Final Second Five-Year Review Report for the Hudson River PCBs Superfund Site

APPENDIX 5

PCB Aroclors Data Treatment

Prepared by:
Louis Berger US, Inc.

April 2019

FINAL SECOND FIVE-YEAR REVIEW REPORT FOR THE HUDSON RIVER PCBs SUPERFUND SITE

TABLE OF CONTENTS

LIST OF TABLES

FINAL SECOND FIVE-YEAR REVIEW REPORT FOR THE HUDSON RIVER PCBs SUPERFUND SITE

LIST OF FIGURES

1 INTRODUCTION TO PCB MEASUREMENT CONSIDERATIONS

Measurement of polychlorinated biphenyl (PCB) concentrations has evolved over time as analytical technology has improved, increasing the understanding of their physical properties and their fate in the environment. A dramatic change in analytical methods has occurred between more recent data, obtained using state-of-the-art, capillary-column, PCB congener analyses, and older analyses based on packed-column quantitation of Aroclor equivalents. A valid interpretation of trends in PCB concentrations cannot be made without consideration of the changes in analytical methods which have occurred over time. That is, a comparison is valid only when there is consistency in what is being measured. Most of the data collected for the Hudson River were analyzed using the Aroclor-based analytical method. However, the United States Environmental Protection Agency (EPA), the New York State Department of Environmental Conservation (NYSDEC), and General Electric Company (GE) have collected enough samples analyzed by both Aroclor-based and congener-based methods to develop a translation scheme to convert the Aroclor-based data so that it is consistent with congener-based quantitation. For the Hudson River PCBs Site, translation schemes (or regression equations) were developed for sediment, water, and fish samples to convert the results from the various Aroclor-based analytical methods to a standard metric. By standardizing the PCB measurements to a congener-based equivalent quantitation, it is possible to compare the various data sets over time on a consistent basis. In this manner, variations among the various data sets through time can be attributed to temporal changes in the environment and not to changes in PCB analytical methods. These regressions also provide a standard statistical basis to estimate the uncertainties associated with these conversions when assessing temporal trends.

The purpose of this appendix is to summarize the translation schemes (using a regression factor approach) to convert Aroclor-based measurements to congener-based quantitation in sediment, water, and fish. Section 1 of this appendix discusses PCB chemistry and the different analytical methods to measure PCB concentrations. Sections 2 through 4 discuss translation schemes in sediment, water, and fish, respectively.

1.1 PCB Chemistry

PCBs are a group of industrially-produced organic chemicals consisting of carbon, hydrogen, and chlorine atoms. Each PCB consists of a biphenyl molecule with a specific number of attached chlorine atoms. The number of chlorine atoms and their location on a PCB molecule determines many of its physical and chemical properties. There are 209 distinct PCB compounds, known as congeners. A PCB congener is any single, uniquely-structured chemical compound in the PCB category. The name of a congener specifies the total number of chlorine substituents, and the position of each chlorine. For example: 4,4'-dichlorobiphenyl is a congener comprising the biphenyl structure with two chlorine substituents - one on each of the #4 carbons of the two rings (Figure A5-1). In 1980, a numbering system was developed which assigned a sequential identification number to each of the 209 PCB congeners.

1.2 PCB Metrics and Analytical Methods

PCBs were manufactured and sold as complex mixtures of several congeners with a variety of trade names, including Aroclor (e.g., Aroclor 1242). An Aroclor PCB mixture might consist of over 100 different individual PCB congeners, although 10 to 20 congeners might make up over 50 percent of the mixture. One of the most common ways to measure PCBs is a method based on the original industrial Aroclor mixtures, EPA gas chromatography (GC) Aroclor-based PCB analysis Method 8082 (M8082). When M8082 is employed to determine whether a sample has an Aroclor PCB mixture in it, the analytical chemist looks for a distinctive gas chromatographic pattern that is indicative of one of the Aroclors. There are nine common PCB Aroclor mixtures (1016, 1221, 1232, 1242, 1248, 1254, 1260, 1262, and 1268). Each of them has a distinctive gas chromatographic pattern.

Measuring PCBs as Aroclors relies on a relatively fixed composition of congeners in the mixture. M8082 uses a pattern recognition technique to qualitatively determine whether or not a given Aroclor mixture is present, after which that portion of the spectrum is quantified using a standard which includes the particular Aroclor. This process references certain well-identified PCB peaks and compares them to the standard to determine Aroclor concentrations in the sample. Provided the sample has not been subjected to conditions that might degrade or change the composition of the PCBs, quantitation of PCB Aroclor using M8082 will give information on the total concentration of PCBs (TPCBs[1]) in the sample, but not the identity or the concentrations of the individual PCB congeners present. However, if an environmental sample has been subject to degradation, weathering or dechlorination, Aroclor-based analysis may over- or under-estimate the actual Aroclor-related PCB concentrations since the apparent Aroclor mixture in the environmental sample may not contain the same suite of congeners or the same proportions of congeners as the standard Aroclor. In such a case, even if PCB congeners originally present in an Aroclor are present, that particular PCB Aroclor may be reported as not detected due to a lack of pattern recognition and/or the mixture may be quantified as a different Aroclor. It can also be difficult to determine a total PCB concentration using the Aroclor approach when environmental degradation or weathering has occurred. This is especially true when more than one Aroclor is determined to be present. As the individual Aroclors represent overlapping mixtures of PCB congeners, there is a possibility that "double counting" of PCBs could also occur. However, since Aroclor-based analyses do not quantitate all PCB-related peaks in the sample chromatogram, it is also possible that Aroclor-based analysis can under-report PCB concentrations. Thus, analytical Aroclor quantitations on environmental samples are not directly comparable to actual concentrations of PCB congeners.

Environmental samples such as air samples, sediment samples, water samples, and biota samples are most likely to have had their congener composition changed by environmental conditions, compared to non-environmental samples. This happens because the PCB congeners with fewer chlorine atoms tend to partition into air and water more readily than those with more chlorine atoms. Biota samples can also be subjected to biodegradation with some congeners being selectively degraded or eliminated and others persisting in the animal tissue. For samples whose congener compositions have been substantively altered,

[1] Total PCBs represents the sum of all measured PCB congeners. PCBs are a group of chemicals consisting of 209 individual compounds known as congeners. The congeners can have from one to ten chlorine atoms per molecule, each with its own set of chemical properties.

analytical testing for PCB homologues will give more reliable results than testing for Aroclors. Homologues are a way of grouping PCB congeners based on the number of chlorine substituents, which vary from one to ten. The PCB congeners that have the same number of chlorine atoms belong to the same homologue group. For example, there are 24 different tri-chloro congeners in the 3-chlorine homologue group. Laboratory results for PCB homologues will report the amount of PCBs present in the sample for each homologue group.

EPA Method 680 (M680) (PCB Homologues) is a gas chromatography/mass spectrometry (GC/MS) procedure that quantifies PCB homologues. This method has the advantage of quantifying PCBs that may not be in the form of Aroclors, as well as original Aroclor mixtures that have been weathered, or may have been misidentified or otherwise not detected by M8082. While M608 provides an advantage over M8082 in this regard, in some circumstances, more detailed quantification of specific congeners is needed. In those cases, PCB congeners can be accurately quantified by EPA Method 1668 (M1668) which determines the concentrations of individual congeners by a sophisticated analytical method using high-resolution gas chromatography/high-resolution mass spectrometry (HRGC/HRMS) combined with isotope dilution techniques. This method requires no presumptions regarding the PCB source material; M1668 identifies the presence and concentration of each of the 209 PCB congeners in a sample.

For the Hudson River PCBs Superfund Site, during the Phase 2 Remedial Investigation/Feasibility Study (RI/FS) investigation in the 1990s, EPA developed a program-specific method which used a dual capillary column gas chromatograph with electron capture detectors (GC/ECD) to analyze PCB congeners. This method was based on the NYSDEC Analytical Services Protocol Method 91-11 (NYSDEC, 1989) for PCB congeners. This technique employs the use of two independent capillary columns with unique resolution capabilities for PCB congener separation, allowing for coelution on the first GC analytical column to be potentially resolved on the second GC analytical column. The dual column GC/ECD allows separation of a larger number of PCB congeners. EPA Phase 2 PCB analyses quantified 126 individual PCB congeners out of 209 possible PCB congeners (General Electric 1997).

GE employed an analytical technique which lies somewhere between the relatively simple M8082 and the highly sophisticated M1668. The GE technique involves the extraction of PCBs from the sample matrix, GC separation of PCB congeners on a DB-1 capillary column, and quantification with an electron capture detector. Calibration of the DB-1 column is based on the method developed by the EPA under the Green Bay Mass Balance Study (EPA, 1987). GE's modification of the original Green Bay Method (mGBM) involves GC standardization using a 25:18:18 mixture of Aroclors 1232, 1248, and 1262. Individual DB-1 peak response factors (RFs) are calculated based on standard peak weight percent values originally developed by EPA (EPA, 1987). These RFs are then used to calculate PCB content of environmental samples. The DB-1 column separates PCBs into 118 unique chromatographic peaks. Several of these peaks contain multiple (coeluting) congeners. DB-1 PCB peaks have been mapped to the corresponding PCB congeners; see Table A5-1 for chromatograph peaks and the corresponding PCB congeners. Peak compositions are based on Aroclor data published in Frame et al. (1996). Another modification is due to the comparison of water column PCB concentrations in samples collected by GE in 1993 with those measured by EPA as part of the Phase 2 study from the Fort Edward and Thompson Island Dam monitoring stations, which suggested that the Green Bay Method results are biased low (General Electric 1997). The GE study in 1997, which examined the dechlorination products, suggested that the analytical biases were manifested in individual PCB congeners, predominantly in Peak 5 (which consists of PCB congeners 2,2'-dichlorobiphenyl (BZ#4)[2] and 2,6-dichlorobiphenyl (BZ#10). The congener distribution (predominantly peak 5 components) within the Green Bay mixed Aroclor standard was apparently miscalculated, as a revision to the calibration was later published (EPA, 1994). This error introduced systematic analytical biases in the GE data because underestimation of the Peak 5 weight percent in the DB-1 calibration standard caused measured Peak 5 values in Hudson River environmental samples to be

[2] BZ refers to the numbering system for PCB congeners developed by Ballschmitter and Zell (1980)

underestimated (i.e., biased low). Since the error is in the calibration standard composition, not the PCB mass, it affects data for all DB-1 peaks (i.e., low bias in peak 5 requires that other peaks are biased high). GE analysis showed that biases are evident in DB-1 peaks 5, 8, and 14. Therefore, GE made a revision for the DB-1 peak response factors used to calculate PCB mass in each detected chromatographic peak of an environmental sample. GE developed calibration error correction factors for water, sediment, and biota samples separately (General Electric 1997). Later in 2010, as directed by EPA, these correction factors were no longer used in the Phase 2 Remedial Action Monitoring Program (RAMP) for water column samples. Instead, the mGBM has been updated to include a second column (CP-SIL5-C18) analysis for the dichlorobiphenyl congeners BZ#4 and BZ#10. The second column analysis was used to achieve a more accurate quantification for PCB congeners BZ#4 and BZ#10 (that co-elute in mGBM peak 5) by achieving full resolution and individual measurement for these two congeners. Correction factors were no longer utilized for mGBM DB-1 Peaks 8 and 14 due to their relatively minor contribution to TPCB (General Electric 2011j). GE's mGBM was used by GE through 2016 to provide congener information on a subset of samples collected for water column and fish tissue monitoring programs.

Among the analytical methods described above, the Aroclor method (M8082) is readily available from many commercial labs and relatively inexpensive. EPA M680 can be a cost-effective option for characterizing contaminated samples for PCB that have undergone environmental degradation; this homologue-based method provides a more representative option for the determination of TPCBs than the Aroclor method. GE's mGBM method is available on a limited basis; only one laboratory can perform the analysis. EPA M1668 is a highly specialized analytical method and can achieve lower quantitation limits, but generally at a higher cost (about five times more expensive than M8082).

1.3 Environmental Considerations

Several studies conducted in the 1980s documented PCB dechlorination in the sediments of the Upper Hudson (for example, Brown et al. 1984). Dechlorination is the chemical process of removing one or more chlorine atoms from a chemical structure, in this case a

PCB molecule. The dechlorination process largely serves to change the nature of the PCBs, affecting both the geochemical and toxicological properties of the mixture. In general, dechlorination converts the PCBs to a more soluble form. The dechlorination transforms PCB congeners from those commonly found in Aroclors to lighter congeners that were virtually absent from the Aroclor composition, as originally manufactured.

Because of the dechlorination process at the Hudson River PCBs Site, PCB contamination within the sediments, water, and biota cannot be accurately quantified using an Aroclor-based analytical method alone. Additionally, it was recognized that individual PCB congeners have different geochemical properties and thus would redistribute themselves among sediments, water, and biota, tending to confound the original Aroclor-related distribution. From these considerations, a congener-specific analytical method is clearly preferred, since variations in the detected congener mixtures could be used to identify PCB sources, as well as important geochemical transformations. As stated previously, most of the data collected for the Hudson River is in the form of Aroclors, therefore, a translation scheme is required to make the Aroclor data consistent with congener-based quantitation. Since both GE and EPA have collected and analyzed a significant number of both Aroclor-based and congener-based samples, translation schemes to convert the Aroclor-based data to the congener-based quantitation were developed. The Aroclor samples were subsequently corrected to a common, congener-based metric so that data could be accurately compared across time.

The following sections discuss the translation schemes (using a regression factor approach) developed to convert Aroclor-based measurements to congener-based quantitation in sediment, water, and fish. Note that selection of PCB metrics for measurement, modeling, and evaluation was based on risk assessment considerations. The tri-homologue and higher (i.e., Tri+) group of compounds are expected to include the PCB congeners that are most toxic to fish, wildlife, and humans, and is therefore a metric that captures most of the toxicity associated with PCBs. An analysis of the historical Aroclor data show that the sum of particular Aroclors is equivalent to the trichloro and higher congeners (Tri+ PCB) (EPA

2000b) and that Tri+ PCBs[3] is essentially the same as the sum of all PCBs (TPCBs) that are bioaccumulated in fish tissue, the main pathway for human exposure. As a result, estimates of both TPCBs and Tri+ PCBs obtained from both congener (directly analyzed) and Aroclor-based analyses (converted as described in this appendix) were used as the basis to assess PCB contamination in sediment, water, and fish for the Hudson River.

[3] Tri + PCBs represents the sum of all measured PCB congeners with three or more chlorine atoms per molecule. PCBs are a group of chemicals consisting of 209 individual compounds known as congeners. The congeners can have from one to ten chlorine atoms per molecule, each with its own set of chemical properties.

2 SEDIMENT

The Hudson River is one of the most extensively monitored PCB contamination sites. The system has been studied extensively and monitored over a period of more than 30 years. The various monitoring studies provided numerous sediment data sets. This section discusses translation schemes to convert Aroclor-based measurements to congener-based quantitation for various sediment data sets.

2.1 Historical Sediment Data Sets

Two historical large-scale sediment investigations were conducted by NYSDEC; one in 1976 to 1978 (reported in Tofflemire and Quinn, 1979), and one in 1984 (reported by Brown et al., 1988). The 1976 to 1978 sampling covered the area from Fort Edward to Troy (RM 194.8 to RM 154); whereas the 1984 sampling was restricted to the Thompson Island Pool (TIP) (RM 194.6 to RM 188.5). The 1984 sediment survey (Brown et al., 1988) represents the most comprehensive database on PCB concentrations in Thompson Island Pool sediments prior to the 2002-2005 GE remedial design investigation, with over 1,200 samples collected on a triangular grid with 125-foot centers. The spatial coverage of that effort was adequate for the purpose of estimating PCB sediment mass inventory, as discussed in Brown et al., 1988 and EPA 1997b. The 1976-1978 and 1984 data were needed to identify likely areas for remediation and provide estimates of PCB mass and sediment volumes for the 2002 Record of Decision (ROD) for the Upper Hudson River.

Prior to using the 1976-1978 NYSDEC data, the comparability of these sediment data was evaluated as part of the Phase 2 RI/FS investigation. A good level of agreement between the 1976-1978 and 1994 Phase 2 conditions serves to support internal consistency of the various sediment classification data sets. TPCBs were reported by O'Brien and Gere for the 1976-1978 sediment data set. These were based on Aroclor analysis using a limited number of packed column peaks, which tended to miss the mono- and di-homologues. Based on reconstruction of the 1976-1978 total PCB results from USEPA Phase 2 sediment congener data, a regression between the Tri+ concentration and the 1977-1978 total PCB concentrations produced a zero-intercept model with which to estimate Tri+ concentrations from these data (Equation 2-1). Details of this analysis are presented in EPA 1998 and

Butcher (2000). The following relationship was used to calculate Tri+ PCB for the 1976-1978 NYSDEC data (EPA, 2000b):

$$Tri + (1977) = 1.131 \times [Aroclor\ 1016 + 1254] \qquad \text{Eqn. 2-1}$$

PCB concentrations reported by NYSDEC for the 1984 Thompson Island Pool sediment survey were dependent on the Aroclor quantitation methods used and were not equivalent to results which would be obtained using capillary column GC analysis for PCB congeners. It was thus crucial to understand what is reported in these data and estimate how well the NYSDEC-reported total represented actual total PCBs that would have been calculated by measuring and summing congener concentrations. As part of the Phase 2 RI/FS investigation, a study was made of the differences between the two techniques. This is documented in Appendix E of the Low-Resolution Sediment Coring Report (EPA 1998), which describes the quantitation issues relating the 1994 Phase 2 and 1984 NYSDEC PCB data. The recommendation of this analysis was to use the 1984 quantitation of total PCB as representative of the sum of congeners in the trichloro through decachloro homologue groups. A linear relationship was developed to adjust (or correct) the 1984 NYSDEC data to a basis consistent with the sum of tri- and higher-chlorinated congeners (Tri+) in the 1994 EPA data. The following equation was used to calculate Tri+ PCB for the 1984 NYSDEC data (EPA 1998):

$$Tri+ (1984) = 0.944 \times 1984\ Aroclor\ Sum \qquad \text{Eqn. 2-2}$$

In 1991 and 1998, GE also conducted sediment surveys and the data were reported as PCB congeners based on the mGBM. The GE 1991 survey sampled the upper river from Fort Edward to the Federal Dam at Troy, while the 1998 survey sampled the TIP only. The sediment survey conducted in the TIP by GE in 1998 attempted to 'repeat' portions of the 1991 O'Brien and Gere and 1994 EPA sediment surveys (GE 1999). GE 1991 and 1998 data were used in the modeling effort during Phase 2 RI/FS investigation.

During Phase 2 RI/FS investigation, EPA collected low-resolution cores in 1994. These data were analyzed based on congener-specific standards as described in Section 1.2. Rather than resurvey the entire TI Pool, the 1994 low-resolution coring effort focused on replicating a representative subset of the 1984 locations. Since the EPA 1994, GE 1991, and 1998 data were reported as PCB congeners, no translation scheme was needed to calculate the Tri+ PCB concentrations.

2.2 Remedial Design Investigation, Remediation and OM&M Sediment Data Sets

During the remedial design investigation, GE collected sediment samples under the Sediment Sampling and Analysis Program (SSAP) in 2002-2005. These data were collected for delineating the final areas for removal. Tri+ PCB concentrations were not measured directly on the SSAP samples as a cost-saving measure for GE, since the M8082 (Aroclor-based) analysis is substantially less expensive than the M680 (homologue-based) analysis. Given that the SSAP comprised some 30,000 sediment samples, EPA allowed GE to analyze all samples by M8082 and to also analyze a subset of the samples by the EPA-approved M680, and then use the comparison of the paired analytical results to develop a relationship to estimate the Tri+ PCB fraction in the remaining samples analyzed only via M8082.

Given the extensive use of M8082 during the SSAP, as well as the short analytical turn-around time for M8082, GE continued to use M8082 throughout the remediation itself to satisfy the residual performance standard sampling requirements. M8082 was also used to conduct the Downstream PCB Deposition Study (DDS) during 2011 through 2013. To maintain consistency from the SSAP to the present, GE continued to use M8082 to establish the baseline sediment conditions at the beginning of the OM&M period in 2016.

In all applications of M8082, the determination of the TPCB concentration in a sample was simply the summation of all detected Aroclors in a sample. This sum was shown to be well correlated with the sum of PCB congeners determined by the mGBM as part of the SSAP (General Electric 2004e). Also as part of the SSAP, GE standardized the reporting of Aroclors under M8082 to four specific Aroclors (Aroclors 1221, 1242, 1254 and 1260) and maintained consistent reporting throughout the SSAP, remediation, DDS and OM&M sampling efforts. The basis for estimating Tri+ PCB concentrations from M8082 throughout the period 2002 to 2016 was developed from a regression relating M8082 and M680 (based on matched sample pairs obtained during the SSAP) and is described below. Data quality was also tracked throughout this period and is also described below.

The quality of the regression for Tri+ determination is contingent upon the underlying data – i.e., the M680 (homologue) and M8082 (Aroclor) data. To assess the accuracy of the M680 and M8082 methods, laboratory control samples (LCS) and performance evaluation (PE) samples were analyzed by laboratories used by GE during the remedial design investigation. There are two measurements of "known" TPCB (sum of Aroclors) concentrations in the suite of Quality Assurance/Quality Control (QA/QC) measurements performed by GE. One is the LCS sample analysis (Aroclor 1242 spiked into a clean matrix), and the other is the analysis of PE-5 (comprised of known concentrations of Aroclor 1221 and Aroclor 1242). Since the LCS and PE-5 results are of known composition, they can be used to evaluate the Aroclor distribution and to assess the accuracy of the analytical methods. PE-5 and LCS samples were analyzed using both M680 and M8082. In addition to the LCS and PE-5 samples, GE also prepared and analyzed additional PE samples. Sediment samples from the Upper Hudson River at 4 different locations (PE-1, PE-2, PE-3, and PE-4) were used to provide a range of PCB concentrations. Sections 2.2.1 and 2.2.2 summarize the results of QA/QC data analyzed during the remedial design investigation period.

2.2.1 M680 Homologue Data

The critical component of the M680 data is how well the reported homologue distribution (or more specifically, the Tri+ fraction) accurately reflects that of the sample analyzed and how they are compared to a known concentration sample. The M680 results are discussed below.

<u>PE Sample Results</u>

PE samples were analyzed concurrently with the Year 1 and Year 2 SSAP. As noted above, PE-5 is a manufactured standard, so there is information on its true or known composition.

Fifteen replicates of PE-5 were analyzed with the Year 1 and Year 2 SSAP. The standard deviation (SD) and relative standard deviation (RSD) were calculated for TPCBs, Tri+ PCBs, fraction Tri+, and for each homologue group (see Table A5-2). The SD and RSD are both measures of precision; the more precise data will result in smaller SDs and RSDs. As shown in Table A5-2 (M680-all samples), the RSDs are typically about 0.16 (16 percent) for most parameters. However, the RSD for monochlorobiphenyl (MonoCB) is much higher (0.345); by inspection, it can be seen that there are two anomalous low values for MonoCB (about 3 milligrams per kilogram (mg/kg); with the results for the other 13 samples ranging from 7 to 16 mg/kg). When these two samples are excluded (M680 - Two Samples Excluded), the MonoCB RSD becomes similar to that of the other parameters; and the Tri+ RSD reduces to 0.090.

<u>LCS Sample Results</u>

A total of 79 Lab Control Samples (LCSs) were analyzed for PCB homologues; all homologue analyses were performed by Lab 15 (Northeast Analytical [NEA]). Of the 79 LCS samples, 28 are Year 1 and 51 are Year 2 LCS analyses. There was some difference by year, with the median Year 1 sum of homologue recovery being about 83 percent, and the median Year 2 recovery about 73 percent. The LCS data were reviewed to see if the homologue distribution might account for the low recovery of total homologues, relative to TPCBs. In virtually all the Year 1 LCS samples analyzed by GE, the homologues reported as present were di, tri, and tetra PCBs; and the median Tri+ fraction was 0.84. This fraction is in very good agreement with literature values, which suggest that about 85 percent of Aroclor 1242 is Tri+. (Aroclor 1242 has one percent or less monoCB, so the fact that mono was not reported as present by GE/NEA is not significant.) The homologue distribution of the Year 2 LCS samples was spot-checked; the Tri+ fraction ranged from 80 to 87 percent, with a median of 84 percent, in the nine samples checked. The overall median recovery (Year 1 and Year 2 combined) for LCS homologues is 75.8 percent.

<u>Field Duplicate Results</u>

The field duplicate results are presented here to show the precision of the method. A total of 79 field duplicate pairs were analyzed for homologues. As only samples considered likely to have detectable concentrations of PCBs were selected for homologue analysis (and in fact, PCBs were detected in all the samples), all of the field duplicate pairs were included in the review. Of the 79 field duplicate pairs, 65 (about 82 percent) met GE's precision criterion of ≤ 40 percent RPD (relative percent difference) for TPCBs (sum of detected homologues). The RPD for TPCBs was less than 100 percent in all but four of the samples (i.e., in 95 percent); and the median RPD for TPCBs is about 17 percent. About 78 percent of the Tri+ PCB (concentration data) met the ≤ 40 percent RPD criterion, with a median value of about 17 percent (the same as for TPCBs). Precision data for individual homologues groups were also calculated for mono- through tetra-chlorobiphenyls. The median precision for each of the homologues ranged from about 16 to 20 percent RPD.

In addition to TPCBs and Tri+ PCB concentrations, the fraction Tri+ was calculated for each of the samples and the precision of the Tri+ fraction data was assessed. The data show that the Tri+ fraction is quite reproducible. The highest RPD for Tri+ fraction was 41 percent (this was the only one of the 79 samples with an RPD over 40 percent); and the RPDs were less than 20 percent for 95 percent of the samples, with a median RPD for Tri+ fraction of less than 6 percent.

Summary

In summary, in every instance where the Tri+ mass fraction can be calculated from sample or QC data, the results show a higher degree of precision in the estimate of this fraction than in the absolute estimate of its concentration as measured by M680. This result supports the choice of the Tri+ conversion model described previously. As described previously, LCS and PE-5 sample results indicate a systematically low recovery (low reported value) relative to the known concentrations for these samples. Despite the low bias, the results for these samples as well as other QA/QC samples and the field duplicates appear to be precise with mean RSDs or RPDs on the order of 0.16 to 0.20 (16 to 20 percent) for individual parameters. Better precision for the Tri+ mass fraction as well as the major homologue fractions DiCB and TriCB was evidenced by the lower RSDs and RPDs, on the order of 0.05 (5 percent).

2.2.2 M8082 Aroclor Data

For the Aroclor data, the critical components are the accuracy and precision of TPCB concentration, and the correct assignment (identification) of Aroclors. The other critical component of the M8082 data is the assignment of PCBs to individual Aroclors.

<u>QA/QC Sample Results</u>

The LCS samples analyzed for this program consisted of a low concentration (typically about 1.2 mg/kg) Aroclors spike into a clean matrix. Overall, GE laboratories recovered close to 100 percent of the spiked concentration based on the median recovery value (see Table A5-3). There was some variability by laboratory, with median LCS recovery values ranging from 89 percent for Lab 1 to 117 percent for Lab 16 (Lab 16 also had the fewest LCS data points - only 35); the median Aroclor LCS recovery for Lab 15 (the lab which did all the homologue analyses) was 101 percent. The Aroclor LCS data were reviewed by year. The aggregate median recovery for the Aroclor LCS recoveries is 95 percent for Year 1 (n= 471) and 100 percent for Year 2 (n = 1528). The overall median recovery of LCS Aroclors (Year 1 and Year 2 combined) is 99.2 percent (n = 1999).

Initial analyses of PE-5 were reported in the Inter-laboratory Comparison Study (ILCS). The "grand mean" (mean of final pool) reported in Table A5-4 is actually about 12 percent higher than the known value of PE-5. However, calculations performed by EPA, using only the reported PE data from the six laboratories which were ultimately accepted into GE's analytical program, were closer to the known value (about 4 percent high). The mean value calculated from the Year 1 and Year 2 PE-5 data (see Table A5-4) is within 1 percent of the known value. (Lab 4 data are excluded; no detected values of Aroclors have been used by GE from Lab 4. This issue is addressed in greater detail in the Year 1 Data Summary Report (DSR) [Section 6.1.1] [General Electric 2003a].)

As with the total Aroclor quantitation, only samples of known composition can be assessed quantitatively. As the LCS was comprised of a single commercial Aroclor, the LCS data provide no information regarding the accuracy of Aroclor identification. Therefore, only the data from PE-5 provide any useful means of quantitatively assessing Aroclor identification in samples with more than one Aroclor.

Table A5-4 shows the Year 1 and Year 2 PE-5 Aroclor data. The observed mean ratio of 1221 to 1242 (just over 2.5:1) is a bit lower than the known ratio (just under 3.0:1). On an Aroclor basis, the mean concentration of Aroclor 1221 is about 5 percent lower than the known value, whereas the Aroclor 1242 mean is slightly more than 10 percent higher than the known value. These relatively low errors are environmentally protective; i.e., the slight high bias toward identification of Aroclor 1242 will result in a slight high bias in Tri+ calculations (assuming that the environmental samples exhibit the same phenomena as PE-5, an assumption which is not necessarily true).

In addition to comparing the overall Aroclor concentration data to known values (to assess possible systematic bias), the precision of the data were also reviewed to judge how much error is likely to be present in any individual result. For this assessment, PE-2 data were used along with PE-5. (It was expected that there would be more variability in PE-2 results due to the fact that the Aroclor composition may not match up as well with peak patterns of Aroclor standards).

For PE-5, the RSD for TPCBs is about 0.17 (as shown in Table A5-4). Somewhat surprisingly, the precision on PE-2 was slightly better, with an RSD of about 0.15 (Table A5-5). The precision of the Aroclor identification (as "fraction Aroclor 1221") was also assessed. For PE-5, the RSD is 0.042, and for PE-2, the RSD is 0.059. An example of the significance of these RSDs can be presented, assuming that the data are normally distributed and that 95 percent of the values with fall within two standard deviations of the mean. For PE-2 (for example), a sample reported to have the mean Aroclor 1221 fraction of 0.65 has a 95 percent chance that the 'true' fraction is 0.65 ± 2*(0.059*0.65) or 0.66 ± 0.077 (i.e., the true value probably is between about 0.57 and 0.73).

Field Duplicate Results

The data discussed above represent results of Quality Assurance (QA) sample analyses. While the true values were not always known to the laboratory, it was generally evident that the laboratory knew they were QA samples and may have (consciously or unconsciously) taken extra care in the analysis and reporting of those samples. In addition, even the PE samples that were prepared from Hudson River sediment (PE-1 through PE-4) were well-homogenized and adjusted to known moisture content, and may not have posed some of the analytical challenges that the environmental samples presented. It is difficult, if not impossible, to assess the degree to which this may be the case. The only other available measure is the performance of the laboratories on blind field duplicates; while these provide no information on the accuracy of the reported results, they do provide an indication of the precision of the results.

Field duplicate data were reported in summary form in the GE DSRs (Year 1 DSR, Phase 1 DSR, and Phase 2 DSR). Overall, about 80 percent of the detected values (3,024 out of 3,789 data points, including TPCBs and data for each individual Aroclor [1221, 1242, 1254, and 1260], met the established criteria (relative percent difference of 40 percent or less; or absolute difference less than two times the reporting limit for concentrations less than five times the reporting limit). A 40 percent RPD is analogous to ±20 percent of the average of the values.

A more detailed review of the PCB (Aroclor) field duplicate data was conducted. First, the data were sorted by concentration, with the greatest focus on the samples with TPCB (sum of Aroclors) concentrations greater than or equal to 5 mg/kg in the original sample. The precision of the TPCB analysis, precision of data for Aroclor 1221 and Aroclor 1242 individually, and the reproducibility of the fraction of the total represented by Aroclor 1221 were all reviewed.

