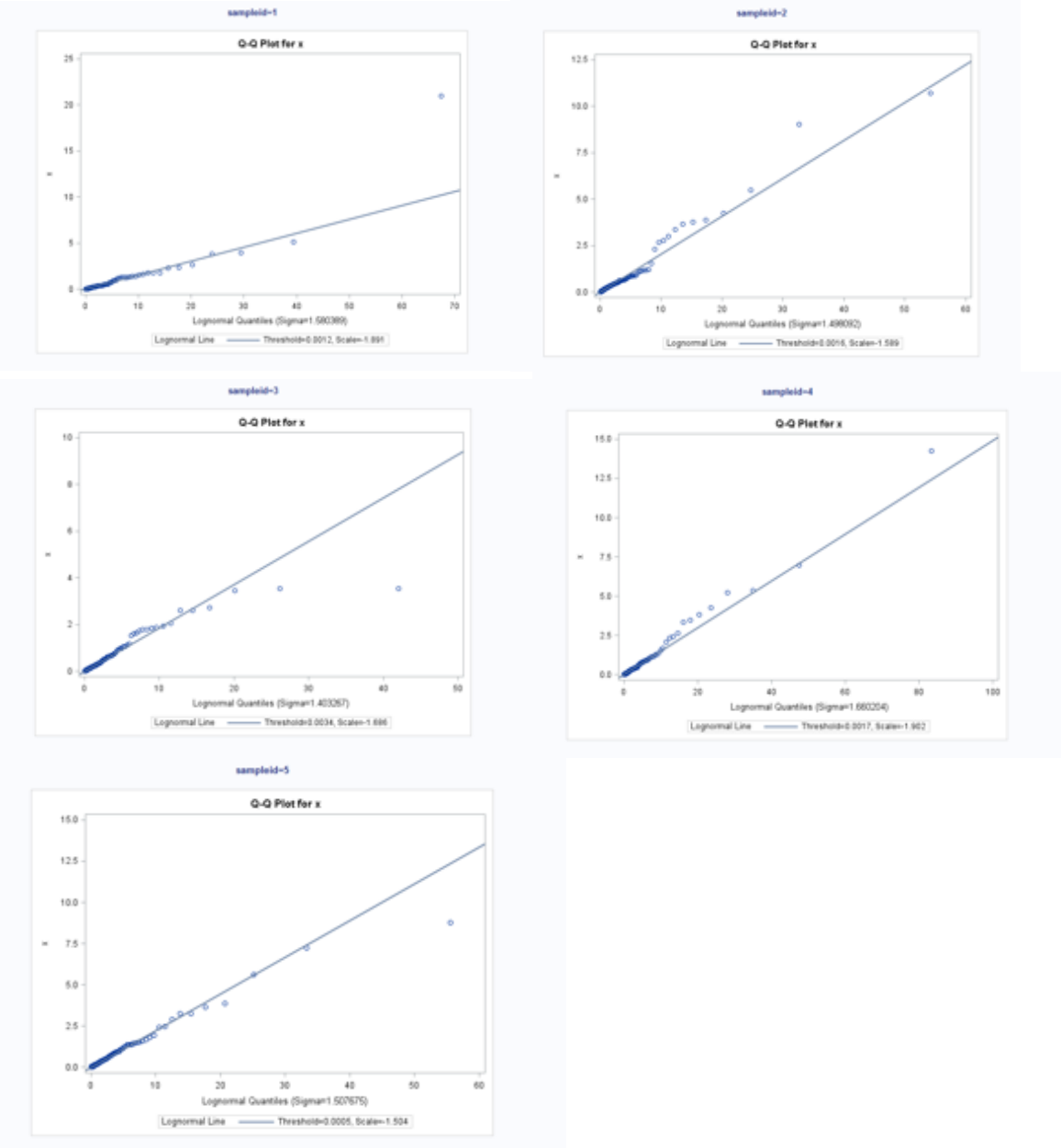
Chemical Name	Chemical Class
PCB 108	PCBs
PCB 115	PCBs
PCB 118	PCBs
PCB 138	PCBs
PCB 158	PCBs
PCB 163	PCBs
PCB 164	PCBs
PCB 170	PCBs
PCB 180	PCBs
PCB 190	PCBs
PCB 193	PCBs
PCB 28	PCBs
PCB 31	PCBs
PCB 47	PCBs
PCB 48	PCBs
PCB 5	PCBs
PCB 56	PCBs
PCB 60	PCBs
PCB 63/76	PCBs
PCB 8	PCBs
PCB 87	PCBs
PCB 88/95	PCBs
PCB 92/84/89	PCBs

Appendix II. Percent of total measurements within each analyte that are considered reportable data.

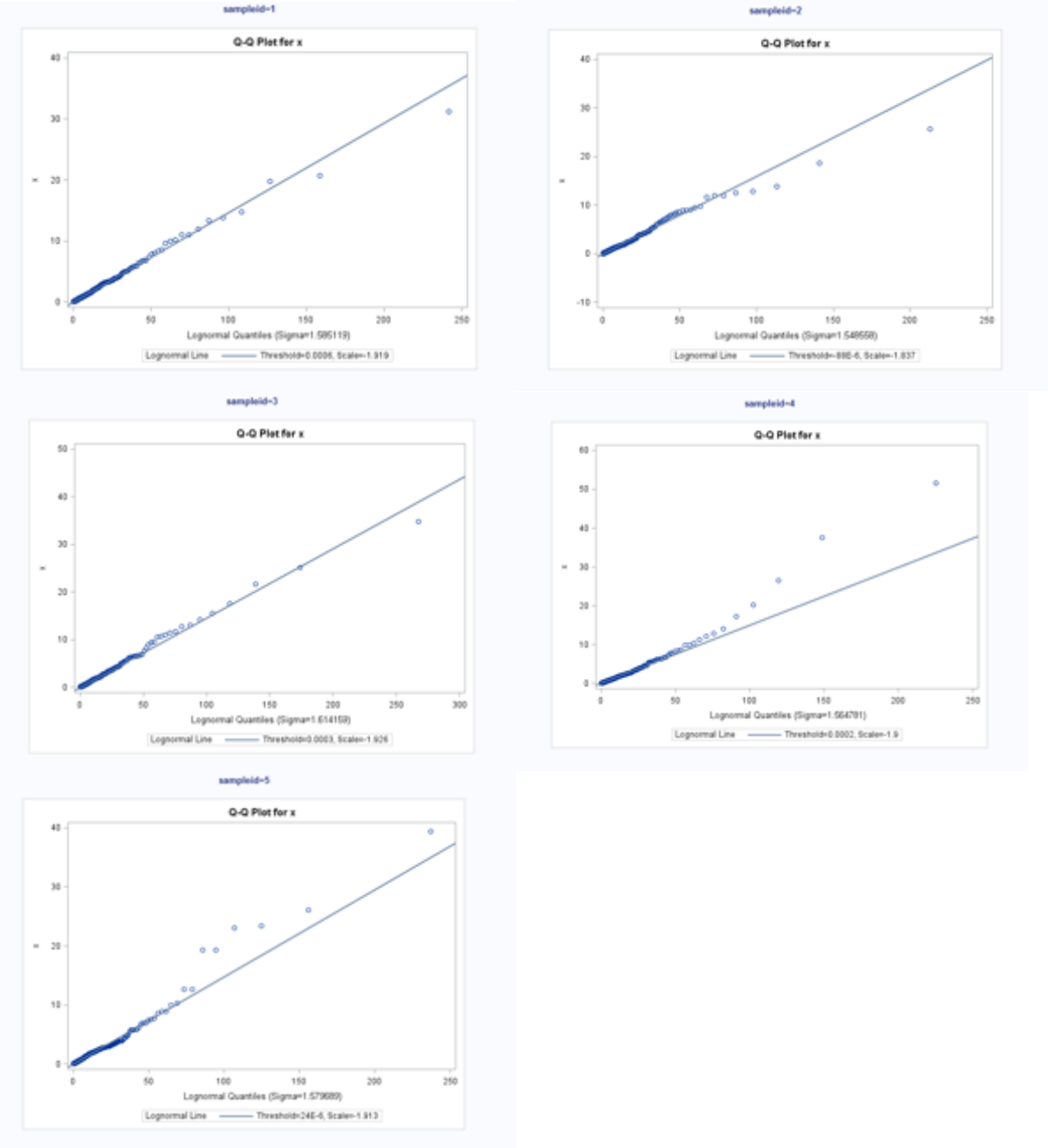
Percent of Total Analyte Measurements that are Reportable	Percent of all Analytes
0%	30.99%
5.56%	4.09%
8.33%	1.75%
11.11%	7.02%
16.67%	3.51%
22.22%	1.17%
25%	0.58%
27.78%	0.58%
33.33%	6.43%
38.89%	1.17%
41.67%	0.58%
44.44%	2.92%
50%	2.34%
55.56%	2.92%
58.33%	0.58%
61.11%	2.34%
66.67%	3.51%
72.22%	1.17%
77.78%	1.17%
83.33%	0.58%
88.89%	2.92%
91.67%	1.17%
94.44%	2.92%
100%	17.54%

Appendix III. Simulated results from a log-normal distribution.

