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

APPENDIX 1

EVALUATION OF WATER COLUMN PCB CONCENTRATIONS AND LOADS

Prepared by:
WSP USA Solutions Inc.

July 2024

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

TABLE OF CONTENTS

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

LIST OF TABLES

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

LIST OF FIGURES

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

LIST OF ATTACHMENTS

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

EXECUTIVE SUMMARY

Background

The purpose of this appendix is to evaluate water column polychlorinated biphenyls (PCBs) concentration data collected during the post-dredging or natural recovery period after the implementation of the remedial action selected by the United States Environmental Protection Agency (EPA) in the 2002 Record of Decision for Operable Unit 2 (OU2) (ROD; EPA, 2002) of the Hudson River PCBs Superfund Site (Site). As part of this evaluation, from 2016 to 2021 water column samples were collected at consistent intervals over time at five long-term monitoring locations within the Upper Hudson River area of the Site, providing a long-term dataset to track changes in water column PCB concentrations over time and assess whether the remedial action objectives (RAOs) established in the ROD are being achieved. The relevant RAOs for water column PCBs are: (1) to reduce PCB levels in sediments in order to reduce PCB concentrations in river (surface) water that are above surface water Applicable or Relevant and Appropriate Requirements (ARARs), and (2) to minimize the long-term downstream transport of PCBs at the Site.

Analyses

Data collected between 2004 and 2021 (encompassing three time periods: pre-dredging – 2004 to 2008; dredging – 2009 to 2015; and post-dredging – 2016 to 2021) are presented in this appendix, but the focus is the post-dredging period and the progress toward achieving the RAOs established in the ROD.

The frequency of the water column data collection ranged from sub-daily to monthly, depending on the monitoring station, time period, and sampling program objectives (e.g., routine or high-flow sampling). This appendix provides a summary of post-dredging water column PCB concentrations at the long-term monitoring stations. Assessment of progress toward meeting the above RAOs includes: (1) comparing water column PCB concentrations to the most protective water column ARAR (14 nanograms per liter [ng/L] Total PCBs (TPCBs), the Criteria Continuous Concentration - Federal Water Quality Criterion for protection of aquatic life in freshwater [CCC-FWQC]); and (2) calculating the mass load of PCBs transported from the Upper Hudson River to the Lower Hudson River and assessing whether the mass load of PCBs is being minimized over time. Other analyses include a statistical comparison of pre- and post-dredging data and an assessment of the minimum number of years of data necessary to estimate accurate time trends.

Key Results

- For the post-dredging period (2016 to 2021), the percentage of samples with TPCB concentrations less than 14 ng/L at Thompson Island Dam, Schuylerville, and Waterford was 76 percent, 44 percent, and 57 percent, respectively. These values represent substantial improvements over the pre-dredging period, when only 10 percent, 16 percent, and 18 percent of samples were less than 14 ng/L at the Thompson Island Dam, Schuylerville, and Waterford stations, respectively.
- During the post-dredging period, annual Tri+ PCB[1] loads transported to the Lower Hudson River ranged from 34 kilograms (kg) in 2020 to 101 kg in 2019. As expected, annual loads in part reflect the magnitude of flows within a year, with annual loads higher in years with higher median flows. PCB loads during the post-dredging period are lower than the pre-dredging period, consistent with observed changes in water column PCB concentration. Finally, by removing variability in the water column Tri+ PCB loads introduced by flow and seasonality, the results indicate that Tri+ PCB loads to the Lower Hudson River are decreasing.
- Statistical analysis indicates the current six years of post-dredging water column data are insufficient to accurately determine a long-term time trend. Utilizing the pre-dredging water column data, it was determined that eight or more years of data are needed to estimate time trends that accurately reflect the long-term time trend in PCB concentration. When using only six years of data (the current number of years of post-dredging data), time trends exhibit substantial variability (as measured by deviation from the long-term time trend), with trend estimates falling well outside the 95 percent confidence interval of the long-term time trend. Thus, to determine a meaningful and accurate time trend in routine water column PCB concentrations, at least eight years of routine water column data are needed. The results of this analysis are consistent with results from the Second Five-Year Review Comment Response (EPA, 2019b) using pre-dredging fish tissue data. Comparison of the variability in the pre- and post-dredging data indicates that the results based on pre-dredging data are transferrable to the post-dredging period.

[1] 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 10 chlorine atoms per molecule, each with its own set of chemical properties.

1 INTRODUCTION

1.1 Background and Overview

As determined by the Remedial Investigation conducted by the United States Environmental Protection Agency (EPA) for Operable Unit 2 (OU2) of the Site, polychlorinated biphenyls (PCB) in the water column of the Upper Hudson River (defined as the approximately 40-mile portion of the river between Fort Edward and the Federal Dam in Troy) are linked to PCBs in sediments and fish. PCBs in the water column, including particulate-bound and dissolved phases, are transported throughout the Upper Hudson River and into the Lower Hudson River (defined as the portion of the river between the Federal Dam [River Mile (RM) 153.9] and the Battery [RM 0.0]). Water column samples have been collected consistently through time at multiple locations within the Upper Hudson River since 2004, providing a long-term dataset to track changes in water column PCB concentrations and assess whether the remedial action objectives (RAOs) established by EPA in the 2002 Record of Decision (ROD; EPA, 2002) are being achieved. The ROD for OU2 called for a two-part remedy: dredging (conducted between 2009 and 2015, with no dredging occurring in 2010) followed by monitored natural recovery. This appendix evaluates water column PCB concentration data collected during the post-dredging or natural recovery phase of the remedial action for which data are currently available, 2016 to 2021 (referred to herein as the post-dredging period).