About 723 duplicate sample pairs with concentrations greater than or equal to 5 mg/kg PCBs (sum of Aroclors) were located in the database. Not all pairs could be used for all the evaluations due to anomalies in the data. (For example, some statistics cannot be calculated for values of zero; and there were eight samples [of the 723] in which Aroclor 1221 was not detected in one or both of the analyses. There were also seven additional suspect data pairs, for which the agreement was 'too good' - the exact same result was entered for each detected Aroclor and TPCBs.)

For TPCBs, the RPD was greater than 100 percent in about 3 percent of the samples; and the RPD is less than 88 percent in 95 percent of the samples. With respect to the stated GE QA/QC criterion of 40 percent RPD, the TPCB RPD for these sample pairs was less than this value in 81 percent of the pairs. The median TPCB RPD was 16 percent. Precision was similar for Aroclor 1221 and Aroclor 1242 individually.

In addition to the data for TPCB concentrations, the identification of Aroclors (e.g., Aroclor 1221 vs. Aroclor 1242) is an important factor in the regression analysis. Even if replicate analyses are in agreement on the TPCB concentration, the Tri+ fraction will not be reproducible if the Aroclor composition of the analysis is not reproducible. To evaluate this aspect, the Aroclor 1221 fraction was calculated for each sample, and the precision of duplicate pairs was assessed for this parameter. Precision for this parameter was good (even with the few poor precision samples due to Aroclor 1221 not being detected in one of the analyses); the RPD was less than 50 percent in 99 percent of the samples, and was 20.1 percent or less in 95 percent of the samples. The median RPD for the fraction Aroclor 1221 data is only 3.5 percent.

Similar assessments were also made on the set of 259 field duplicate pairs in which the TPCB (sum of Aroclors) concentration ranged from 1 to less than 5 mg/kg (see Table A5-6). Precision was somewhat less for this dataset as compared to the duplicate pairs with 5 mg/kg or more TPCBs. About 70 percent of the duplicate pairs in this PCB concentration range met the ≤ 40 percent RPD criterion for TPCBs (as opposed to slightly over 80 percent in the higher concentration. The median RPD for fraction 1221 in this lower-concentration group was 5.3 percent, with 95 percent of the samples having a fraction 1221 RPD of less than 31 percent.

Summary

The M8082 results agreed well with the expected values for the two "known" measurements, the LCS and PE-5. The available data suggest that the quantitation of TPCBs by M8082 is likely to be close to the true value, taken as a whole, based on available metrics. In terms of Aroclor identification, the M8082 results indicate that Aroclors have, in general, been consistently identified, with an RSD for the fraction of Aroclor 1221 in QC samples on the order of 0.05. The RPD for Aroclor identification in field duplicates is similar, with a median RPD of 5 percent or less, depending on the TPCB concentration (lower RPDs were found at higher concentrations). For both QA/QC samples as well as field duplicates, the results show a higher degree of precision in the estimation of Aroclor mass fractions relative to absolute quantitation. Mean RSDs for the sum of Aroclors in QA/QC samples (0.15 to 0.17) were greater than those estimated for the Aroclor mass fractions (0.05 to 0.06). Nonetheless, both are acceptable for the planned use of the data. The median RPD for field duplicates for TPCBs (sum of Aroclors) in field duplicates greater than or equal to 5 parts per million (ppm) (16 percent) is similarly larger than the median RPD for the Aroclor mass fraction RPD of 3.5 percent. These RPDs are roughly one third higher for concentrations below 5 ppm. With respect to the M680 results, both the field duplicate precision, as well as the mass fraction precision estimates for M680, are very similar to those of M8082, further supporting the contention that the M680 bias is primarily an absolute mass underestimate and not a bias resulting from PCB identification.

Based on the LCS and PE samples results discussed above, GE’s M680 TPCB results showed a systematic low bias compared to those of M8082. As a result, in the March 25, 2004 comments on GE’s Draft Phase 1 Dredge Area Delineation (DAD) Report, EPA required GE to apply a regression-based correction factor to account for the systematic low bias observed in the M680 homologue PCB data relative to M8082 and mGBM results. A small subset of SSAP data was analyzed using mGBM. EPA also required additional analyses of the Tri+ PCBs regression, including an assessment of the data quality and the uncertainty in the predicted Tri+ PCB concentrations. EPA performed additional statistical analyses, which further indicated that the M680 data were biased low relative to M8082 and mGBM. In response to this concern, GE prepared a two-part correction procedure, dated June 14, 2004 (General Electric 2004e), which was submitted to (and approved by) the EPA.

2.2.3 Data Ranges of Interest

As stated in the 2002 ROD, the remedy is based on the removal of PCBs with 3 or more chlorine atoms (Tri+ PCBs). The removal of sediments in the Hudson River was based primarily on a mass per unit area (MPA) estimate (i.e., grams of PCBs per square meter). PCB inventory in sediment is represented by an MPA value calculated for each sediment core . The ROD makes clear that, for River Sections (RSs) 1 and 2, the MPA of Tri+ PCBs is to be the primary basis for identifying specific areas for dredging. The removal criteria as stated in the ROD are as follows:

- MPA of 3 g/m^2 Tri+ PCBs or greater from RS 1;
- MPA of 10 g/m^2 Tri+ PCBs or greater from RS 2;
- Removal of selected sediments with high concentrations of PCBs and high erosional potential (NYSDEC Hot Spots 36, 37, and the southern portion of 39) from RS 3

Since the Tri+ PCBs concentration was estimated via a regression-based correction factor from Aroclors measurements (see Section 2.2.4 below for the regression-based correction factor), there could be biases in the estimated Tri+ PCBs concentration ranges. To address this, EPA examined three concentration ranges to assess uncertainty and potential bias and their impacts on area selection for the remedial design: 0 to 5 ppm, 5 to 80 ppm and 80 ppm or higher Tri+ PCBs. The 0 to 5 ppm Tri+ PCB (homologue) sample pairs were identified because this concentration range was unlikely to result in exceedance of removal criteria and, in the Year 1 Data Summary Report, GE described the accuracy of the homologue analyses as decaying for the lower concentration range. The 5-80 ppm range was selected because this range was critical to determining if a sample location would exceed either the surface concentration or the MPA threshold levels. Table A5-7 shows the range of critical concentrations.

2.2.4 M680 Tri+ PCB correction factor

This section presents the development of an equation to estimate Tri+ PCB concentrations from the M680 and M8082 paired analytical results. Because of the low bias in the M680 results discussed in Section 2.2.1, a bias correction was first developed for the M680 data before the Tri+ PCB equation was developed. Several approaches were explored for correcting for the low bias in the M680 PCB homologue data. After the bias in M680 data was corrected, a multiple regression approach was used to develop the relationship between Tri+ PCB and the detected PCB Aroclors. The final model to estimate Tri+ PCB concentrations from the detected Aroclors by M8082 was:

$$Tri+\ PCB = 0.03*A1221 + 1.16*(A1242+A1254) \qquad \text{Eqn. 2-3}$$

where A1221, A1242, and A1254 are Aroclor 1221, 1242 and 1254 concentrations reported by M8082, respectively. Uncertainty associated with the selected model and the data used to develop the model was also estimated. An uncertainty estimate based on the field duplicate precision (median variability of $\pm$16 to 20 percent, 95th percentile at $\pm$90 percent) was used in the weight of evidence approach for the dredge area delineation.

As noted above, the paired SSAP data show a low bias for TPCBs by M680 relative to TPCBs by M8082 in all concentration ranges. Thus, the correction factor is applied to all concentration ranges in the paired data. The development of the correction factors was based on point-by-point comparison of M680 and M8082 TPCB concentrations. The correction factor is calculated for each data point as follows:

$$Corr_i = \frac{Total_{8082_i}}{Total_{680_i}} \qquad \text{Eqn. 2-4}$$

where

$Corr_i$ = Correction factor for sample *i*

Total8082 i = TPCBs by M8082 for sample i

Total680 i = TPCBs by M680 for sample i

The correction factor is then applied to the Tri+ PCB M680 concentration to yield the corrected Tri+ PCB for each sample, using the following equation:

$$Tri_{corrected_i} = Corr_i \times Tri_{680_i} \qquad \text{Eqn. 2-5}$$

where

$Tri_{corrected_i}$ = the corrected Tri+ PCB value for sample *i*

$Corr_i$ = correction factor for sample *i*

Tri_{680_i} = the original Tri+ PCBs value by M680 for sample *i*

This approach presumes that the M680 results are accurate in their estimate of the fraction of Tri+ PCB present in the sample but biased low in the overall estimate of TPCB. Thus the approach increases the Tri+ PCB concentration in direct proportion to the TPCB correction.

2.2.4.1 Regression Methodology

After completing the first step of bias correction for the Tri+ PCB concentration, the second step involved the development of the regression model correlating the corrected Tri+ PCB concentrations with the M8082 Aroclor results. The regression formula represents the GE 2 draft model, which includes Aroclor 1221 and the sum of Aroclors 1242 and 1254:

$$Tri_{corrected_i} = a1 \times (A1221) + a2 \times (A1242 + A1254) \qquad \text{Eqn. 2-6}$$

where the a1 and a2 are regression coefficients and the other terms are defined as above.

The following sections describe the details regarding calculation of the regression coefficients and metrics of accuracy.

2.2.4.2 Algorithm

A robust regression approach based on the bisquare influence function (Holland and Welsch, 1977) was used to develop the regression equations. The robust regression method was used because it automatically identifies outliers (both regression and leverage points) and down-weights them in the derivation of the regression coefficients by minimizing the weighted error sum of squares. The computational approach used for the weighted regression in this analysis is based on an iteratively reweighted least squares regression.

The damped leverage approach used by GE is similar to the weighted robust regression method. The difference is that the weights are not calculated iteratively. Ideally, the damped leverage approach assumes that the weights are known in advance but this very seldom happens in practice. Using the available data, these weights can be computed in more than one way using classical methods. GE used the damped leverage approach to calculate the weights. However, it should be noted that when outliers are present, the weights (and all other statistics including regression coefficients) obtained using a damped leverage procedure can get distorted; therefore, it is preferable to use robust methods which automatically identify outliers and down-weight them accordingly in the regression process.

2.2.5 Data Selected

Samples with paired data in the February 2004 database were selected for this analysis. There are 1,346 samples with paired M8082 and M680 analyses.[4] A total of 23 sample pairs where the M8082 TPCB or M680 Tri+ PCB concentrations were nondetect or inconsistent were eliminated, and the remaining 1,323 samples were included in the regression analysis. Statistical outliers were not eliminated because the procedure is able to handle them directly. The 23 samples eliminated from the analysis included:

- 18 samples with nondetect Tri+ PCB by M680,
- 2 samples with nondetect Aroclor 1221 and Aroclor 1242 by M8082; and
- 3 samples where Tri+ PCB by M680 was greater than TPCB by M8082.

---

[4] This large data set represented the available data for analysis at the time of the negotiations between EPA and GE. Subsequent data collection efforts were reviewed and found to be consistent with the regression developed based on the February 2004 database.

2.2.5.1 Training and Testing Data Sets

Two-thirds of the data were selected for the training set and used to develop the regression equations. The full set of paired data was sorted by laboratory, river section, and concentration range prior to selecting a representative data subset (note that GE employed five different laboratories during the SSAP). The remaining data were used to test the regression equations. This approach allows the predictive ability of the regression equations to be tested on an independent dataset.

2.2.5.2 All-Laboratories and Laboratory-Specific Equations

The influence of the different individual laboratories on the regression model was evaluated. Development of the draft Tri+ PCB regression equations indicated that coefficients for individual laboratories were generally similar to the coefficients for the single, all-laboratories equation. Therefore, the regression model was constructed as a single model for the entire set of data ("all-laboratories"), because the uncertainty in the correction of M680 Tri+ PCB results is expected to be larger than the differences among the laboratories.

2.2.6 Model Results

The final model to determine Tri+ PCB concentrations based on M8082-detected Aroclors was:

$$Tri+\ PCB = 0.03*A1221 + 1.16*(A1242+A1254) \qquad \text{Eqn. 2-7}$$

This formula integrates the bias correction in the M680 data and the conversion of Aroclor results to Tri+ PCBs as described above. This model became the basis for estimating Tri+ PCB concentrations for dredge area delineation and other remedial activities. However, as described in the Consent Decree, it was anticipated that GE would update this regression based on paired mGBM and M8082 results when additional Tri+ PCB and Aroclor paired results became available (EPA, 2004e).

A scatter plot of the bias-corrected measured versus the predicted Tri+ PCBs concentrations can be found in Figure A5-2 and a scatter plot of the bias-corrected measured versus the predicted Tri+ PCBs fraction can be found in Figure A5-3. The slope for the corrected measured vs. predicted Tri+ PCB fraction is close to one (with a value of 1.02) as shown in Figure A5-2, but there is a scatter around the 1 to 1 line with R^2 of 0.8.

Plots of the Tri+ PCB concentrations show good agreement between the predicted and corrected measured values (Figure A5-2), but the 0 to 5 ppm Tri+ PCB range is more often overestimated. The overestimation in the concentration is environmentally protective because the overestimated Tri+ PCB concentration will result in a location being identified for removal as opposed to leaving it in place. The scatter for the 5 to 80 ppm and 80 ppm or higher Tri+ PCB concentrations is more balanced. These same conclusions can be drawn from review of the residual plots in Figure A5-4.

The uncertainty in the models was estimated by the 95 percent individual confidence curves on the log of the measured and predicted Tri+ PCB concentrations. The measured and predicted results are shown in Figure A5-5 with the confidence curves by concentration range. The graphs show that, generally, the uncertainty in the data increases with increasing concentration and can be approximated by a percentage of the average concentration. To estimate this percentage, the confidence curves were approximated by linear fits by concentration range. These results are summarized in Table A5-8. For the 0 to 5 ppm Tri+ PCB concentrations, curvature in the confidence curves is evident. Ordinary least squares linear regressions on these upper and lower confidence curves, provide a poor fit to the data near 0 ppm. The trend of the data is through 0 ppm, but the linear fits are above this concentration. No precise estimate of uncertainty was made for the 0 to 5 ppm Tri+ ppm range on this basis, but a rough approximation from examination of the data would indicate an uncertainty of approximately +70 percent and -20 percent. Estimates of uncertainty can be made on the basis of the confidence curves for the 5 to 80 ppm Tri+ PCB range (-29 percent and +38 percent) and the 80 ppm or higher Tri+ PCB range (-44 percent and +41 percent). As expected from data exhibiting heteroscedasticity (i.e., data with unequal variability [scatter] across a set of second, predictor variables), the range of uncertainty is greater at higher concentrations.

2.2.7 Update to Tri+ PCB Regression during Remedial Action Monitoring Phase

As stated in the Dispute Resolution, the regression equation developed to estimate Tri+ PCB using M8082 Aroclor data (see Section 2.2.6 above) was allowed to be updated using paired mGBM and M8082 results if GE performed a Tri+ PCB Study (EPA, 2004e). In 2005, GE conducted a laboratory methods comparison study to compare the relative accuracy and precision of M680, M8082, and mGBM for the sediment data. That study found that the mGBM has better accuracy and precision relative to M680 and M8082 in measuring TPCB and Tri+ PCB (General Electric 2005d). In addition, the 2005 GE study found a strong correlation between Tri+ PCB concentrations measured by the mGBM and the Aroclor concentrations measured by M8082. Using paired analyses on 150 sample extracts and 30 archived sediment samples, the following regression equation was developed (General Electric 2009d):

$$Tri{+}\ PCB = 0.13*A1221 + 0.94*(A1242+A1254) \qquad \text{Eqn. 2-8}$$

The EPA reviewed this regression and requested that GE analyze 100 additional paired samples to provide more data at low concentrations to support using an mGBM-based regression to calculate Tri+ PCB concentrations from M8082 Aroclor concentrations measured on residual sediments post-dredging, which were expected to have relatively low PCB concentrations. Out of the 100 paired additional samples, 71 samples were selected from those having a M8082 Total PCB concentration less than 7 ppm, and 29 samples were selected from those having a M8082 Total PCB concentration greater than 57 ppm (General Electric 2009d).

The data for the original 180 paired samples and 98 of the 100 supplemental paired samples (two samples were excluded because they did not have results for Aroclor 1221) were used to develop a refined regression model (General Electric 2009d). Regression analysis was performed using the statistical software package R® (http://www.R-project.org). The basic regression methodology consisted of applying a weighted least squares (i.e., "damped-leveraged) algorithm to calculate the coefficients, summary statistics and goodness of fit measures (General Electric 2009d). The regression equation is:

$$Tri+\ PCB = 0.14*A1221 + 0.91*(A1242+A1254) \qquad \text{Eqn. 2-9}$$

The equation (Eqn. 2-9) is nearly identical to the equation developed from the original subset of 180 samples (Eqn. 2-8). The additional data result in tighter bounds on the coefficients (i.e., lower standard errors) and approximately the same goodness of fit statistics (Table A5-9; General Electric2009d).

During the remedial action period, GE continued collecting paired M8082-mGBM sediment data at a rate of 4 percent of the overall residual sediment samples to allow continued evaluation of the regression equation relating the Aroclor PCB concentrations measured using M8082 to the Tri+ PCB concentrations. The regression coefficients were recalculated each time the overall data set was increased by 5 percent and these updated regression coefficients were used in subsequent CU evaluations (General Electric, 2009d).

As documented in Corrective Action Memorandum (CAM) No. 3 – Modification of Sediment Residual Monitoring Program – Discontinuing mGBM Analysis of Sediment Samples and Updates of the Regression Coefficients (General Electric 2011j), GE performed four rounds of sampling in 2009 and 1 round of sampling in 2011 (Table A5-10) which resulted in a total of 445 paired M8082-mGBM sediment samples. The regression analysis was then repeated following the methodology described in the 2009 Quality Assurance Project Plan (QAPP) (General Electric 2009d). The coefficients derived from the update, as well as the coefficients from the original regression and subsequent rounds of updates in 2009, are provided in Table A5-10. The coefficients changed very little as a result of the updates. The additional data result in tighter uncertainty bounds on the coefficients (i.e., lower standard errors) and approximately the same goodness of fit statistics (Table A5-11). Because the coefficients have stabilized, GE recommended discontinuing the mGBM analysis of sediment samples and subsequent update of the regression coefficients.

The latest coefficients developed in 2011 are as follows:

$$Tri+\ PCB\ = 0.13*A1221 + 0.89*(A1242+A1254) \qquad \text{Eqn. 2-10}$$

Eqn. 2-10 was applied to any samples collected in all dredging 'certification units' that had not yet been sampled for first pass of the residual samples as of August 5, 2011.

2.2.8 Congener-Specific Measurements during the 2016 OM&M Sampling

As part of the sediment collection conducted by GE in 2016, a subset of the samples (approximately 10 percent) were provided to EPA as split samples. At the request of the federal natural resources trustees, these samples were analyzed for PCB congeners via method 1668C (M1668C). This was done, in part, due to the anticipated transition from mGBM to a congener-specific method for the long-term OM&M program. EPA is further analyzing the relationship between the M8082 results obtained by GE and the M1668C results obtained by EPA. EPA also anticipates collecting more matched pairs of these analyses as part of future OM&M sampling. Additionally, EPA plans to meet with the federal natural resources trustees, NYSDEC, and GE to further discuss consideration related to various analytical testing methods and ongoing OM&M sampling.

2.3 Summary

This section summarizes the translation schemes (regression equations) for estimating Tri+ PCBs in the sediment for the historical data that support the EPA decisions contained in the ROD, as well as the development of relationships between Tri+ PCB and Aroclors during the remedial investigation period. The sediment data sets and the translation schemes used to support the ROD, remedial investigation, and remedial action phases are summarized below:

- The 1976-1978 NYSDEC data covered the area from Fort Edward to Troy (RM 194.8 to RM 154) and were used to identify likely areas for remediation and provide estimates of PCB mass and sediment volumes for the 2002 ROD for the Upper Hudson River. The data were analyzed by Aroclor-based method. The following equation was used to calculate Tri+ PCB: $Tri + (1977) = 1.131 \times [Aroclor\ 1016 + 1254]$

- The 1984 NYSDEC data were restricted to TIP. Similar to the 1976-1978 NYSDEC data, the 1984 data were used support the ROD. These datasets were also dependent on the Aroclor quantitation method. The following equation was used to calculate Tri+ PCB for the 1984 NYSDEC data: *Tri+ (1984) = 0.944 × 1984 Aroclor Sum*
- GE conducted sediment surveys in 1991 and 1998. The 1991 survey sampled the upper river from Fort Edward to Federal Dam, while the 1998 survey sampled the TIP only. These datasets were used in the modeling effort during Phase 2 RI/FS investigation. Both 1991 and 1998 data were reported as PCB congeners based on mGBM. Therefore, the Tri+ PCB concentrations can be calculated directly and no translation scheme was needed.
- In 1994, the EPA collected low-resolution cores and the data were analyzed based on the congener-specific method. The 1994 low-resolution coring effort focused on replicating a representative subset of the 1984 locations. Since these data were reported as PCB congeners, no translation scheme was needed to calculate the Tri+ PCB concentrations.
- During the remedial design investigation phase, GE collected sediment samples in 2002-2005. These data were collected for delineating the final areas for removal. Given the large number of samples (approximately 30,000 sediment samples), the samples were analyzed using M8082. A subset of the samples was analyzed by the EPA-approved M680. Using the matched-pair samples, a regression equation was developed to estimate the Tri+ PCBs from the Aroclor-based measurement. As discussed in Section 2.2, the regression equation to estimate the Tri+ PCB concentration integrates the bias correction in the M680 data and the conversion of Aroclor results. The following regression equation was used to estimate the Tri+ PCB concentration: *Tri+ PCB = 0.03*A1221 + 1.16*(A1242+A1254).* This regression equation developed during the remedial investigation period was used to delineate dredge areas, to examine Tri+ PCB concentrations in residual core samples, to examine Tri+ PCB concentrations in special studies such as the downstream deposition study and sediment traps during remedial action period, and for the OM&M study results.

- During the remedial action phase, GE conducted a laboratory methods comparison study to compare the relative accuracy and precision of M680, M8082, and the mGBM. Based on this study, the following equation was developed to estimate the Tri+ PCB concentrations in the sediment data: *Tri+ PCB = 0.13*A1221 + 0.94*(A1242+A1254)*. Throughout the remedial action phase, GE conducted multiple rounds of paired mGBM and M8082 PCB data collection and evaluation which were used to assess the regression equation relating the Aroclor PCB concentrations measured using M8082 to the Tri+ PCB concentration. In 2011, GE recommended discontinuing the mGBM analysis of sediment samples and subsequent update of the regression coefficients because the coefficients do not change significantly and have been stabilized (General Electric 2011j). The 2011 regression equation (*Tri+ PCB = 0.13*A1221 + 0.89*(A1242+A1254)*) was used for sediment samples collected after August 5, 2011.

3 WATER

The concentration of PCBs has been measured in surface water in the Upper Hudson River since 1977 by different organizations. Over the period of measurement, different lab analytical techniques have been used. Valid comparisons of the trends in surface water PCB concentrations among various datasets can only be made when the various analytical results are converted to a consistent, congener-based quantitation.

This section describes the analytical conversions developed for water column data in the Upper Hudson River. Since 1977, PCBs have been measured in the water column in the Upper Hudson River by the United States Geologic Survey (USGS), GE, and EPA. While the water column analyses by EPA (in 1993) and GE (starting in 1991) were routinely based on congener quantitation, the long-term record from the USGS (1977-1995) used different analytical methods that required adjustments in order to make the datasets comparable for use in the data evaluation and model calibration. It was determined that an equivalent TPCB quantitation could not be determined from the USGS data and therefore Tri+ PCB, which could be calculated in all datasets, was used for data evaluation and model calibration.

As stated in the Revised Baseline Modeling Report (RBMR) for the Upper Hudson River (EPA, 2000b), the USGS water column data represented whole water analyses, with PCBs quantified using Aroclor standards. Packed column analysis was used until 1987, when data began to be analyzed with capillary columns. Split sample analysis between USGS and EPA Phase 2 data supported use of the USGS-reported TPCB concentration from the packed column analysis as a direct measure of the Tri+ PCB. A regression relating USGS TPCB to the Tri+ sum gave a good linear fit with an intercept not significantly different from zero (EPA 1997b). Thus, the USGS packed-column TPCB results were used directly as Tri+ concentrations through 1987. Further, the re-analysis of 60 USGS capillary column sample chromatograms by QEA (Rhea and Werth, 1999) supported use of the USGS-reported Aroclor 1242 results or, when 1242 results are not available, use of Aroclor 1248, as the best representation of the Tri+ PCB concentration in the USGS data after 1987.

Beginning in 1991, water samples collected from the Hudson River by GE were routinely analyzed for PCB congeners by the mGBM. EPA conducted a short-term series of studies on PCB levels in Hudson River water in 1993 using its dual phase PCB congener analysis. Thus, for both GE and EPA data, congener-specific information was available to directly determine Tri+ PCB and TPCB water column concentrations. In 2011, GE asserted that the mGBM was unnecessarily specialized for compliance purposes, and they began an evaluation of the efficacy of using a standard Aroclor method (M8082) for estimating Tri+ PCB for water samples collected during the remedial construction and off-season monitoring period. Notably, the M8082 method is also to be used during the OM&M period after the construction. Therefore, it is important that an accurate methodology be determined to estimate Tri+ PCB from the M8082 results so that future data can be compared to existing data and recovery trends can be accurately discerned. The procedure to determine Tri+ PCB concentrations from Aroclor data was based on the same regression methodology used to estimate Tri+ PCBs in sediments (see Section 2 above). In 2011, GE analyzed the paired data available and produced preliminary regression equations, and has used additional data collected after 2012 to refine these equations (General Electric 2012i, 2012k, 2013j, 2014h, and 2016a). The remainder of this section describes the analyses conducted by the EPA to determine the best regression equation(s) to estimate water column concentrations of Tri+ PCBs and the associated uncertainties using all the available paired data. Different regression relationships were explored for cases where only a single Aroclor was reported, when multiple Aroclors were reported, and also for the PCB concentrations determined in the samples. EPA's analysis was performed using the R® statistical package. The models employed were Ordinary Least Square Regression (OLS), Least Absolute Values Regression (LAV), M-Estimation with Huber Weight (M-Huber), M-Estimation with Bi-square Weight (M-Bisquare), Least Trimmed Squares Regression (LTS), and Least Median Squares Regression (LMS).

3.1 Paired GE mGB and Aroclor Water Column Data Set

A total of 249 water samples collected using automatic samplers at far-field dredging monitoring stations in 2011 to 2013 and 2015 were analyzed by both the mGBM and the

Aroclor methods. These paired data were used to develop the relationship between Aroclor and Tri+ PCB concentrations. Among the 249 samples, EPA excluded 10 from the analysis because, as documented in CAM 5 and CAM 8 (General Electric 2013j and 2014h, respectively), they were judged not representative of the Hudson River or did not meet analytical QA/QC requirements. Duplicate and triplicate samples were averaged prior to incorporation into the data analysis.

3.2 Test for Regression Outliers and High Leverage Samples

Multiple linear regression analysis based on OLS is sensitive to data points with large residuals (regression[5]) and/or high leverage.[6] This evaluation considers whether there were influential samples which may have affected the OLS results based on Cook's distance (or Cook's D). Cook's D combines the information of leverage and residual of the observation, and is commonly used to estimate the influence of individual data points. A simple operational guideline of Cook's D greater than 1 has been suggested for spotting highly influential data points (Cook and Weisberg, 1982). Figure A5-6 shows the Cook's D of the OLS regression for the three groups of samples separately and all the samples together. The results show that all the samples have Cook's D less than 1, indicating that there are no influential samples on the OLS regression. While these results show that there are no influential data that will affect the regression, EPA explored both OLS regression and several commonly-used robust regression approaches (see Section 3.4) to develop the best equation(s) for estimating Tri+ PCB in the water column data.

3.3 Validity of Using Aroclors as Predictable Variables

In the paired data sets available for this evaluation, Aroclor 1221 was detected by M8082 in all 239 water samples, while Aroclor 1242 and Aroclor 1254 were detected in 71.5 percent and 7.1 percent of the samples, respectively; no other Aroclors were detected. For

[5] A regression outlier is an observation with large residual. In other words, it is an observation whose dependent-variable value is unusual given its value for the predictor variables.

[6] Leverage is a measure of how far an independent variable deviates from its mean. An observation with an extreme value on a predictor variable is a point with high leverage. High leverage points can have an inordinate effect on the estimate of regression coefficients.

samples where Aroclor 1221 was the only detectable Aroclor by M8082, EPA used Aroclor 1221 as the only independent variable in the regression relating A221 and Tri+ PCB. In all other cases with multiple Aroclor detections, EPA used the concentrations of Aroclor 1221 and the sum of Aroclor 1242 and Aroclor 1254 in the regression. This section evaluates whether using individual Aroclors or total Aroclor as independent variables could produce a better correlation.

For simplicity, a test was performed using OLS linear regression for all the samples. In this approach, the best model is selected based on several feature selection measures, specifically the leave-one-out cross-validation (LOOCV) statistic (also known as PRESS, i.e., prediction residual sum of squares[7]), Akaike's Information Criterion (AIC), corrected AIC (AICc), Schwarz's Bayesian Information Criteria (BIC), and adjusted R^2. A small value of PRESS, AIC, AICc, and BIC, or a large value of adjusted R^2, indicates a model with a low test error. Table A5-12 shows the combination of various linear regression model formulations considered by EPA to test the efficacy of using individual Aroclors as predictor variables. The results of this test indicate that the lowest values for PRESS, AIC, AICc and BIC and the highest value for adjusted R^2 can be obtained for a regression model that is similar to remedial design sediment conversion model, formulated as follows:

$$Tri + PCB \ = \ a \ \times \ A1221 \ + \ b \ \times \ (A1242 + A1254) \qquad \text{Eqn. 3-1}$$

where a and b are regression model regression coefficients and Aroclor 1221, Aroclor 1242, and Aroclor 1254 are the Aroclor concentrations reported by M8082.

In the application of this model to determine the model coefficients for the Upper Hudson River, regression analyses for three groups of data were performed as follows:

- Group 1: Samples for which Aroclor 1221 was the only detected mixture.

[7] PRESS is the sum of squared prediction residuals. It is the same as the sum of the squared "leave-one-out" residuals.

- Group 2: Samples with multiple Aroclors detected but limited to those for which TPCB is less than or equal to 150 nanogram/Liter (ng/L), or with Aroclor ratios less than or equal to 0.2, where Aroclor ratio = (Aroclor 1242 + Aroclor 1254)/Aroclor 1221.
- Group 3: Samples with multiple Aroclors detected but limited to data with TPCB greater than 150 ng/L and Aroclor ratios (see above) greater than 0.2.

3.4 Regression Model Results

In addition to the OLS regression, EPA explored several robust regression approaches to determine the best values of the model coefficients a and b. In general, robust regression models give less weight to unusual observations that would otherwise have undue influence on the regression line. Table A5-13 lists the OLS and various robust regression methods, as well as the functions applied in the statistical R® package to determine the regression coefficients. Table A5-14 presents the regression coefficients for each model. Figure A5-7 compares the predicted and measured Tri+ PCB concentrations for the test models. As illustrated in Figure A5-7, a comparison between predicted and measured concentrations using the training data set does not necessarily readily identify the model with the best performance. To select the best model, EPA used a cross-validation approach.

3.5 Model Selection

Two model evaluation methodologies were used to determine the best model coefficients applicable to the water samples in the Hudson River. The first was a cross-validation using all available data. The second involved model validation using test data to evaluate the performance of the various regression models.

Cross-validation is primarily a way of measuring the predictive performance of a statistical model. LOOCV was used in this study to evaluate the prediction accuracy of the regression models. LOOCV means that the regression is performed on all the data except for one point and a prediction is made for that point. This process is repeated such that every data point is included in a test set exactly once, and in a training set *n*-1 times (where *n* denotes the total number of data points). The process creates a prediction error[8] for each point, and the average error is used to evaluate the model. Table A5-15 compares the average prediction errors among the regression models. In general, the lower the average prediction error, the better the model. The results indicate that average prediction error values do not show much difference among the models. The best model with the lowest average prediction errors for each group of regression is highlighted in bold in Table A5-15. The relative prediction error of the best model was calculated as mean prediction error divided by the mean of the measured Tri+ PCB concentration. The results indicate that the relative prediction error is 27, 18, and 23 percent for Group 1, Group 2 and Group 3 samples, respectively.