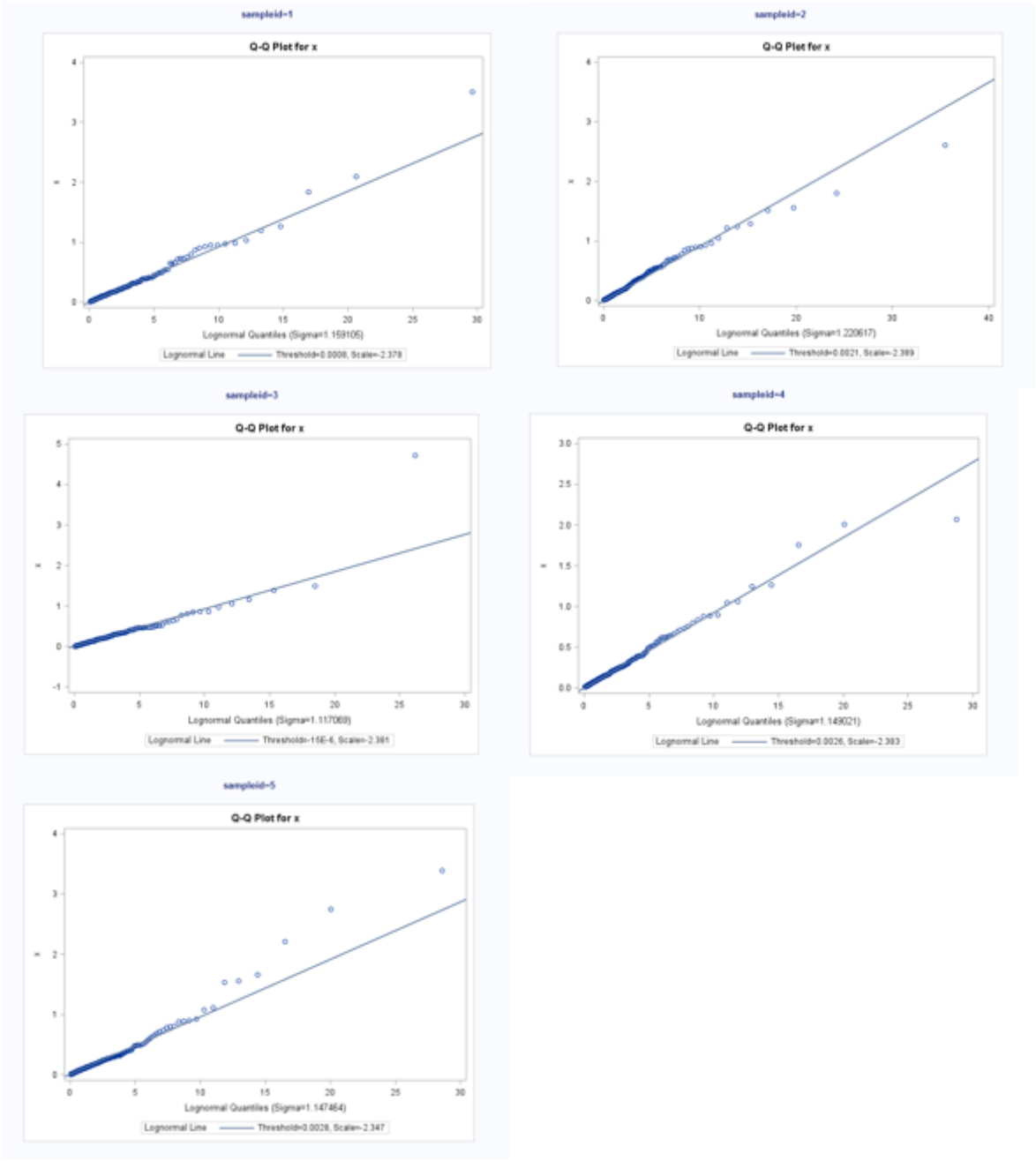
DDTs



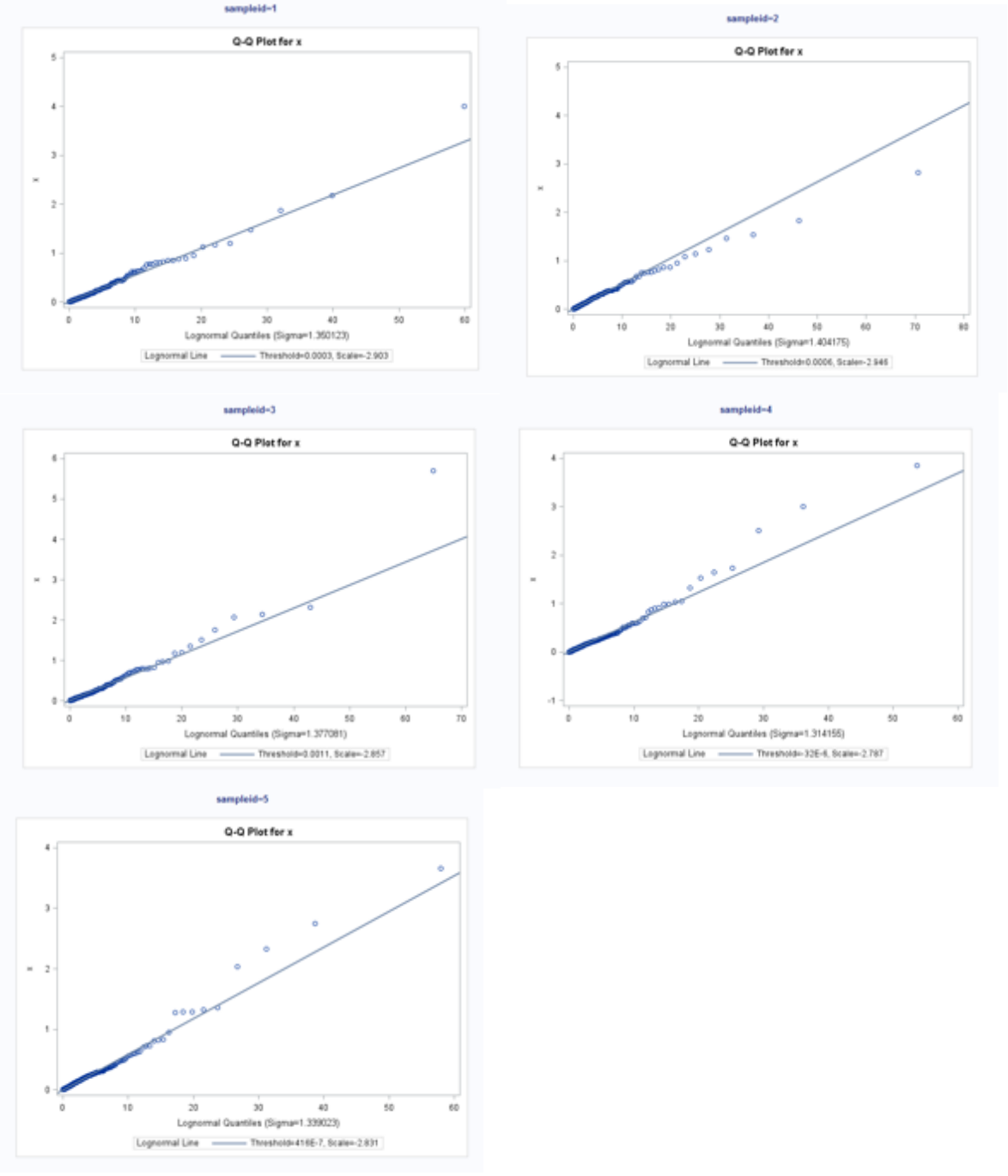
PCBs


sampleid=1
Q-Q Plot for x
Lognormal Quantiles (Sigma=1.585119)
Lognormal Line
Threshold=0.0006, Scale=-1.919
sampleid=2
Q-Q Plot for x
Lognormal Quantiles (Sigma=1.548558)
Lognormal Line
Threshold=-886-6, Scale=-1.837
sampleid=3
Q-Q Plot for x
Lognormal Quantiles (Sigma=1.614159)
Lognormal Line
Threshold=0.0003, Scale=-1.926
sampleid=4
Q-Q Plot for x
Lognormal Quantiles (Sigma=1.564781)
Lognormal Line
Threshold=0.0002, Scale=-1.9
sampleid=5
Q-Q Plot for x
Lognormal Quantiles (Sigma=1.579689)
Lognormal Line
Threshold=24E-6, Scale=-1.913

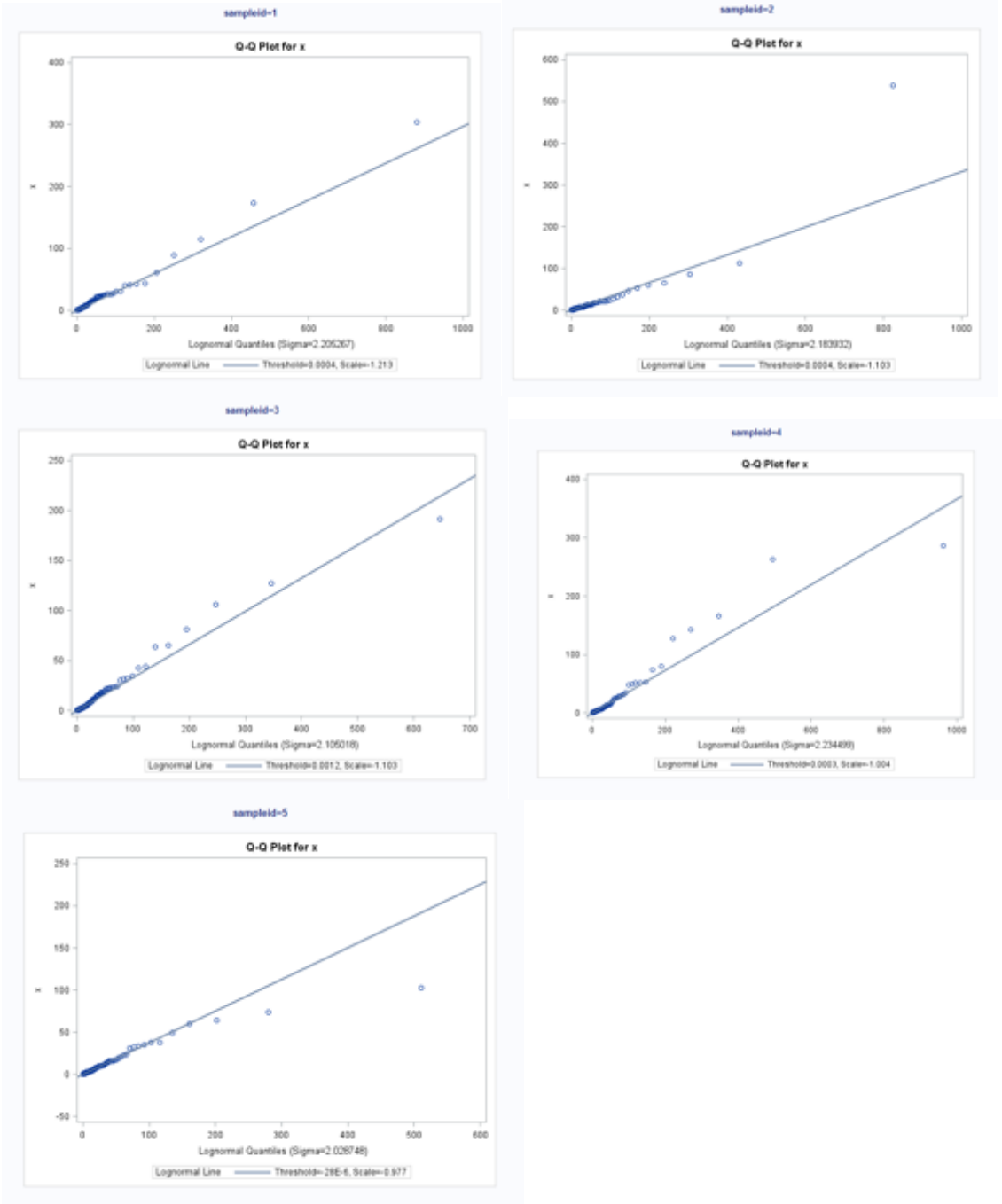
PBDEs


sampleid=1
Q-Q Plot for x
Lognormal Quantiles (Sigma=1.159105)
Lognormal Line
Threshold=0.0008, Scale=-2.378
sampleid=2
Q-Q Plot for x
Lognormal Quantiles (Sigma=1.220617)
Lognormal Line
Threshold=0.0021, Scale=-2.389
sampleid=3
Q-Q Plot for x
Lognormal Quantiles (Sigma=1.117069)
Lognormal Line
Threshold=-15E-6, Scale=-2.381
sampleid=4
Q-Q Plot for x
Lognormal Quantiles (Sigma=1.149021)
Lognormal Line
Threshold=0.0026, Scale=-2.383
sampleid=5
Q-Q Plot for x
Lognormal Quantiles (Sigma=1.147464)
Lognormal Line
Threshold=0.0028, Scale=-2.347

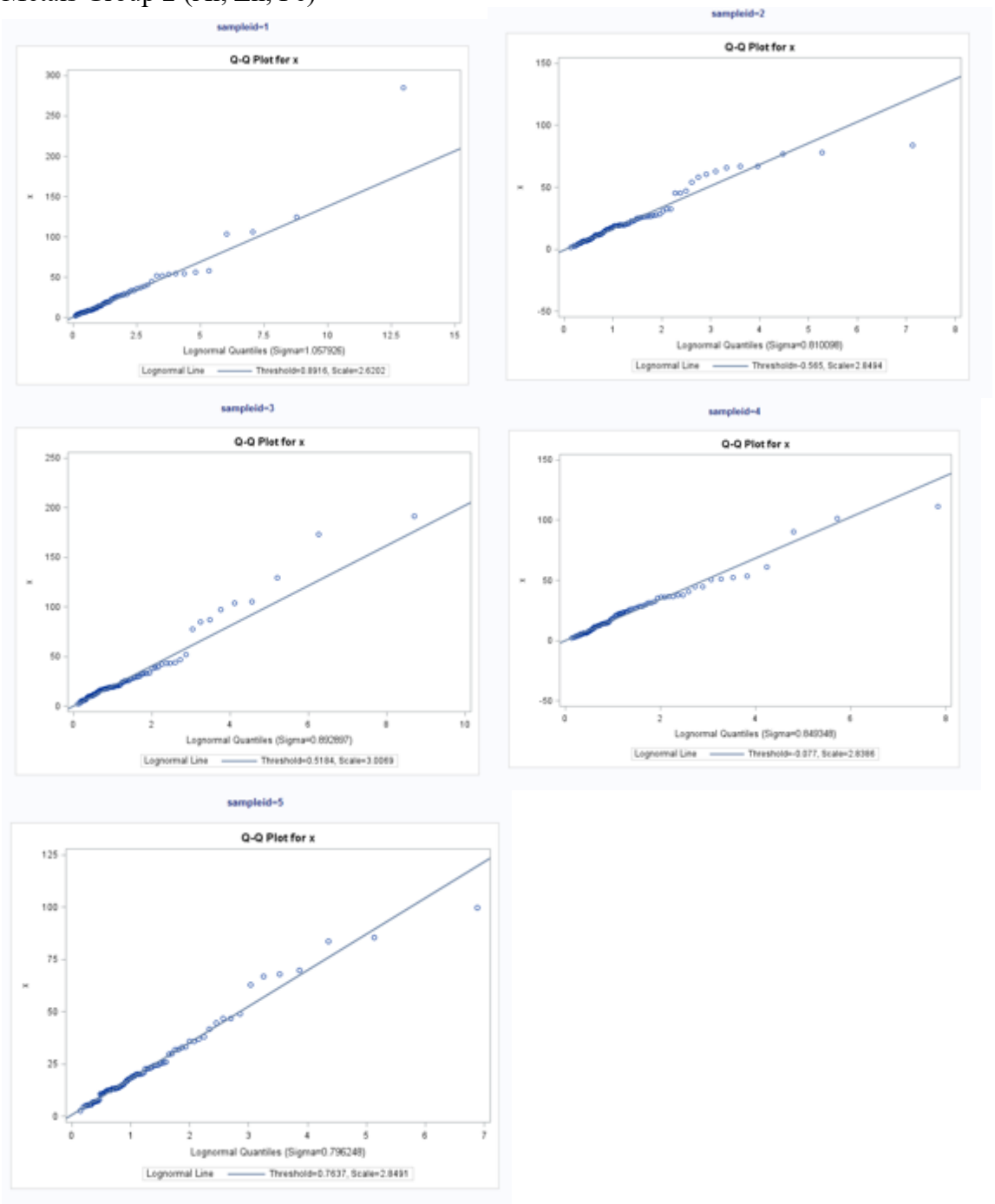
Pesticides


sampleid=1
Q-Q Plot for x
Lognormal Quantiles (Sigma=1.350123)
Lognormal Line
Threshold=0.0003, Scale=-2.903
sampleid=2
Q-Q Plot for x
Lognormal Quantiles (Sigma=1.404175)
Lognormal Line
Threshold=0.0006, Scale=-2.946
sampleid=3
Q-Q Plot for x
Lognormal Quantiles (Sigma=1.377081)
Lognormal Line
Threshold=0.0011, Scale=-2.857
sampleid=4
Q-Q Plot for x
Lognormal Quantiles (Sigma=1.314155)
Lognormal Line
Threshold=.32E-6, Scale=-2.787
sampleid=5
Q-Q Plot for x
Lognormal Quantiles (Sigma=1.339023)
Lognormal Line
Threshold=.416E-7, Scale=-2.831