1.2 Purpose and Objectives of the Water Column Sampling Program

The purpose and objectives of the Upper Hudson River water column sampling program have varied over time, reflecting the different stages of remedial activities for OU2. There are broadly three major periods of data collection following the 2002 ROD: the pre-dredging baseline period (2004 to 2008) that includes the Baseline Monitoring Program (BMP), the dredging period (2009 to 2015) that includes the Remedial Action Monitoring Program (RAMP) and the post-dredging period (2016 to present) that includes the Operations, Monitoring and Maintenance (OM&M) Program. The following provides a brief overview of the sampling during each of these time periods.

1.2.1 Pre-Dredging Baseline Period: 2004 to 2008

The years immediately preceding the remedy (referred to herein as the pre-dredging period, which occurred from 2004 to 2008) included the BMP, which spanned from 2004 to early 2009 and established baseline water column PCB concentrations for all three river sections prior to the in-water remedial activities. This baseline dataset was used to establish requirements for determining compliance with resuspension performance standards during dredging (EPA, 2004, 2010). Additional objectives of the BMP included establishing a baseline PCB load at the Waterford monitoring station and a background water column PCB concentration upstream of the GE Hudson Falls facility (GE, 2004).

1.2.2 Dredging Period (RAMP): 2009 to 2015

During active dredging, which occurred from the spring of 2009 to the fall of 2015, with no dredging occurring in 2010 (referred to hereinafter, as "the dredging period"), water column PCB concentrations were monitored under the procedures established in the RAMP (GE, 2006a, 2009a, 2011a, 2012a) to assess compliance with the requirements of the Phase 1 and Phase 2 Resuspension Engineering Performance Standard (EPA, 2004, 2010). Substantially more sampling/monitoring occurred during the RAMP compared with the BMP period. The majority of monitoring during the RAMP occurred during the dredging season, typically April to November during each construction year, to evaluate potential impacts associated with the in-water remedial activities. However, sampling was also performed in the off-season when in-water remedial activities temporarily ceased, typically from December to March, according to the procedures of the off-season monitoring program encompassed within the RAMP (GE, 2006a, 2009a, 2011a, 2012a).

1.2.3 Post-Dredging Period: 2016 to Present

Since the completion of dredging in the fall of 2015, water column monitoring has continued under the RAMP off-season monitoring program, which started in 2016 and continues to the present (referred to herein as the "post-dredging period"). Sampling during this period is being used to assess the recovery of the river and progress towards the ROD RAOs. The directly applicable RAOs related to water include:

- Reduce PCB concentrations in river (surface) water that are above surface water Applicable or Relevant and Appropriate Requirements (ARARs).
- Minimize the long-term downstream transport of PCBs in the river.

1.3 Document Organization

This appendix is organized into the following sections:

- Section 1 (Introduction): Provides the purpose and objectives for monitoring PCBs in the water column.
- Section 2 (Program Description): Presents an overview of the water column monitoring program, sampling locations, and the analytical methods used.
- Section 3 (Analysis Methods): Describes the data handling performed in the evaluation presented in this appendix, as well as the methods for evaluating PCB concentrations and loads.
- Section 4 (Results and Discussion): Presents the results of the evaluation of PCB concentrations measured in the Upper Hudson River, estimates of PCB loads at the Waterford monitoring station, and implications of these evaluations.
- Section 5 (Conclusions): Summarizes appendix findings.

- Section 6 (Abbreviations and Acronyms): Defines the acronyms and abbreviations used in this appendix.
- Section 7 (References): Provides the complete references for documents cited in this appendix.

2 PROGRAM DESCRIPTION

The water column PCB concentration dataset presented in this appendix was collected under the sampling programs implemented between 2004 and 2021 (see Section 1.2). During the pre-dredging and post-dredging periods, sampling is conducted under either the routine (or off-season) or high-flow sampling programs. As discussed previously, during the dredging period, additional sampling was conducted for the purpose of monitoring resuspension associated with dredging activities (i.e., far-field monitoring). Routine sampling consists of the collection of samples on a weekly to monthly basis that represent non-storm event conditions (see Section 2.1). The high-flow sampling program occurs when specific high-flow thresholds are exceeded at the United States Geological Survey's (USGS) gauging stations at Fort Edward or Waterford, and involves the collection of samples during the rising, peak, and falling limb of the high-flow hydrograph (see Section 2.2).

Water column data were consistently collected since 2004 at five monitoring station locations in the Upper Hudson River: Bakers Falls, Rogers Island, Thompson Island Dam, Schuylerville, and Waterford (Figure A1-1). The data from these five stations are the focus of this Five-Year Review appendix. While other station locations were sampled, they were either not sampled consistently for the purpose of long-term monitoring or were not sampled within the last five years and, therefore, the data from them is not included in this appendix. Additionally, data are collected in the Lower Hudson River. However, those data is being evaluated in the context of the Lower Hudson River as its own OU and consequently are not evaluated in this appendix.