In the second approach, selecting the best model involved determination of the predictive performance of each statistical model using a test data set.[9] Similar to the cross-validation analysis presented above, the predictive accuracy of a model can be measured by the average prediction error on a test set. In this evaluation, the test data set contains far-field grab, near-field background and off-season in-river samples. This dataset contains 12 samples from Group 1 and 23 samples from Group 2, and therefore, only Groups 1 and 2 regression could be tested. Table A5-16 provides the average prediction error for each regression model and the results indicate that the LMS and LTS models performed better for Groups 1 and 2 regressions, respectively.

Overall, based on the LOOCV results, M-Huber, OLS, and LAV appear to be the best models for Group 1, Group 2, and Group 3 samples, respectively. Based on the test set results, LMS and LTS appear to be the best models for Group 1 and Group 2 samples, respectively. Since the test data sample size is much smaller than that for the LOOCV, the best models for Group 2 and Group 3 are selected as OLS and LAV, based on the LOOCV results.

---

[8] Prediction error or prediction residual is defined as the difference between the predicted value and the actual value for a test data point. It is generally greater than the residual from the model, which is based on the training data.

[9] A test data set includes samples which are not used in the development of regression models.

3.6 Summary

The ability to predict the concentrations of Tri+ PCB accurately from Aroclor data was essential for daily compliance evaluations at the far-field monitoring stations, and will continue to be required for determining future trends in water column concentrations during the OM&M period. In this evaluation, several linear regression approaches were considered for the prediction of Tri+ PCB concentrations based on Aroclor measurements in water column samples. The models employed were Ordinary Least Square Regression (OLS), Least Absolute Values Regression (LAV), M-Estimation with Huber Weight (M-Huber), M-Estimation with Bi-square Weight (M-Bisquare), Least Trimmed Squares Regression (LTS), and Least Median Squares Regression (LMS). The models were developed using a training data set containing 239 water samples, and their predictive performance was assessed using LOOCV and an independent test data set. The mean absolute prediction error and the relative prediction error were calculated to compare the predictive performance of the models. Based on the available data, LMS, OLS, and LAV were the best prediction models for Group 1, Group 2 and Group 3 samples, respectively. Table A5-17 provides the regression equations developed by EPA for use with the water column data. For comparison purposes, the table also provides the regression that GE developed and refined as documented in applicable corrective action memoranda ([CAMs] General Electric 2013j and 2014h, respectively).

4 FISH

Similar to the two previous sections, this section presents the analyses conducted to reconcile various historical and current fish tissue sampling and analysis efforts into a single, internally consistent series of measurements. For both sediment and water, it was necessary to separately estimate TPCB and Tri+ PCB, since both measures were important to understanding fate, transport and biological exposure. For fish however, this separation is not necessary since TPCB is approximately equal to Tri+ PCB (due to preferential fish biouptake/bioaccumulation for these homologues). Based on congener data obtained by EPA in the 1990s, Tri+ PCB was found to represent 90 percent or more of the total PCB burden in fish samples (EPA 1999a). This observation continued to hold true until the start of dredging in 2009, which is discussed further below. In all cases, the conversions developed were intended to estimate the Homolog Equivalent TPCB ($TPCB_{HE}$) mass in the tissue samples, regardless of whether the mixture is solely Tri+ PCB or includes substantial amounts of monochloro or dichloro homologues.

This section presents the procedure followed to calculate the $TPCB_{HE}$ concentrations, based on the sum of Aroclor concentrations ($TPCB_{Aroclor}$) reported in fish samples collected and analyzed by NYSDEC and GE from the Hudson River. This procedure follows the calculation process first described in the Hudson River PCBs Site remedial investigation reports and in Butcher, et al., 1998, and is necessary to facilitate year-to-year comparisons of PCB levels in fish using an equivalent basis of measurement through time. As mentioned previously, TPCB analysis by Aroclors, typically M8082, is a relatively subjective analysis, relying on the discretion of the analyst as to the selection of Aroclors to be reported and quantitated in a sample. As a result, the mixture of Aroclors reported, as well as the sum of Aroclors, will vary from laboratory to laboratory for equivalent samples. This is evident in the historical record of Aroclor-based measurements in fish for the Hudson River.

In addition to concerns regarding Aroclor mixture variations attributable to lab differences, a measured congener pattern in a fish sample would not be expected to directly replicate an exact Aroclor congener distribution. This is because variations in historical Aroclor use, weathering of released PCBs in the environment prior to exposure to biota, dechlorination, and also preferential biological uptake of Tri+ congeners over others all serve to extensively modify the actual congener distribution present in fish (as described in Section 1). Thus, quantitation of PCBs in fish tissue by Aroclor-based analysis can only be an approximation of the actual TPCB mass ($TPCB_{HE}$). This concern applies to both the NYSDEC data and the GE data described below. To address this concern, EPA examined fish samples analyzed by both congener-based and Aroclor-based methods to develop correction factors for the Aroclor-based results. These are described in sections to follow.

4.1 Discussion of NYSDEC Fish Data

The NYSDEC fish data examined here span the period 1990 to 2015. Prior analyses by NYSDEC are not discussed here but are covered at length in Butcher, et al. 1998. For the entire period examined, fish tissue data are reported by NYSDEC on an Aroclor basis. As presented in Figure A5-8, the selection of reported Aroclors (the Aroclor composition) in NYSDEC fish samples clearly varies through time. Annual variation is apparent through 1998 and 2013 to 2015, while the patterns are relatively similar from 1999 through 2011. Changes in the suite of reported Aroclors largely correspond to changes in the analytical labs performing the procedures. Table A5-18 summarizes the appearance of various Aroclor mixtures in the quantitation.

Given the dominance of the GE source of PCBs in the Upper Hudson and the large reservoir of PCB contamination in the sediments, it is unlikely that the pattern of PCBs exhibited in fish was actually varying in a significant fashion through time. This assertion is directly supported by the 1999 to 2011 period, when a single lab (Mississippi State Chemical Laboratories) was used by NYSDEC, and a uniform set of Aroclors in roughly similar proportions was consistently reported over the full period.

4.2 Discussion of GE Fish Data

Aroclor concentrations reported in GE fish samples also clearly vary through time, although in a much more systematic fashion than in the NYSDEC data. As shown in Table A5-19, GE employed a single lab for PCB Aroclor analysis over the entire period

(2004-2016). Prior to 2009, the GE lab consistently reported Aroclors 1248 and 1254 as the dominant fractions, with minor amounts of Aroclors 1242 and 1260. Aroclor 1221 is essentially absent during this period (see Figure A5-9). Beginning in 2009, however, the Aroclor 1221 fraction began increasing, reaching 20 percent of the total mixture for 2012 and 2015. This period is highlighted in Table A5-19.

In this instance, the change of the Aroclor distribution in fish tissues presumably is not due to a change in analytical procedures or judgment, since only one laboratory was used by GE, as shown in Table A5-19. Rather, it is most likely a direct reflection of the increased exposure of fish to congeners associated with Aroclor 1221, or more specifically to the presence of congener BZ#4 (peak 5 based on GE's mGBM) released into the water column when remedial dredging operations started in the Hudson River in 2009. An increase in the proportion of lighter congeners in the water column was extensively observed and documented in the various dredging reports issued by EPA and GE.

4.3 Aroclor Measurements and Estimation of TPCB Concentration

As noted above, the mixtures of Aroclors reported through time varied widely but, prior to the onset of dredging, it is unlikely that the actual distribution of congeners varied so extensively with time. Indeed, even for the period where NYSDEC and GE each consistently using their own dedicated laboratory (2004 to 2011), the Aroclor mixtures reported by NYSDEC's and GE's laboratories differ significantly. This can be readily observed by comparing Figures A5-8 and A5-9 for this period. While the reporting of different Aroclor mixtures by different laboratories is clearly problematic in trying to discern changes in PCB patterns through time, it also presents a much larger issue for long-term monitoring purposes. Specifically, in reporting different Aroclors, it is highly likely that different laboratories will obtain different estimates for the mass of PCBs contained in the sample; the mass of total PCBs is estimated by summing the individual reported Aroclors.

The issue with this estimate of TPCB mass in a sample arises from disconnect between the actual PCB pattern present in the fish sample and the "standard" congener pattern present in the Aroclor standard mixtures. Included in this problem is the extensive overlap in congener content among different Aroclor mixtures. That is, many congeners are observed in several different Aroclors (*e.g.*, Aroclors 1242, 1248, 1254, and 1260 all have multiple congeners in common). As a result, reporting of multiple Aroclors with overlapping congener spectra potentially leads to "double counting" of individual congener concentrations. Different laboratories have different means of avoiding "double counting," again similarly leading to differences in reported concentrations; this time TPCBs as the sum of Aroclors. The presence of overlapping Aroclor spectra was extensively documented in EPA's Data Evaluation and Interpretation Report (EPA 1997b) as well as in Frame et al. (1996).

Alternatively, the actual congener distributions in the fish tissue may differ widely from those found in the Aroclor standards. Variations in analytical response factors, relative contributions to total PCB mass and simply lack of pattern similarity can all result in inaccurate quantitation of the actual PCB mass, yielding both underestimates and overestimates of the TPCB concentration.

Fortunately, for all of the data sets examined here, an alternate means of quantifying PCB mass is available for a subset of samples obtained by each laboratory. Specifically, TPCBs were reported by homologue or by congener (equivalent to EPA M680 or GE's mGBM, hereinafter referred to as congener-based methods) for each laboratory data set, along with the Aroclor-based analysis. Totals PCBs by congener-based methods are much less subject to the analyst's judgment as to the mixture of PCBs present as a basis for quantitation of Aroclors. These homologue and congener methods simply require the quantitation and summation of all congeners in the mixture. As a result, use of these methods should provide a consistent basis of TPCB measurement across time and laboratories.

4.4 Development of the Homologue Equivalent Basis

As noted above, congener-based PCB data are available for all of the individual laboratories but exist for only a subset of the total number of samples from each laboratory. To estimate TPCB concentrations as the sum of homologues ($TPCB_{HE}$) for the vast majority of the fish data, a series of correlation analyses were conducted using the available Aroclor and homologue-based analytical pairs in both NYSDEC and GE data sets. Ideally, a simple proportionality constant should relate the Aroclor and homologue sums. Given that both the sum of Aroclors and the sum of homologues should approach zero when a sample does not contain PCBs, the translations were conducted using the following formulations:

$$TPCB_{HE} = \alpha \times TPCB_{Aroclor} \qquad \text{Eqn. 4-1}$$

where α is determined by a regression on the $TPCB_{HE}$ and $TPCB_{Aroclor}$ pairs for a given laboratory.

$$\alpha_i = \frac{TPCB_{HE}}{TPCB_{Aroclor}} \qquad \text{Eqn.4- 2}$$

where the α_i are determined for the individual samples with both $TPCB_{HE}$ and $TPCB_{Aroclor}$ values.

The α_i were then used to determine the arithmetic mean, median and geometric mean ratio (α) for a given laboratory and period of time. The α_i were also used in a bootstrap analysis to determine confidence intervals around the geometric mean, discussed further below.

These equations were applied for the various data sets for the period from 1998 to 2011. For NYSDEC data for the period from 1990 to 1997, the regression equations as described in the RBMR (EPA 2000b) were used without modification. For the 1998 to 2015 data, the regression was prepared as follows: The limited number of nondetect results (approximately 4 percent of sample pairs across all post-1998 data) were eliminated from this analysis and only detected results for both Aroclor and congener-specific data were used. That is, if either the sum of Aroclors or sum of homologues was nondetect, the sample pair was not included in the regression. In all instances, the geometric mean ratio appeared to best represent the relationship between $TPCB_{HE}$ and $TPCB_{Aroclor}$ values across the entire range of concentrations. A summary of the resulting factors to be applied to the Aroclor sums is shown in Table A5-20, along with a brief description of the origins of the individual factors.

To provide an example of the importance of these factors in adjusting the $TPCB_{Aroclor}$ to yield $TPCB_{HE}$, the NYSDEC's matched pairs of $TPCB_{HE}$ and $TPCB_{Aroclor}$ values were plotted against each other in Figure A5-10, for the period 1991 to 2000. The great degree of scatter reflects the level of variation in the relationship between the two metrics for individual samples. However, also note that the data tend to cluster by year, as indicated by the colored symbols. For example, 1991 samples consistently lie below the 1-to-1 line, indicating that $TPCB_{Aroclor}$ is consistently greater than $TPCB_{HE}$ for this period. For the 1999 and 2000 data, the samples consistently plot above the 1-to-1 line, indicating $TPCB_{Aroclor}$ is consistently less than $TPCB_{HE}$ for this period. For the data collected before 1998, the EPA developed relationships between $TPCB_{HE}$ and $TPCB_{Aroclor}$ which were included in the documents for the ROD. Since 1998, additional data has been obtained by the NYSDEC and GE that warrants similar treatment. The development of four new relationships for the period 1998 to 2015 between $TPCB_{HE}$ and $TPCB_{Aroclor}$ and their associated statistical support are described below.

4.5 NYSDEC Data Factors

The currently available NYSDEC data on PCB concentrations in fish tissue post-1997 span the period 1998 to 2015. These data were analyzed by three laboratories. Enchem Environmental Laboratories was used for all fish samples collected in 1998, during which NYSDEC obtained 82 samples analyzed for both $TPCB_{HE}$ and $TPCB_{Aroclor}$. These data were used to develop the formula given in Table A5-20. Table A5-21 summarizes the additional statistics examined for the 1998 data. The data are plotted as $TPCB_{Aroclor}$ *vs.* $TPCB_{HE}$ in Figure A5-11. In the figure, the data generally lie below the 1-to-1 line, indicating that the $TPCB_{Aroclor}$ results tend to be higher than those for $TPCB_{HE}$. As a result, the estimate for α is less than 1 at 0.741. This value is similar to that obtained by arithmetic mean (0.826), median (0.783) or by the regression (0.784).

In general, the similarities among geometric mean, median, and average all yielded similar values for α within a single laboratory-time period for both the NYSDEC and GE data. However, given the reduced sensitivity of the geometric mean to the occurrence of outliers relative to the average, and that statistical calculations using the geometric mean are generally simpler than those for the median, the geometric mean was selected as the basis to select α and was further examined for its uncertainty via a bootstrap technique. While the arithmetic mean and the median generally agreed well with the geometric mean, the estimate of α by regression analysis was inconsistent in its agreement across all the four single laboratory-time period model. For this reason, the regression analysis was discontinued and is not discussed further.

The uncertainty in the geometric mean was estimated by multiple resampling (a statistical technique referred to as "bootstrapping") of the population of matched $TPCB_{Aroclor}$ and $TPCB_{HE}$ pairs. Table A5-21 also contains the bootstrap output for the 2.5, 5, 95, and 97.5 percentile confidence limits for the geometric mean. Additionally, the table contains an estimate of the uncertainty of the geometric mean expressed as a percentage of the geometric mean value itself, using the 5th and 95th percentile values. The 95 percent confidence limits on the geometric mean value for 1998 of 0.741 is $\pm$ 8 percent. This confidence interval includes the median value, indicating that these values agree within the uncertainty. The statistical agreement between geometric mean and median occurs for all four single laboratory-time period models described here.

Besides illustrating the general correlation between $TPCB_{Aroclor}$ and $TPCB_{HE}$, the symbols in Figure A5-11 are color-coded to identify the fish species analyzed for each individual sample pair. For 1998, the majority of the samples are brown bullhead and largemouth bass. Both species are similarly distributed over the range of values, indicating that the conversion from $TPCB_{Aroclor}$ to $TPCB_{HE}$ is similar for both species, and no species-specific conversion is needed. This observation based on the 1998 data is also supported by the 1999 to 2000 data, which also show consistent behavior between $TPCB_{HE}$ and $TPCB_{Aroclor}$ across the various species. The similar behavior between $TPCB_{HE}$ and $TPCB_{Aroclor}$ across species indicates that congener patterns of PCBs in tissues are similar across species, yielding similar Aroclor percentages across species. This consistency was also noted in the congener-specific analysis of PCB patterns in fish conducted for the ecological risk assessment for the ROD (EPA 1999b).

Figures A5-12, A5-13, and A5-14 provide additional statistical background on the estimate of the ratio (α) for the NYSDEC 1998 data. Figure A5-12 shows the detailed statistics on the range of α for the individual samples. The figure also shows that α skews right and may not be normally distributed. Figure A5-13 represents several statistical tests of significance comparing the paired $TPCB_{Aroclor}$ to $TPCB_{HE}$ values. The results in this figure indicate that for 1998 data, $TPCB_{Aroclor}$ is significantly higher than $TPCB_{HE}$ by all tests including the Wilcoxon signed rank test, which has no requirement regarding an underlying normal distribution. Based on this test, the use of an adjustment factor (i.e., α) is statistically justified. Figure A5-14 compares the logs of $TPCB_{Aroclor}$ and $TPCB_{HE}$. In this test, showing that the log values have a significant difference is mathematically equivalent to showing that the ratio of $TPCB_{Aroclor}$ to $TPCB_{HE}$ is statistically different than unity (1). In this test, the logs are significantly different but the difference appears relatively constant with increasing log value (*e.g.,* concentration). This result actually indicates that the ratio of $TPCB_{HE}$ to $TPCB_{Aroclor}$ is relatively constant, and independent of concentration. This further justifies the use of a single coefficient α to estimate $TPCB_{HE}$ from $TPCB_{Aroclor}$.

The 1999 to 2000 NYSDEC data were analyzed in a parallel fashion, yielding the statistics given in Table A5-21 for this period, as well as Figures A5-15 through A5-18. These figures parallel the presentation described above for the 1998 result. Figure A5-16 presents the NYSDEC fish results, with symbols and color-coding to indicate the year of collection as well as species type. Of particular note, the data appear sufficiently consistent across the more than 2 orders of magnitude range of detected concentrations to suggest that neither year (*i.e.,* 1999 *vs.* 2000) nor fish species are important factors to consider in the development of α for this period. Figures A5-16 through A5-18 present the statistical support for developing a correction factor for this period. In this instance, however, the geometric mean factor is greater than 1 (i.e., 1.17), indicating that $TPCB_{Aroclor}$ is an underestimate of $TPCB_{HE}$. The signed rank tests confirm a statistically significant difference between $TPCB_{HE}$ and $TPCB_{Aroclor}$. Again, the log-based analysis, shown in Figure A5-18, shows no trend with increasing concentration, although there is greater variability than for the 1998 results. This last analysis again supports the use of a single coefficient α to estimate $TPCB_{HE}$ from $TPCB_{Aroclor}$, similar to the 1998 data.

One remaining concern for the NYSDEC data is the lack of further matched pairs of $TPCB_{HE}$ and $TPCB_{Aroclor}$ in the post-2000 period. Thus, there are no additional data from which to develop these factors. However, NYSDEC continued to utilize the same laboratory from 1999 through 2011, a condition that would be expected to maintain the $TPCB_{HE}$ to $TPCB_{Aroclor}$ relationship observed in 1999-2000. Consistent with this, the Aroclor mixtures reported by the NYSDEC laboratory during this period have remained similar through the period (see Table A5-18 and Figure A5-8 for the period 1999 to 2011). These considerations justify the continued use of the 1999 to 2000 coefficient α value of 1.17 to estimate $TPCB_{HE}$ from $TPCB_{Aroclor}$ for the post-2000 period.

The NYSDEC data factors represent the Upper Hudson and near-Albany Lower Hudson conditions. It is unclear if these factors will be applicable to Lower Hudson conditions characterized by relatively lower PCB concentrations in fish.

4.6 GE Data Factors

The GE data for 2004 to 2016 were analyzed in a manner parallel to the NYSDEC data described above. Note that analysis by the mGBM was performed on 10 percent of the total number of fish samples during the baseline monitoring program (2004-2008) and 5 percent of the total number of fish samples during odd years (2009, 2011, etc.) during the remedial action monitoring period. GE did not perform mGBM in 2015, but performed it in 2016, as requested by the EPA (General Electric 2016k). The shift in 2016 was to allow sample-pair data one year after dredging was finished. Since GE has only employed a single laboratory for fish analysis to date, it was anticipated that a single coefficient α would be needed to estimate $TPCB_{HE}$ from $TPCB_{Aroclor}$ for the entire period. However, initial review of the GE data showed that the mixture of Aroclors reported by the GE laboratory was not constant, but rather changed substantially over the period of record, particularly after the start of the Phase 2 dredging program in 2011. Specifically, the reported laboratory results changed from trace amounts of Aroclor 1221 prior to dredging in 2009 to an average of 12 to 20 percent of the total Aroclor mixture during dredging (see Figure A5-9 and Table A5-19). The relationship of this change with the more extensive dredging effort and a substantial increase in monochloro- and dichloro-homologue water column loads is almost certainly not a coincidence. The likely causative relationship between water column loads and fish body burdens of Aroclor 1221 is further illustrated in Table A5-22, which presents the average Aroclor 1221 content in fish tissue as a function of both year and river mile.

In particular, later years of dredging tend to have higher fractions of Aroclor 1221. However, the table has been highlighted to show the correlation between the major areas of dredging during the 2009 to 2016 periods and fish sampling. Evident from the table are the higher fractions of reported Aroclor 1221 in the fish tissue in the areas of dredging and in areas downstream. This evidence suggests that the change in Aroclor mixture noted in the GE fish tissue data is likely a reflection of an increase in the 'true' monochloro- and dichloro-homologue fractions in fish tissue, resulting in higher reported Aroclor 1221 fractions. However, the higher Aroclor 1221 fraction reported in 2009 and later would also indicate that α is likely to require a change in response. Since the publication of the Proposed Second Five-Year Review Report in June 2017, EPA received the 2016 fish tissue data from GE. The 2016 fish tissue data showed that the Aroclor 1221 fraction, on average, is now under 5 percent, which is three to four times lower than the values detected during dredging (see Table A5-19). Based on this, the GE data were split into two periods for analysis: 2004 to 2008 and 2009 to 2013. The 2016 data were not used in the correction factor development at this time. However, EPA anticipates collecting more matched pairs of samples for Aroclors and congener-based analyses as part of future OM&M sampling. EPA will further analyze the relationship between the Aroclors and congener-based methods in fish tissue as they become available. Estimates of α for these periods are given in Table A5-21 along with the associated statistics and uncertainty estimates.

The GE data for the period 2004 to 2008 are shown in Figures A5-19 to A5-22. These figures parallel the same analyses performed for the NYSDEC data, and yield similar conclusions concerning the statistical significance of α. Given the apparent lack of species-specific differences in the relationship between $TPCB_{HE}$ from $TPCB_{Aroclor}$ for the NYSDEC data, EPA did not repeat the analysis for the GE data. Rather, for the GE data, the data in Figure A5-19 (and Figure A5-23 discussed below) are color-coded by year of collection. The data examined in this fashion do not suggest substantial variation in α from year to year within the 2004 to 2008 interval.

One additional observation is worth noting, however. Specifically, as can be observed in Figure A5-19, the correlation between $TPCB_{HE}$ and $TPCB_{Aroclor}$ is much stronger for the GE data relative to the NYSDEC data. This is reflected in the narrower range of variability in the GE data, as well as in smaller estimates of variance, *i.e.*, the standard deviation and the geometric mean confidence levels (see Table A5-21).

The analysis for the period 2009 to 2013 is shown in Figures A5-23 to A5-26, again paralleling the analyses conducted for the NYSDEC data. Like the 2004 to 2008 period, the GE results show reduced variance as compared to the NYSDEC data. Similar to the previous results, the GE data for 2009 to 2013 support the use of α for estimating $TPCB_{HE}$ from $TPCB_{Aroclor}$. Of note, α for the 2009 to 2013 period based on the geometric mean 0.784, is similar to the value obtained for the 2004 to 2008 period (0.831). However, as can be inferred from the lack of overlap between the 5^{th} and 95^{th} percent confidence intervals, the difference in α between the two periods is considered to be statistically different. This analysis confirms EPA's decision to treat these periods differently, although the difference in the ratios is less than 6 percent.

Similar to the NYSDEC data factors, the GE data factors represent the Upper Hudson and near-Albany Lower Hudson conditions. It is unclear if these factors will be applicable to Lower Hudson conditions, which are characterized by relatively lower PCB concentrations in fish.

4.7 Fish Species Distribution

The analyses performed on NYSDEC data described above considered the possibility of variations in α related to fish species but did not find any strong evidence for the need to consider fish species in the analysis. For completeness, the available distribution of fish species as a percentage of the total number of paired samples for both the NYSDEC and GE data are summarized in Figures A5-27 and A5-28. Notably, the NYSDEC paired data for 1998 to 2000 represent fewer species than the GE data (2004 to 2015).

4.8 Summary

This section summarized the development of a relationship to estimate $TPCB_{HE}$ from $TPCB_{Aroclor}$ for the historical fish tissue data that supported the EPA decisions contained in the ROD, as well as the development of new relationships between $TPCB_{HE}$ and $TPCB_{Aroclor}$ for fish data collected by NYSDEC and GE during the post-ROD period. The following conclusions were drawn from this analysis.

- PCB analysis by Aroclor is subject to significant variations (including the analyst's professional judgment) in the types of Aroclors reported as well as in the sum of Aroclors, all of which impact the estimation of total PCB mass in a sample. This is readily apparent in the NYSDEC data, which were generated by multiple laboratories, but is an issue for all measurements based on Aroclors ($TPCB_{Aroclor}$).
- When compared to estimates of total PCB mass obtained by homologue or congener-based methods ($TPCB_{HE}$), the sum of Aroclors ($TPCB_{Aroclor}$) can either exceed or fall below the $TPCB_{HE}$ value.
- Despite this limitation, internal to a single laboratory, the $TPCB_{Aroclor}$ can represent a sufficiently precise measurement so as to permit the estimation of $TPCB_{HE}$ for a given period of time.
- Based on these observations, coefficients (α) to estimate $TPCB_{HE}$ from $TPCB_{Aroclor}$ were developed for four lab-period pairs.
  - For NYSDEC, these were 1998 data from the Enchem laboratory, and 1999-2000 data from the Mississippi State Chemical Laboratories. The results for the latter period (1999 to 2000) were applied to all subsequent NYSDEC

data since no additional $TPCB_{HE}$ / $TPCB_{Aroclor}$ sample pairs were available and NYSDEC did not change laboratories during the subsequent period.

  - For GE data, the laboratory-period pairs were 2004 to 2008 from NEA and 2009 to 2013 from NEA. In this instance, the data were divided into two subsets because the Aroclor pattern in fish changed between periods, likely due to the increased presence of lighter congeners in the water column released via dredging associated with the remediation.

- While there is variability in the ratio of $TPCB_{HE}$ to $TPCB_{Aroclor}$ in all the matched pairs evaluated, the data sets are sufficiently large as to provide well-constrained estimates of the ratio (expressed as α). The maximum uncertainty (95 percent confidence interval) obtained by a bootstrap analysis of variation in the geometric mean estimate expressed as a percentage of α was ± 8 percent.
- In all instances, the α values were shown to be statistically significant and therefore, their application as a basis to estimate $TPCB_{HE}$ from $TPCB_{Aroclor}$ is statistically supported.
- Finally, the absolute magnitude of the α values is sufficient that they represent a substantive, as well as a statistically significant adjustment to the $TPCB_{Aroclor}$ in order to obtain a more accurate estimate of $TPCB_{HE}$. Failure to account for variations in the $TPCB_{HE}$ to $TPCB_{Aroclor}$ is likely to introduce significant uncertainties and potential trend artifacts unrelated to actual changes in fish body burdens in any time-based trend analysis.
- While these factors represent Upper Hudson and near-Albany Lower Hudson conditions, there may be an issue with extrapolation of these coefficients to Lower Hudson conditions, which are characterized by relatively lower PCB concentrations in fish.
- The range of the geometric mean of α across datasets was 0.74 to 1.17. These values represented the two NYSDEC lab-period pairs. The range of α values for the GE laboratory-period pairs was much tighter (i.e., 0.79 to 0.854).

5 REFERENCES

Brown Jr., J.F., R.E. Wagner, and D.L. Bedard. 1984. PCB Transformations in Upper Hudson Sediments. *Northeast. Environ. Sci.* 3: 184-189

Butcher, J.B., T.D. Gauthier, and E.A. Garvey, 1998. Use of Historical PCB Aroclor Measurements: Hudson River Fish Data. Environmental Toxicology and Chemistry, Vol. 16, No. 8, pp. 1618-1623.

Butcher, J.B. 2000. Memorandum from J.B. Butcher of Tetra Tech to V. Bierman and S. Hinz of Limno-Tech, Inc. and E. Garvey and A. DiBernardo of TAMS Consultants, Inc. regarding Historic Sediment PCB Data Conversions dated June 9, 1998.

Cook, R. Dennis and S.Weisberg. 1982. Residuals and Influence in Regression. New York, NY: Chapman & Hall. ISBN 0-412-24280-X.

EPA Great Lakes National Program Office. 1987. *Quality Assurance Plan: Green Bay Mass Balance Study. I. PCBs and Dieldrin.* Prepared by D.L. Swackhamer.

EPA. 1994. Memorandum from M.D. Mullin of the U.S. EPA Environmental Research Laboratory - Duluth, Large Lakes Research Station to G. Frame of GE Corporate Research and Development dated 21 November 1994.

_______. 1997b. Further Site Characterization and Analysis, Volume 2C – Data Interpretation and Evaluation Report (DEIR), Hudson River PCBs Reassessment RI/FS. Prepared for USEPA Region 2 and USACE by TAMs Consultants, Inc. and Gradient Corporation. February 1997.

________. 1998. *Phase 2 Report - Review Copy Further Site Characterization and Analysis Volume 2C-A Low Resolution Sediment Coring Report Addendum to the Data Evaluation and Interpretation Report Hudson River PCBs Reassessment RI/FS*. Prepared for USEPA Region 2 and USACE by TAMS Consultants, Gradient Corporation, and Tetra-Tech, Inc. July 1998.

________. 1999a. Further Site Characterization and Analysis, Volume 2D - Baseline Modeling Report (BMR), Hudson River PCBs Reassessment RI/FS. Prepared for USEPA Region 2 and USACE May 1999.

_________. 1999b. Further Site Characterization and Analysis, Volume 2E – Ecological Risk Assessment for the Upper Hudson River, Hudson River PCBs Reassessment RI/FS. Prepared for USEPA Region 2 and USACE. August 1999.

________. 2000b. Further Site Characterization and Analysis, Revised Baseline Modeling Report (RBMR), Hudson River PCBs Reassessment RI/FS Volume 2D. Prepared for EPA Region 2 and USACE, Kansas City District by TAMS Consultants, Limno Tech, Menzie-Cura & Associates, and Tetra-Tech, Inc. January 2000.

____________. 2004b. Resolution of GE Disputed Issues since GE's May 21, 2004 Presentation to the Regional Administrator. July 22, 2004.

Frame, G.M., R.E. Wagner, J.C. Carnahan, J.F. Brown, R.J. May, L.J. Smullen and D.L. Bedard. 1996. Comprehensive, Quantitative, Congener-Specific Analyses of Eight Aroclors and Complete PCB Congener Assignments on DB-1 Capillary Columns. Chemosphere, 33(4):603-623.

General Electric. 1997. *Development of Corrections for Analytical Biases in the 1991-1997 GE Hudson River PCB Database.* Prepared by HydroQual, Inc. for General Electric.

General Electric. 1999. "PCBs in the Upper Hudson River; Volume 2; A Model of PCB Fate, Transport, and Bioaccumulation." Prepared for General Electric by QEA, LLC. . May 1999.

General Electric. 2003a. Hudson River Sediment Sampling and Analysis Program Year 1 Data Summary Report. Prepared by QEA for General Electric. May 24, 2003.