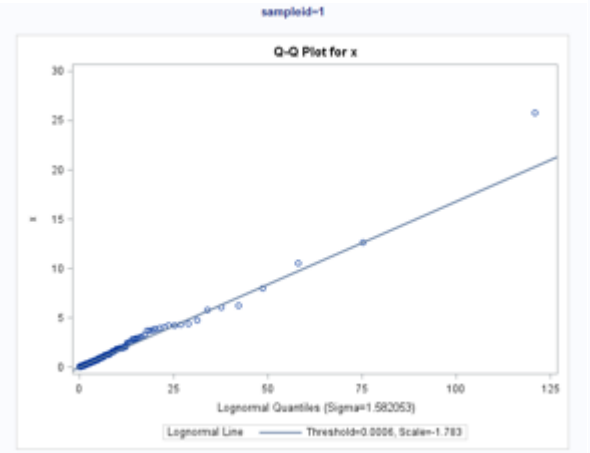
Metals Group 1 (All)

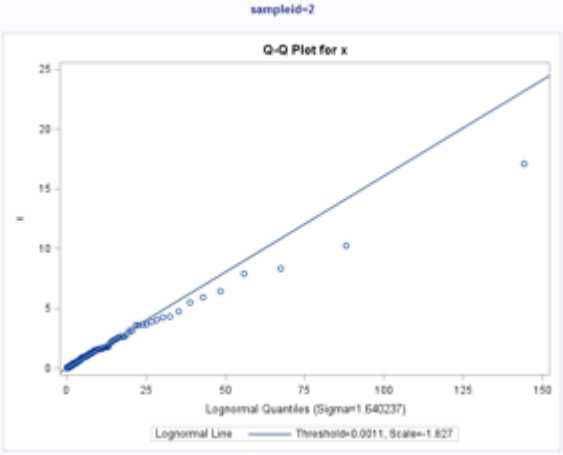


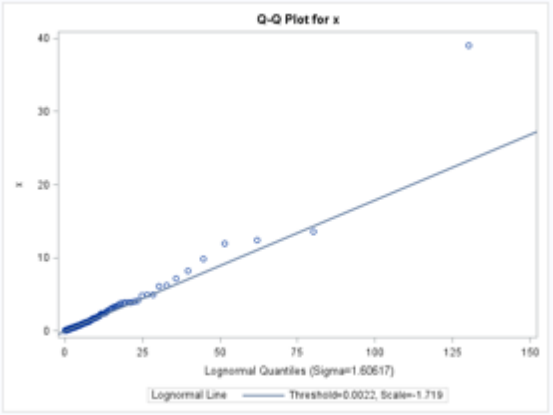
Metals Group 2 (Al, Zn, Fe)

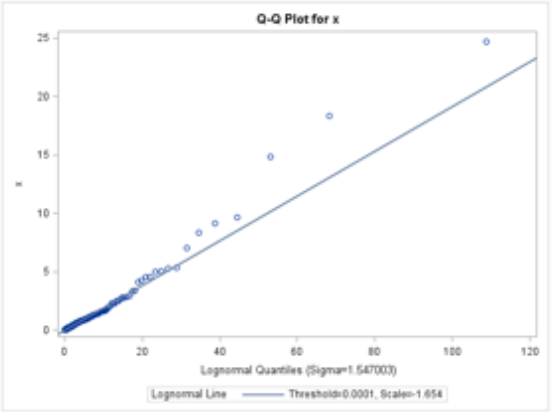


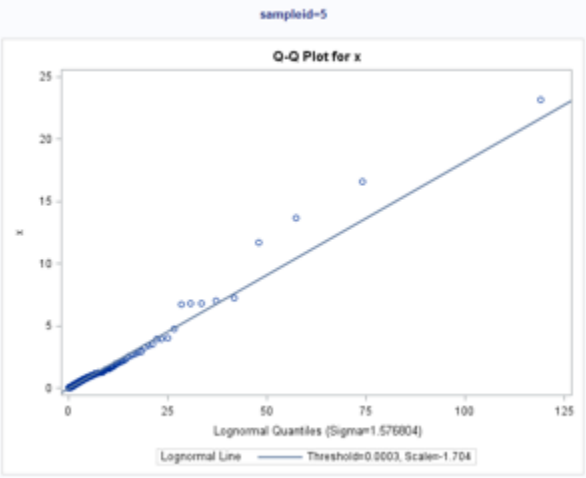
Metals Group 3





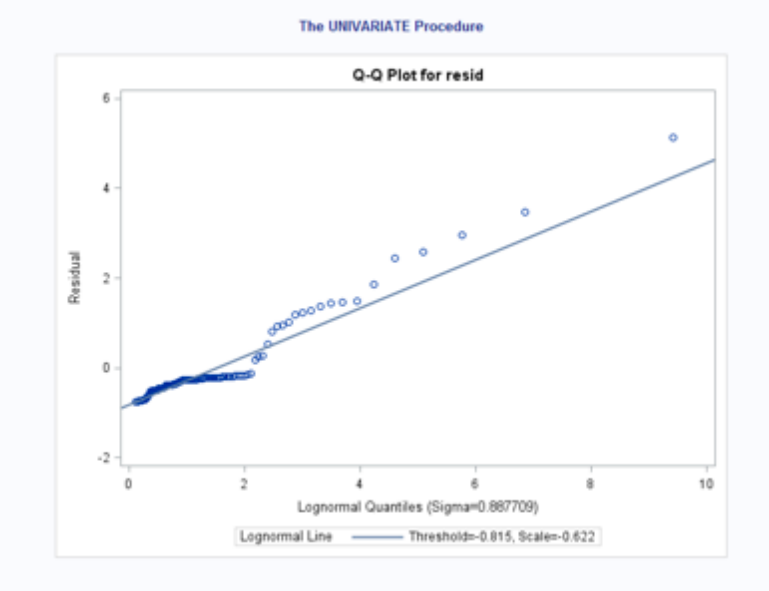




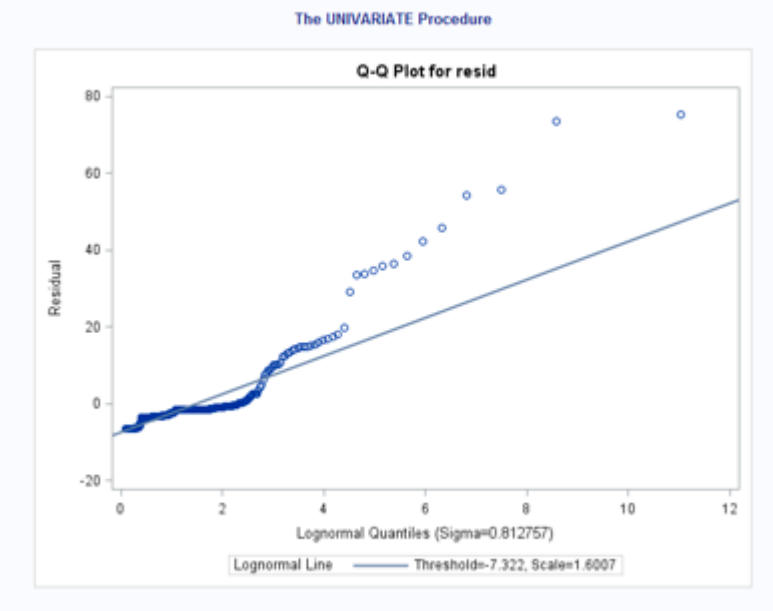


Appendix IV. Residual plots using half of the LOD as the outcome.

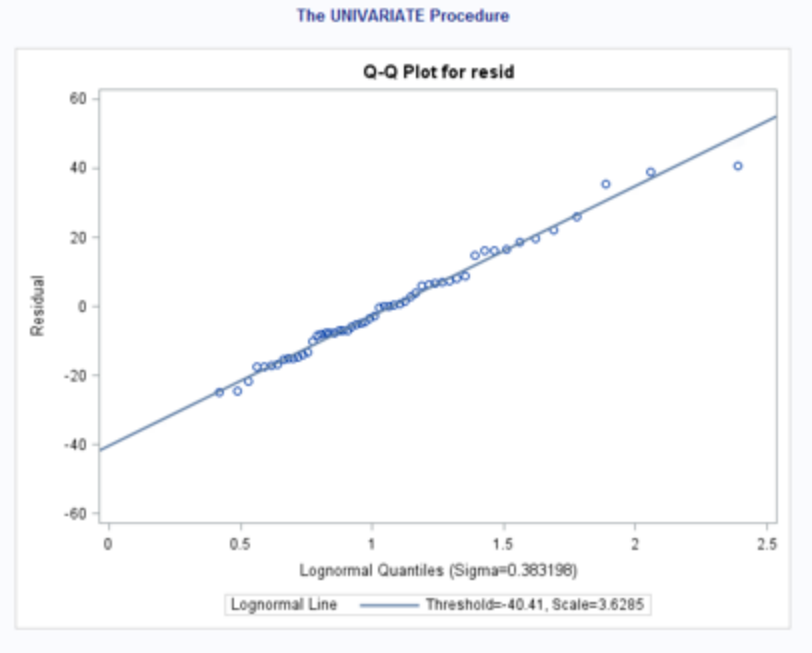
DDTs



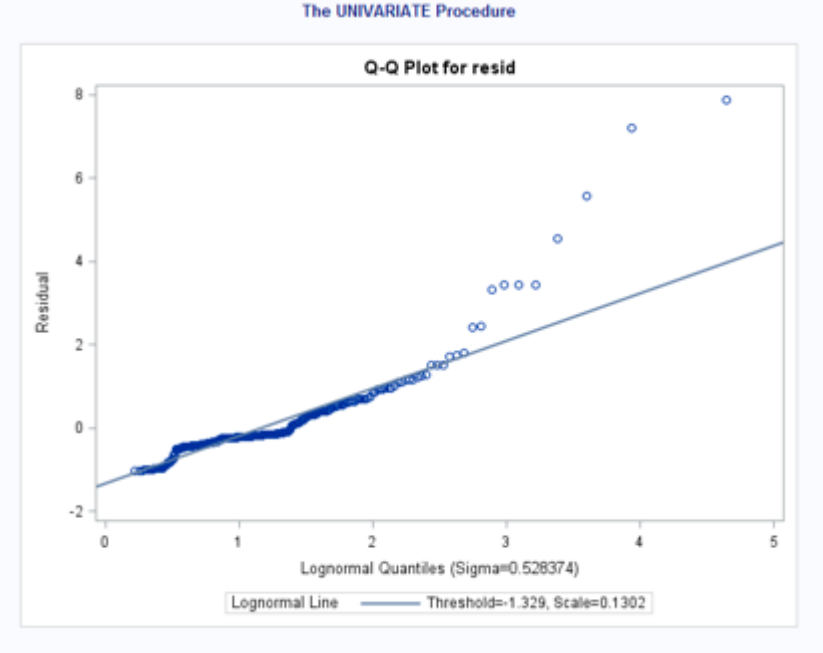
Metals Group 1 (All)