The data from the above five monitoring stations were collected under different sampling programs with different data quality objectives. Thus, the frequency, sample collection methods, and analytical methods varied across programs and years. Although there are differences in the programs, the project initiated various special studies and evaluations to provide confidence that the data are comparable across time. A brief overview of the different programs was previously described in Section 1.2, and a summary of the number of PCB samples collected at each station during the three periods is provided in Table A1-1. Further details regarding sampling procedures and analytical methods are found in annual Data Summary Reports (DSRs) submitted by the General Electric Company (GE) (GE, 2005, 2006b, 2007, 2008, 2009b, 2010, 2011b, 2012b, 2013, 2014, 2015, 2016a, 2017, 2019a, 2019b, 2020, 2021).

2.1 Routine Sampling Program

The routine sampling/monitoring program is designed to monitor PCB concentrations in the water column, for which samples are collected on a pre-determined schedule that spans seasonal fluctuations in PCB concentrations and flows. Routine water column monitoring is conducted at the five long-term monitoring stations in the Upper Hudson River. Three of these monitoring stations are located within the portion of the OU2 area that was dredged between 2009 and 2015 (referred to herein as the project area), and two are located upstream of the OU2 project area.

Starting with the northern-most station, the approximate RM, the primary purpose for the station, and the relationship to the three OU2 River Sections (RS) are as follows:

- Bakers Falls (RM 197, background station located upstream of GE's plant sites);
- Rogers Island (RM 194, background station for OU2 located downstream of GE's plant sites and upstream of the project area boundary);
- Thompson Island Dam (RM 187.5, project area monitoring, downstream boundary of RS 1);
- Schuylerville (RM 181.4, project area monitoring, downstream boundary of RS 2);
- Waterford (RM 156, project area monitoring, downstream boundary of RS 3, used to assess load to the Lower Hudson River).

The following is a summary of the sample collection at each of the five long-term monitoring stations.

2.1.1 Bakers Falls

The Bakers Falls location is located upstream of GE's Fort Edward and Hudson Falls plant sites and, therefore, represents the background for the Site as this location is unaffected by PCB releases associated with the two plants. Samples collected at Bakers Falls are taken manually at the approximate centroid of the river from the downstream side of Bakers Falls Bridge (County Route 27 Bridge) using a multiple aliquot depth integrating sampler (MADIS) at approximately RM 197. Samples were collected weekly between 2004 and 2008. In September 2008, sampling was reduced to monthly with limited sampling between December and March. Sampling continues monthly at Bakers Falls between April and November.

2.1.2 Rogers Island

Rogers Island is located downstream of GE's former plants and the Remnant Deposits but upstream of the OU2 project area, making it a suitable location for monitoring PCB levels entering the OU2 project area from the north. Between 2004 and 2008, samples were collected by boat as two surface grab aliquots from the center of the east and west channels at Rogers Island. Samples were then composited using a volumetric ratio that is consistent with the flow ratio in the east and west channels (i.e., 60:40 ratio). In 2009, the location was moved approximately 1,500 feet upstream to a single point near the center of the channel, upstream of all dredging activities.

Samples at Rogers Island were collected weekly between 2004 and 2008. Between 2009 and 2011, samples were collected weekly during active remediation and monthly during the off-season (note: no dredging was completed in 2010 so only monthly sampling was conducted), with sampling primarily between April and November of that year. In 2012, sampling was reduced to a monthly frequency. Sampling continues monthly at Rogers Island between April and November.

2.1.3 Thompson Island Dam

The Thompson Island Dam station is located just downstream of the Thompson Island Dam and is used to monitor water column PCB concentrations exiting RS1. Between 2004 and 2008, grab sampling at this station was conducted from a boat at six equal discharge increment (EDI) stations placed along a transect located downstream of the southern tip of Thompson Island Dam using a programmable, variable speed crane that lowered a custom designed MADIS. Between 2009 and 2014, an automated station was also used to collect samples from this location for monitoring in-water activities associated with dredging. The automated sample collection system was located in close proximity to the BMP transect and consisted of a sampler located along the western shore of the river with piping that extended into the river, forming an EDI transect consisting of five intake ports. Beginning in 2015, sampling reverted to the manual EDI transect used during the BMP.

Between 2004 and 2009, samples at Thompson Island Dam were collected weekly between March and November. Between 2009 and 2015, sampling was conducted either daily while dredging was ongoing or weekly. In 2016, after the completion of dredging, weekly monitoring resumed during the months of April to November.

2.1.4 Schuylerville

The Schuylerville station is located approximately three-quarters of a river mile downstream of the Champlain Canal Lock 5 and the confluence with the Batten Kill and is used to monitor water column PCB concentrations exiting RS2. Between 2004 and 2008, grab sampling at this station was conducted by boat from the upstream side of the Route 29 Bridge at six EDI stations using the same procedures as at Thompson Island Dam. Between 2009 and 2015, an automated station was also used to collect samples from this location for monitoring in-water activities associated with dredging. The automated sample collection system was located in close proximity to the BMP transect and consisted of a sampler located along the western shore of the river with piping that extended into the river, forming an EDI transect consisting of five intake ports. Beginning in 2016, sampling reverted to the manual EDI transect used during the BMP.

Samples at Schuylerville were collected weekly between 2004 and 2009. Between 2009 and 2015, sampling was conducted either daily while dredging was ongoing or weekly. In 2016, after the completion of dredging, weekly monitoring resumed primarily during the months of April to November.