______. 2004e. Memo - Additional Analyses of Tri+ PCB Regression for use in the Revised Phase 1 Dredge Area Delineation Report. Prepared by QEA for General Electric. June 14, 2004.

______. 2005d. Data Summary Report: PCB Methods Comparison Study Congener Analytical Standards Analysis. Prepared for the General Electric Company by Environmental Standards, Inc., and Quantitative Environmental Analysis, LLC, May 20, 2005.

______. 2009d. Hudson River PCBs Site Phase 1 Remedial Action Monitoring Program Quality Assurance Project Plan. Prepared for the General Electric Company by Anchor QEA, LLC and Environmental Standards, Inc. May 2009.

_______. 2011c. 2010 Data Summary Report on Hudson River Water and Fish – Hudson River PCBs Superfund Site. Prepared by Anchor QEA, LLC in conjunction with Environmental Standards, Inc. for General Electric.

___________. 2011j. Corrective Action Memorandum (CAM) No. 3 – Modification of Sediment Residual Monitoring Program – Discontinuing mGBM Analysis of Sediment Samples and Updates of the Regression Coefficients. Memorandum from Emily Chen, Anchor QEA to Bob Gibson, General Electric. August 15, 2011.

____________. 2012i. Using an Aroclor Method to Quantify Total PCB Concentrations and Calculate Tri+ PCB Concentrations in Water. Prepared for the General Electric Company by Anchor QEA, LLC. May 2012.

___________. 2012k. Technical Memorandum: Refinement of Water Column Tri+ PCB Regressions for Far-field. August 2012. Prepared by Anchor QEA, LLC.

___________. 2013j. Updated Water Column Tri+ PCB Correlations for Far-field Stations. Remedial Action Monitoring Program Corrective Action Memorandum (CAM) No. 5. Prepared for the General Electric Company by Anchor QEA, LLC. June 2013.

___________. 2014h. Water Column Tri+ PCB Correlations Refined with 2013 Data. Remedial Action Monitoring Program CAM No. 8. Prepared for the General Electric Company by Anchor QEA, LLC. April 2014.

__________. 2016a. Water Column Tri+ PCB Regressions for Far-field Stations-Refined with 2015 Data. Remedial Action Monitoring Program CAM No. 11. Prepared for the General Electric Company by Anchor QEA, LLC. January 2016

_________. 2016k. Addendum to 2015 Data Summary Report – Fish Sampling Data. Prepared for the General Electric Company by Anchor QEA, LLC. December 2016.

Holland, P.W. and Welsch R.E. 1977. Robust regression using iterative reweighted least-squares. In *Communications in Statistics – Theory and Methods*, Vol 6(9): 813-827.

New York State Department of Conservation (NYSDEC). 1988. Brown, M.P., M.B. Werner, C.R. Carusone, and M. Klein. Distribution of PCBs in the Thompson Island Pool of the Hudson River, Final Report of the Hudson River PCB Reclamation Demonstration Project Survey. New York State Department of Environmental Conservation, Albany, NY.

________. 1989. Analytical Service Protocols. Issued September 1989, revised December 1991 and September 1993, Method 91-11, pp D-XXVIII, 5-59. New York State Department of Environmental Conservation, Bureau of Technical Services and Research, Albany, New York.

Rhea, J. and M. Werth. 1999. "Phase 2 Evaluation of Analytical Bias in the USGS Water Column Database." Technical Memorandum from Quantitative Environmental Analysis to John Haggard, General Electric. March 22, 1999. 12 pp.

Tofflemire, T.J., and S.O. Quinn. 1979. PCB in the Upper Hudson River: Mapping and Sediment Relationships. NYSDEC Technical Paper No. 56. March 1979. NYSDEC, Albany, New York.

Final Second Five-Year Review Report for the Hudson River PCBs Superfund Site

APPENDIX 5

PCB Aroclors Data Treatment

Tables and Figures

Prepared by:
Louis Berger US, Inc.

April 2019

Table A5-1
DB-1 Chromatograph Peaks and Corresponding PCB Congeners for Modified Green Bay Method

DB-1 Peak #	Congener BZ #	Chlorination Structure	DB-1 Peak #	Congener BZ #	Chlorination Structure	DB-1 Peak #	Congener BZ #	Chlorination Structure	DB-1 Peak #	Congener BZ #	Chlorination Structure	DB-1 Peak #	Congener BZ #	Chlorination Structure
1	0	biphenyl	31	52	2 2'5 5'	53	90	2 2'3 4'5	74	105	2 3 3'4 4'	109	201	2 2'3 3'4 5 5'6'
2	1	2	31	73	2 3'5'6	53	101	2 2'4 5 5'	74	132	2 2'3 3'4 6'	110	196	2 2'3 3'4 4'5 6'
3	2	3	32	49	2 2'4 5'	54	99	2 2'4 4'5	75	153	2 2'4 4'5 5'	110	203	2 2'3 4 4'5 5'6
4	3	4	33	47	2 2'4 4'	55	112	2 3 3'5 6	76	168	2 3'4 4'5'6	111	189	2 3 3'4 4'5 5'
5	4	2 2'	34	48	2 2'4 5	55	119	2 3'4 4'6	77	141	2 2'3 4 5 5'	112	195	2 2'3 3'4 4'5 6
5	10	2 6	34	75	2 4 4'6	55	150	2 2'3 4'6 6'	78	179	2 2'3 3'5 6 6'	113	208	2 2'3 3'4 5 5'6 6'
6	7	2 4	35	62	2 3 4 6	56	83	2 2'3 3'5	79	130	2 2'3 3'4 5'	114	207	2 2'3 3'4 4'5 6 6'
6	9	2 5	35	65	2 3 5 6	56	109	2 3 3'4 6	80	137	2 2'3 4 4'5	115	194	2 2'3 3'4 4'5 5'
7	6	2 3'	36	35	3 3'4	57	86	2 2'3 4 5	81	176	2 2'3 3'4 6 6'	116	205	2 3 3'4 4'5 5'6
8	5	2 3	37	44	2 2'3 5'	57	97	2 2'3 4'5'	82	138	2 2'3 4 4'5'	117	206	2 2'3 3'4 4'5 5'6
8	8	2 4'	37	104	2 2'4 6 6'	57	152	2 2'3 5 6 6'	82	163	2 3 3'4'5 6	118	209	2 2'3 3'4 4'5 5'6 6'
9	14	3 5	38	37	3 4 4'	58	87	2 2'3 4 5'	83	158	2 3 3'4 4'6	NQ	20	2 3 3'
10	19	2 2'6	38	42	2 2'3 4'	58	111	2 3 3'5 5'	84	129	2 2'3 3'4 5	NQ	38	3 4 5
11	30	2 4 6	38	59	2 3 3'6	58	115	2 3 4 4'6	85	178	2 2'3 3'5 5'6	NQ	41	2 2'3 4
12	11	3 3'	39	64	2 3 4'6	59	85	2 2'3 4 4'	86	166	2 3 4 4'5 6	NQ	43	2 2'3 5
13	12	3 4	39	71	2 3'4'6	59	116	2 3 4 5 6	87	175	2 2'3 3'4 5'6	NQ	69	2 3'4 6
13	13	3 4'	40	68	2 3'4 5'	60	136	2 2'3 3'6 6'	88	182	2 2'3 4 4'5 6'	NQ	72	2 3'5 5'
14	15	4 4'	41	96	2 2'3 6 6'	61	77	3 3'4 4'	88	187	2 2'3 4'5 5'6	NQ	78	3 3'4 5
14	18	2 2'5	42	40	2 2'3 3'	61	110	2 3 3'4'6	89	128	2 2'3 3'4 4'	NQ	79	3 3'4 5'
15	17	2 2'4	43	57	2 3 3'5	62	154	2 2'4 4'5 6'	90	183	2 2'3 4 4'5'6	NQ	80	3 3'5 5'
16	24	2 3 6	43	103	2 2'4 5'6	63	82	2 2'3 3'4	91	167	2 3'4 4'5 5'	NQ	81	3 4 4'5
16	27	2 3'6	44	67	2 3'4 5	64	151	2 2'3 5 5'6	92	185	2 2'3 4 5'5 6	NQ	88	2 2'3 4 6
17	16	2 2'3	44	100	2 2'4 4'6	65	124	2 3'4'5 5'	93	174	2 2'3 3'4 5 6'	NQ	102	2 2'4 5 6'
17	32	2 4'6	45	58	2 3 3'5'	65	135	2 2'3 3'5 6'	93	181	2 2'3 4 4'5 6	NQ	113	2 3 3'5'6
18	23	2 3 5	45	63	2 3 4'5	66	144	2 2'3 4 5'6	94	177	2 2'3 3'4 5'6'	NQ	117	2 3 4'5 6
19	34	2 3'5'	46	74	2 4 4'5	67	107	2 3 3'4'5	95	156	2 3 3'4 4'5	NQ	120	2 3'4 5 5'
19	54	2 2'6 6'	46	94	2 2'3 5 6'	67	108	2 3 3'4 5'	95	171	2 2'3 3'4 4'6	NQ	121	2 3'4 5'6
20	29	2 4 5	47	61	2 3 4 5	67	147	2 2'3 4'5 6	96	202	2 2'3 3'5 5'6 6'	NQ	125	2 3'4'5'6
21	26	2 3'5	47	70	2 3'4'5	68	123	2 3 4 4'5'	97	157	2 3 3'4 4'5'	NQ	126	3 3'4 4'5
22	25	2 3'4	47	76	2 3'4'5'	69	106	2 3 3'4 5	98	173	2 2'3 3'4 5 6	NQ	127	3 3'4 5 5'
23	31	2 4'5	48	66	2 3'4 4'	69	118	2 3'4 4'5	99	200	2 2'3 3'4 5'6 6'	NQ	142	2 2'3 4 5 6
24	28	2 4 4'	48	93	2 2'3 5 6	69	149	2 2'3 4'5'6	99	204	2 2'3 4 4'5 6 6'	NQ	145	2 2'3 4 6 6'
24	50	2 2'4 6	48	95	2 2'3 5'6	70	139	2 2'3 4 4'6	100	172	2 2'3 3'4 5 5'	NQ	148	2 2'3 4'5 6'
25	21	2 3 4	49	55	2 3 3'4	70	140	2 2'3 4 4'6'	100	192	2 3 3'4 5 5'6	NQ	159	2 3 3'4 5'5'
25	33	2 3'4'	49	91	2 2'3 4'6	71	114	2 3 4 4'5	101	197	2 2'3 3'4 4'6 6'	NQ	160	2 3 3'4 5 6
25	53	2 2'5 6'	49	98	2 2'3 4'6'	71	134	2 2'3 3'5 6	102	180	2 2'3 4 4'5 5'	NQ	162	2 3 3'4'5 5'
26	22	2 3 4'	50	56	2 3 3'4'	71	143	2 2'3 4 5 6'	103	193	2 3 3'4'5 5'6	NQ	164	2 3 3'4'5'6
26	51	2 2'4 6'	50	60	2 3 4 4'	72	122	2 3 3'4'5'	104	191	2 3 3'4 4'5'6	NQ	165	2 3 3'5 5'6
27	45	2 2'3 6	51	84	2 2'3 3'6	72	131	2 2'3 3'4 6	105	199	2 2'3 3'4 5 6 6'	NQ	169	3 3'4 4'5 5'
28	36	3 3'5	51	92	2 2'3 5 5'	72	133	2 2'3 3'5 5'	106	170	2 2'3 3'4 4'5	NQ	184	2 2'3 4 4'6 6
29	46	2 2'3 6'	51	155	2 2'4 4'6 6'	73	146	2 2'3 4'5 5'	107	190	2 3 3'4 4'5 6	NQ	186	2 2'3 4 5 6 6'
30	39	3 4'5	52	89	2 2'3 4 6'	73	161	2 3 3'4 5'6	108	198	2 2'3 3'4 5 5'6	NQ	188	2 2'3 4'5 6 6'

Note: NQ = Not quantified in DB-1 method

Source: GE, 1997

Development of Corrections for Analytical Biases in the 1991-1997 GE Hudson River PCB Database. 1997. Prepared by HydroQual, Inc. for General Electric.

Table A5-2
GE PE sample results
PE 5 - Homolog Data

Method 680 Results

FIELD_SAMPLE_ID	PE ID	LAB	Total PCBs	Tri+ PCBs	% Tri+	MonoCB	DiCB	TriCB	TetraCB	PentaCB	HexaCB
RS1-8988-PE030827-E01	PE5	Lab 15	29.15	8.35	28.6%	12	8.8	4.3	3.5	0.55	
RS1-8988-PE030827-E02	PE5	Lab 15	25.74	7.74	30.1%	11	7.0	3.8	3.4	0.54	
RS1-8988-PE030827-E03	PE5	Lab 15	26.58	7.68	28.9%	11	7.9	4.2	3.3	0.18	
RS1-8988-PE030827-E04	PE5	Lab 15	32.90	10.20	31.0%	13	9.7	4.7	4.5	1.00	
RS1-9089-PE021030-C01	PE5	Lab 15	20.79	9.69	46.6%	2.9	8.2	4.6	4.1	0.99	
RS1-9291-PE021024-C02	PE5	Lab 15	30.52	11.52	37.7%	7	12	6.1	5.1	0.32	
RS1-9291-PE030625-C01	PE5	Lab 15	29.23	8.53	29.2%	12	8.7	4.2	3.6	0.73	
RS2-8483-PE030605-A02	PE5	Lab 15	20.15	8.85	43.9%	3.1	8.2	4.2	3.8	0.82	0.034
RS2-8887-PE030918-B01	PE5	Lab 15	37.23	11.23	30.2%	16	10	5.7	4.8	0.73	
RS2-8887-PE030918-C01	PE5	Lab 15	29.59	8.69	29.4%	12	8.9	4.3	3.7	0.69	
RS3-7170-PE030717-B01	PE5	Lab 15	28.62	7.72	27.0%	14	6.9	3.8	3.3	0.62	
RS3-7271-PE031009-A04	PE5	Lab 15	26.48	7.48	28.2%	11	8.0	4.0	3.2	0.28	
RS3-7271-PE031023-A01	PE5	Lab 15	24.14	6.94	28.7%	9.8	7.4	3.6	3.0	0.34	
RS3-7978-PE030806-C01	PE5	Lab 15	32.57	10.07	30.9%	13	9.5	4.6	4.5	0.97	
RS3-7978-PE030806-C02	PE5	Lab 15	28.34	7.74	27.3%	12	8.6	4.2	3.2	0.34	
		Mean	28.14	8.83	31.9%	10.65	8.65	4.42	3.80	0.61	
		Median	28.62	8.53	29.4%	12.00	8.60	4.20	3.60	0.62	
		Standard Deviation	4.471	1.410	6.0%	3.681	1.301	0.678	0.650	0.272	
		Relative Std Dev	0.159	0.160	0.189	0.345	0.150	0.153	0.171	0.449	

Method 680 Results-two samples (low MonoCB values) excluded

FIELD_SAMPLE_ID	PE ID	LAB	Total PCBs	Tri+ PCBs	% Tri+	MonoCB	DiCB	TriCB	TetraCB	PentaCB	HexaCB
RS1-8988-PE030827-E01	PE5	Lab 15	29.15	8.35	28.6%	12	8.8	4.3	3.5	0.55	
RS1-8988-PE030827-E02	PE5	Lab 15	25.74	7.74	30.1%	11	7.0	3.8	3.4	0.54	
RS1-8988-PE030827-E03	PE5	Lab 15	26.58	7.68	28.9%	11	7.9	4.2	3.3	0.18	
RS1-8988-PE030827-E04	PE5	Lab 15	32.90	10.20	31.0%	13	9.7	4.7	4.5	1.00	
RS1-9291-PE021024-C02	PE5	Lab 15	30.52	11.52	37.7%	7	12	6.1	5.1	0.32	
RS1-9291-PE030625-C01	PE5	Lab 15	29.23	8.53	29.2%	12	8.7	4.2	3.6	0.73	
RS2-8887-PE030918-B01	PE5	Lab 15	37.23	11.23	30.2%	16	10	5.7	4.8	0.73	
RS2-8887-PE030918-C01	PE5	Lab 15	29.59	8.69	29.4%	12	8.9	4.3	3.7	0.69	
RS3-7170-PE030717-B01	PE5	Lab 15	28.62	7.72	27.0%	14	6.9	3.8	3.3	0.62	
RS3-7271-PE031009-A04	PE5	Lab 15	26.48	7.48	28.2%	11	8.0	4.0	3.2	0.28	
RS3-7271-PE031023-A01	PE5	Lab 15	24.14	6.94	28.7%	9.8	7.4	3.6	3.0	0.34	
RS3-7978-PE030806-C01	PE5	Lab 15	32.57	10.07	30.9%	13	9.5	4.6	4.5	0.97	
RS3-7978-PE030806-C02	PE5	Lab 15	28.34	7.74	27.3%	12	8.6	4.2	3.2	0.34	
		Mean	29.31	8.76	29.8%	11.83	8.72	4.42	3.78	0.56	
		Median	29.15	8.35	29.2%	12.00	8.70	4.20	3.50	0.55	
		Standard Deviation	3.466	1.500	2.7%	2.130	1.391	0.728	0.697	0.261	
		Relative Std Dev	0.118	0.171	0.090	0.180	0.159	0.165	0.184	0.466	

Table A5-3
Summary of GE Method 680 LCS Recovery

LAB_SAMPLE_ID	Type	Method	Analysis Date	Unit	LCS Conc	LCS rec	GE Rec %	TAMS Rec %	Count
Year 1 LCS Samples									
RS1-9594-LABQCCOC02110052-AF10920L	LCS	GEHR680	11/9/2002	mg/Kg	1.36	0.93	68.4	68.4%	
RS1-9493-LABQCCOC02110055-AF10974L	LCS	GEHR680	11/14/2002	mg/Kg	1.24	1.10	88.7	88.7%	
RS1-9392-LABQCCOC02110056-AF10991L	LCS	GEHR680	11/15/2002	mg/Kg	0.82	0.53	64.6	64.6%	
RS1-9392-LABQCCOC02110054-AF10953L	LCS	GEHR680	11/17/2002	mg/Kg	1.24	1.10	88.7	88.7%	
RS1-9392-LABQCCOC02110088-AF11187L	LCS	GEHR680	11/19/2002	mg/Kg	0.82	0.61	74.0	74.4%	
RS1-9493-LABQCCOC02110089-AF11211L	LCS	GEHR680	11/21/2002	mg/Kg	1.24	1.00	80.6	80.6%	
RS1-9493-LABQCCOC02110091-AF11259L	LCS	GEHR680	11/21/2002	mg/Kg	1.24	0.87	70.2	70.2%	
RS1-9392-LABQCCOC02110090-AF11235L	LCS	GEHR680	11/22/2002	mg/Kg	1.24	1.20	96.8	96.8%	
RS1-9493-LABQCCOC02110106-AF11402L	LCS	GEHR680	11/29/2002	mg/Kg	1.33	1.20	90.2	90.2%	
RS1-9392-LABQCCOC02110109-AF11473L	LCS	GEHR680	12/1/2002	mg/Kg	1.24	1.20	96.8	96.8%	
RS1-9493-LABQCCOC02110128-AF11662L	LCS	GEHR680	12/1/2002	mg/Kg	1.21	0.90	74.4	74.4%	
RS2-8584-LABQCCOC02110129-AF11684L	LCS	GEHR680	12/2/2002	mg/Kg	1.2	0.91	75.8	75.8%	
RS2-8685-LABQCCOC02110107-AF11425L	LCS	GEHR680	12/2/2002	mg/Kg	1.3	1.10	84.6	84.6%	
RS1-9291-LABQCCOC02110104-AF11355L	LCS	GEHR680	12/4/2002	mg/Kg	1.24	0.99	79.8	79.8%	
RS1-9392-LABQCCOC02110110-AF11496L	LCS	GEHR680	12/4/2002	mg/Kg	1.24	0.78	62.9	62.9%	
RS2-8584-LABQCCOC02120012-AF12215L	LCS	GEHR680	12/4/2002	mg/Kg	1.35	1.00	74.1	74.1%	
RS1-9392-LABQCCOC02110111-AF11507L	LCS	GEHR680	12/6/2002	mg/Kg	1.24	1.10	88.7	88.7%	
RS1-9190-LABQCCOC02120014-AF12248L	LCS	GEHR680	12/7/2002	mg/Kg	1.24	1.60	129.0	129.0%	
RS2-8584-LABQCCOC02120022-AF12301L	LCS	GEHR680	12/8/2002	mg/Kg	1.33	1.20	90.2	90.2%	
RS2-8584-LABQCCOC02120023-AF12325L	LCS	GEHR680	12/10/2002	mg/Kg	1.23	1.20	97.6	97.6%	
RS2-8483-LABQCCOC02120024-AF12348L	LCS	GEHR680	12/11/2002	mg/Kg	1.3	1.20	92.3	92.3%	
RS1-9190-LABQCCOC02120025-AF12371L	LCS	GEHR680	12/13/2002	mg/Kg	1.24	1.10	88.7	88.7%	
RS1-9392-LABQCCOC02120026-AF12394L	LCS	GEHR680	12/13/2002	mg/Kg	1.32	1.00	75.8	75.8%	
RS1-9089-LABQCCOC02120028-AF12441L	LCS	GEHR680	12/14/2002	mg/Kg	1.24	1.00	80.6	80.6%	
RS2-8584-LABQCCOC02120027-AF12417L	LCS	GEHR680	12/14/2002	mg/Kg	1.29	1.20	93.0	93.0%	
RS1-9190-LABQCCOC02120029-AF12464L	LCS	GEHR680	12/16/2002	mg/Kg	1.24	1.20	96.8	96.8%	
RS1-9291-LABQCCOC02120030-AF12487L	LCS	GEHR680	12/16/2002	mg/Kg	1.24	0.92	74.2	74.2%	
RS2-8986-LABQCCOC03010033-AG00169L	LCS	GEHR680	1/15/2003	mg/Kg	1.34	1.00	74.6	74.6%	
Year 1 LCS - Median					1.24	1.05	82.6	82.6%	28
Year 2 LCS Analyses - First Half (4/5 through 9/18/03)									
RS1-9392-LABQCCOC03040003-AG01635L	LCS	GEHR680	4/5/2003	mg/Kg	1.32	0.85	64.4	64.4%	
RS2-8483-LABQCCOC03040004-AG01653L	LCS	GEHR680	4/7/2003	mg/Kg	1.33	1.20	90.0	90.2%	
RS2-8584-LABQCCOC03060235-AG06101L	LCS	GEHR680	6/26/2003	mg/Kg	1.36	0.97	71.3	71.3%	
RS2-8584-LABQCCOC03060236-AG06115L	LCS	GEHR680	7/1/2003	mg/Kg	1.24	0.95	76.6	76.6%	
RS2-8483-LABQCCOC03070063-AG07346L	LCS	GEHR680	7/11/2003	mg/Kg	1.24	0.88	71.0	71.0%	
RS1-9291-LABQCCOC03070064-AG07369L	LCS	GEHR680	7/13/2003	mg/Kg	1.33	1.10	82.7	82.7%	
RS1-9089-LABQCCOC03070065-AG07392L	LCS	GEHR680	7/14/2003	mg/Kg	1.24	0.90	72.6	72.6%	
RS1-8988-LABQCCOC03070129-AG07913L	LCS	GEHR680	7/21/2003	mg/Kg	1.32	1.00	80.0	75.8%	

Table A5-3
Summary of GE Method 680 LCS Recovery

LAB_SAMPLE_ID	Type	Method	Analysis Date	Unit	LCS Conc	LCS rec	GE Rec %	TAMS Rec %	Count
RS1-8988-LABQCCOC03070130-AG07937L	LCS	GEHR680	7/21/2003	mg/Kg	1.24	0.96	77.4	77.4%	
RS1-8988-LABQCCOC03070203-AG08680L	LCS	GEHR680	7/26/2003	mg/Kg	1.24	0.72	58.1	**58.1%**	
RS1-9089-LABQCCOC03070195-AG08554L	LCS	GEHR680	7/26/2003	mg/Kg	1.24	1.10	88.7	88.7%	
RS1-9190-LABQCCOC03070128-AG07890L	LCS	GEHR680	7/26/2003	mg/Kg	1.24	1.60	129.0	129.0%	
RS1-9190-LABQCCOC03070202-AG08656L	LCS	GEHR680	7/28/2003	mg/Kg	1.24	0.45	36.3	**36.3%**	
RS1-9291-LABQCCOC03070240-AG09069L	LCS	GEHR680	7/29/2003	mg/Kg	1.24	0.34	27.4	**27.4%**	
RS1-9291-LABQCCOC03070241-AG09092L	LCS	GEHR680	7/30/2003	mg/Kg	1.24	0.39	31.5	**31.5%**	
RS3-6766-LABQCCOC03070242-AG09116L	LCS	GEHR680	7/30/2003	mg/Kg	1.24	0.98	79.0	79.0%	
RS3-7877-LABQCCOC03070328-AG10030L	LCS	GEHR680	8/7/2003	mg/Kg	1.24	1.20	96.8	96.8%	
RS3-7978-LABQCCOC03070329-AG10053L	LCS	GEHR680	8/10/2003	mg/Kg	1.3	0.96	73.8	73.8%	
RS3-7069-LABQCCOC03080065-AG10744L	LCS	GEHR680	8/11/2003	mg/Kg	1.24	0.90	72.6	72.6%	
RS3-8281-LABQCCOC03080066-AG10768L	LCS	GEHR680	8/12/2003	mg/Kg	1.24	0.98	79.0	79.0%	
RS3-7473-LABQCCOC03080134-AG11428L	LCS	GEHR680	8/24/2003	mg/Kg	1.24	0.98	79.0	79.0%	
RS3-7776-LABQCCOC03080135-AG11452L	LCS	GEHR680	8/24/2003	mg/Kg	1.36	0.83	61.0	61.0%	
RS3-7069-LABQCCOC03080201-AG12173L	LCS	GEHR680	8/26/2003	mg/Kg	1.29	0.94	72.9	72.9%	
RS3-7069-LABQCCOC03080204-AG12207L	LCS	GEHR680	8/27/2003	mg/Kg	1.24	0.74	59.7	**59.7%**	
RS3-7170-LABQCCOC03080136-AG11475L	LCS	GEHR680	8/27/2003	mg/Kg	1.24	1.00	80.6	80.6%	
RS3-5958-LABQCCOC03080205-AG12230L	LCS	GEHR680	8/28/2003	mg/Kg	1.24	0.90	72.6	72.6%	
RS3-6463-LABQCCOC03080241-AG12687L	LCS	GEHR680	9/2/2003	mg/Kg	1.24	0.98	79.0	79.0%	
RS3-6968-LABQCCOC03080242-AG12711L	LCS	GEHR680	9/5/2003	mg/Kg	1.24	1.10	88.7	88.7%	
RS3-5958-LABQCCOC03080243-AG12734L	LCS	GEHR680	9/6/2003	mg/Kg	1.24	0.95	76.6	76.6%	
RS3-6766-LABQCCOC03090032-AG13156L	LCS	GEHR680	9/9/2003	mg/Kg	1.24	0.97	78.2	78.2%	
RS3-7069-LABQCCOC03090031-AG13133L	LCS	GEHR680	9/9/2003	mg/Kg	1.35	1.10	81.5	81.5%	
RS3-6766-LABQCCOC03090033-AG13180L	LCS	GEHR680	9/13/2003	mg/Kg	1.24	0.98	79.0	79.0%	
RS3-7069-LABQCCOC03090082-AG13789L	LCS	GEHR680	9/15/2003	mg/Kg	1.24	0.99	79.8	79.8%	
RS3-6463-LABQCCOC03090083-AG13813L	LCS	GEHR680	9/18/2003	mg/Kg	1.24	0.97	78.2	78.2%	
Year 2 - First half - Median								77.0%	34
Year 2 LCS Analyses - Second Half (9/1/9 through 11/12/03)									
RS3-6766-LABQCCOC03090084-AG13837L	LCS	GEHR680	9/19/2003	mg/Kg	1.24	0.80	64.5	64.5%	
RS2-8685-LABQCCOC03090138-AG14571L	LCS	GEHR680	9/22/2003	mg/Kg	1.36	0.88	64.7	64.7%	
RS2-8685-LABQCCOC03090139-AG14596L	LCS	GEHR680	9/26/2003	mg/Kg	0.99	0.72	72.7	72.7%	
RS2-8382-LABQCCOC03090140-AG14620L	LCS	GEHR680	9/30/2003	mg/Kg	1.24	0.93	75.0	75.0%	
RS2-8685-LABQCCOC03090185-AG15145L	LCS	GEHR680	10/5/2003	mg/Kg	1.35	1.00	74.1	74.1%	
RS1-9291-LABQCCOC03090186-AG15168L	LCS	GEHR680	10/7/2003	mg/Kg	1.35	0.84	62.2	62.2%	
RS3-6766-LABQCCOC03090187-AG15190L	LCS	GEHR680	10/9/2003	mg/Kg	1.24	0.86	69.4	69.4%	
RS3-7170-LABQCCOC03100026-AG15968L	LCS	GEHR680	10/11/2003	mg/Kg	1.24	0.86	69.4	69.4%	
RS3-7372-LABQCCOC03100027-AG15992L	LCS	GEHR680	10/11/2003	mg/Kg	1.32	0.96	72.7	72.7%	
RS3-6059-LABQCCOC03100028-AG16015L	LCS	GEHR680	10/14/2003	mg/Kg	1.24	0.89	71.8	71.8%	
RS1-9291-LABQCCOC03100072-AG16535L	LCS	GEHR680	10/21/2003	mg/Kg	1.32	0.84	63.6	63.6%	

Table A5-3

Summary of GE Method 680 LCS Recovery

LAB_SAMPLE_ID	Type	Method	Analysis Date	Unit	LCS Conc	LCS rec	GE Rec %	TAMS Rec %	Count
RS2-8483-LABQCCOC03100144-AG17460L	LCS	GEHR680	10/24/2003	mg/Kg	1.3	0.79	60.8	60.8%	
RS2-8887-LABQCCOC03100143-AG17436L	LCS	GEHR680	10/24/2003	mg/Kg	1.24	0.89	71.8	71.8%	
RS2-8786-LABQCCOC03100145-AG17484L	LCS	GEHR680	10/26/2003	mg/Kg	1.24	0.80	64.5	64.5%	
RS2-8887-LABQCCOC03100176-AG17870L	LCS	GEHR680	10/29/2003	mg/Kg	1.3	0.63	48.5	**48.5%**	
RS2-8786-LABQCCOC03100177-AG17894L	LCS	GEHR680	11/8/2003	mg/Kg	1.24	0.59	47.6	**47.6%**	
RS3-6463-LABQCCOC03110017-AG18506L	LCS	GEHR680	11/12/2003	mg/Kg	1.24	0.67	54.0	**54.0%**	
Year 2 - second half only - Median								64.7%	17
Year 2 - complete - Median					1.24	0.93	72.7	72.7%	79
								TAMS/Earthtech	