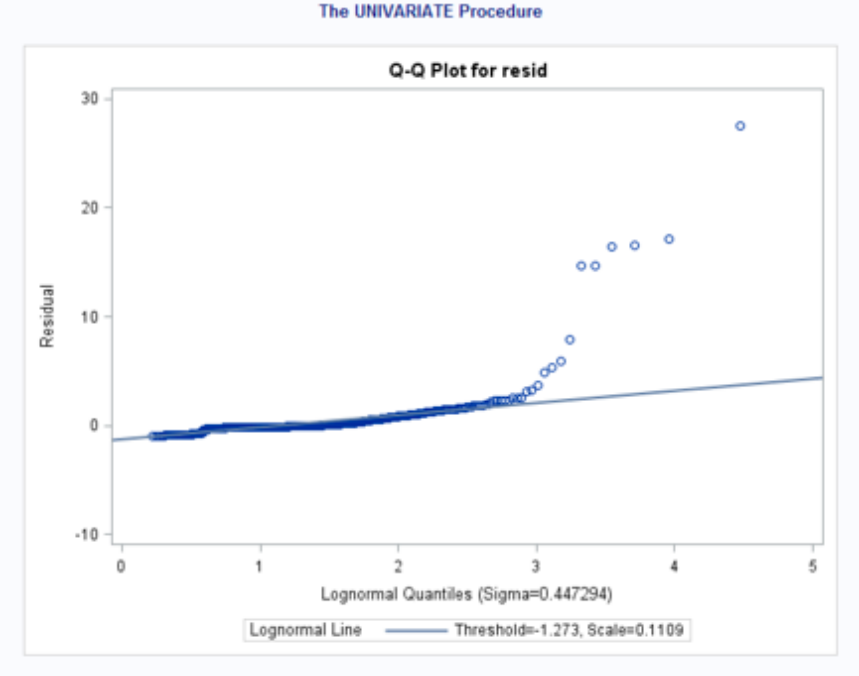
Metals Group 2 (Al, Zn, Fe)



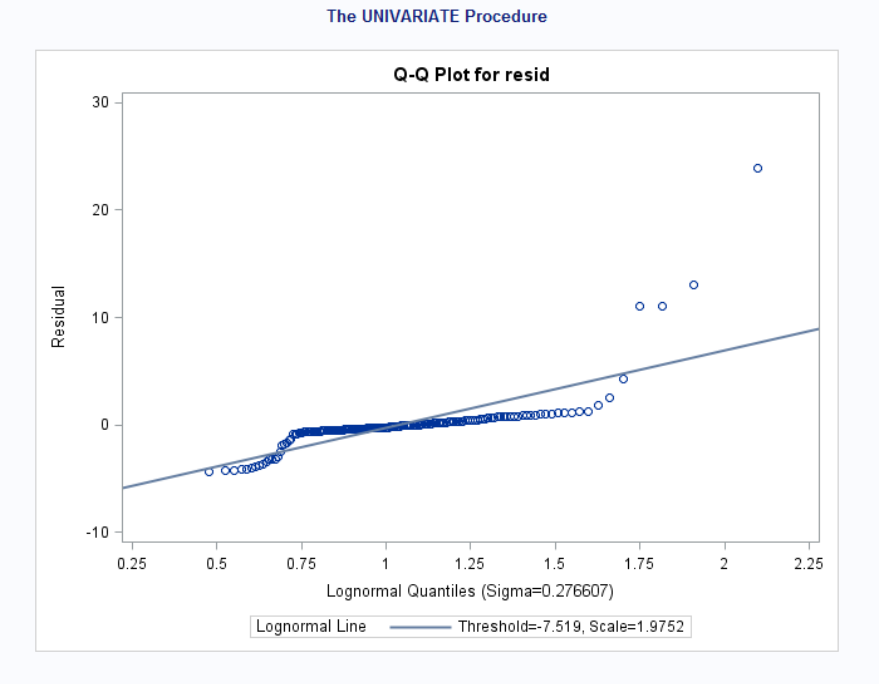
Metals Group 3



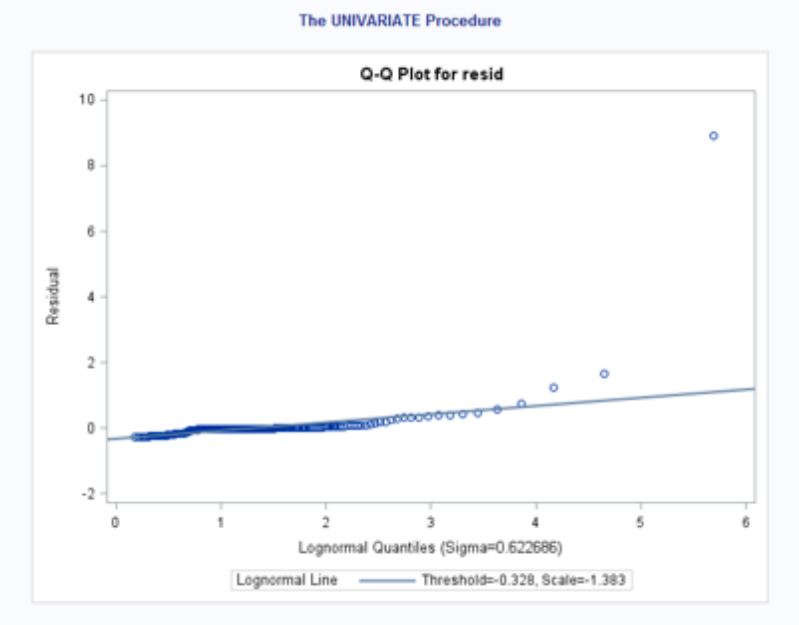
PCBs



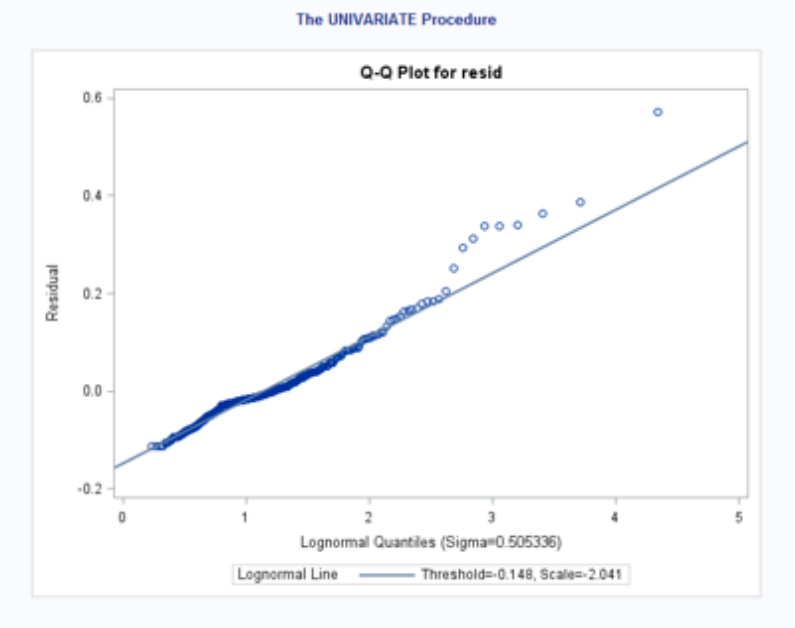
Aroclor 1268



PBDEs

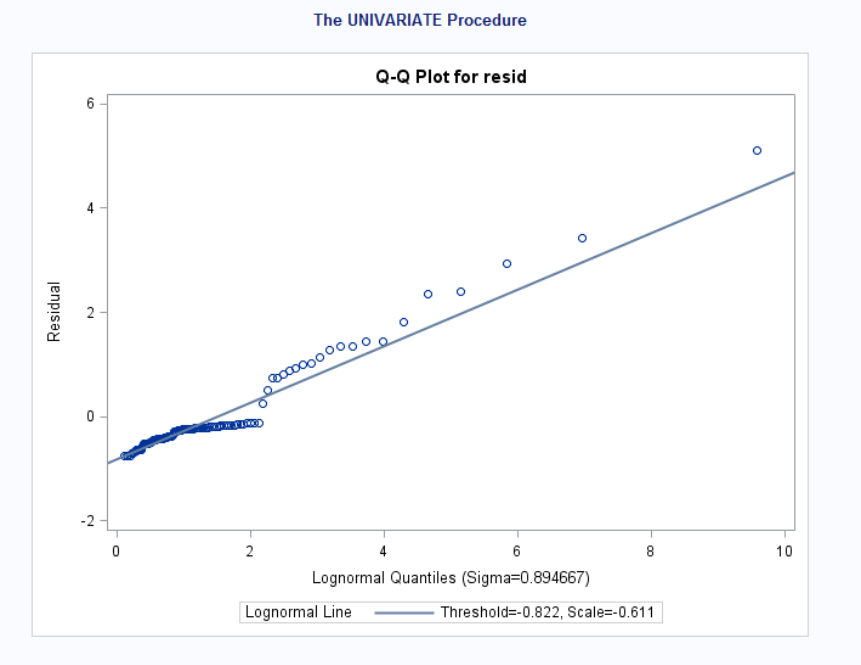


Pesticides

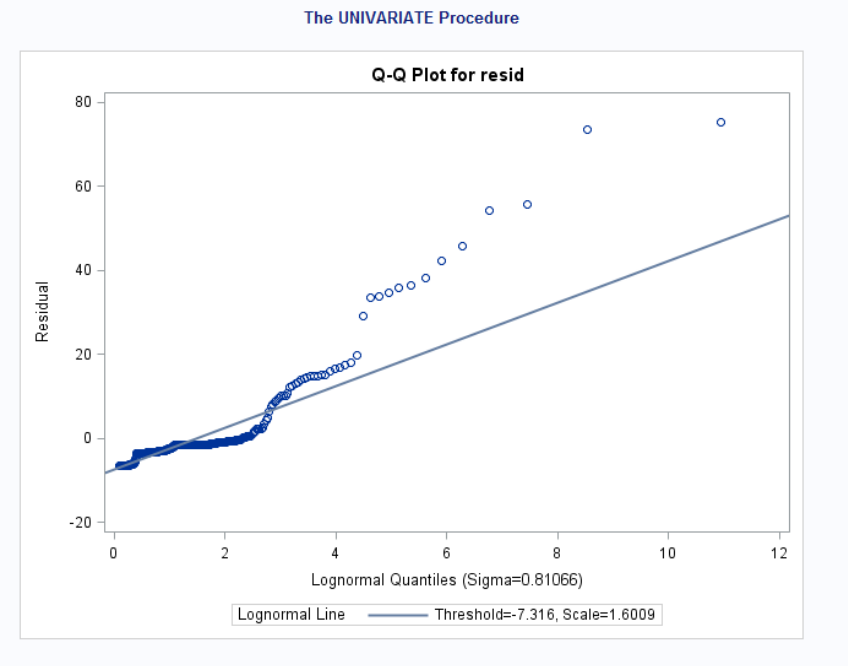


Appendix V. Residual plots using the LOD as the outcome variable.

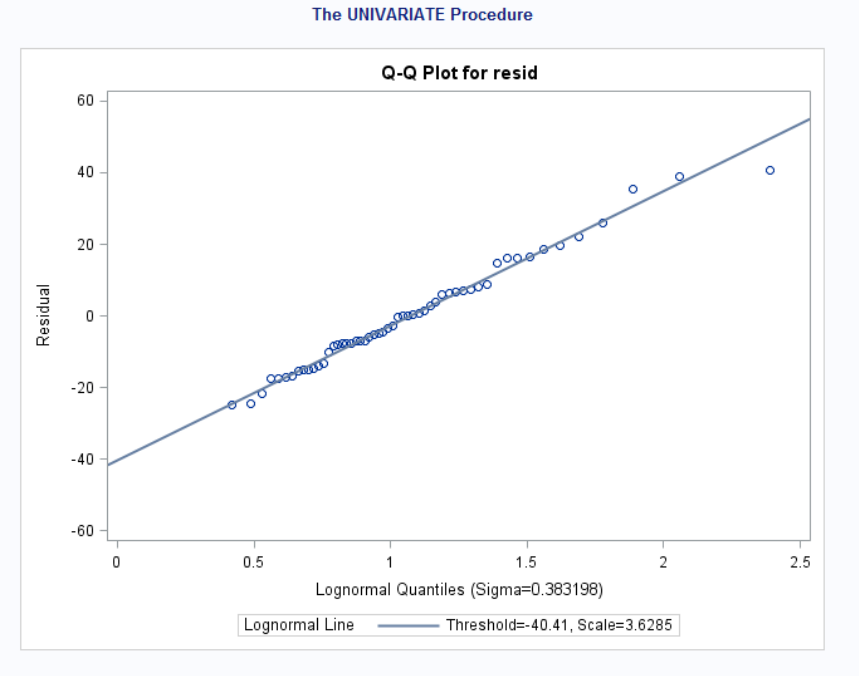
DDTs



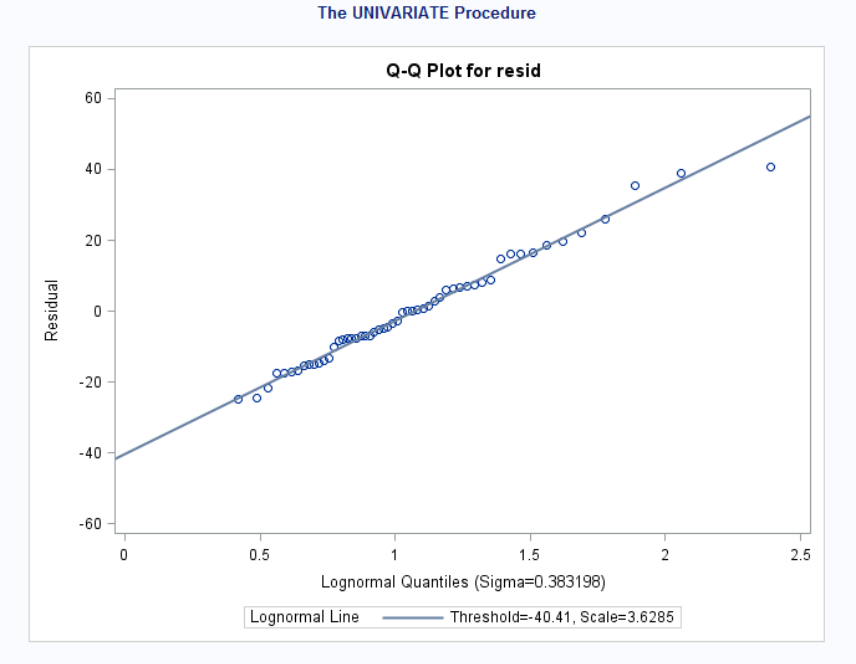
Metals Group 1 (All)