2.1.5 Waterford

The Waterford station is located 2 miles upstream of the Federal Dam and is above the confluence with the Mohawk River. The station is located such that any influence from the Mohawk River is minimal. This station represents the downstream boundary of RS 3 and is used to monitor PCB concentration and load entering the Lower Hudson River from OU2. Between 2004 and 2008, grab sampling at this station was conducted by boat from the upstream side of the Route 4 Bridge at five EDI locations placed along a transect using the same procedure as at Thompson Island Dam.

Between 2009 and 2015, an automated station was used to collect samples from this location for monitoring in-water activities associated with dredging. The automated sample collection system was located upstream of the Route 4 Bridge and consisted of piping that extended from the sampling house on the west bank of the river to approximately the center of the river channel, though outside of the navigation channel. Beginning in 2016, sampling reverted to the manual EDI transect used during the BMP.

Samples at Waterford were collected weekly between 2004 and 2009. Between 2009 and 2015, sampling was conducted either daily while dredging was ongoing or weekly. In 2016, after the completion of dredging, weekly monitoring resumed.

2.2 High-Flow Sampling Program

Water column PCB concentrations in the Upper Hudson River are directly influenced by river flows. At lower flows, as flows increase the water column PCB concentrations decrease; however, as flows continue to increase, there is a change point at which water column PCB concentrations begin to increase with increasing flow (Figures A1-2 and A1-3, Section 4.1.1, and Attachment A, *Tri+ PCB Load Calculation Technical Memorandum* for additional discussion regarding the relationship between the water column PCB concentration and river flow). The high-flow program is designed to monitor PCB concentrations during high-flow events when water column PCB concentrations are increasing with increasing flows. The high-flow sampling program targets sample collection during the rising, peak, and falling portions of the hydrograph, to the extent possible. Sampling for this program is triggered by the following river flows monitored at the Fort Edward and Waterford USGS gauging stations:

- Schuylerville: When river flows at the USGS gauging station at Fort Edward, New York (Station No. 01327750) exceed 15,000 cubic feet per second (cfs).
- Waterford: When river flows at the USGS gauging station in Waterford, New York (Station No. 01335754) exceed 22,500 cfs.

Post-dredging high-flow monitoring is conducted at the Schuylerville and Waterford long-term water column monitoring stations. The following is a brief description of the high-flow sampling program at these two stations. High-flow sampling occurred for two years during the dredging period at the Thompson Island Dam automated station (2010 and 2011), however, high-flow sampling is no longer implemented at this station. Additional information can be found in the corresponding DSRs.

2.2.1 Schuylerville

During the BMP, no high-flow samples were collected at the Schuylerville station. During the dredging period, high-flow sampling was conducted from 2010 to 2011 and 2013 at the Schuylerville automated station. During the post-dredging period, high-flow sampling events were conducted at Schuylerville from 2017 through 2021. During this period, high-flow sampling at

Schuylerville was conducted using a MADIS from Dix Bridge. Dix Bridge is located immediately to the east of Lock 5 in Schuylerville, approximately 1.5 miles north of the routine Schuylerville in-river location at the Route 29 Bridge. This station is above the confluence of the Batten Kill, and thus any influence from the Batten Kill would be minimal. Dix Bridge is used due to safety concerns associated with high-flow sampling from the Route 29 Bridge, however, these two locations are anticipated to have very similar PCB concentrations.

2.2.2 Waterford

During the BMP, high-flow sampling was conducted at the Waterford station between 2005 and 2009. High-flow sampling continued during the dredging period, with samples collected in 2009 to 2011, and 2013 to 2014 at the Waterford automated station. During the post-dredging period, high-flow sampling was conducted between 2016 and 2021. High-flow sampling was conducted from the Route 4 Bridge using a MADIS in the approximate centroid of the channel.

2.3 Laboratory Analytical Methods

Surface water samples collected during the BMP were analyzed by the congener-specific modified Green Bay Method (mGBM) (Pace Method 207, 294 and subsequent revisions). Information on the mGBM and quality assurance/quality control (QA/QC) procedures for PCB sample handling and analysis, including data management and validation, are provided in the annual DSRs submitted by GE during the BMP (GE, 2005, 2006b, 2007, 2008, 2009b).

The congener-specific mGBM continued to be used to analyze PCBs in water column samples during the first two years of the RAMP (2009 and 2010). Beginning in 2011, water column PCBs were analyzed using either the congener-specific mGBM or an Aroclor-based method (Pace Method 231, 273, and subsequent revisions), depending on the intended purpose of the sample collection. Information on the QA/QC procedures for PCB sample handling and analysis, including data management and validation, are provided in the annual DSRs submitted by GE during the RAMP (GE, 2010, 2011b, 2012b, 2013, 2014, 2015, 2016a), and can be made available upon request.

In 2016, samples collected in January and early February were analyzed for PCBs via the same Aroclor-based method used during the RAMP. Beginning in late February 2016, water column PCBs were analyzed using the mGBM to provide better resolution at lower concentrations. In 2017, due to the closure of the laboratory that had historically supported the monitoring activities at OU2, the PCB analytical method was changed to EPA Method 1668C for all stations, which remains the PCB analytical method used at the Site for both routine and high-flow surface water PCB quantitation.

2.4 Data Used in Current Five-Year Review

A summary of the number of PCB samples collected at each monitoring station between 2004 and 2021 is provided in Table A1-1. Data used in the current Five-Year Review was collected under the following programs:

- Between 2004 and 2009, data were collected under the BMP.
- Between 2009 and 2015, data were collected under the RAMP.
- Beginning in 2016 through the present, data continue to be collected under the RAMP off-season and high-flow monitoring program.