Table A5-4
GE PE 5 Samples
Aroclor Data

LAB_SAMPLE_ID	PE ID	LAB	Aroclor Concentrations 1221	1242	1254	Total PCB	Fraction 1221	Calculated Tri+ PCBs	Literature Tri+
RS1-8988-PE030827-N46939-21	PE5	Lab 1	27.7	11.2		38.9	0.712	0.355	0.316
RS1-9089-PE021029-N25559-22	PE5	Lab 1	27.2	11.6		38.8	0.701	0.368	0.324
RS1-9291-PE021022-N25115-17	PE5	Lab 1	39.9	12.3		52.2	0.764	0.296	0.277
RS1-9291-PE030625-N42266-9	PE5	Lab 1	23.9	9.78		33.7	0.710	0.358	0.318
RS1-9594-PE021015-N24646-4	PE5	Lab 1	32.8	11.4		44.2	0.742	0.321	0.293
RS2-8483-PE030605-N40750-1	PE5	Lab 1	27.9	13.1		41	0.680	0.391	0.340
RS2-8887-PE030918-N48618-16	PE5	Lab 1	17.2	6.64		23.8	0.721	0.345	0.309
RS3-7170-PE030717-N43957-13	PE5	Lab 1	26.1	11		37.1	0.704	0.365	0.322
RS3-7271-PE031009-N50386-4	PE5	Lab 1	20.6	8.16		28.8	0.716	0.351	0.313
RS3-7271-PE031023-N51499-1	PE5	Lab 1	26.2	10.8		37	0.708	0.360	0.319
RS3-7978-PE030806-N45396-1	PE5	Lab 1	26.5	10.6		37.1	0.714	0.353	0.314
RS1-8988-PE030827-4111286	PE5	Lab 14	31	12		43	0.721	0.345	0.309
RS1-9190-PE030605-4057651	PE5	Lab 14	24	12		36	0.667	0.407	0.350
RS1-9291-PE021101-3932816	PE5	Lab 14	21	9.6		30.6	0.686	0.385	0.335
RS1-9291-PE030625-4072109	PE5	Lab 14	28	10		38	0.737	0.327	0.297
RS1-9392-PE021018-3922707	PE5	Lab 14	36	14		50	0.720	0.346	0.310
RS1-9493-PE021011-3917819	PE5	Lab 14	43	17		60	0.717	0.350	0.313
RS2-8483-PE021025-3928521	PE5	Lab 14	29	13		42	0.690	0.380	0.332
RS2-8887-PE030918-4125239	PE5	Lab 14	31	12		43	0.721	0.345	0.309
RS3-6766-PE030806-4096032	PE5	Lab 14	26	10		36	0.722	0.344	0.308
RS3-7170-PE030717-4084503	PE5	Lab 14	33	13		46	0.717	0.349	0.312
RS3-7271-PE031009-4142257	PE5	Lab 14	27	11		38	0.711	0.357	0.317
RS3-7271-PE031023-4150066	PE5	Lab 14	27	11		38	0.711	0.357	0.317
RS1-8988-PE030827-AG12595	PE5	Lab 15	33	11		44	0.750	0.313	0.288
RS1-9089-PE030606-AG04382	PE5	Lab 15	23	11		34	0.676	0.396	0.343
RS1-9291-PE021101-AF10570	PE5	Lab 15	29	10		39	0.744	0.320	0.292
RS1-9291-PE030625-AG06003	PE5	Lab 15	33	11		44	0.750	0.313	0.288
RS1-9392-PE021018-AF09487	PE5	Lab 15	35	12		47	0.745	0.319	0.291
RS1-9493-PE021004-AF08274	PE5	Lab 15	28	10		38	0.737	0.327	0.297
RS1-9493-PE021011-AF08745	PE5	Lab 15	31	11		42	0.738	0.326	0.296
RS1-9493-PE030317-AG01181	PE5	Lab 15	29	10		39	0.744	0.320	0.292
RS2-8483-PE021025-AF09981	PE5	Lab 15	27	10		37	0.730	0.335	0.303
RS2-8887-PE030918-AG14534	PE5	Lab 15	30	10		40	0.750	0.313	0.288
RS2-8986-PE030417-AG02189	PE5	Lab 15	25	11		36	0.694	0.375	0.329
RS3-7170-PE030717-AG08613	PE5	Lab 15	31	10		41	0.756	0.306	0.283
RS3-7271-PE031009-AG16592	PE5	Lab 15	30	11		41	0.732	0.333	0.301
RS3-7271-PE031023-AG17951	PE5	Lab 15	30	11		41	0.732	0.333	0.301
RS3-7877-PE031118-AG19392	PE5	Lab 15	26	10		36	0.722	0.344	0.308
RS3-7978-PE030806-AG10660	PE5	Lab 15	33	12		45	0.733	0.331	0.300
RS1-9291-PE021024-386099	PE5	Lab 16	28	12		40	0.700	0.369	0.325
RS2-8584-PE021031-387723	PE5	Lab 16	21	9.2		30.2	0.695	0.374	0.328
RS1-8988-PE030827-C3H280124022	PE5	Lab 6	27	11		38	0.711	0.357	0.317
RS1-9089-PE021028-C2J290172016	PE5	Lab 6	22	13	1.7	36.7	0.599	0.483	0.407
RS1-9291-PE030625-C3F260341011	PE5	Lab 6	32	13	1.3	46.3	0.691	0.379	0.336
RS1-9392-PE021021-C2J220181016	PE5	Lab 6	28	12		40	0.700	0.369	0.325
RS1-9493-PE021007-C2J080284011	PE5	Lab 6	28	14		42	0.667	0.407	0.350
RS1-9594-PE021014-C2J150280012	PE5	Lab 6	38	15		53	0.717	0.350	0.312
RS2-8483-PE030605-C3F060207001	PE5	Lab 6	28	14	0.95	43.0	0.652	0.423	0.364
RS2-8887-PE030918-C3I190390017	PE5	Lab 6	28	9.5		37.5	0.747	0.316	0.290
RS3-7271-PE031009-C3J110201011	PE5	Lab 6	17	8.9		25.9	0.656	0.418	0.358
RS3-7271-PE031023-C3J240278001	PE5	Lab 6	21	9		30	0.700	0.369	0.325
RS3-7776-PE030717-C3G180160009	PE5	Lab 6	33	14		47	0.702	0.367	0.323
RS3-7877-PE030806-C3H070274004	PE5	Lab 6	31	13		44	0.705	0.364	0.322
		Mean	28.43	11.32	1.32	39.83	0.713	0.35	0.32
		Median	28.00	11.00	1.30	39.00	0.716	0.35	0.31
		Standard Deviation	5.110	1.808	0.375	6.520	0.030	0.034	0.023
		Relative Std Dev	0.182	0.164	0.289	0.167	0.042	0.097	0.074

Lab 4 Data (excluded)

RS1-9089-PE021030-TA2J0P830018	PE5	Lab 4	29	15		44	0.659	15.65
RS1-9392-PE021016-TA2J0P496018	PE5	Lab 4	15	6.5		21.5	0.698	7.025
RS1-9594-PE021009-TA2J0P300004	PE5	Lab 4	19	9.3		28.3	0.671	9.805
RS2-8483-PE021023-TA2J0P679001	PE5	Lab 4	24	12		36	0.667	12.6

"Calculated" Tri + based on (.03*Ar1221+(1.16*(Ar1242+1254))/(Ar1221+1242+1254)
Literature Tri+ based on (0.1 *Ar1221)+(.85*Ar1242)+(.99*Ar1254)

Table A5-5
GE PE 2 Samples
Aroclor Data

LAB_SAMPLE_ID	PE ID	LAB	Aroclor Concentrations 1221	1242	1254	Total PCB	Fraction 1221	Calculated Tri+ PCBs	Literature Tri +
RS1-9089-PE021030-N25643-21	PE2	Lab 1	17.6	8.21		25.81	0.682	0.389	0.339
RS1-9392-PE021016-N24752-12	PE2	Lab 1	11.1	5.8		16.90	0.657	0.418	0.357
RS1-9493-PE021009-N24212-10	PE2	Lab 1	14	7.41	0.618	22.03	0.636	0.442	0.377
RS2-8584-PE030529-N40160-1	PE2	Lab 1	11.2	5.58		16.78	0.667	0.406	0.349
RS3-6968-PE030730-N44844-11	PE2	Lab 1	11.1	5.62	0.577	17.30	0.642	0.435	0.373
RS3-7978-PE031002-N49731-1	PE2	Lab 1	9.82	5.5		15.32	0.641	0.436	0.369
RS1-9089-PE021029-3929926	PE2	Lab 14	9.8	5.3		15.10	0.649	0.427	0.363
RS1-9392-PE021015-3919355	PE2	Lab 14	9.4	5.3		14.70	0.639	0.437	0.370
RS1-9392-PE021022-3924871	PE2	Lab 14	11	5.8		16.80	0.655	0.420	0.359
RS1-9493-PE021008-3914767	PE2	Lab 14	11	6.5		17.50	0.629	0.450	0.379
RS2-8584-PE030529-4053262	PE2	Lab 14	11	6.5		17.50	0.629	0.450	0.379
RS3-7069-PE030730-4091703	PE2	Lab 14	11	6.5		17.50	0.629	0.450	0.379
RS3-7978-PE031002-4136357	PE2	Lab 14	9	5.6		14.60	0.616	0.463	0.388
RS1-9190-PE021028-AF10153	PE2	Lab 15	13	5.2		18.20	0.714	0.353	0.314
RS1-9392-PE021021-AF09641	PE2	Lab 15	15	7.5		22.50	0.667	0.407	0.350
RS1-9392-PE030317-AG01052	PE2	Lab 15	14	5.8		19.80	0.707	0.361	0.320
RS1-9392-PE030529-AG03661	PE2	Lab 15	12	5.4		17.40	0.690	0.381	0.333
RS1-9493-PE021007-AF08331	PE2	Lab 15	11	5.2		16.20	0.679	0.393	0.341
RS1-9594-PE021014-AF08824	PE2	Lab 15	12	5.6		17.60	0.682	0.390	0.339
RS2-8483-PE030317-AG01263	PE2	Lab 15	13	5.3		18.30	0.710	0.357	0.317
RS2-8986-PE021211-AF12606	PE2	Lab 15	15	6.1		21.10	0.711	0.357	0.317
RS2-8986-PE021218-AF12814	PE2	Lab 15	15	6.3		21.30	0.704	0.364	0.322
RS3-6968-PE030730-AG09977	PE2	Lab 15	12	6		18.00	0.667	0.407	0.350
RS3-7877-PE031211-AG19678	PE2	Lab 15	11	5		16.00	0.688	0.383	0.334
RS3-7978-PE031002-AG15884	PE2	Lab 15	13	5.8		18.80	0.691	0.379	0.331
RS1-9190-PE021025-386567	PE2	Lab 16	11	5.5	0.69	17.19	0.640	0.437	0.376
RS1-9392-PE021101-388327	PE2	Lab 16	8.5	4	0.48	12.98	0.655	0.420	0.364
RS1-9493-PE021011-383281	PE2	Lab 16	11	5.1	0.58	16.68	0.659	0.415	0.360
RS1-9392-PE021017-C2J180225018	PE2	Lab 6	9.6	6.2	0.7	16.50	0.582	0.503	0.420
RS1-9392-PE021031-C2K010155013	PE2	Lab 6	15	7.7		22.70	0.661	0.413	0.354
RS1-9392-PE030529-C3E300201007	PE2	Lab 6	12	6.6	0.74	19.34	0.620	0.459	0.390
RS1-9493-PE021010-C2J110279021	PE2	Lab 6	12	6.3		18.30	0.656	0.419	0.358
RS2-8584-PE021024-C2J250174001	PE2	Lab 6	9.6	6.5	0.71	16.81	0.571	0.515	0.428
RS3-6766-PE030730-C3G310148023	PE2	Lab 6	12	8.3	0.73	21.03	0.571	0.515	0.427
RS3-7877-PE031002-C3J030129006	PE2	Lab 6	9	5.5	0.68	15.18	0.593	0.490	0.412
		Mean	11.79	6.01	0.65	17.99	0.65	0.42	0.36
		Median	11.10	5.80	0.69	17.50	0.66	0.42	0.36
		Standard Deviation	2.067	0.929	0.084	2.687	0.038	0.043	0.030
		Relative Std Dev	0.186	0.160	0.122	0.154	0.059	0.104	0.085

Lab 4 Data (excluded)

RS1-9190-PE021028-TA2J0P768018	PE2	Lab 4	14	8.7		22.70	0.617	0.463
RS1-9392-PE021021-TA2J0P599010	PE2	Lab 4	9	4.6		13.60	0.662	0.412
RS1-9594-PE021007-TA2J0P236006	PE2	Lab 4	8.7	4.9		13.60	0.640	0.437
RS1-9594-PE021014-TA2J0P418015	PE2	Lab 4	8	4.8		12.80	0.625	0.454

"Calculated" Tri + based on (.03*Ar1221+(1.16*(Ar1242+1254))/(Ar1221+1242+1254)
Literature Tri+ based on (0.1 *Ar1221)+(.85*Ar1242)+(.99*Ar1254)

Table A5-6
Field Duplicate RPDs

	Aroclors						Homologs		
	> 5 mg/kg			1 - 4.9 mg/kg			All Data		
	Total PCB	Ar1221	% Ar1221	Total PCB	Ar1221	% Ar1221	Total PCB	Tri+ PCB	% Tri +
Number of pairs	723	720	720	259	259	259	79	79	79
95th percentile RPD	88	97.7	20.1	115	129	30.5	90	94	25
90th percentile RPD	60	69.8	13.4	78	84.4	21	60	65	15
80th percentile RPD	39	42.9	8.0	52.7	57.6	13.5	37	43	11
Median (50th) RPD	16.3	17.5	2.8	22.2	25.4	5.3	16.8	16.8	5.7
Percent RPD <40	81	77.6	99	70.7	64.3	96.1	82.3	78.5	99

Table A5-7

Critical Tri+ PCBs Concentration Range

		6-in Segment	22-in Segment
River Section 1 - MPA Threshold 3 g/ m^2			
Finer Grained Areas			
Bulk Density	g/cc	0.85	0.85
Tri+ PCB needed to exceed MPA criteria	ppm	23	6.4
Coarser Grained Areas			
Bulk Density	g/cc	1.8	1.8
Tri+ PCB needed to exceed MPA criteria	ppm	11	***3***
River Sections 2 & 3 - MPA Threshold 10 g/ m^2			
Finer Grained Areas			
Bulk Density	g/cc	0.85	0.85
Tri+ PCB needed to exceed MPA criteria	ppm	***77***	21

Notes:

1. The MPA is the product of the Tri+ PCB concentration, length and bulk density. The concentration in finer-grained areas must be higher to have the same MPA as a coarse grained area. The table above shows the range of critical concentrations.
2. In River Section 1, the MPA threshold is 3 g/m^2. For coarse-grained areas of River Section 1 and thicker segments, the concentration would only need to be 3 ppm or higher to exceed the threshold.
3. In River Sections 2 and 3, the MPA threshold is 10 g/m^2. For finer-grained areas and shorter segments, the concentration would need to be 77 ppm or higher to exceed the threshold. 80 ppm or higher data tend to be widely scattered, but are likely to result in exceedance of the removal criteria unless there is a pronounced bias in the predicted values.

Table A5-8
Lower and Upper Limits for Predicted Sediment Tri+ PCB using the Point-by-Point Correction Regression Model

Corrected Tri+ PCB Range (mg/kg)	Lower Limit (%)	Upper Limit (%)	
0-5	NA	NA	
5-80	-29	38	
80+	-44	41	

Model:

1) Tri+ Concentration = a[A21]+b[A42+A54]

Table A5-9
Statistics of the 2009 Regression Models

Statistic	Model using original 180	Model using full data set
A1221 Coefficient	0.126	0.140
Std. Error of A1221 Coefficient	0.0114	0.0105
A1242+A1254 Coefficient	0.944	0.913
Std. Error of A1242+A1254 Coefficient	0.0293	0.0253
Multiple R^2	0.970	0.961

Table A5-10
Regression Coefficients Update

Round	Regression Date	Number of Samples		Tri+ PCB = a[A1221] + b[A1242 + A1254]	
		Additional	Total	a	b
Original	In QAPP 2009	---	278	0.14	0.91
2009 Round 1	8/14/2009	21	299	0.14	0.91
2009 Round 2	9/3/2009	63	362	0.13	0.90
2009 Round 3	9/17/2009	21	383	0.13	0.90
2009 Round 4	10/26/2009	20	403	0.13	0.89
2011	7/29/2011	42	445	0.13	0.89

Table A5-11
Statistics of the 2009 and 2011 Regression Models

Statistic	Original	2009				2011
		Round 1	Round 2	Round 3	Round 4	
A1221 Coefficient	0.140	0.140	0.134	0.128	0.127	0.125
Std. Error of A1221 Coefficient	0.0105	0.0099	0.0083	0.0079	0.0076	0.0071
A1242+A1254 Coefficient	0.913	0.911	0.893	0.895	0.892	0.886
Std. Error of A1242+A1254 Coefficient	0.0253	0.0235	0.0209	0.0203	0.0195	0.0183
Multiple R2	0.961	0.962	0.959	0.959	0.959	0.960

Table A5-12
Feature Selection Measures for Several OLS Linear Regression Models for Water Data

Models	Linear Regression Formula	LOOCV - PRESS	AIC	AICc	BIC	Adjusted R^2
GE Model	Tri+PCB = a × A1221 + b × (A1242+A1254)	100	791	791	801	0.96
Test Model 1	Tri+PCB = a × A1221 + b × A1242+ c × A1254	101	1101	1102	1115	0.95
Test Model 2	Tri+PCB = a × (A1221+A1242+A1254)	188	1252	1252	1258	0.91

Note:
LOOCV: leave-one-out cross-validation
PRESS: prediction residual sum of squares
AIC: Akaike’s Information Criterion
AICc: corrected AIC
BIC: Schwarz’s Bayesian Information Criteria

Table A5-13
List of OLS and Several Robust Regression Models of Water Data

Method Acronym	Method Name	R function
OLS	Ordinary Least Square Regression	lm()
LAV	Least Absolute Values Regression	rq() in package quantreg
M-Huber	M-Estimation with Huber Weight	rlm() in package MASS
M-Bisquare	M-Estimation with Bisquare Weight	rlm(,psi = psi.bisquare) in package MASS
LTS	Least Trimmed Squares Regression	ltsreg() in package lqs
LMS	Least Median Squares Regression	lmsreg() in package lqs

Table A5-14
Regression Results for OLS and Several Robust Regression Models for Water Data

	Group 1 (n=68) [1]	Group 2 (n=136) [2]		Group 3 (n=35) [3]		All Samples (n=239) [4]	
Model	**a**	**a**	**b**	**a**	**b**	**a**	**b**
OLS	0.236	0.126	0.757	0.191	0.696	0.143	0.749
LAV	0.247	0.129	0.709	0.188	0.721	0.146	0.694
M-Huber	0.240	0.128	0.722	0.185	0.721	0.144	0.721
M-Bisquare	0.245	0.128	0.708	0.185	0.720	0.145	0.699
LTS	0.258	0.168	0.384	0.185	0.749	0.223	-0.037
LMS	0.274	0.134	0.612	0.188	0.721	0.227	-0.021

Note: all the regression analysis was based on the formula of Tri+PCB = a × A1221 + b × (A1242+A1254)

(1) Group 1 corresponds to the first correlation in GE’s model.
(2) Group 2 corresponds to the second correlation in GE’s model.
(3) Group 3 corresponds to the third correlation in GE’s model.
(4) Models were developed using all available samples.

Table A5-15
Cross Validation Results and the Model Prediction Error of Water Data

Model		Group 1 (n=68) [1]	Group 2 (n=136) [2]	Group 3 (n=35) [3]	All Samples (n=239) [4]
OLS	CV (ng/L)	4.912	**7.160**	16.343	**9.988**
LAV	CV (ng/L)	4.924	7.363	**15.912**	9.992
M-Huber	CV (ng/L)	**4.897**	7.167	16.297	10.011
M-Bisquare	CV (ng/L)	4.915	7.195	16.306	10.048
LTS	CV (ng/L)	5.060	9.064	15.952	17.583
LMS	CV (ng/L)	5.738	7.668	16.440	16.625
Mean of Measured Tri+ PCB (ng/L)		18	41	68	38
Mean of Prediction Error for the Best Model (ng/L)		4.897	7.160	15.912	9.988
Relative Prediction Error for the Best Model		27%	18%	23%	26%

Note: CV-mean of the estimated prediction errors.

(1) Group 1 corresponds to the first correlation in GE's model.
(2) Group 2 corresponds to the second correlation in GE's model.
(3) Group 3 corresponds to the third correlation in GE's model.
(4) Models were developed using all available samples.

Table A5-16
Water Column Data Model Validation Results using Test Data Set

Model		Group 1 (n=12)	Group 2 (n=23)	All Samples (n=35)
GE Model	Average prediction error	2.447	5.085	NA
OLS	Average prediction error	3.174	5.337	5.761
LAV	Average prediction error	2.892	5.154	**5.545**
M-Huber	Average prediction error	3.054	5.192	5.638
M-Bisquare	Average prediction error	2.944	5.144	5.555
LTS	Average prediction error	2.630	**4.661**	6.069
LMS	Average prediction error	**2.396**	5.413	6.608
Mean of Measured Tri+ PCB (ng/L)		10	34	26
Mean of Prediction Error for the Best Model (ng/L)		2.396	4.661	5.545
Relative Prediction Error for the Best Model		24%	14%	22%

Table A5-17

Regression Equations for EPA's Best Model.

Model	First Regression	Second Regression	Third Regression
EPA's Best Model [1]	0.274×A1221 (LMS Model)	0.126×A1221+0.757×(A1242+A1254) (OLS Model)	0.188×A1221+0.721×(A1242+A1254) (LAV Model)
GE weighted least squares (i.e., "damped-leveraged) algorithm	0.27×A1221	0.12×A1221+0.73×(A1242+A1254)	0.16×A1221+0.85×(A1242+A1254)

Note: (1) Best Model was selected among OLS and several robust regression models discussed in this work (LAV, M-Huber, M-Bisquare, LTS, LMS)

Table A5-18

Aroclors Reported By Year as an Average Percentage of $TPCB_{Aroclor}$ – NYSDEC Data

	Aroclors Reported							
Year	**1016**	**1221**	**1242**	**1248**	**1254**	**1260**	**1254/ 1260**	**Laboratory as reported in NYSDEC Database**
1990	53%						47%	Hale Creek
	28%	0.2%			72%			Hazleton Laboratories
1991	53%				10%		36%	Dept. of Health
	48%				17%		35%	Hale Creek
1992	54%				44%	0.3%	2%	Hale Creek
			4%	12%	74%	11%		Hazleton Laboratories
1993	58%					42%		Hale Creek
			3%	32%	53%	12%		Hazleton Laboratories
1994			2%	27%	59%	12%		Hazleton Laboratories
1995				29%	53%	17%		Hazleton Laboratories
1996	86%						14%	Hale Creek
				31%	54%	16%		Hazleton Laboratories
1997	0.004%			31%	46%	23%		EnChem Environmental Laboratories
				5%	70%	25%		Hazleton Laboratories
1998				41%	38%	21%		EnChem Environmental Laboratories
1999			21%	30%	28%	21%		Mississippi St. Chem. Laboratories
2000			19%	27%	28%	25%		Mississippi St. Chem. Laboratories
2001			18%	23%	30%	29%		Mississippi St. Chem. Laboratories
2002			20%	24%	33%	23%		Mississippi St. Chem. Laboratories
2003			22%	26%	32%	20%		Mississippi St. Chem. Laboratories
2004			21%	24%	35%	21%		Mississippi St. Chem. Laboratories
2005			10%	14%	41%	35%		Mississippi St. Chem. Laboratories
2006			10%	13%	42%	35%		Mississippi St. Chem. Laboratories
2007			13%	18%	40%	29%		Mississippi St. Chem. Laboratories
2008			5%	16%	44%	34%		Mississippi St. Chem. Laboratories
2009			12%	21%	32%	35%		Mississippi St. Chem. Laboratories
2010			10%	20%	37%	32%		Mississippi St. Chem. Laboratories
2011			12%	18%	38%	31%		Mississippi St. Chem. Laboratories
2013			42%				58%	Hale Creek
2014			44%				56%	Hale Creek
2015			43%				57%	Hale Creek

Note, for example, the different mixtures reported by different laboratories in the same year of collection. Additionally, Aroclor 1016 and an unresolved mixture of Aroclors 1254/1260 are reported only in 1996 and prior (primarily by the Hale Creek lab), while Aroclor 1242 appears very infrequently among all labs through 1998. Aroclor 1221, which can be inferred to indicate the presence of monochloro and dichloro homologues in the samples, was only reported in 1990 by Hazleton Laboratories.[1]

[1] The lack of reported Aroclor 1221 in these data are further evidence of the lack of monochloro and dichlorohomologues in fish tissue, as reported in the various documents supporting the EPA ROD.

Table A5-19

Aroclors Reported by Years as an Average Percentage of $TPCB_{Aroclor}$ - GE Data

	Aroclors Reported					
Year	**1221**	**1242**	**1248**	**1254**	**1260**	**Laboratory**
2004	1.5%	0.66%	57%	38%	3.1%	Northeast Analytical Inc.
2005	1.1%	0.38%	56%	40%	2.7%	
2006	0.38%	0.32%	56%	38%	4.8%	
2007	0.49%	0.32%	58%	36%	4.9%	
2008	0.74%	0.23%	60%	38%	1.4%	
2009	4.3%	1.7%	54%	38%	2.3%	
2010	3.3%	1.88%	46%	44%	4.8%	
2011	12%	4.2%	45%	35%	3.5%	
2012	20%	0.15%	43%	31%	6.6%	
2013	20%	0.56%	44%	27%	8.7%	
2014	16%	0.094%	47%	30%	6.8%	
2015	12%	0.14%	44%	32%	11%	
2016	3.9%	0%	50%	34%	12%	

Notes:

1. Highlighted area represents the remedial dredging period and a change in Aroclors as reported by GE

2. Numbers are rounded to 2 significant figures. Therefore, the sum of Aroclors percentage may not be exactly 100%.

Table A5-20
Fish Tissue Regression Equations

Data Source	Period of Available Data	Applicable Laboratory Codes	Equation to Obtain the Homologue Equivalent Total PCB Concentration ($TPCB_{HE}$)	Equation Source	Period of Application
NYSDEC	1990	HES, U	1.3070 * (Aroclor 1016 + Aroclor 1254)	USEPA. 2000. Further Site Characterization and Analysis, Revised Baseline Modeling Report (RBMR), Hudson River PCBs Reassessment RI/FS Volume 2D. Prepared for USEPA Region 2 and USACE, Kansas City District by TAMS Consultants, Inc., Limno-Tech, Inc., Menzie-Cura & Associates, Inc., and Tetra-Tech, Inc. January 2000.	1990
	1990-1992, 1996	HC, DOH	1.4157 * (Aroclor 1016+ Aroclor 1254/1260)		1990-1992, 1996
	1992-1997	ENC, HES	0.8754 * (Aroclor 1248 + Aroclor 1254 + Aroclor 1260)		1992-1997
	1998	ENC	0.7407 * $TPCB_{Aroclor}$	Geometric mean of $TPCB_{HE}/TPCB_{Aroclor}$ 1998 NYSDEC. See Figure A5-11.	1998
	1999-2000	MSC	1.1743 * $TPCB_{Aroclor}$	Geometric mean of $TPCB_{HE}/TPCB_{Aroclor}$ 1999-2000 NYSDEC. See Figure A5-15.	1999-2015
GE	2004-2008	NEA	0.8542 * $TPCB_{Aroclor}$	Geometric mean of $TPCB_{HE}/TPCB_{Aroclor}$ 2004-2008 GE data. See Figure A5-19.	2004-2008, 2016
	2009-2013	NEA	0.788 * $TPCB_{Aroclor}$	Geometric mean of $TPCB_{HE}/TPCB_{Aroclor}$ 2009-2013 GE data. See Figure A5-23.	2009-2015

Notes:

$TPCB_{Aroclor}$ refers to the sum of detected Aroclor concentrations in the sample. See text for discussion.

Lab Code Key

DOH Dept. of Health
ENC Enchem Environmental Laboratories
HES Hazelton Laboratories
HC Hale Creek
MSC Mississippi St. Chem. Laboratories
NEA Northeast Analyical Laboratories
U Unknown

Table A5-21
Coefficient (α) Summary Statistics

Data Source	Period	Statistics for Determination of α								Confidence Interval for the Geometric Mean (Bootstrap Analysis)				Geometric Mean 95% CI expressed as a Percentage of the Mean	
		Number of Samples	Arithmetic Mean	Std. Dev.	Median	Geometric Mean	Regression Estimate	Minimum Observed	Maximum Observed	2.5%	5% [1]	95% [1]	97.5%		
NYSDEC	1998	82	0.826	0.535	0.783 *	0.741	0.784 *	0.129	4.840	0.672	0.684	0.802	0.813	-7.7%	8.2%
NYSDEC	1999-2000	173	1.284	0.637	1.181 *	1.174	0.871	0.285	5.996	1.104	1.114	1.236	1.248	-5.1%	5.3%
GE	2004-2008	259	0.893	0.334	0.831 *	0.854	0.703	0.151	4.210	0.826	0.830	0.880	0.885	-2.8%	3.0%
GE	2009-2013	140	0.823 **	0.244	0.784 *	0.788	0.851	0.101	2.226	0.746	0.752	0.822	0.828	-4.5%	4.3%

Note:

1 Value used in estimated 95% CI as percentage of the geometric mean.
* Value falls between 5th and 95th percent confidence interval for the geometric mean.
** Value falls between 2.5th and 97.5th percent confidence interval for the geometric mean.

Table A5-22
Average Aroclor 1221 as Percentage of $TPCB_{Aroclor}$ by River Mile by Year

River Mile	Station	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
206	FD1	1.1	0.8	0	0	0	0	0	2.2	4.3	6.1	8.8	0.0	0.0
194	TD1	1.9	1.8	0	0	0	12.2	0	0.6	1.9	2.6	3.5	3.4	0.9
193	TD2	6.7	3.0	0	0	0	4.9	6.0	18.2	8.9	8.1	5.5	0.5	3.2
192	TD3	7.1	0	0	0	0	6.7	2.1	7.8	35.5	2.4	3.3	4.6	4.3
191	TD4	0	0	4.6	2.0	0	2.1	2.1	11.7	26.4	7.2	4.6	5.0	0.2
190	TD5	4.9	0	0	0	0	10.9	3.0	6.5	25.3	15.2	4.4	4.0	0.3
188	ND1	0	0	0	0	NA	0	2.8	16.0	25.7	9.0	16.6	14.2	2.4
187	ND2	0	0	0	0	NA	0	0	13.1	25.0	13.4	7.3	20.3	6.1
186	ND3	0	9.0	0	1.5	2.2	3.1	0	16.3	21.8	21.1	7.7	15.9	4.3
184	ND5	1.5	1.5	0	0	3.0	0.9	9.9	18.3	26.3	36.9	14.1	20.7	5.9
182	SW1	0	0	0	0	0	0	3.8	14.7	15.6	23.1	15.5	19.4	1.0
178.3	SW2	0	0	2.0	2.4	0	7.9	3.7	17.0	22.1	22.4	17.9	19.8	6.7
177.8	SW3	0	1.5	0	0	0	0	0	9.0	10.3	13.0	52.5	10.2	2.5
173	SW4	0	0	2.3	0	3.9	4.8	4.1	21.0	23.2	34.8	21.5	21.3	7.9
169	SW5	0	4.1	0	3.9	0	8.2	7.7	18.6	27.5	34.0	21.6	16.1	9.6
148	AT1	0	0	0	0	0	5.0	5.9	16.7	24.4	28.9	23.6	19.5	5.5
115	CS1	NA	NA	NA	NA	NA	NA	0.3	5.5	4.2	13.8	4.9	2.7	1.3

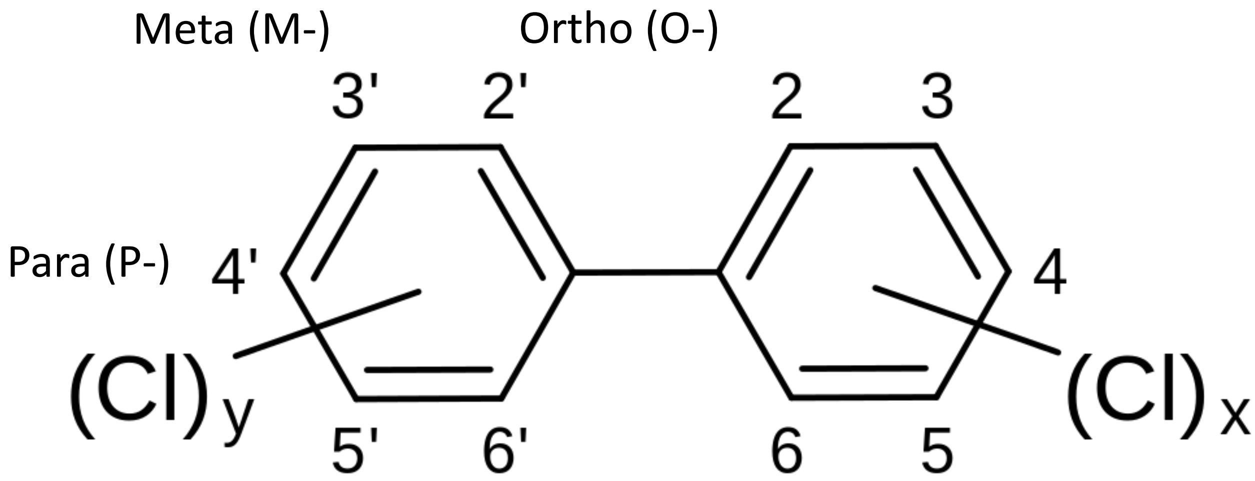
Notes:

[yellow box] Dredging Locations

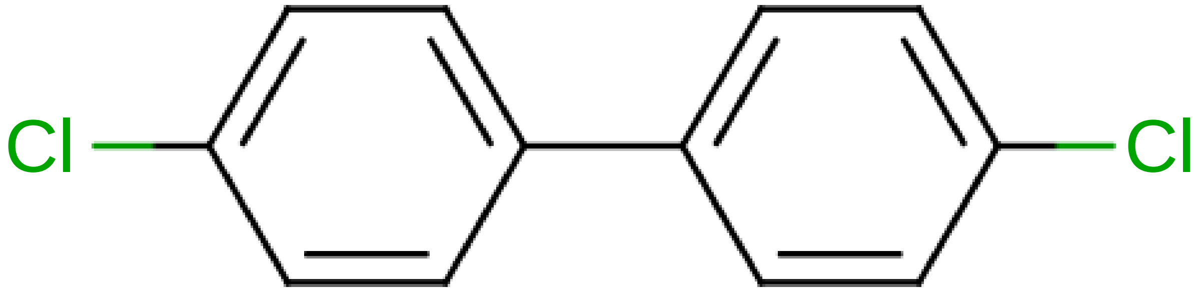
In 2013, dredging also occurred at CU-99 at RM 159 and CU-100 at RM 154.3. These were small areas and are not indicated on the table.