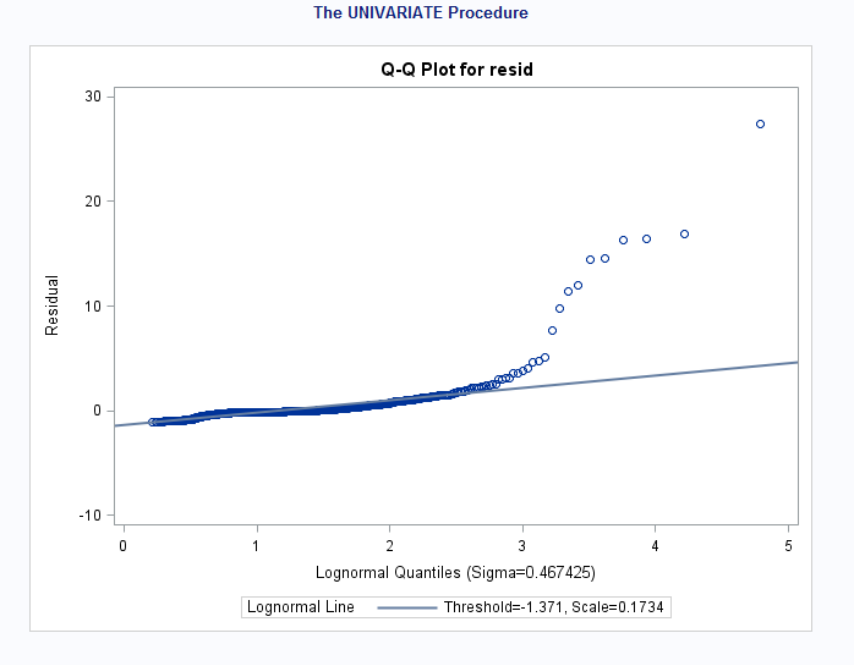
Metals Group 2 (Al, Zn, Fe)



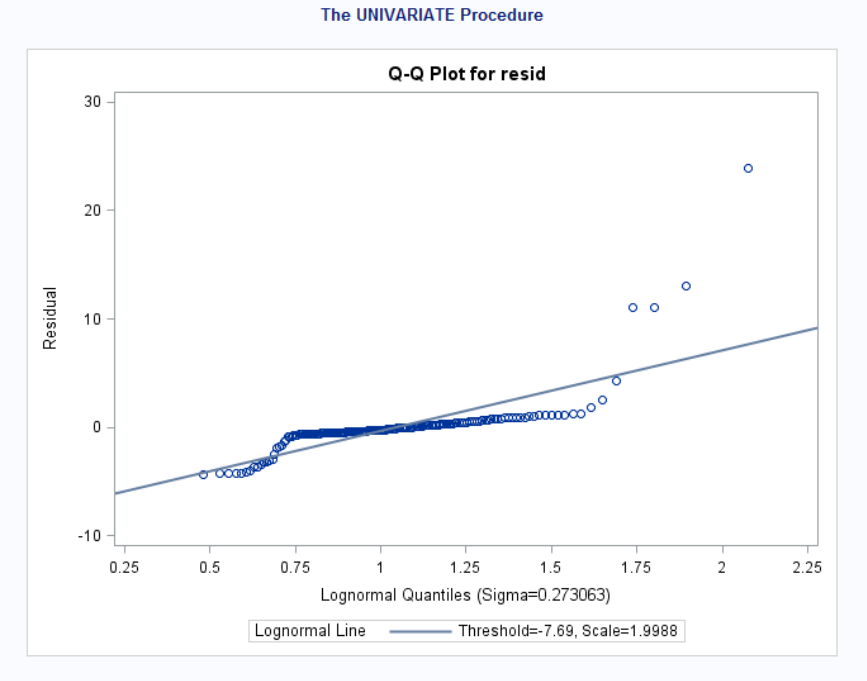
Metals Group 3



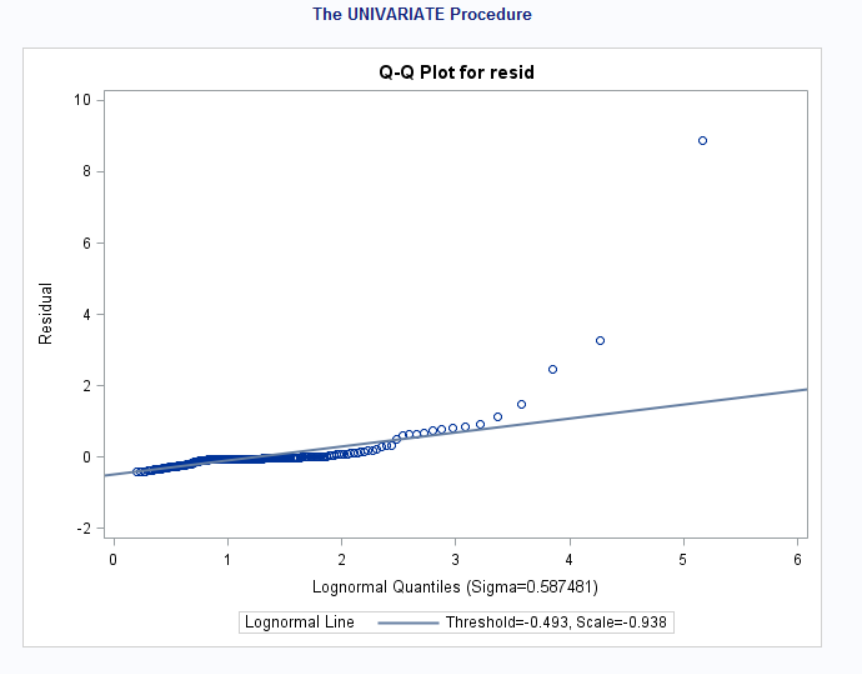
PCBs



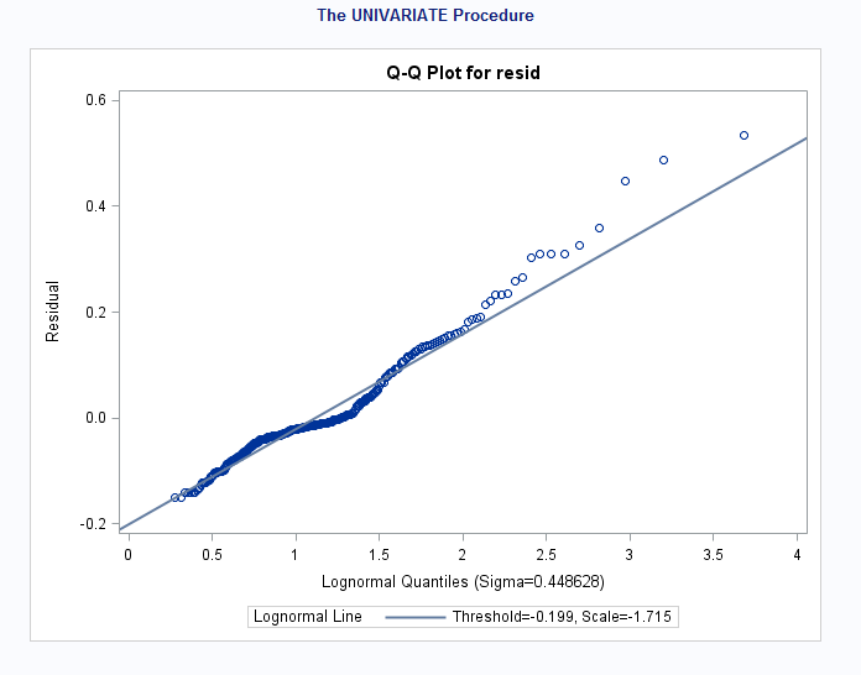
Aroclor 1268



PBDEs



Pesticides

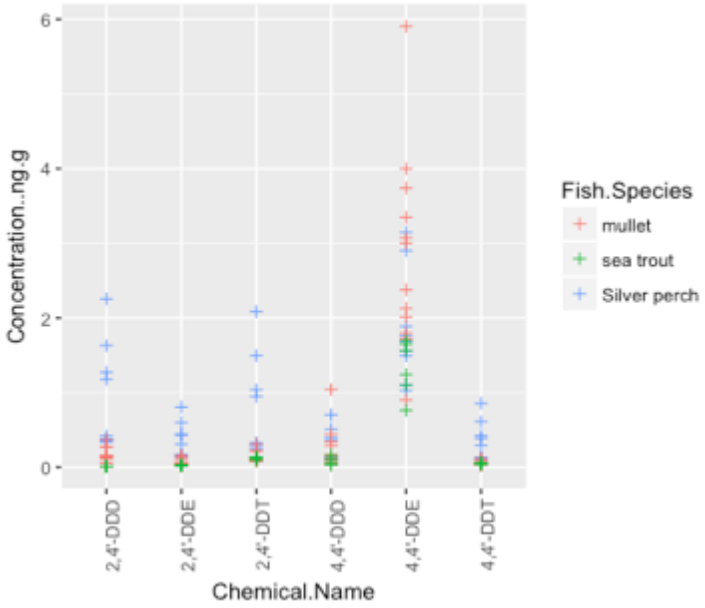


Appendix VI. Example code from SAS® PROC NLMIXED.

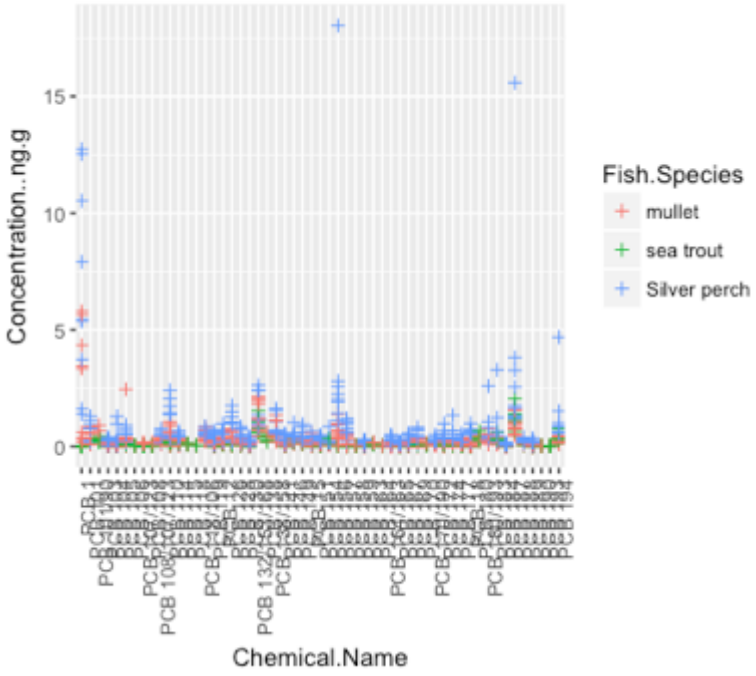
```
%macro mlEstimator(dsn);

PROC NLMIXED data=&dsn maxiter=100 maxfunc=1000 tech=trureg cov;
parms mu=2 alphMullet=2 alphSeaTrout=2 bet3=2 gamm3Mullet=2 gamm3SeaTrout=2
sigsq=2;
bounds sigsq >= 0;
pi = constant("pi");
f1=1;
eta1 = mu + alphMullet*effA + alphSeaTrout*effB + bet3*eff3 + gamm3Mullet*A3ef +
gamm3SeaTrout*B3ef;
if censored=0 then do;
f1 = (1 / sqrt(2*pi*sigsq))*exp(-0.5*((log_concentration-eta1)**2)/sigsq);
end;
else if censored=1 then do;
   f1 = CDF("NORMAL", log_concentration, eta1, sqrt(sigsq));
end;
ll = log(f1);
model log_concentration ~ general(ll);
title "Maximum Likelihood Estimation for &dsn";
run;

%mend mlEstimator;
```