3 ANALYSIS METHODS

3.1 Data Handling

Water column PCB concentration data from the Upper Hudson River monitoring stations presented in this appendix were obtained from GE following data validation, as described in Section 2.3. Parent and field duplicate samples were combined based on the following criteria:

- If concentrations in both the parent and duplicate samples were detected, then the values were averaged.
- If the concentration in one sample was reported as detected and the other was reported as non-detect, then the detected value was used. This is to avoid uncertainty introduced by assigning a specific value for the non-detectable result.
- If concentrations for both samples were reported as non-detect, then the maximum of the two reporting limit values was used.

Summation of individual congeners (for samples analyzed using Method 1668) or peaks (for samples analyzed using the mGBM) for estimation of TPCB and Tri+ PCB (sum of all measured PCB congeners with three or more chlorine atoms per molecule) concentrations was based on both detected values and values flagged with a “J” qualifier (i.e., estimated). For samples analyzed using the Aroclor-based method, the calculation of the water column TPCB concentration used the sum of detected Aroclors, while the water column Tri+ PCB concentration was calculated using site-specific regression equations (GE, 2016b). Aroclor-based calculations also included concentrations based on both detected values and values flagged with a “J” qualifier. In the case of parent-duplicate pairs, calculation of Tri+ PCB and TPCB values was done prior to averaging.

Flow data were obtained from the Fort Edward USGS gauging station (No. 01327750) for the Bakers Falls, Rogers Island, Thompson Island Dam, and Schuylerville monitoring stations and the Waterford USGS gauging station (No. 01335754) for the Waterford monitoring station. The closest 15-minute flow value was assigned to each sample based on the sample date and time. However, if no 15-minute flow value was within 6 hours of a sample date and time, or if the sample did not have a reported sample time, the daily mean flow value for the sample date was used. For the Thompson Island Dam and Schuylerville stations, a flow multiplier was applied to the Fort Edward flow to account for the increased drainage area at these stations relative to the Fort Edward gauge station (1.030919 and 1.067093 for Thompson Island Dam and Schuylerville, respectively [EPA, 2010a]). It should be noted that at the time of writing, USGS flow data at the Fort Edward gauging station were flagged as provisional for the time period October 2021 forward. Provisional flow data are subject to change prior to final approval by the USGS, and therefore, results for 2021 may change after approved data are released. Revisions to the flow data are not expected to substantially change results or conclusions.

Water temperature data for each sample were obtained from water quality measurements collected by GE concurrent with the collection of the water samples. If no water temperature value was available in GE's water quality database for a given sample, a water temperature datapoint from the Albany USGS gauging station (No. 01359139) was assigned to the sample, following similar procedures to the flow data with respect to date and time. The closest 15-minute water temperature value was assigned to each sample based on the sample date and time; however, if no 15-minute temperature value was within 6 hours of a sample date-time, or if the sample did not have a sample time, the daily mean temperature value for the sample date was used. If there was no daily mean temperature value for a sample date, then the long-term average (2002 to 2022) daily mean Albany water temperature on the month and day of the sample collection date was used.

3.2 Methods for Evaluation of PCB Concentrations

3.2.1 PCB Concentrations Through Time

Individual water column Tri+ PCB and TPCB concentrations for the Upper Hudson River monitoring stations were plotted against time for the period 2004 to 2021. Additionally, the annual geometric mean Tri+ PCB and TPCB concentration values for the routine water column samples collected between 2004 and 2021 were plotted (see Section 4.1.2 and 4.1.3). Presenting annual geometric means of PCB concentrations for routine samples aids in visualizing year-to-year changes and better represents the central tendency of a log-normally distributed dataset. The geometric means presented in the figures were calculated using routine water column samples as these samples are collected more consistently at each station and are less influenced by year-to-year variations in the number and magnitude of storm events. Furthermore, high-flow samples are only collected at the Schuylerville and Waterford stations. Presenting results from the routine sampling program allows a more appropriate comparison of concentrations among the five long-term monitoring stations. Uncertainty in the geometric mean was estimated by bootstrapping the data to yield 95 percent confidence limits (2.5th and 97.5th percentile values) from 10,000 bootstrapping runs using the bias-corrected and accelerated bootstrap interval method. The time series plots of the water column TPCB concentration are presented with the ARARs established in the 2002 ROD (EPA, 2002), which are as follows:

- 500 ng/L TPCB, the federal maximum contaminant level (MCL) for drinking water;
- 90 ng/L TPCB, the New York State standard for protection of human health and drinking water sources;
- 30 ng/L TPCB, the CCC-FWQC for protection of aquatic life in saltwater (if applicable);
- 14 ng/L TPCB, the CCC-FWQC for protection of aquatic life in freshwater.

Additionally, 2 ng/L Tri+ PCB is indicated on the Bakers Falls and Rogers Island Tri+ PCB plots. This value is associated with modeling efforts presented in the 2002 ROD (EPA, 2002) that assumed Tri+ PCB concentrations at Rogers Island would decrease to 2 ng/L Tri+ PCBs following source control at GE's Hudson Falls Plant.