NA Not Applicable

Generic Structure:



PCB Congener:



4,4'-dichlorobiphenyl

PCB Structure

Figure A5-1

April 2019

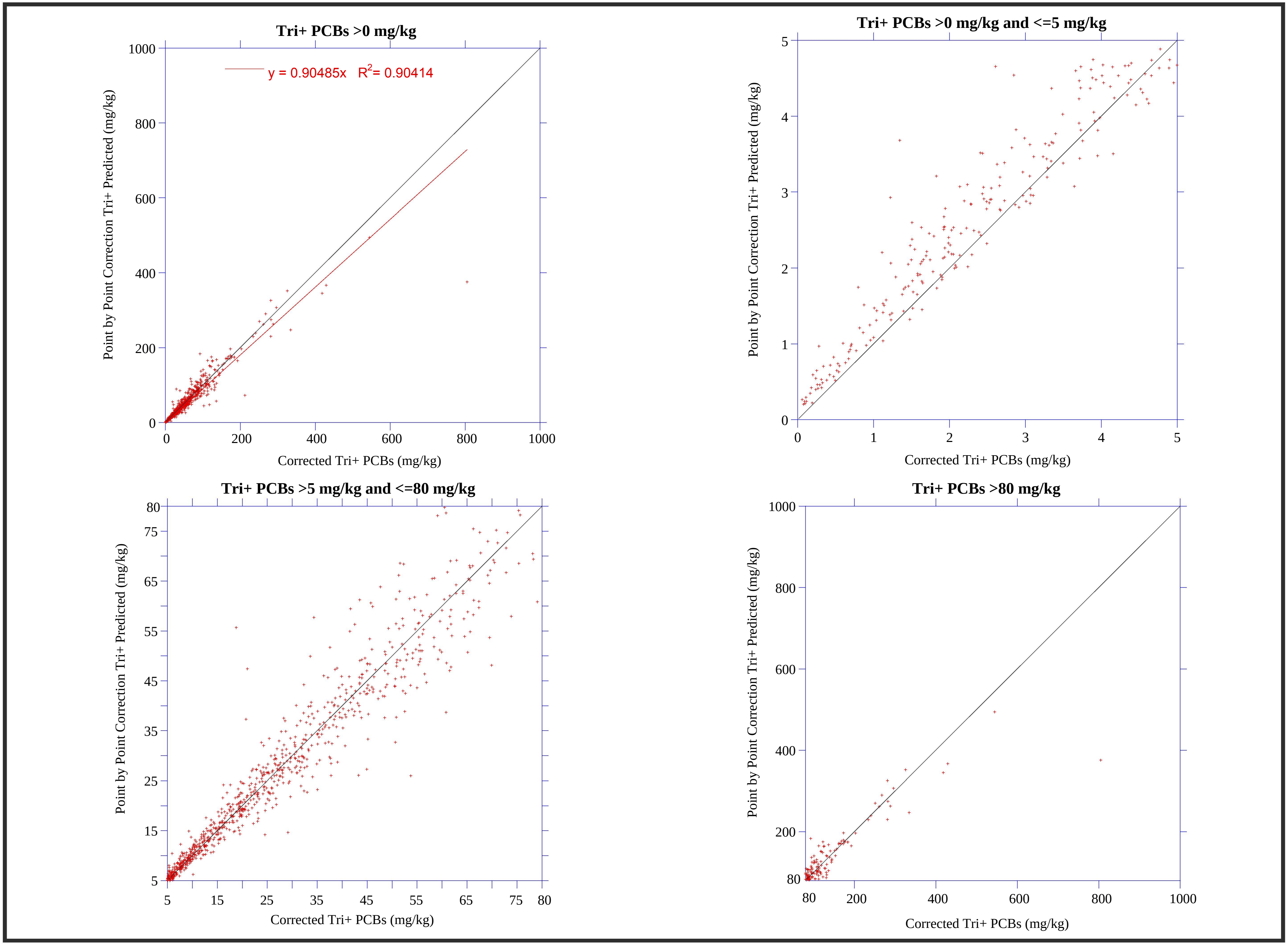


Figure A5-2
Regression Results for GE 2 Model (Point by Point Correction)

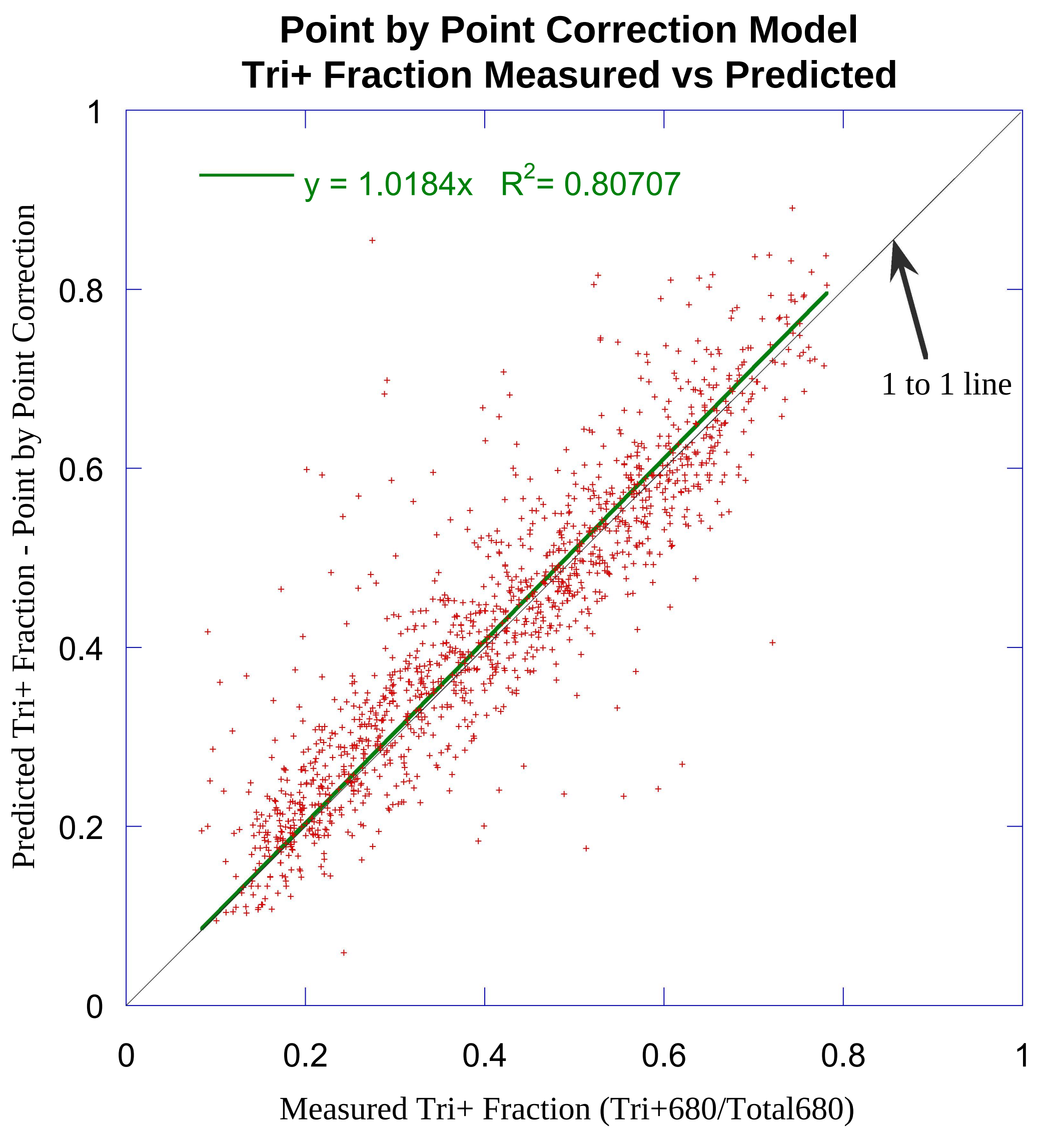


Figure A5-3
Measured vs Predicted Tri+ Fraction for the GE 2 Model (Point by Point Correction)

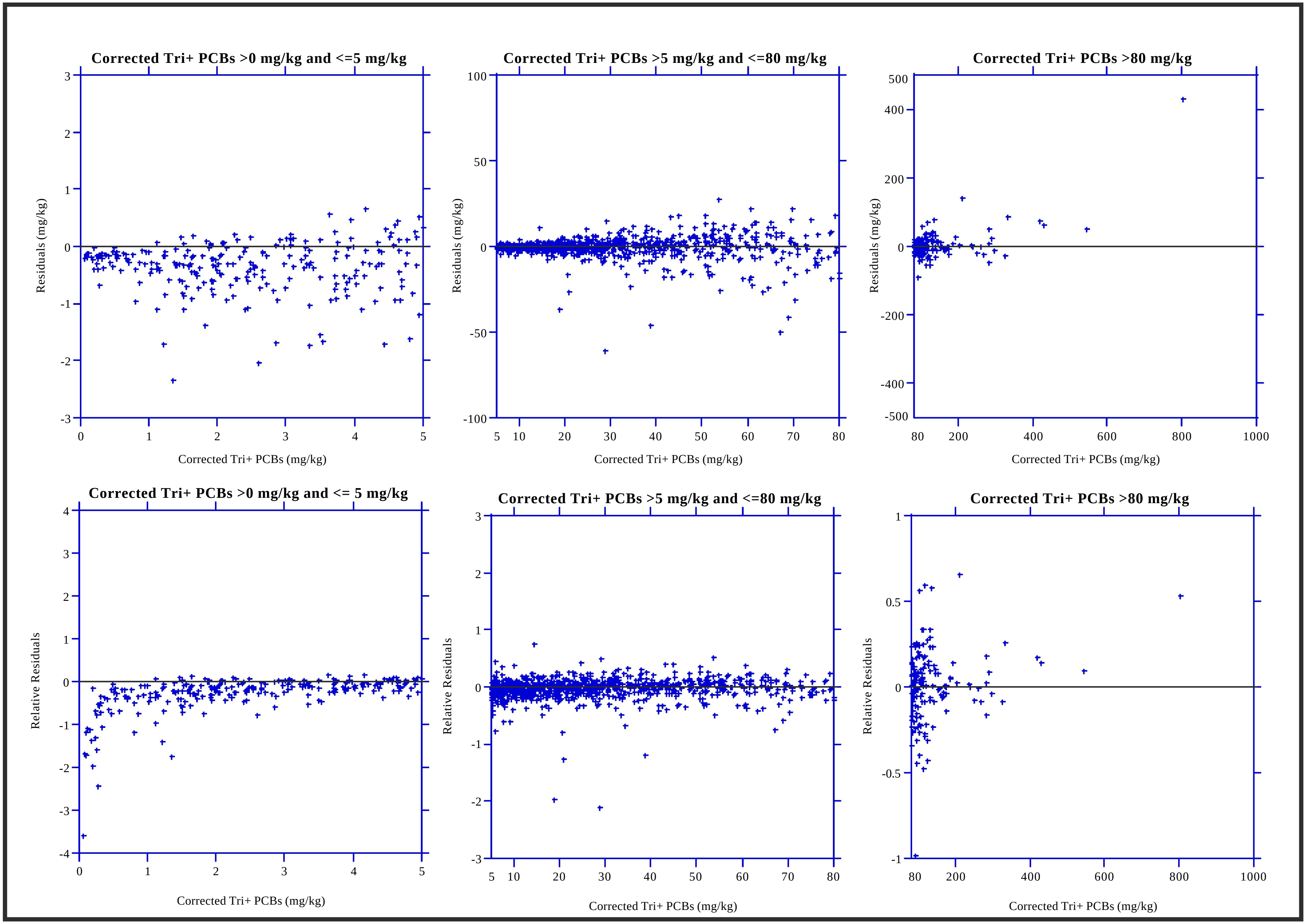


Figure A5-4
Regression Residuals for GE 2 Model (Point by Point Correction)

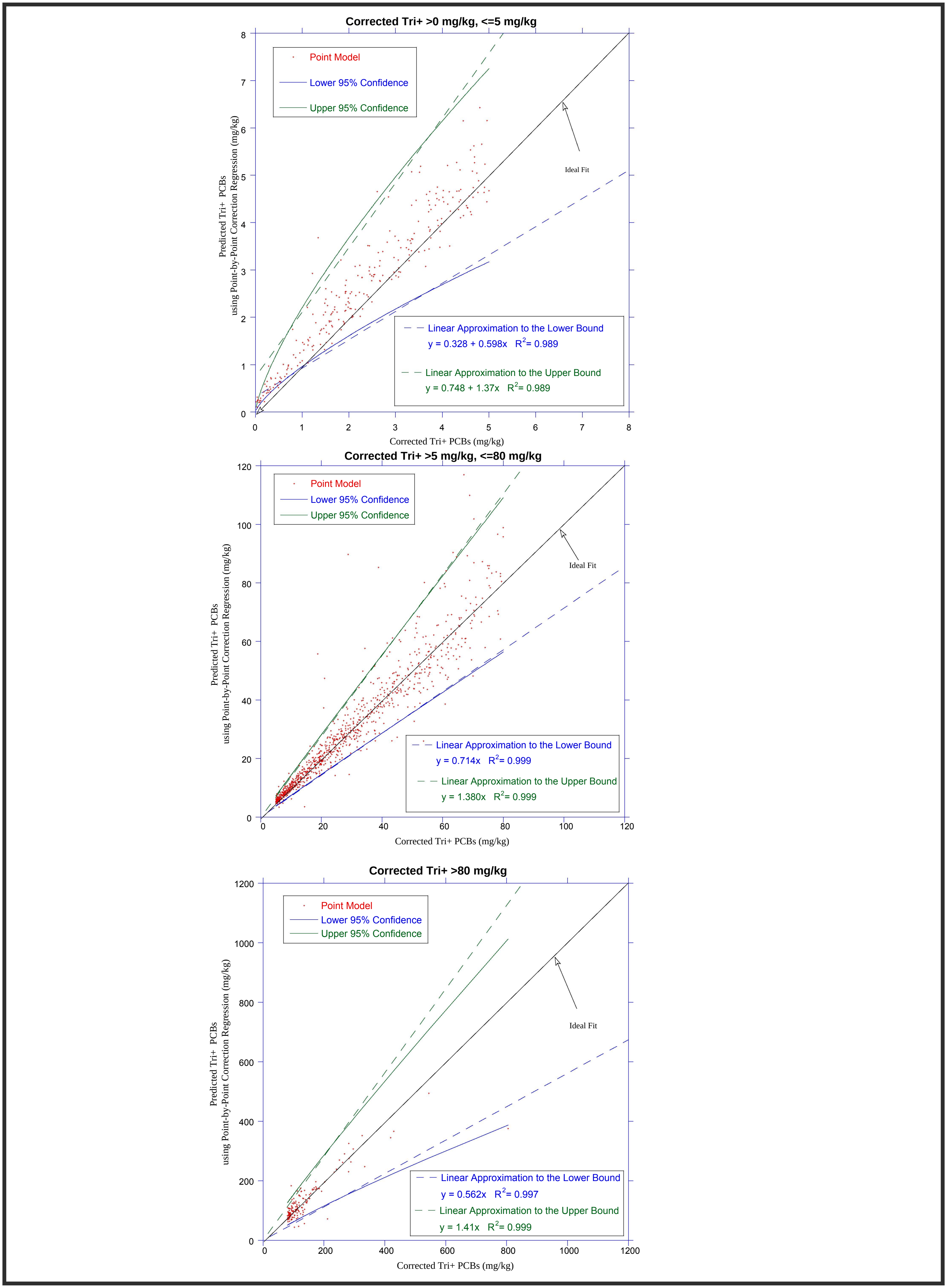
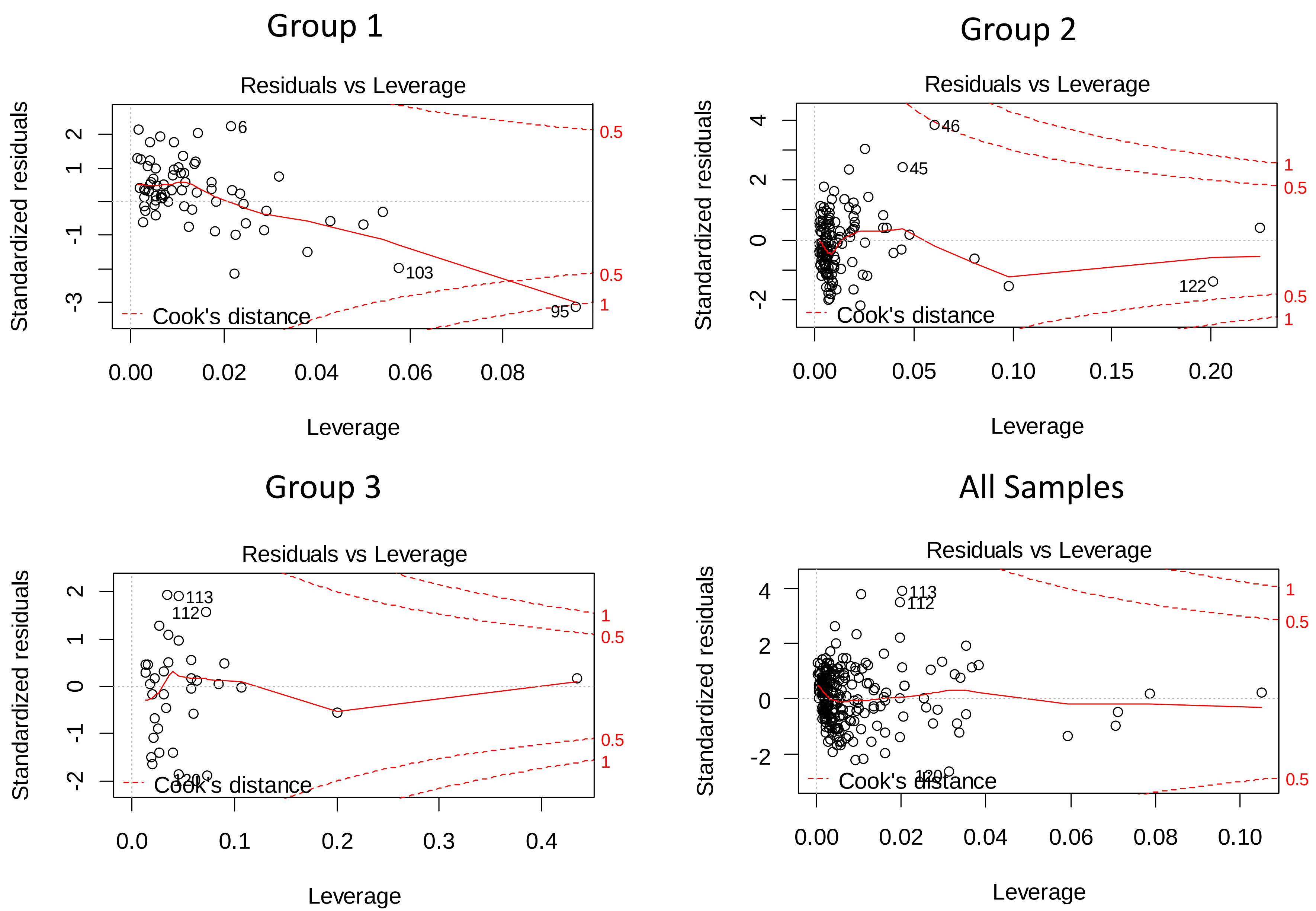


Figure A5-5
Upper and Lower Bound for Corrected vs Predicted Tri+ PCBs using Point-by-Point Correction Regression



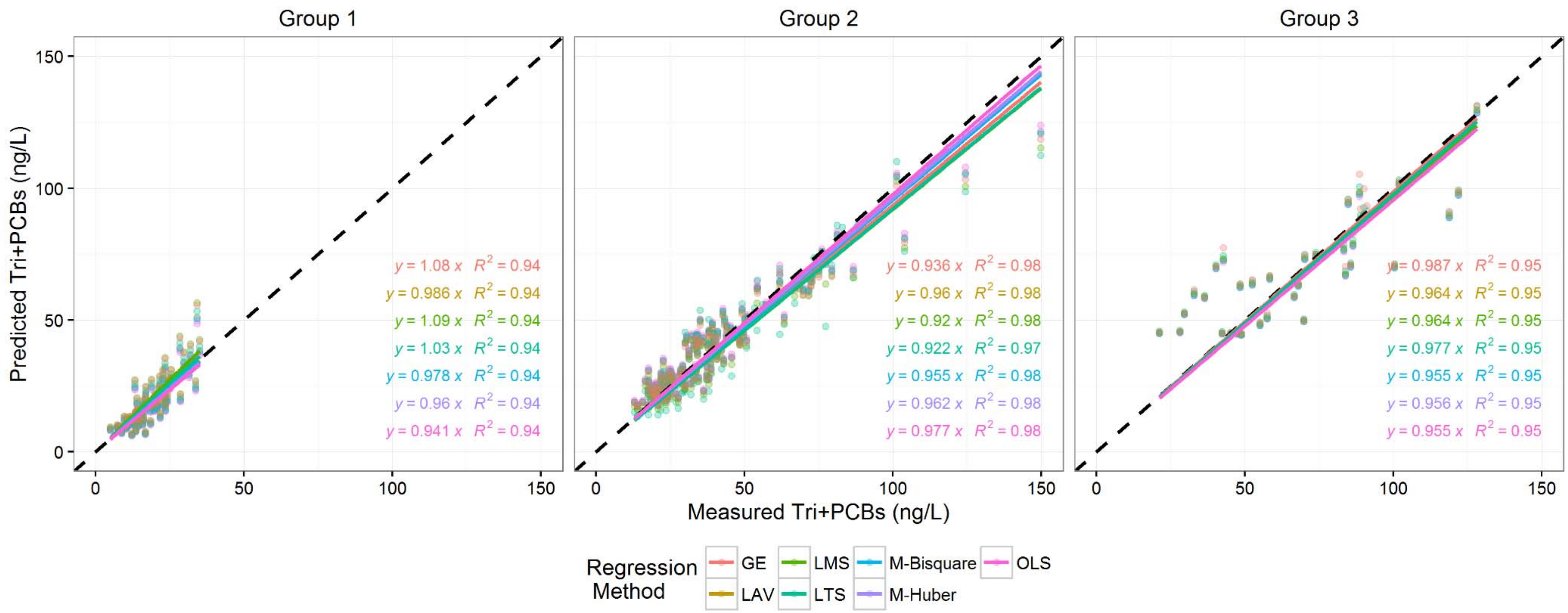
OLS regression diagnostic plots for influential points

OLS Formula: Group 1: Tri+PCB= a × A1221;

Group 2/3 and all samples: Tri+PCB = a × A1221 + b × (A1242+A1254)

Figure A5-6

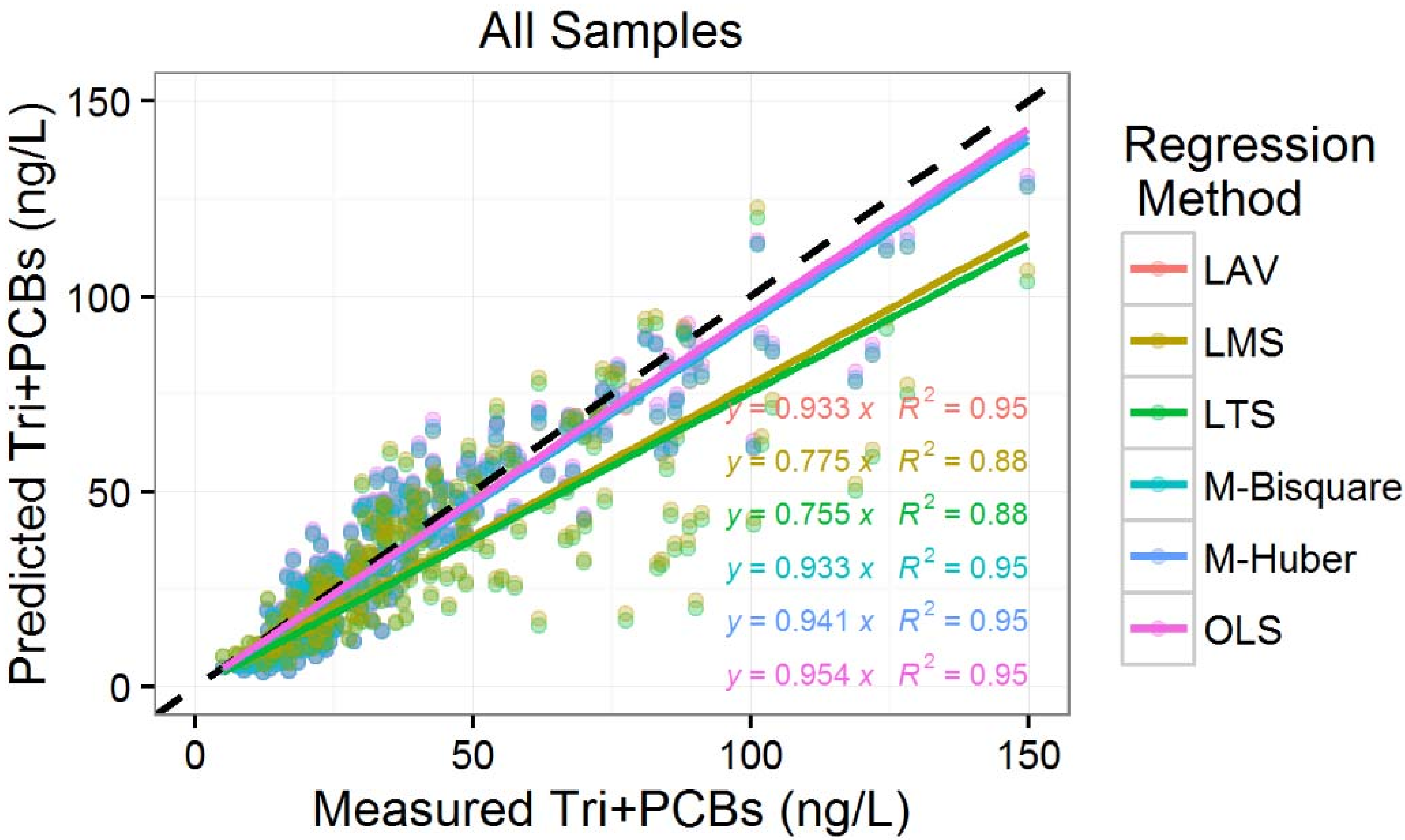
April 2019



Comparison between the predicted and measured Tri+ PCB concentrations for training data set among models using the three correlation approach

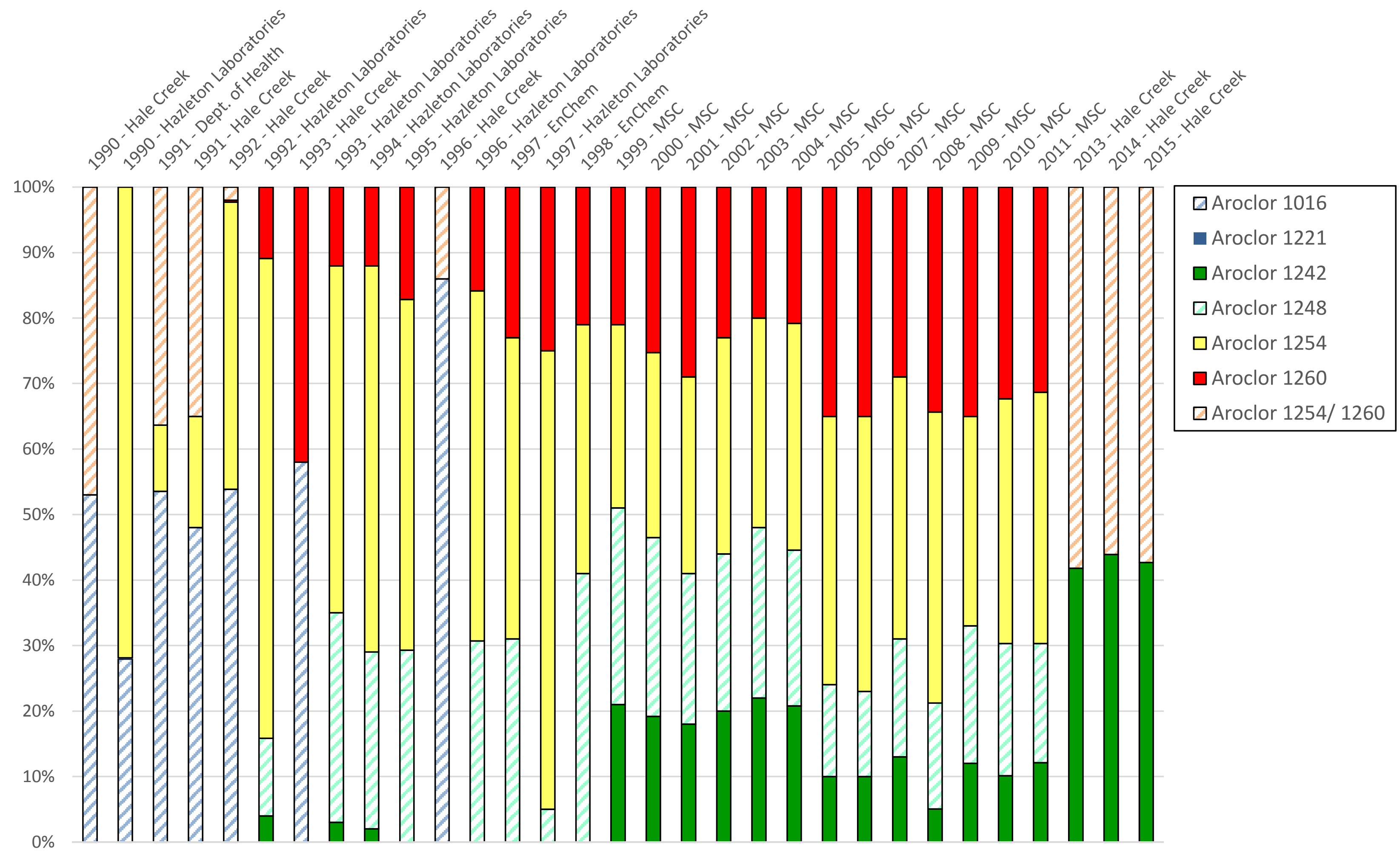
Figure A5-7a

April 2019



Comparison between the predicted and measured Tri+ PCB concentrations for training data set among models using the one correlation approach