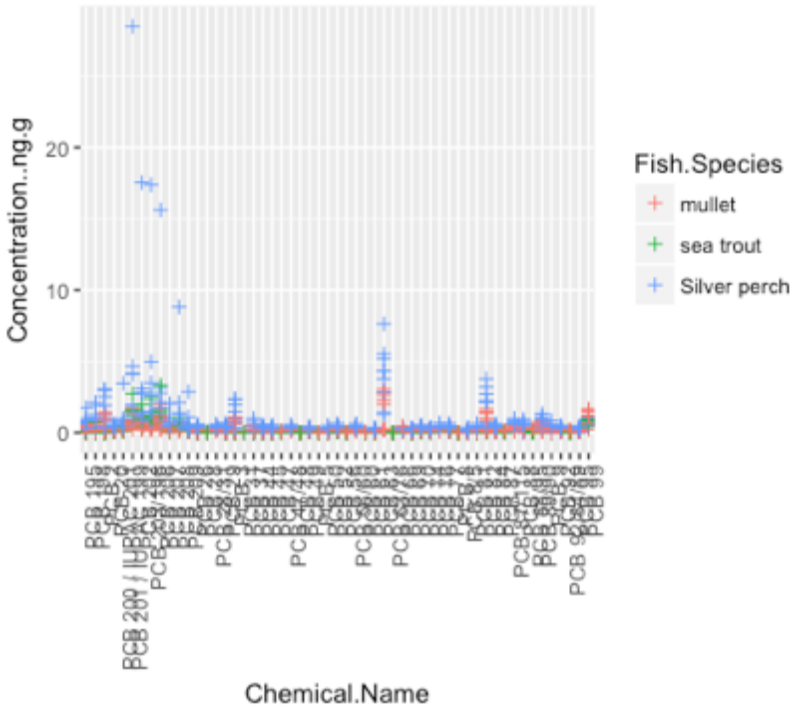
Appendix VII. Comparing across fish species, substituting the LOD for values lower than the detection limit.

DDTs

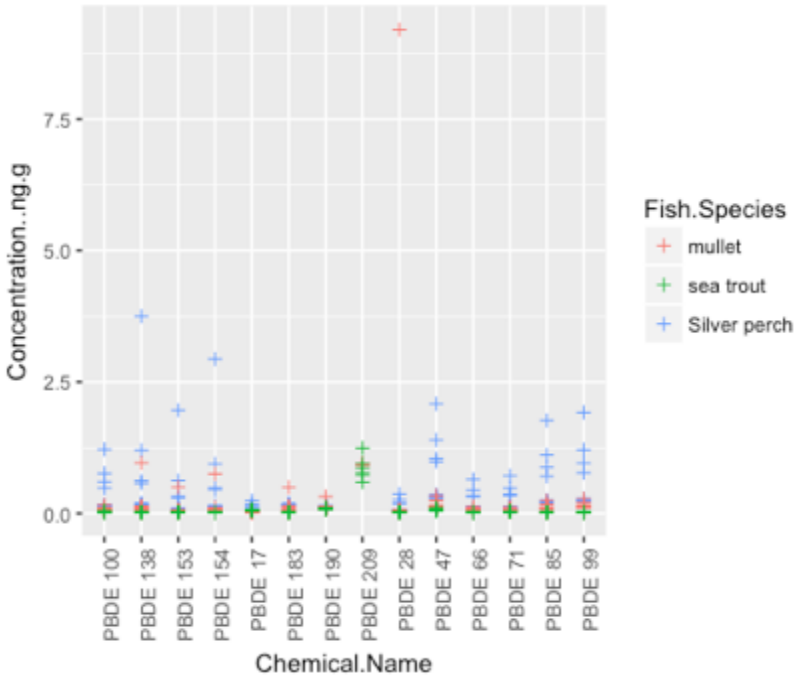


PCBs

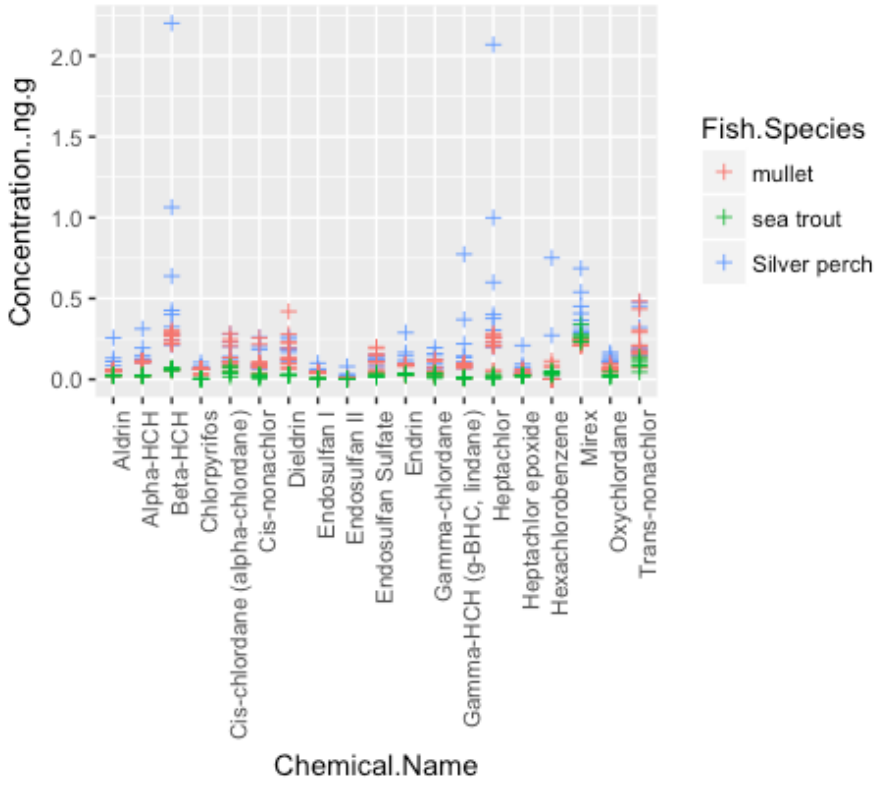

Concentration..ng.g
15
10
5
0
Fish.Species
mullet
sea trout
Silver perch
PCB 10
PCB 13
Chemical.Name


Concentration..ng.g
20
10
0
Fish.Species
mullet
sea trout
Silver perch
PCB 200 / IUPAC
PCB 9
Chemical.Name

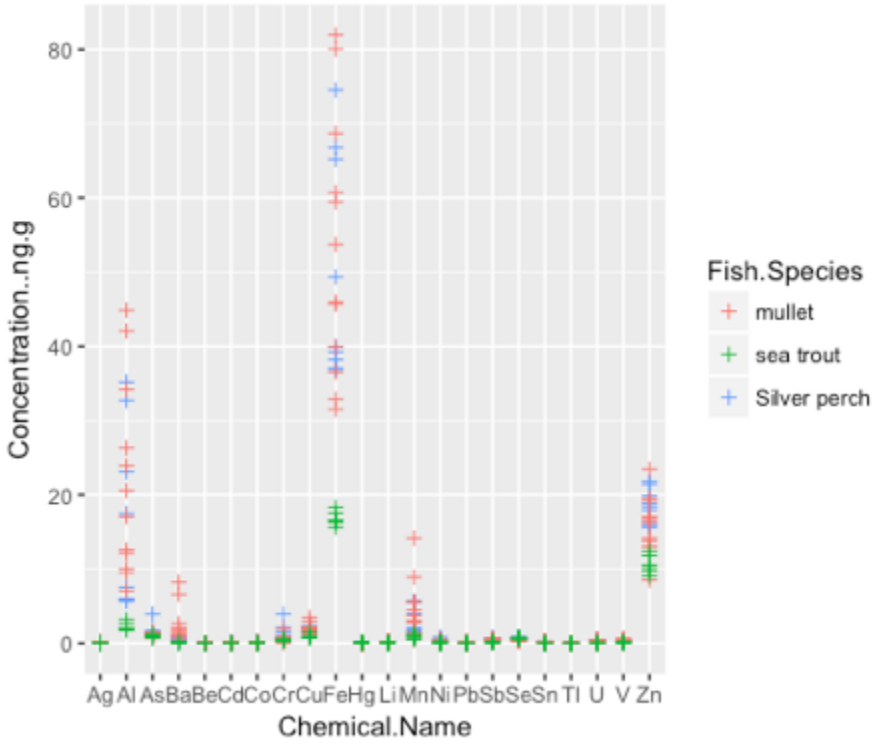
PBDEs



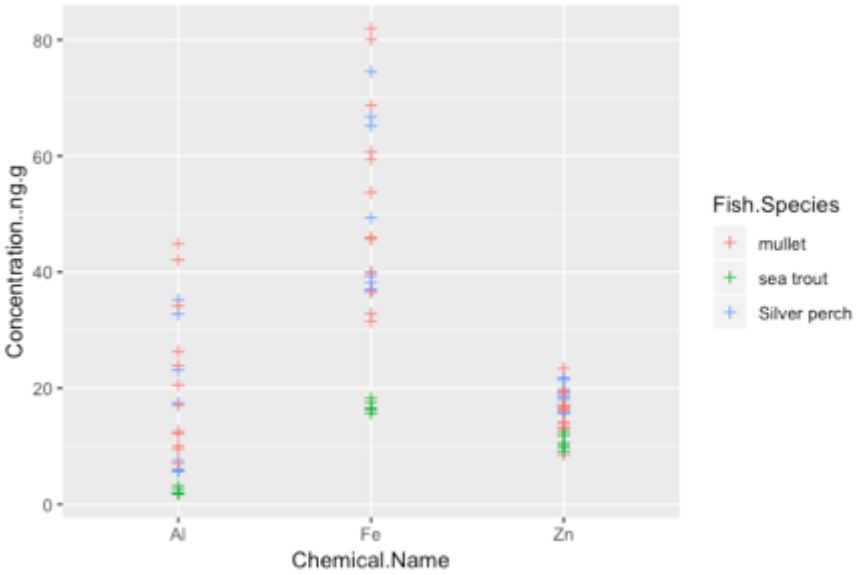
Pesticides



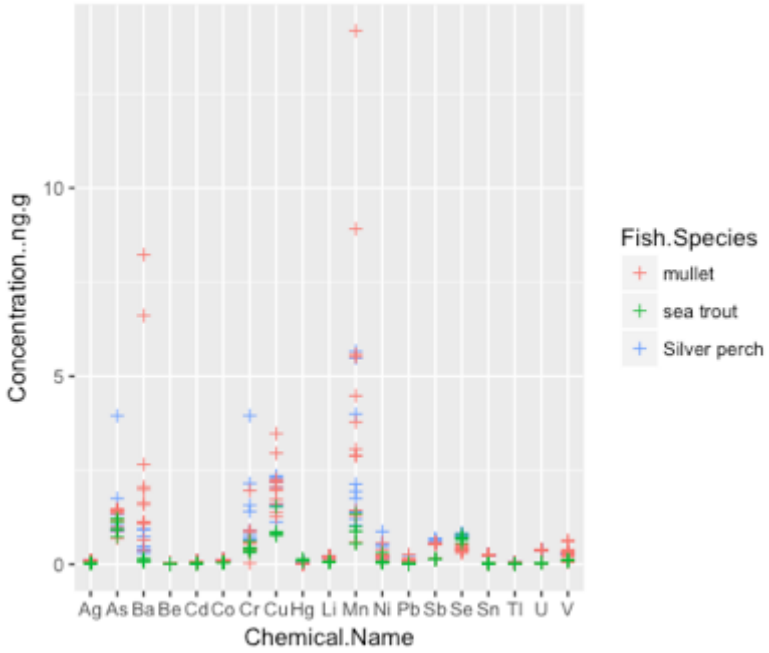
Metals Group 1 (All)



Metals Group 2 (Al, Zn, Fe)