PCB concentrations are also presented spatially by plotting both pre- and post-dredging annual Tri+ PCB and TPCB geometric mean concentrations by monitoring station ordered from upstream to downstream. For this comparison, routine pre- and post-dredging period data were limited to samples collected in May through November, which represents the months when samples were consistently collected at all stations during both the pre- and post-dredging period.

To assess whether the three stations within the project area (Thompson Island Dam, Schuylerville, and Waterford) exhibited a statistically significant decline in PCB concentrations between the pre- and post-dredging periods, an analysis of covariance (ANCOVA) was conducted, which is a simple procedure that combines the features of regression with a statistical test of difference by analysis of variance (ANOVA). The application of ANCOVA allows for the comparison of the pre- and post-dredging data while statistically controlling for variation in flow and temperature between the two periods. As discussed in Section 4.1.1, flow and temperature influence water column concentrations. Therefore, the ANCOVA analysis, which accounts for these variables, provides a robust way to assess changes in concentrations between the two periods. See Section 4.1.3 for the results of this analysis.

3.2.2 Progress Towards Achieving Water Column ARARs

An assessment of the progress towards achieving the water column ARARs at the Thompson Island Dam, Schuylerville, and Waterford stations was performed by calculating the percentage of water samples below the most protective water column ARAR, described previously in Section 3.2.1 (14 ng/L TPCB), on an annual basis and for the entire pre and post-dredging periods. The calculation was done on both the routine samples and all samples including both routine and high-flow samples. Routine samples provide a better metric of the percentage of overall time below the different ARARs, as high-flow events represent a relatively small portion of overall flow values within a given year and have a much higher density in sampling. However, recognizing that water column PCB concentration is strongly influenced by river flow (Attachment A, *Tri+ PCB Load Calculation Technical Memorandum*) and that routine samples do not account for the periods of elevated concentrations associated with high-flow events, this analysis was also performed using all samples. See Section 4.1.5 for the results of this analysis.

3.2.3 Assessment of the Minimum Number of Years of Data Before Accurate Time Trends Can Be Estimated

The estimation of time trends using environmental data is important because it allows for the extrapolation of the data into the future to assess when certain goals may be achieved. However, extrapolation of the data into the future is very sensitive to the time trend estimated from the existing data. Incorrectly estimating the time trend, even by a small amount, can result in very large errors in the time estimated to achieve certain goals. Therefore, before a time trend can be estimated, it is important to determine whether the dataset spans a sufficiently long period of time so that the time trend accurately reflects the true, long-term time trend and is not affected by the short-term natural variability in the dataset. This is particularly relevant for the post-dredging datasets used in the Five-Year Review, for which only six years of data are available. The purpose of this analysis is to assess whether six years of water column data is sufficient to reliably estimate time trends that accurately reflect the true, long-term time trend. In the Second Five-Year Review Comment Response (EPA, 2019b), an analysis (referred to herein as the "moving window" analysis) was presented that indicated eight or more years of data are needed to estimate an accurate time trend in the post-dredging fish tissue data. In this appendix, a similar moving window analysis was conducted using routine water column Tri+ PCB data. See Section 4.1.6 for the results of this analysis.

Because a moving window analysis requires a long-term data set to demonstrate how many years are needed to produce reliable estimates of time trends, the continuous pre-dredging water column data available from 1998 to 2008 at Thompson Island Dam and Schuylerville provide the best basis for this evaluation. The Waterford station does not contain enough continuous years of pre-dredging data to be included in this analysis (Waterford data from 1998 and 1999 was not included in the long-term database due to analytical concerns regarding changes in the analytical method USGS used to quantify PCBs, and no data was collected at this station between 2002 and 2003 (EPA, 2019a). However, results for Waterford are not expected to differ substantially from Schuylerville, given they are connected hydrodynamically and under typical flow conditions are tightly coupled from a PCB fate and transport perspective (EPA, 1997).

As described previously in Section 2, some differences exist in the timing, frequency, and analytical methods used in the various water column programs implemented since 1998 at the Thompson Island Dam and Schuylerville stations. To create consistent datasets across the various programs, data was limited to samples collected in May through November, which represents the months when samples were consistently collected during both the pre- and post-dredging period. Additionally, data collected prior to 2004 (the start of the BMP period) did not differentiate between high-flow and routine sampling. To make pre-BMP data comparable to the BMP and post-dredging routine sampling programs, the pre-BMP samples were restricted to those collected when flows at Fort Edward were below 10,000 cfs consistent with typical routine sampling conditions. This resulted in excluding only one sample collected between 1998 and 2003 at Thompson Island Dam and five samples collected between 1998 and 2003 at Schuylerville.

To conduct the moving window analysis with the pre-dredging water column data, the following steps were performed at each station separately:

1. Calculate the long-term time trend and 95 percent confidence interval on the trend for the full (1998 to 2008) dataset, assuming a first-order rate of decline regression analysis. Determine the percent deviation of the confidence limits from the long-term time trend using the following equation:

$$Deviation\ (\%) = \frac{(Trend_{ST} - Trend_{LT})}{Trend_{LT}} * 100\ (Eq. 1)$$

Where $Trend_{ST}$ is a short-term trend estimated using less than 11 years of data (11 years is the number of years used to estimate the long-term trend), and $Trend_{LT}$ is the long-term trend estimate. When the deviation is calculated for the 95 percent confidence limits on the long-term trend, $Trend_{ST}$ represents either the upper or lower 95 percent confidence limit.