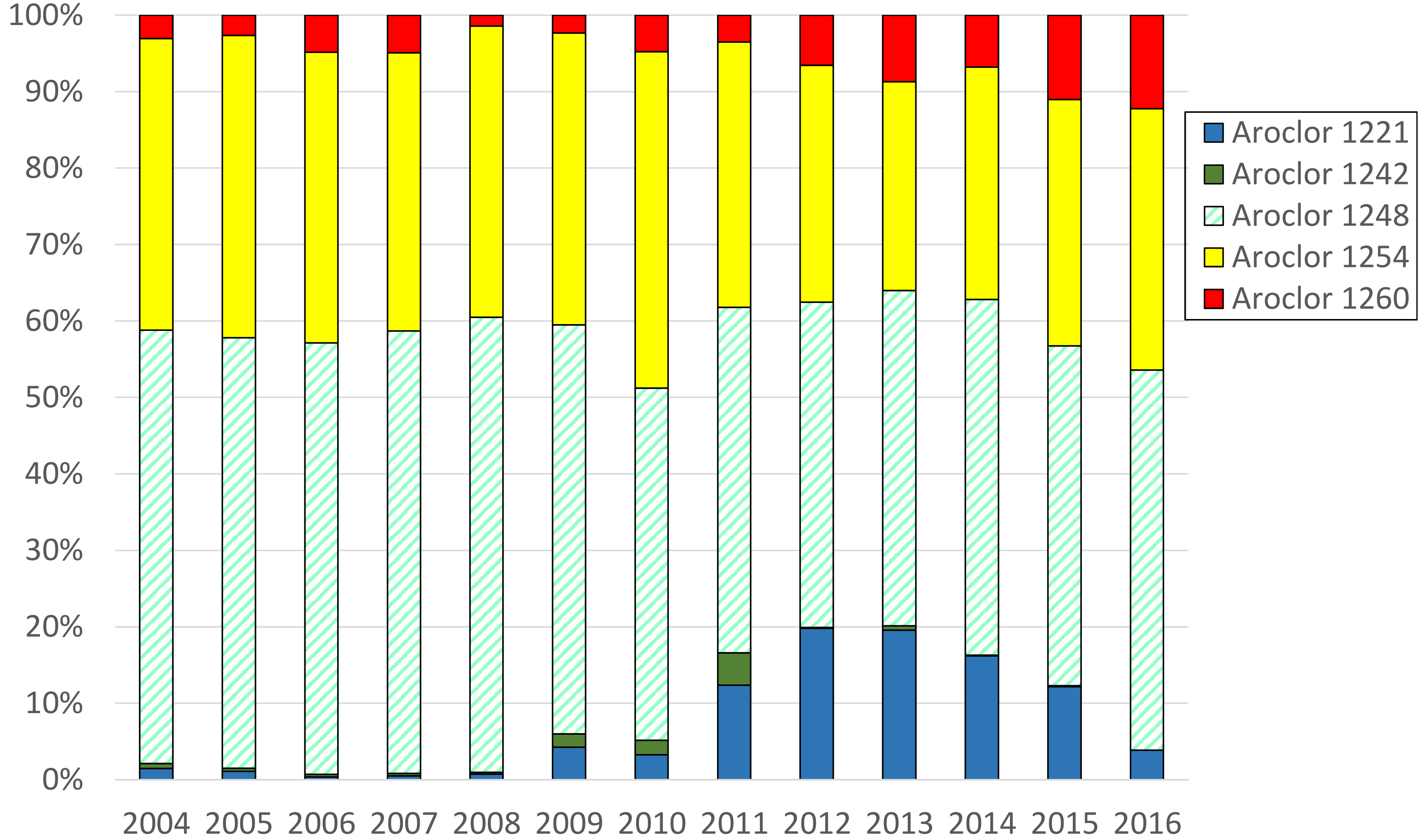
Figure A5-7b



Average Percentage Aroclors Composition in NYSDEC PCB Fish Data between 1990 and 2015

Figure A5-8

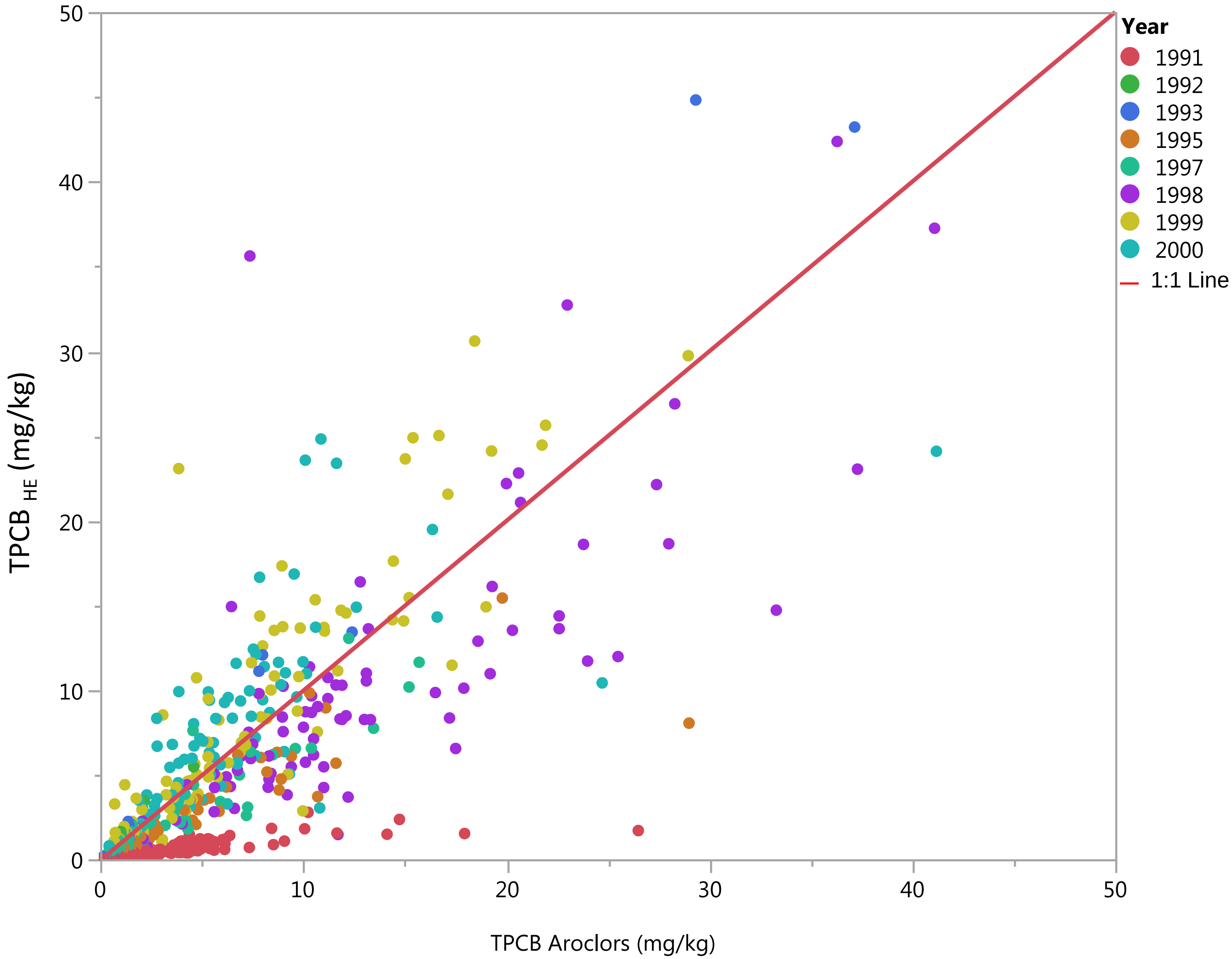
April 2019



Average Percentage Aroclors Composition in GE PCB Fish Data
between 2004 and 2016

Figure A5-9

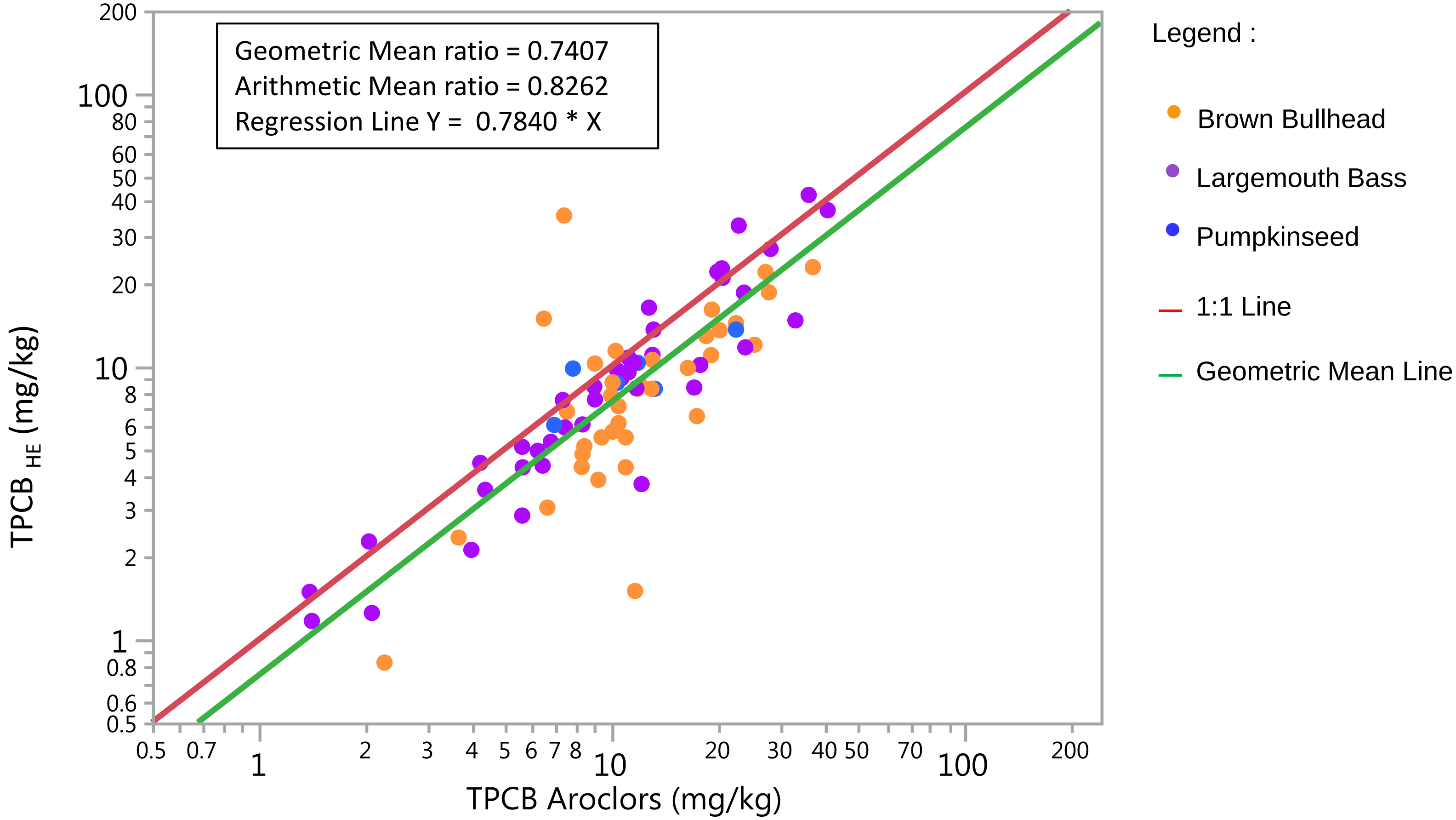
April 2019



TPCB$_{HE}$ vs TPCB $_{Aroclor}$
1991 to 2000 NYSDEC Data

Figure A5-10

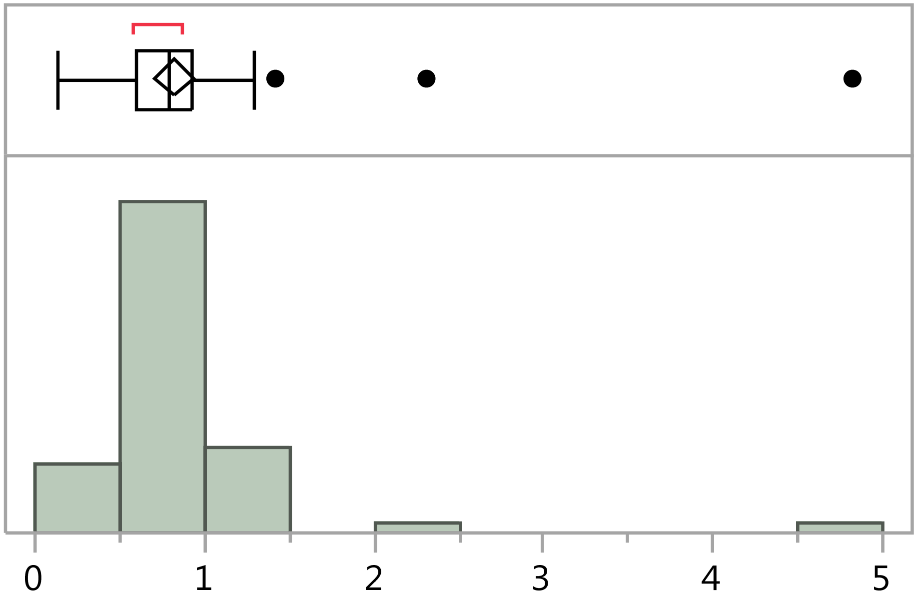
April 2019



TPCB$_{HE}$ vs TPCB $_{Aroclor}$ for Individual Samples
1998 NYSDEC Data - Enchem Environmental Laboratory

Figure A5-11

April 2019



Quantiles

100.0%	maximum	4.84
99.5%		4.84
97.5%		2.25
90.0%		1.11
75.0%	quartile	0.92
50.0%	median	0.78
25.0%	quartile	0.59
10.0%		0.46
2.5%		0.31
0.5%		0.13
0.0%	minimum	0.13

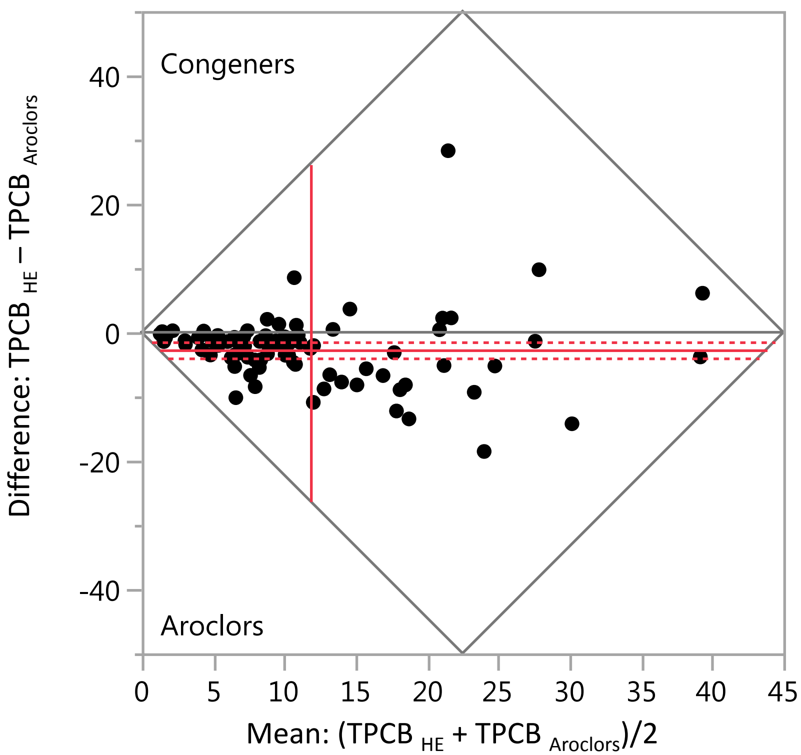
Summary Statistics

Mean	0.83
Std Dev	0.54
Std Err Mean	0.06
Upper 95% Mean	0.94
Lower 95% Mean	0.71
N	82.00

Distribution of NYSDEC 1998 $TPCB_{HE}$ / $TPCB_{Aroclor}$ Ratio for Individual Samples

Figure A5-12

April 2019



Mean TPCB HE	10.4749	t-Ratio	-4.24056
Mean TPCB Aroclors	13.1468	DF	81
Mean Difference	-2.6719	Prob > |t|	<.0001*
Std Error	0.63009	Prob > t	1.0000
Upper 95%	-1.4182	Prob < t	<.0001*
Lower 95%	-3.9256		
N	82		
Correlation	0.76951		

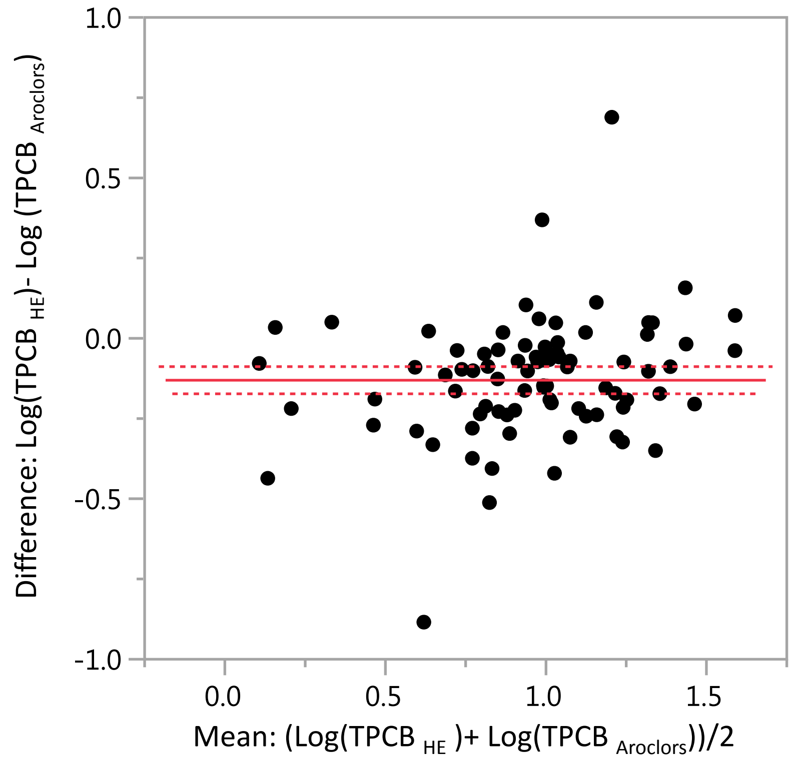
Wilcoxon Signed Rank

	TPCB HE − TPCB Aroclors
Test Statistic S	-1193.5
Prob>|S|	<.0001*
Prob>S	1.0000
Prob<S	<.0001*

Wilcoxon test on 1998 NYSDEC Matched Pairs
($TPCB_{HE}$ - $TPCB_{Aroclor}$) for Individual Samples

Figure A5-13

April 2019



Mean Log (TPCB $_{HE}$)	0.8955	t-Ratio	-6.12821
Mean Log (TPCB $_{Aroclors}$)	1.02586	DF	81
Mean Difference	-0.1304	Prob > |t|	<.0001*
Std Error	0.02127	Prob > t	1.0000
Upper 95%	-0.088	Prob < t	<.0001*
Lower 95%	-0.1727		
N	82		
Correlation	0.83408		

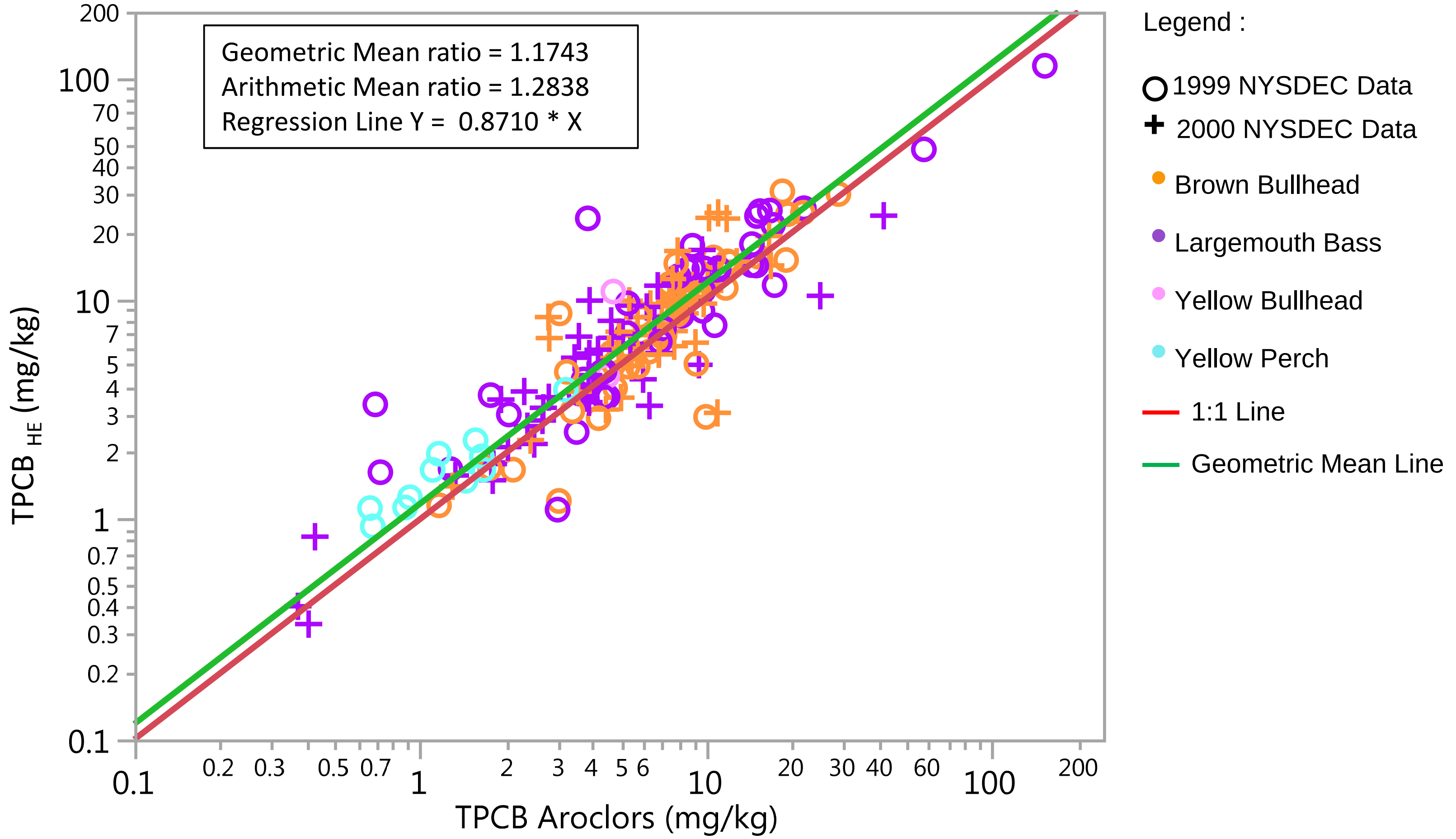
Wilcoxon Signed Rank

	Log (TPCB $_{HE}$)– Log(TPCB $_{Aroclors}$)
Test Statistic S	-1310.5
Prob>|S|	<.0001*
Prob>S	1.0000
Prob<S	<.0001*

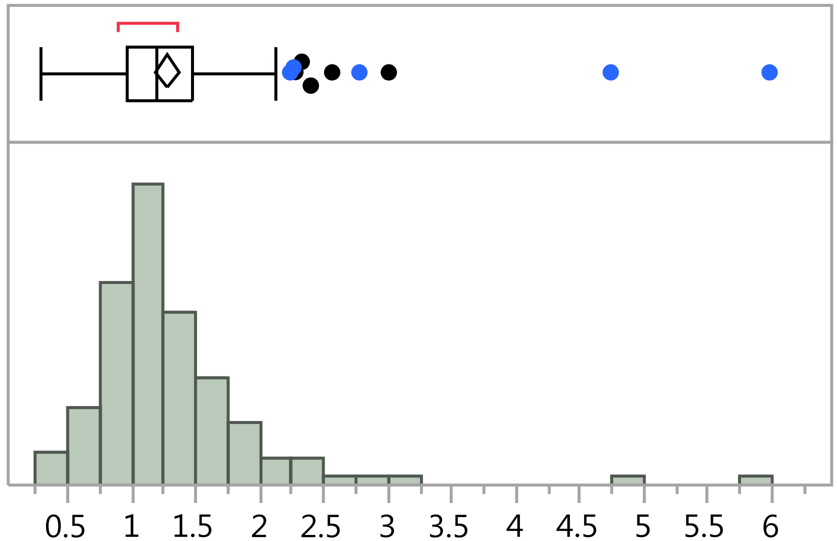
Wilcoxon test on 1998 NYSDEC Matched Pairs
log($TPCB_{HE}$) – log(TPCB $_{Aroclor}$) for Individual Samples

Figure A5-14

April 2019



	$TPCB_{HE}$ vs. $TPCB_{Aroclor}$ for Individual Samples 1999-2000 NYSDEC Data – Mississippi State Chemical Laboratories	Figure A5-15 April 2019



Quantiles

100.0%	maximum	5.996098701
99.5%		5.996098701
97.5%		2.7131001787
90.0%		1.870086091
75.0%	quartile	1.468920854
50.0%	median	1.180856343
25.0%	quartile	0.953664204
10.0%		0.7390527832
2.5%		0.4051755697
0.5%		0.284737037
0.0%	minimum	0.284737037

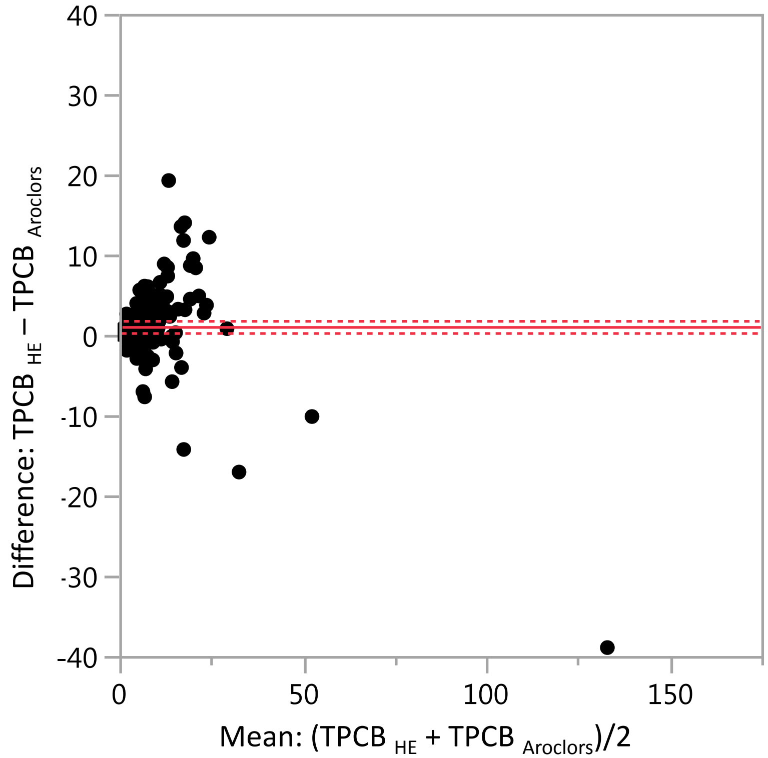
Summary Statistics

Mean	1.2837556
Std Dev	0.6384143
Std Err Mean	0.0485377
Upper 95% Mean	1.3795619
Lower 95% Mean	1.1879493
N	173

Distribution of 1999-2000 NYSDEC $TPCB_{HE}$ / $TPCB_{Aroclor}$ Ratio for Individual Samples

Figure A5-16

April 2019



Mean TPCB $_{HE}$	9.27859	t-Ratio	2.844704
Mean TPCB $_{Aroclors}$	8.18697	DF	172
Mean Difference	1.09162	Prob > |t|	0.0050*
Std Error	0.38374	Prob > t	0.0025*
Upper 95%	1.84907	Prob < t	0.9975
Lower 95%	0.33418		
N	173		
Correlation	0.9268		

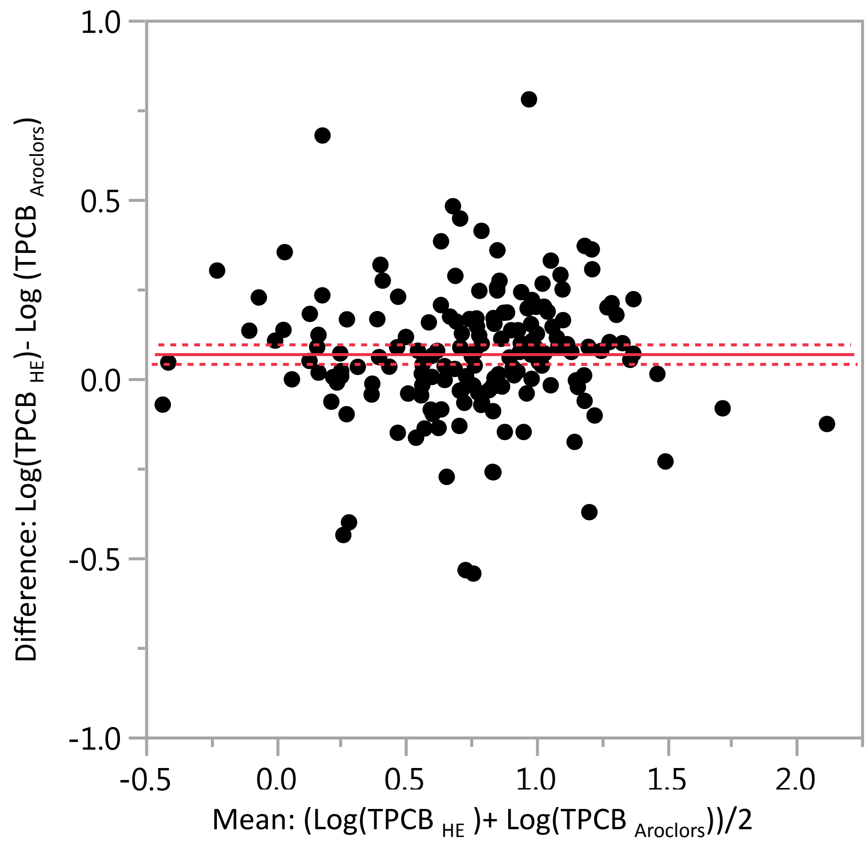
Wilcoxon Signed Rank

	TPCB $_{HE}$ – TPCB $_{Aroclors}$
Test Statistic S	3600.50
Prob>|S|	<.0001*
Prob>S	<.0001*
Prob<S	1.0000

Wilcoxon test on 1999-2000 NYSDEC Matched Pairs
($TPCB_{HE}$ - $TPCB_{Aroclor}$) for Individual Samples

Figure A5-17

April 2019



Mean Log (TPCB $_{HE}$)	0.79532	t-Ratio	5.075368
Mean Log (TPCB $_{Aroclors}$)	0.72556	DF	172
Mean Difference	0.06976	Prob > |t|	<.0001*
Std Error	0.01375	Prob > t	<.0001*
Upper 95%	0.09689	Prob < t	1.0000
Lower 95%	0.04263		
N	173		
Correlation	0.89763		