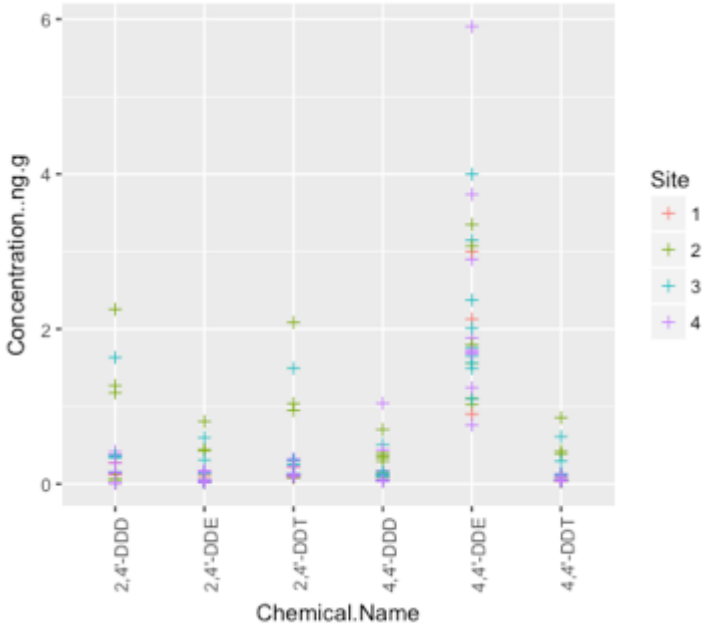


Metals Group 3

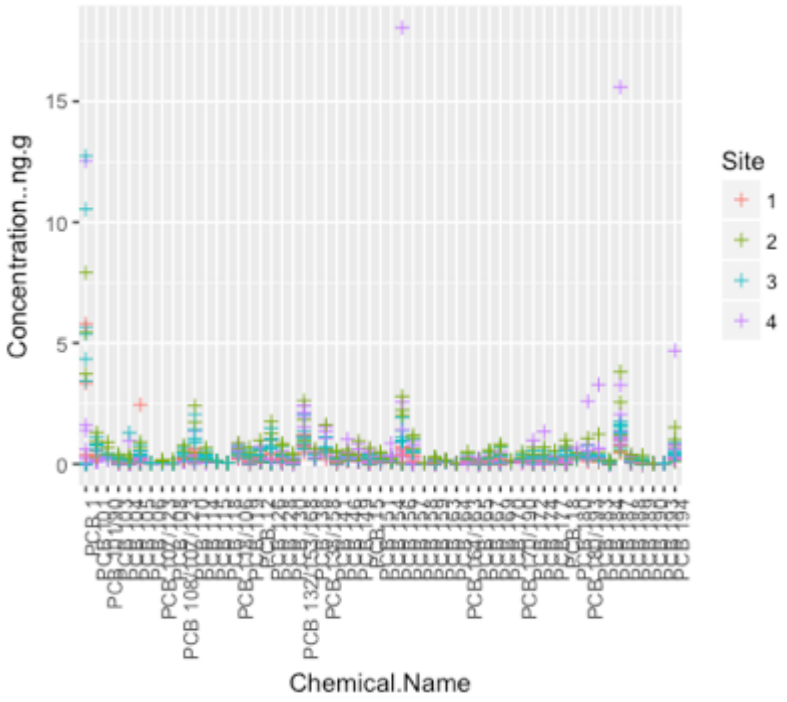


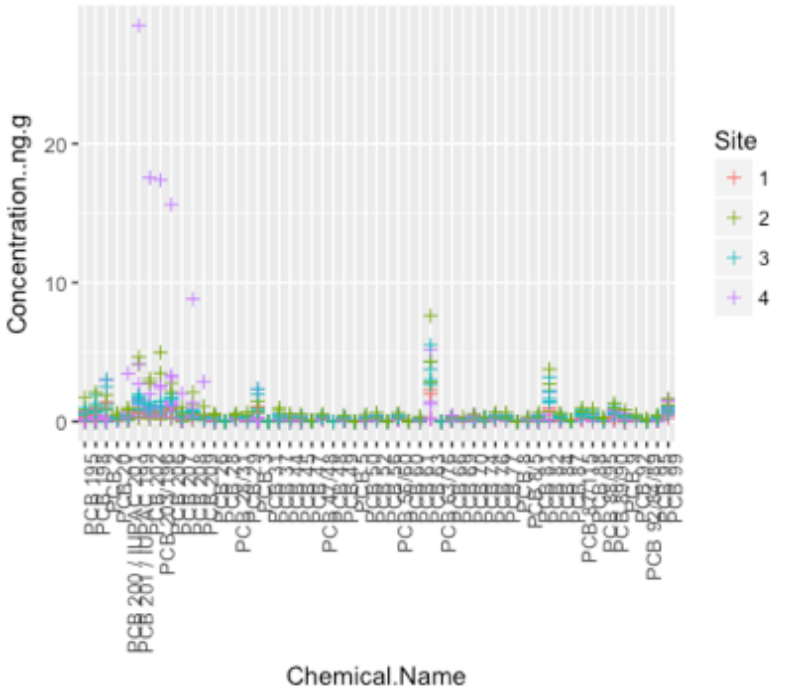
Appendix VIII. Comparing across site, substituting the LOD for values lower than the detection limit.

DDTs

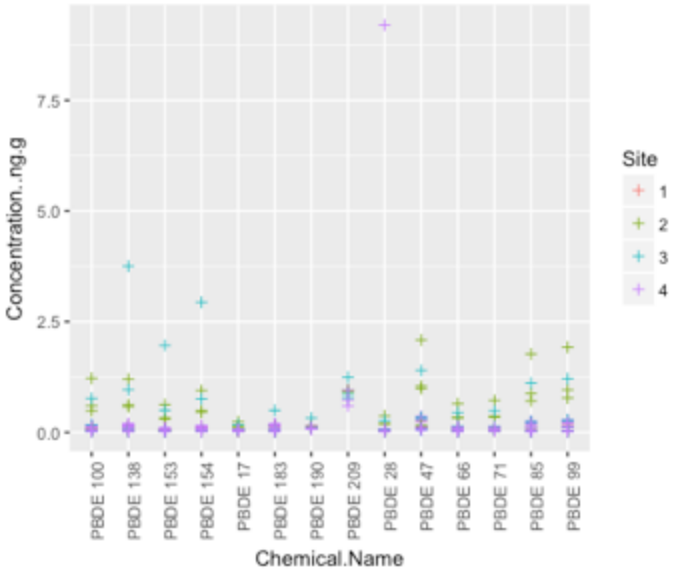


PCBs

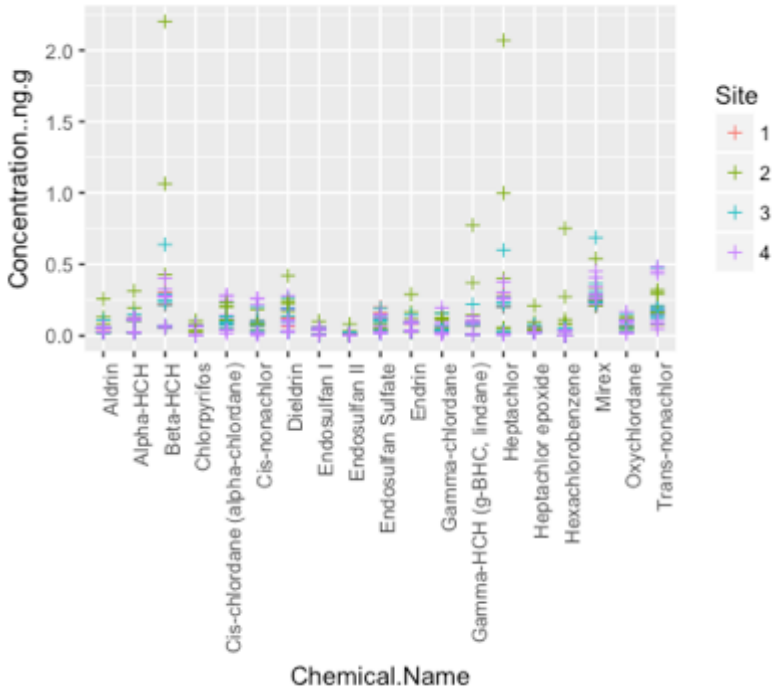

Concentration...ng.g
15
10
5
0
Site
1
2
3
4
Chemical.Name


Concentration...ng.g
20
10
0
Site
1
2
3
4
Chemical.Name

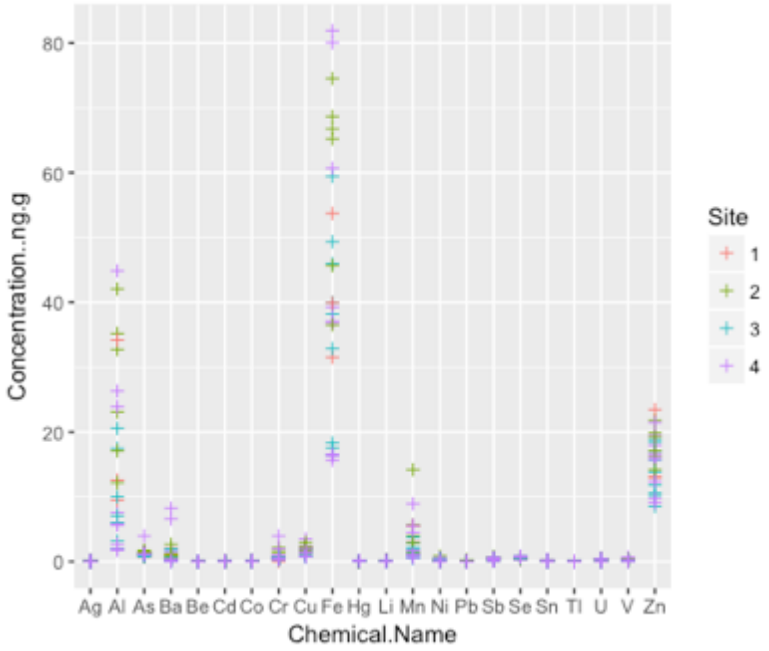
PBDEs



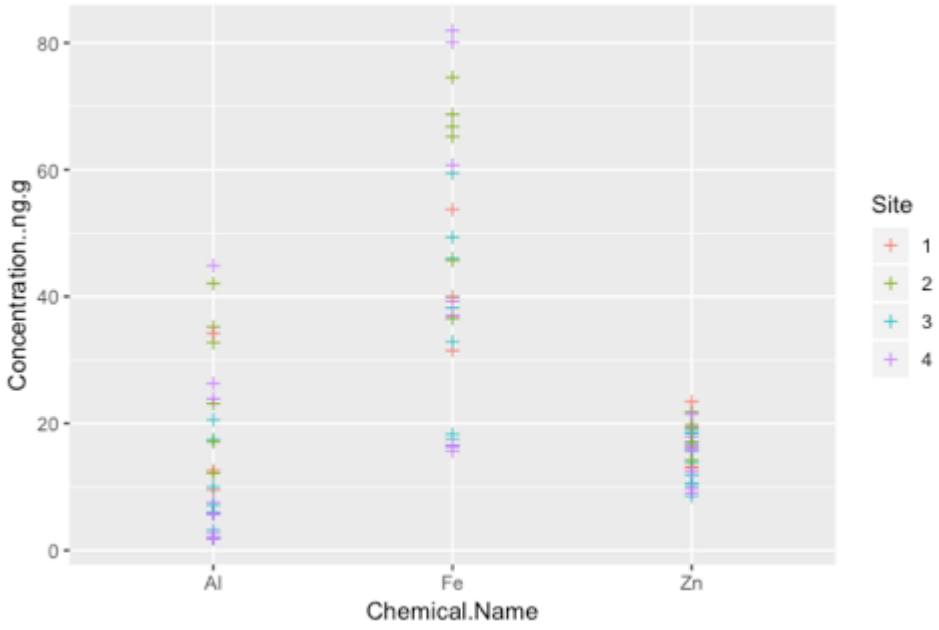
Pesticides



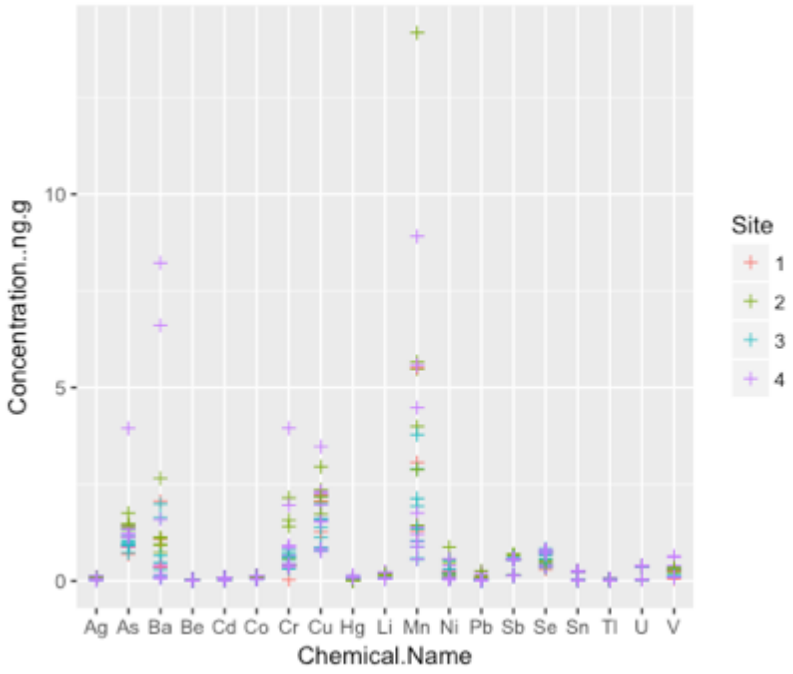
Metals Group 1 (All)