2. Identify all groups of *m* consecutive years between 1998 and 2008, with $3 \leq m \leq 10$. For example, for *m* = 3, the time interval is as follows: 1998 to 2000, 1999 to 2001, 2000 to 2002, 2001 to 2003, 2002 to 2004, 2003 to 2005, 2004 to 2006, 2005 to 2007 and 2006 to 2008.
3. For each grouping of *m* consecutive years identified in Step 2, estimate the time trend using a first-order rate of decline equation. Calculate the deviation of this trend from the long-term trend estimated in Step 1.
4. Repeat Steps 1 to 3 for each value of *m*.
5. Plot the percent deviation as a function of *m*, along with the percent deviation of the confidence bounds on the long-term time trend.
6. Determine the minimum length of time series needed as the window size for which the estimated deviations are contained within the deviation of the 95 percent confidence bounds from the long-term mean trend.

The moving window method described above was performed using the pre-dredging data. The applicability of these results to post-dredging conditions was confirmed by comparing the variability in the pre- and post-dredging datasets. First, the data was log-transformed and then mean-centered on an annual basis. This was done to account for non-normality in the dataset and to remove any variability associated with year-to-year differences in the dataset. Next, the annual standard deviation of the log-transformed mean-centered data was calculated for the pre-dredging and post-dredging periods and compared qualitatively to assess whether the variability was similar across the two periods. Second, the variance of the two periods was compared quantitatively using the Levene Test of homogeneity of variance across the two dredging periods. The Levene Test was used because it is less sensitive to small departures from normality compared with the F-Test or Bartlett's Test. Prior to running the test, the log-transformed mean-centered data was pooled by dredging period to allow the variance of the two periods to be compared.

3.3 Method for Evaluation of PCB Loads to the Lower Hudson River

As previously stated, one of the RAOs for the OU2 remedy is to minimize the long-term downstream transport of PCBs to the Lower Hudson River. To accomplish this, EPA successfully implemented resuspension performance standards during the in-water remedial activities, as discussed in the Second Five-Year Review. To further evaluate the achievement of this RAO, the annual PCB loads at the Waterford station were estimated for the pre-dredging and post-dredging years and compared to confirm loads to the Lower Hudson River have been reduced.

3.3.1 Determination of Annual PCB Loads During Pre- and Post-Dredging Periods to the Lower Hudson River

Annual Tri+ PCB loads at the Waterford monitoring station were estimated using the USGS Load Estimator software package (LOADEST; Runkel et al., 2004). As will be described in Section 4.1.1, flow exerts a strong influence on the water column PCB concentration in the Upper Hudson River, with different relationships at higher and lower flows. Similarly, under routine flow conditions, seasonality in water column PCB concentrations in the Upper Hudson River is well documented, with generally higher concentrations seen during the summer months and lower concentrations during the colder months (EPA, 1997, 1999). A Site-specific rating curve equation was developed to incorporate both phenomena. A rating curve equation is a mathematical equation (commonly a linear regression equation) that describes the relationship between PCB load and river flow (and possibly other covariates). It can be used to estimate PCB loads for time periods when PCB samples were not collected. A detailed discussion of the LOADEST software and the development of the Site-specific rating curve equation used to calculate PCB loads at Waterford is presented in the *Tri+ PCB Load Calculation Technical Memorandum* (Attachment A). The Site-specific rating curve equation, along with the 15-minute flow and temperature data, were used as inputs into the LOADEST software to generate annual PCB loads at Waterford. See Section 4.2.1 for the results of this analysis.

3.3.2 Normalizing Post-Dredge Loads to the Lower Hudson River to Account for Flow and Temperature

Identifying changes in PCB loads attributed to changes in PCB concentration can be difficult, as it is possible that water column PCB concentrations are declining under normal conditions, but that year-to-year variability in flows (and possibly other covariates that influence water column PCB concentrations) may obscure this decline, resulting in annual PCB loads that may appear stationary or appear to increase. To properly identify changes in PCB loads attributed to changes in PCB concentration, variability in annual loads that are a result of year-to-year variability in flow (and possibly other covariates) needs to be removed so that changes in annual load estimates reflect changes in the water column PCB concentration. This can be accomplished using a normalization procedure based on Hirsch and De Cicco (2015). The normalization procedure utilized in this appendix uses the site-specific rating curve equation developed for the LOADEST program with an additional variable to represent time. The modified site-specific rating curve equation is written as:

$$\ln(C) = b_o + b_1 Time + b_2 \ln(Q) + b_3\big(\ln(Q) - CP_Q\big)I_Q + b_4 Temp + \varepsilon \qquad (Eq. 2)$$

Where:

- C is the measured Tri+ PCB concentration;
- b_i are regression coefficients estimated by the equation;

- *Time* is the decimal date;
- *Q* is river flow;
- CP_Q is the estimated change point flow value using segmented regression;
- I_Q is an indicator variable that equals 1 when the measured flow is greater than CP_Q and zero when the flow is less than CP_Q;
- *Temp* is water temperature; and
- ε is the regression equation error term.

The normalization procedure starts with fitting Equation 2 (including estimation of CP_Q) to the full post-dredging period (2016 to 2021) dataset to derive the regression coefficients and change point value. Next, the equation is used to predict six concentration and flux values for each day of the prediction period (defined as January 1, 2016, to December 31, 2021). The number of predicted values calculated for each prediction day is based on the number of years in the prediction period (six years in this analysis).