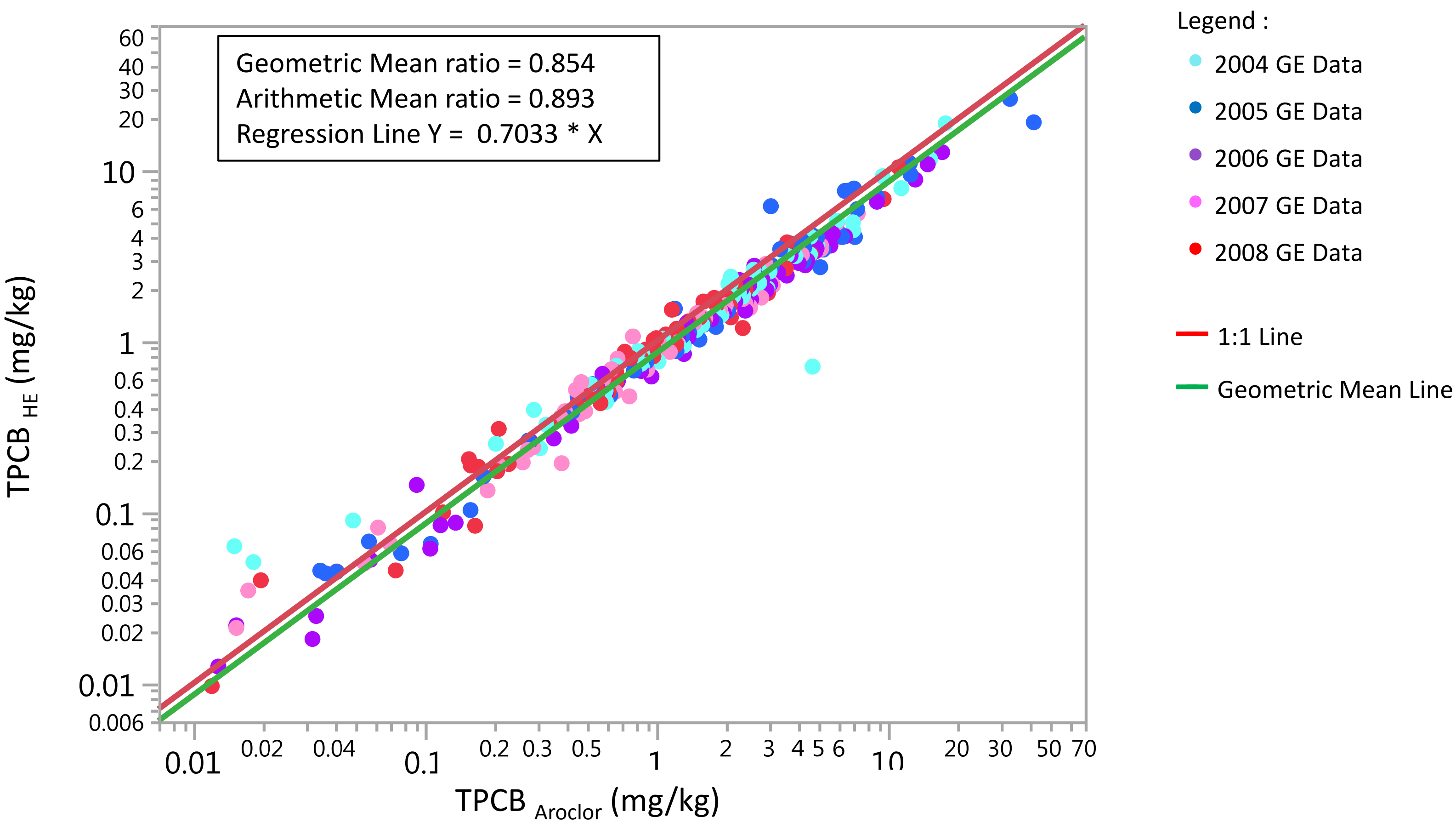
Wilcoxon Signed Rank

	Log (TPCB $_{HE}$)– Log (TPCB $_{Aroclors}$)
Test Statistic S	3681.50
Prob>|S|	<.0001*
Prob>S	<.0001*
Prob<S	1.0000

Wilcoxon test on 1999-2000 NYSDEC Matched Pairs
log(TPCB$_{HE}$) – log(TPCB $_{Aroclor}$) for Individual Samples

Figure A5-18

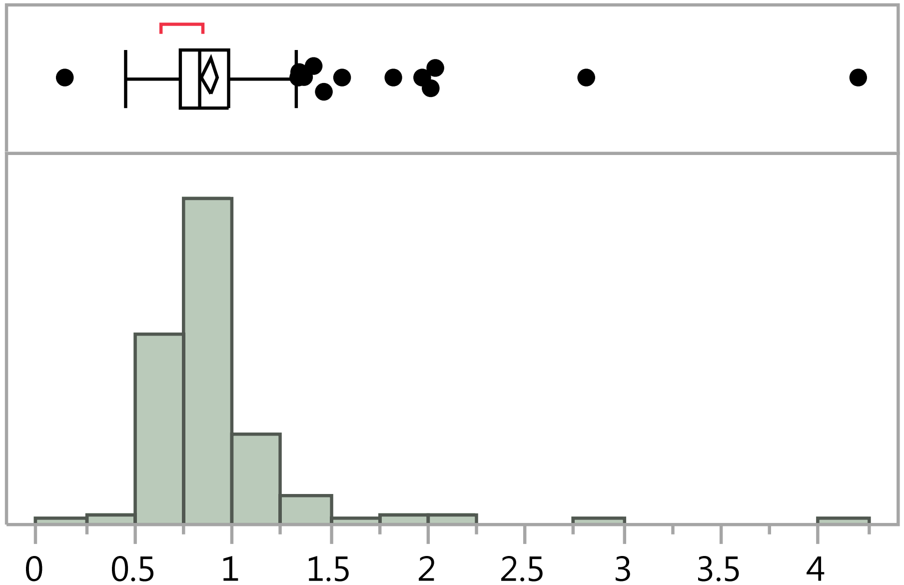
April 2019



$TPCB_{HE}$ vs. $TPCB_{Aroclor}$ for Individual Samples
2004 to 2008 GE Data - Northeast Analytical Laboratory

Figure A5-19

April 2019



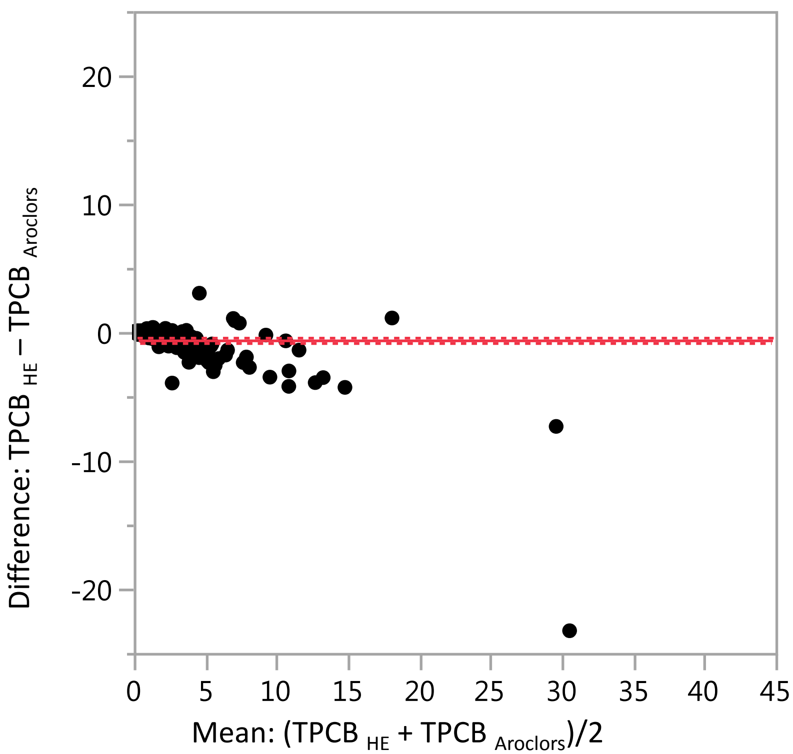
Quantiles		
100.0%	maximum	4.21
99.5%		3.79
97.5%		1.70
90.0%		1.16
75.0%	quartile	0.97
50.0%	median	0.83
25.0%	quartile	0.73
10.0%		0.65
2.5%		0.54
0.5%		0.24
0.0%	minimum	0.15

Summary Statistics	
Mean	0.89
Std Dev	0.33
Std Err Mean	0.02
Upper 95% Mean	0.93
Lower 95% Mean	0.85
N	259.00

Distribution of 2004 to 2008 GE $TPCB_{HE}$ / $TPCB_{Aroclor}$

Figure A5-20

April 2019



Mean TPCB HE	2.12049	t-Ratio	-5.49086
Mean TPCB Aroclors	2.70432	DF	258
Mean Difference	-0.5838	Prob > |t|	<.0001*
Std Error	0.10633	Prob > t	1.0000
Upper 95%	-0.3744	Prob < t	<.0001*
Lower 95%	-0.7932		
N	259		
Correlation	0.94907		

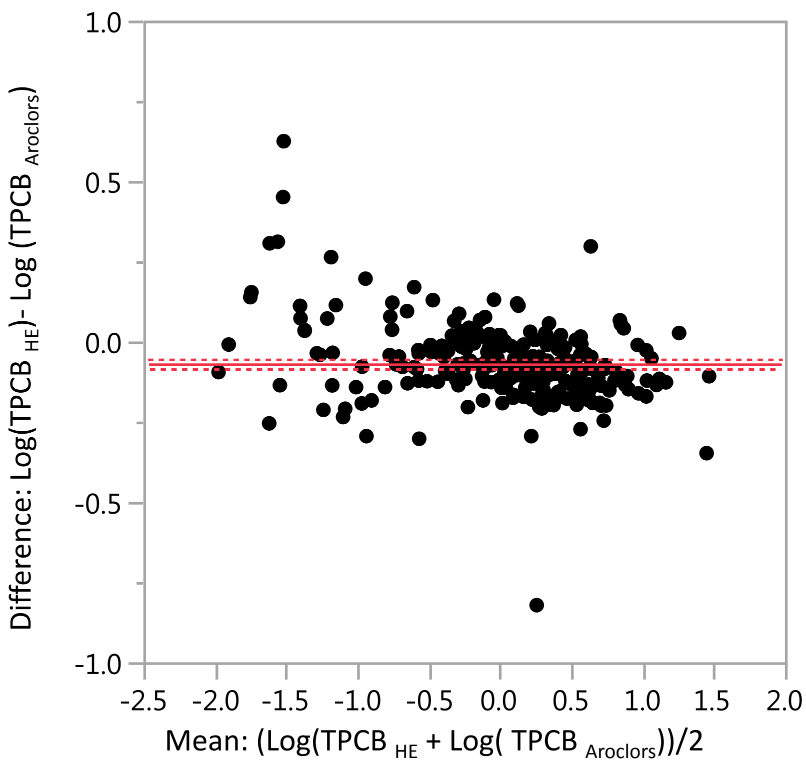
Wilcoxon Signed Rank

	TPCB HE − TPCB Aroclors
Test Statistic S	-12659
Prob>|S|	<.0001*
Prob>S	1.0000
Prob<S	<.0001*

Wilcoxon Test on 2004-2008 GE Matched Pairs
($TPCB_{HE}$ - $TPCB_{Aroclor}$) for Individual Samples

Figure A5-21

April 2019



Mean Log (TPCB $_{HE}$)	-0.0239	t-Ratio	-8.92025
Mean Log (TPCB $_{Aroclors}$)	0.04453	DF	258
Mean Difference	-0.0684	Prob > |t|	<.0001*
Std Error	0.00767	Prob > t	1.0000
Upper 95%	-0.0533	Prob < t	<.0001*
Lower 95%	-0.0835		
N	259		
Correlation	0.98472		

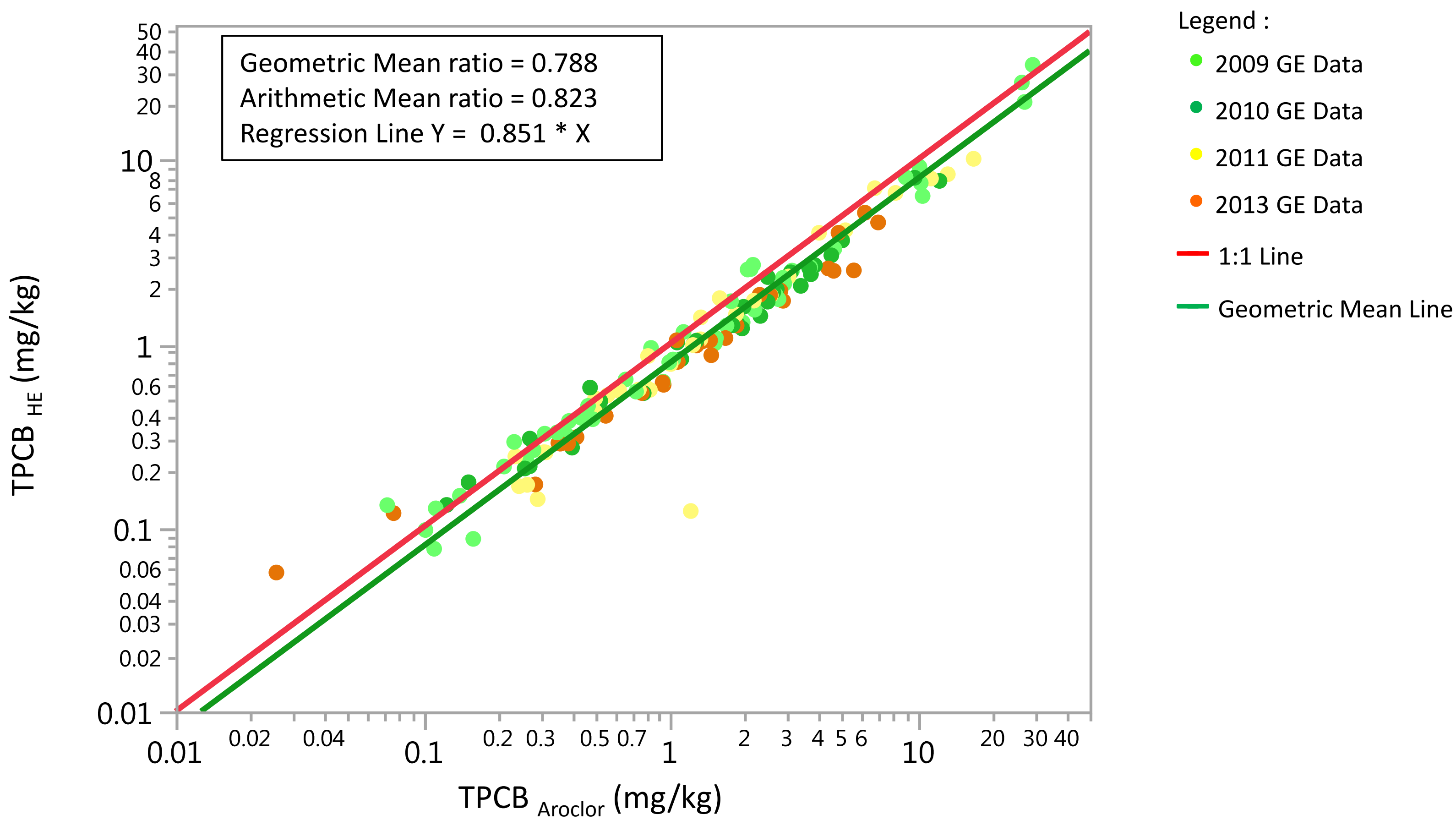
Wilcoxon Signed Rank

	Log (TPCB $_{HE}$)– Log (TPCB $_{Aroclors}$)	
Test Statistic S		-11438
Prob>|S|	<.0001*	
Prob>S	1.0000	
Prob<S	<.0001*	

Wilcoxon Test on 2004 to 2008 GE Matched Pairs
log(TPCB$_{HE}$) – log(TPCB $_{Aroclor}$) for Individual Samples

Figure A5-22

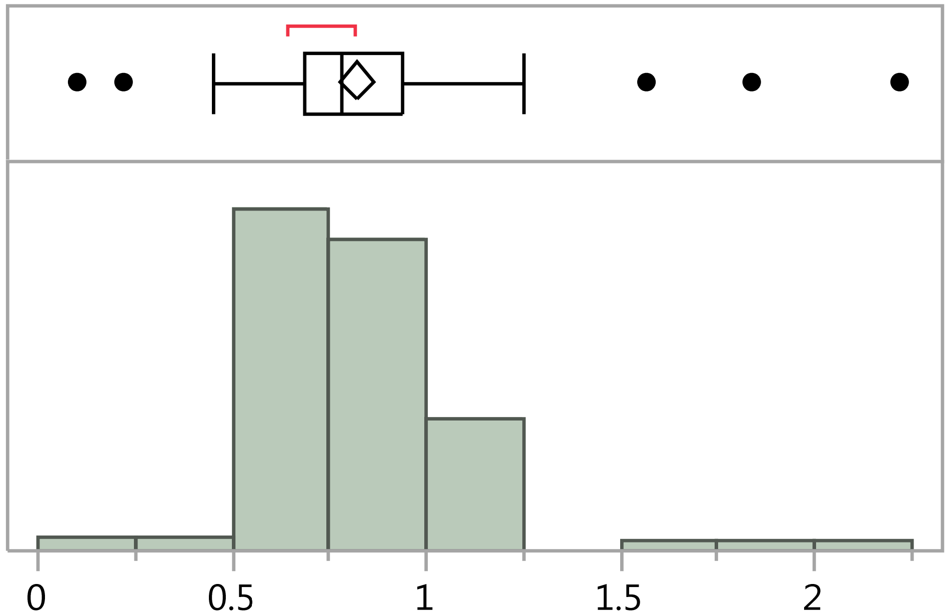
April 2019



$TPCB_{HE}$ vs. $TPCB_{Aroclor}$ for Individual Samples
2009 to 2013 GE Data - Northeast Analytical Laboratory

Figure A5-23

April 2019



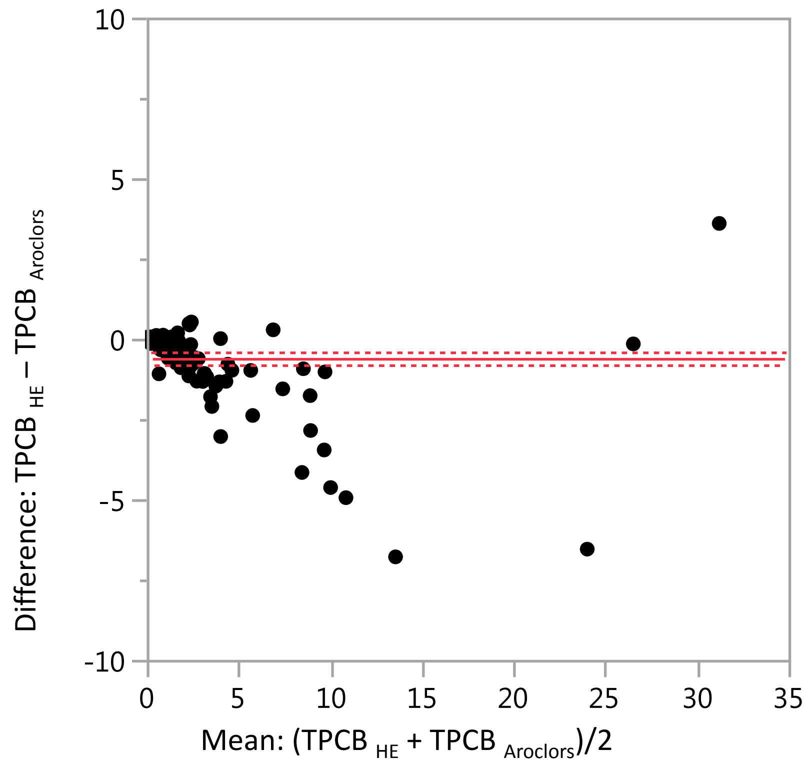
Quantiles		
100.0%	maximum	2.23
99.5%		2.23
97.5%		1.40
90.0%		1.11
75.0%	quartile	0.94
50.0%	median	0.78
25.0%	quartile	0.69
10.0%		0.61
2.5%		0.47
0.5%		0.10
0.0%	minimum	0.10

Summary Statistics	
Mean	0.82
Std Dev	0.24
Std Err Mean	0.02
Upper 95% Mean	0.86
Lower 95% Mean	0.78
N	140.00

Distribution of 2009 to 2013 GE $TPCB_{HE}$ / $TPCB_{Aroclor}$ for Individual Samples

Figure A5-24

April 2019



Mean TPCB HE	2.27202	t-Ratio	-5.92639
Mean TPCB Aroclors	2.87119	DF	139
Mean Difference	-0.5992	Prob > |t|	<.0001*
Std Error	0.1011	Prob > t	1.0000
Upper 95%	-0.3993	Prob < t	<.0001*
Lower 95%	-0.7991		
N	140		
Correlation	0.97011		

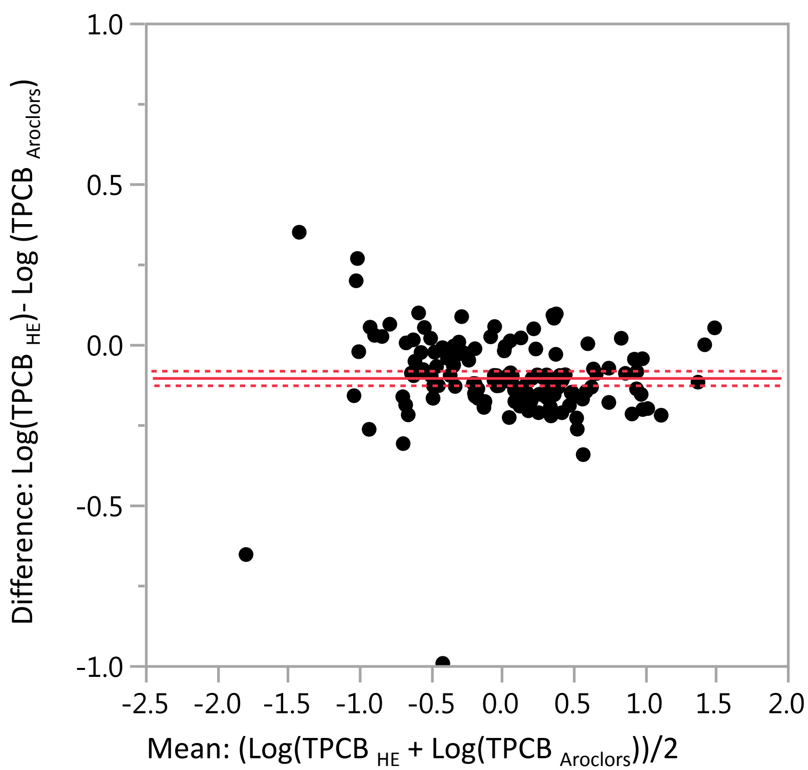
Wilcoxon Signed Rank

	TPCB HE – TPCB Aroclors	
Test Statistic S		-4036.0
Prob>|S|	<.0001*	
Prob>S	1.0000	
Prob<S	<.0001*	

Wilcoxon Test on 2009 to 2013 GE Matched Pairs
($TPCB_{HE}$ - $TPCB_{Aroclor}$) for Individual Samples

Figure A5-25

April 2019



Mean Log (TPCB $_{HE}$)	-0.018	t-Ratio	-8.98789
Mean Log (TPCB $_{Aroclors}$)	0.08546	DF	139
Mean Difference	-0.1035	Prob > |t|	<.0001*
Std Error	0.01151	Prob > t	1.0000
Upper 95%	-0.0807	Prob < t	<.0001*
Lower 95%	-0.1262		
N	140		
Correlation	0.97326		

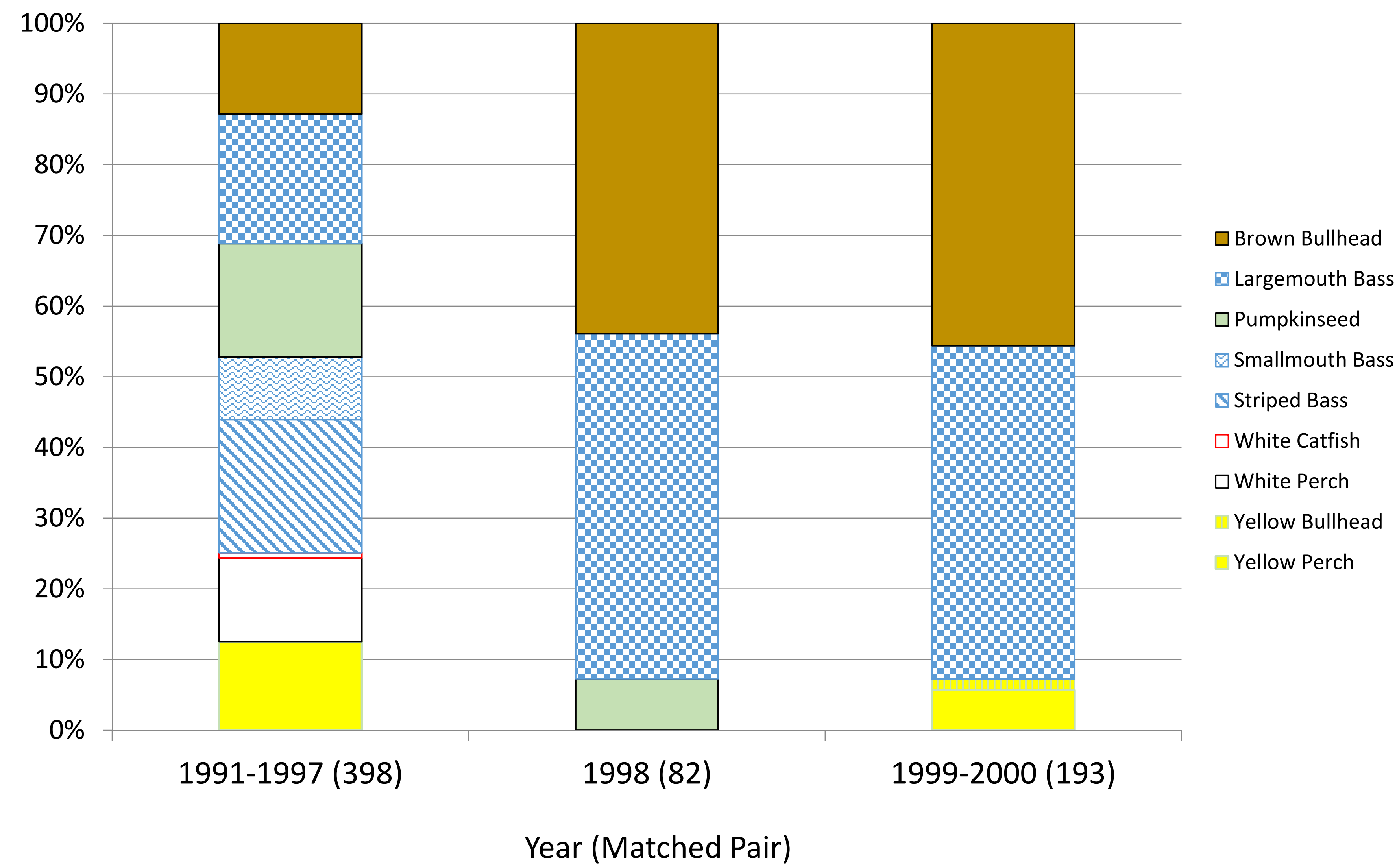
Wilcoxon Signed Rank

	Log (TPCB $_{HE}$)– Log (TPCB $_{Aroclors}$)	
Test Statistic S		-4004.0
Prob>|S|	<.0001*	
Prob>S	1.0000	
Prob<S	<.0001*	

Wilcoxon Test on 2009 to 2013 GE Matched Pairs
log($TPCB_{HE}$) – log(TPCB $_{Aroclor}$) for Individual Samples

Figure A5-26

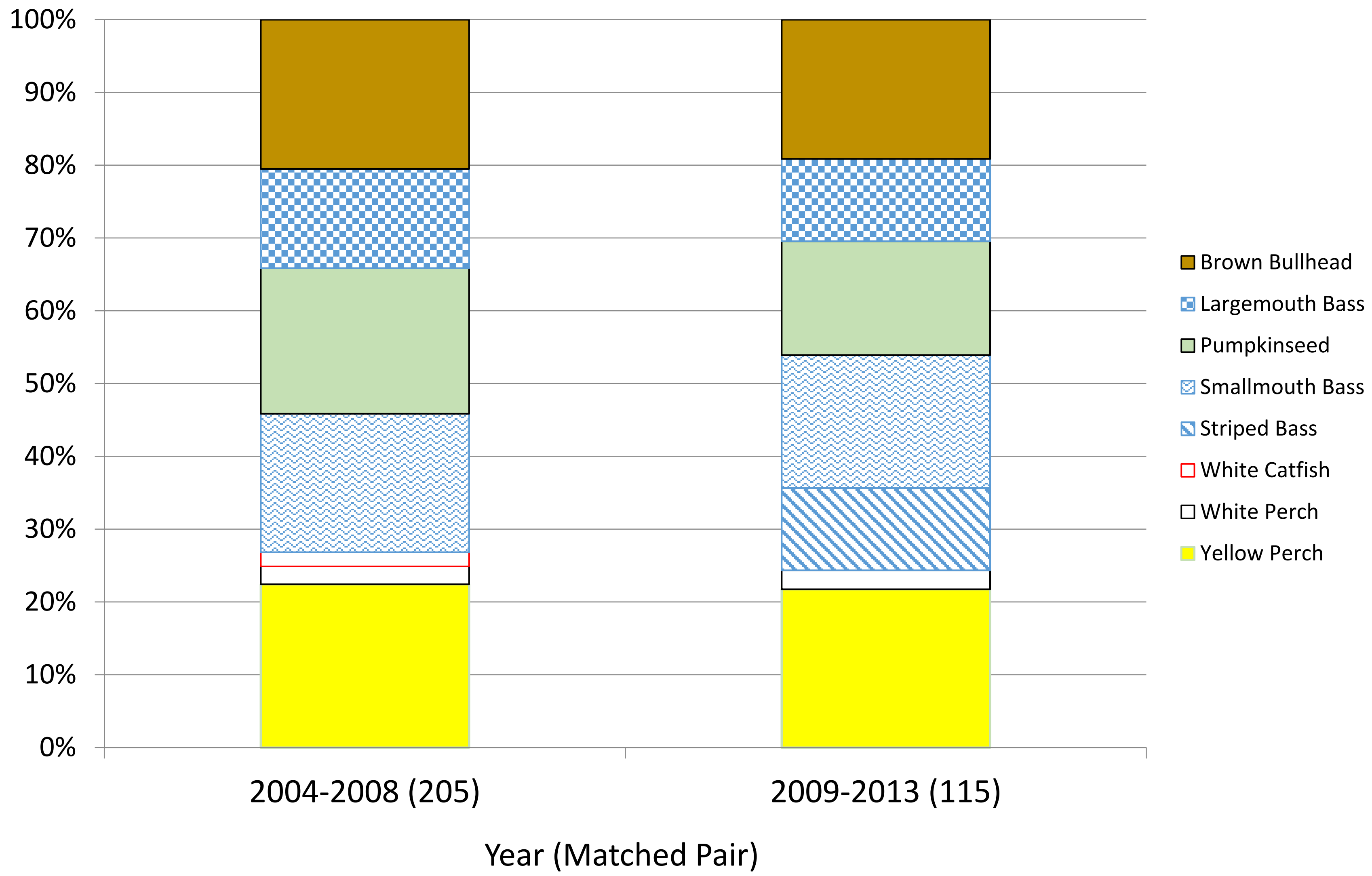
April 2019



Fish Species Distribution in Matched Pair Data Set for NYSDEC Data

Figure A5-27

April 2019



Fish Species Distribution in Matched Pair Data Set for GE Data

Figure A5-28

April 2019