Metals Group 2 (Al, Zn, Fe)

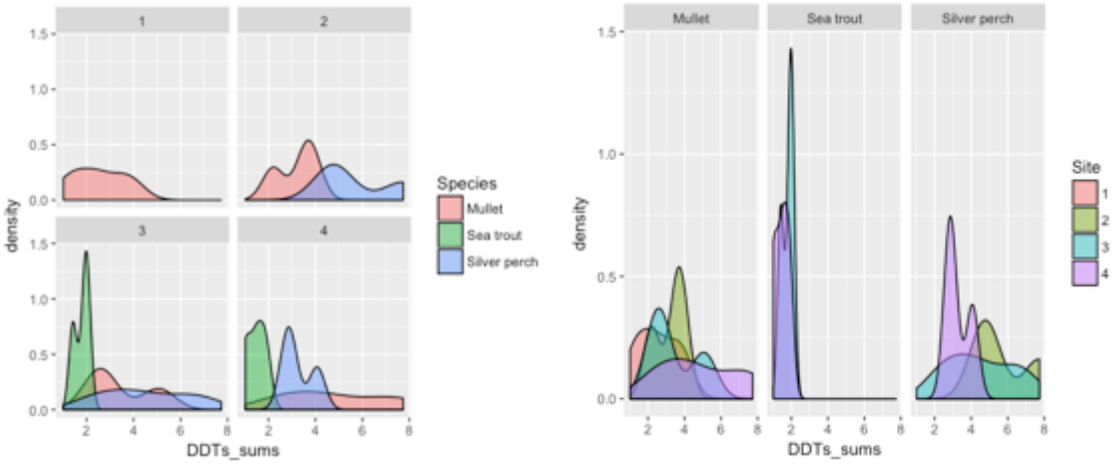


Metals Group 3

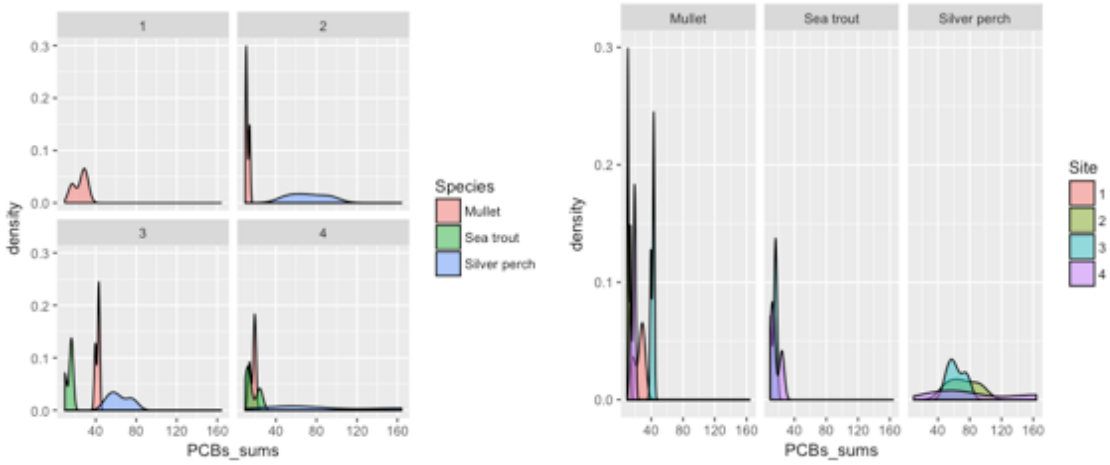


Appendix IX. Comparing summed concentrations of each chemical class by site and species, substituting the LOD for values below the limit of detection.

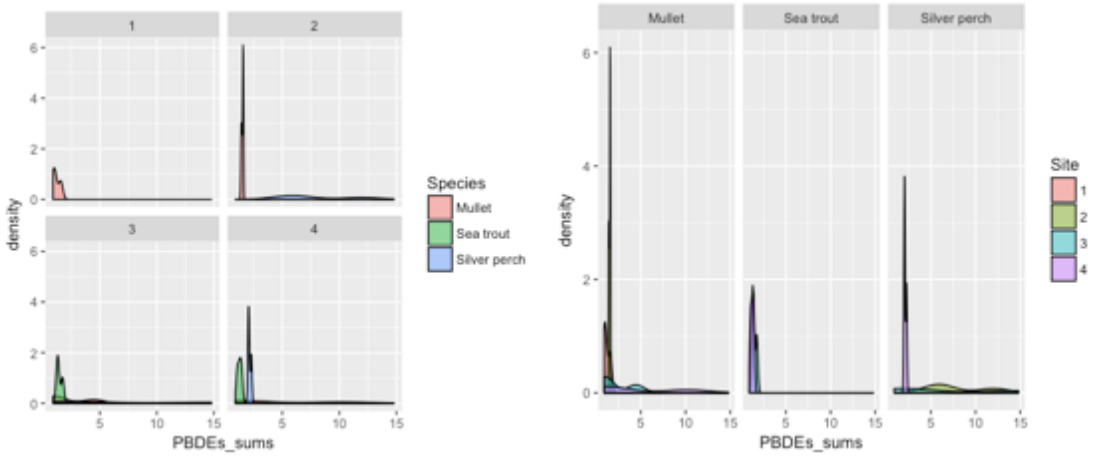
DDTs



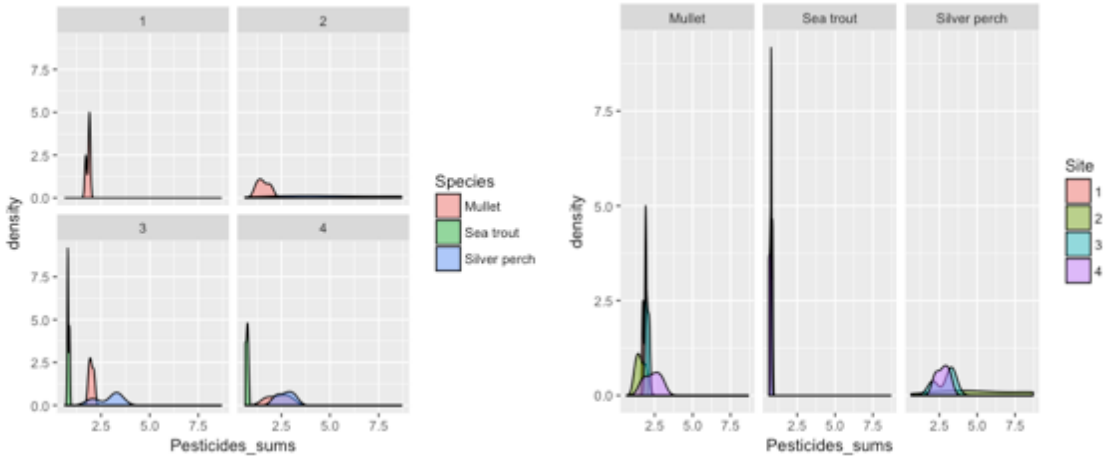
PCBs


1
2
3
4
density
PCBs_sums
Species
Mullet
Sea trout
Silver perch
Site

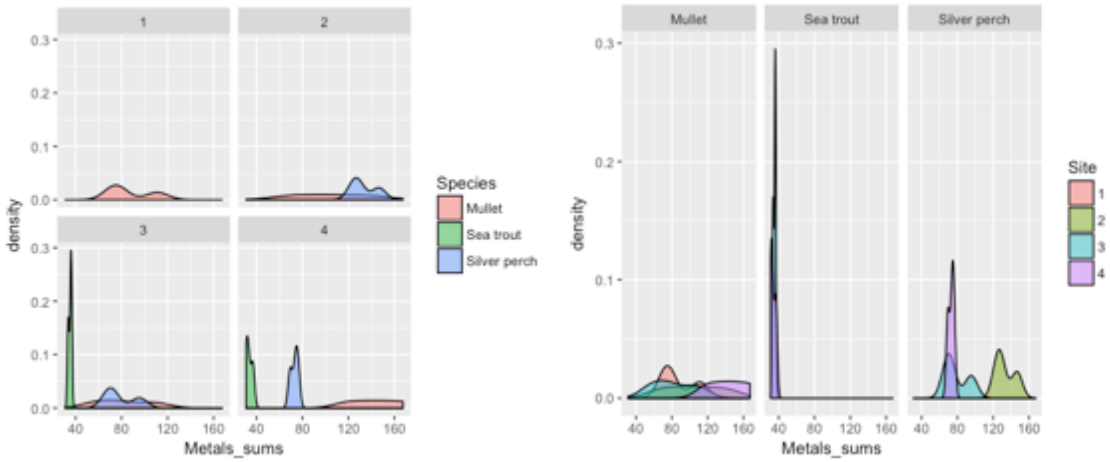
PBDEs


1
2
3
4
density
PBDEs_sums
Species
Mullet
Sea trout
Silver perch
Site

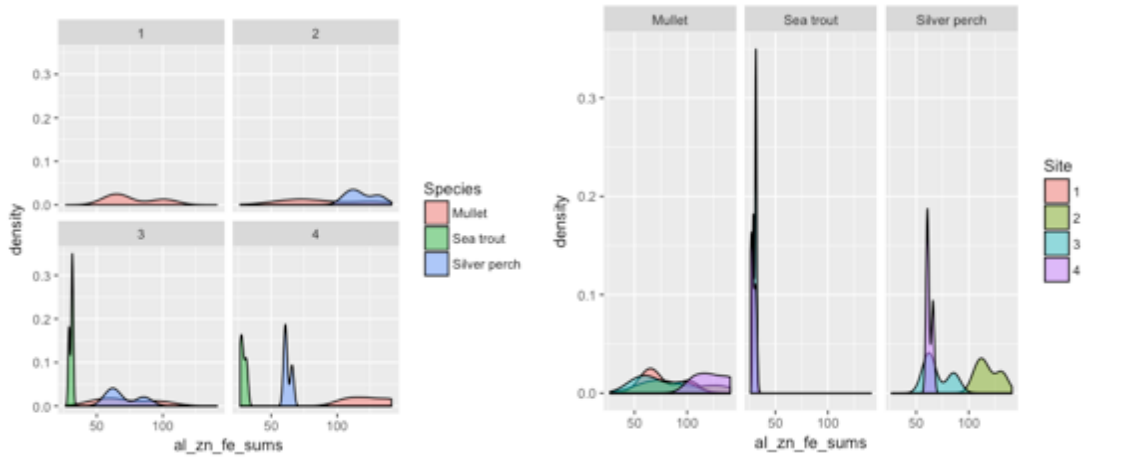
Pesticides



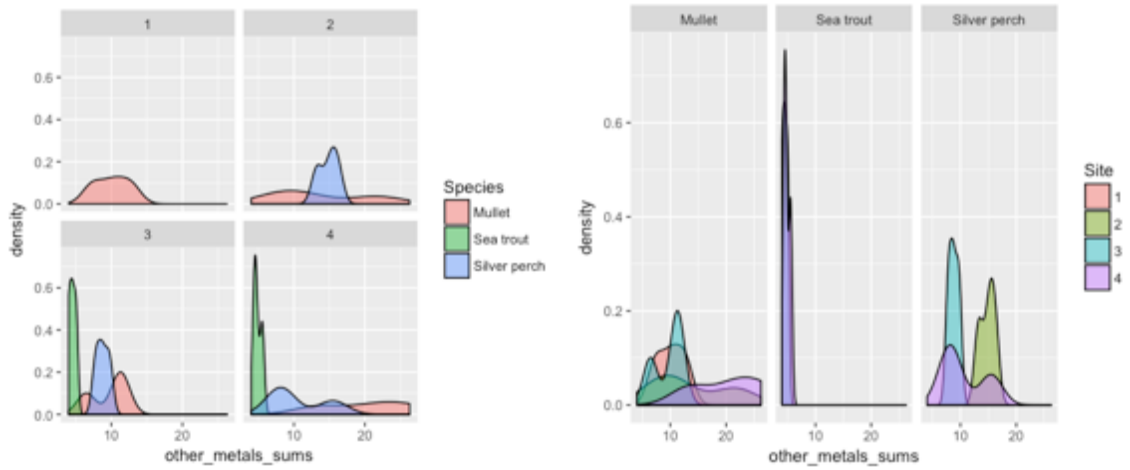
Metals Group 1 (All)



Metals Group 2 (Al, Zn, Fe)



Metals Group 3



Aroclor 1268