The six predicted values for each prediction day are calculated using the following steps:

1. The time variable is set as January 1, 2016, as the starting date, the flow and temperature values for January 1, 2016, are input into the equation and a predicted concentration and flux are calculated.
2. The time variable is held constant as January 1, 2016, but the flow and temperature values for January 1, 2017, are input into the equation and a second predicted flux calculated.
3. The time variable is again held constant as January 1, 2016, but the flow and temperature values for January 1, 2018, are input into the equation and a third predicted flux calculated. This procedure is repeated for the remaining post-dredging years (i.e., 2019 to 2021).
4. The time variable is then advanced one day (January 2, 2016) and the above procedure is repeated using flows and temperature measured on January 2 of each of the post-dredging years.
5. After the six flux values for each day of the prediction period are calculated, the six values are averaged to produce a single flow-temperature-normalized (FTN) PCB flux value for each prediction day.
6. Finally, the daily FTN PCB flux values within each year are summed to obtain an annual FTN PCB load estimate for each post-dredging year.

By removing variability in the water column PCB loads introduced by flow and seasonality, Equation 2 can be used to evaluate changes in the FTN load through time. It should be noted that the FTN PCB load estimates are dependent on the number of years of data included in the analysis, and FTN PCB load estimates and time trend will vary if a longer data record is used in the FTN calculation.

4 RESULTS AND DISCUSSION

4.1 Evaluation of PCB Concentrations

4.1.1 Influence of Seasonality and River Flow on PCB Concentrations Through Time

When interpreting year-to-year variability in PCB concentration, it is important to recognize the processes that can influence the measured concentrations. Seasonality in water column PCB concentrations has been observed in previous studies of PCBs in the Upper Hudson River, with higher concentrations during the warmer summer months and lower concentrations during the colder fall and winter months under routine flow conditions (EPA, 1997, 1999). This observed seasonality is shown on Figures A1-4 through A1-8. The observed seasonality is likely a response to increased biological activity in the near-surface sediment during the warmer months, which enhances the release of sediment-bound PCBs to the water column. Thus, samples collected during the summer months under routine (non-high-flow) conditions would be expected to have higher concentrations than samples collected under similar flow conditions during the winter or early spring months. This phenomenon is further discussed in Section 4.1.3.

Previous studies also observed a “V” shaped, non-linear relationship between PCB concentration and flow in the Upper Hudson River (EPA, 1999). Figures A1-2 and A1-3 present the relationship between PCB concentration and flow for the post-dredging data collected at the Schuylerville and Waterford stations, respectively. These are the two stations where both routine and high-flow samples are collected. Under lower flow conditions, the relationship between concentration and flow is negative with increasing flow values resulting in decreasing concentrations; however, above a certain flow threshold the relationship becomes positive with increasing flow values resulting in increasing concentrations. This non-linear relationship between concentration and flow is likely reflective of a dilution-dominated flow regime at relatively low flows and a resuspension-dominated flow regime at higher flows. As a result of the non-linear relationship between PCB concentration and flow, flow conditions within a given year can influence the PCB concentrations under both routine and high-flow sampling programs. The impact of this flow on observed concentrations is discussed in more detail in Section 4.1.3.

4.1.2 PCB Concentrations Through Time: 2004 to 2021

Figures A1-4 to A1-8 present scatter plot time series of water column Tri+ PCB and TPCB concentrations. The scatterplots display the individual data points and show the variability in PCB concentration both seasonally and during high-flow events. To aid in the comparison of changes in the PCB concentration across years and dredging periods, the geometric mean for each year was plotted in Figures A1-9 to A1-13. As discussed previously in Section 3.2.1, the geometric mean plots only include routine water column data. Summary statistics, including both the arithmetic and geometric means, for pre- and post-dredging period PCB samples collected at the five monitoring stations are presented in Tables A1-2 to A1-5.

The water column Tri+ PCB and TPCB concentrations at monitoring stations upstream of the project area (Bakers Falls and Rogers Island) exhibit temporal changes that differ from stations within the project area. At the Bakers Falls monitoring station, the water column Tri+ PCB concentrations exhibit an increase between 2004 and 2014, from approximately 0.1 ng/L to 1 ng/L. However, beginning in 2015, Tri+ PCB exhibits a decline from approximately 1 ng/L to less than 0.1 ng/L in 2021. It is noted that PCB at Bakers Falls do not exhibit the same increase as Tri+ PCB between 2004 and 2014, and remains consistent at approximately 1 ng/L. However, starting in 2015, TPCB concentrations exhibit a decline from approximately 1 ng/L to less than 0.1 ng/L, similar to Tri+ PCB. A review of the PCB analytical methods used for Bakers Falls samples indicates that the mGBM analytical method was modified in 2011 and EPA Method 1668C was implemented in 2017; however, these changes do not explain the observed temporal variation, particularly the decline in concentrations since 2015. Bakers Falls is located upstream of known GE-related PCB releases and the temporal variations in PCB concentrations may therefore reflect changes to inputs of PCBs from upstream sources or regional background sources. Nonetheless, the absolute PCB concentrations at this station are 2 to 3 orders of magnitude less than those observed in the project area and are well below the 2 ng/L value assumed in the 2002 ROD for Rogers Island following source control activities in the vicinity of the GE Hudson Falls Plant. Therefore, PCB concentrations originating upstream of the GE facilities have a negligible impact on concentrations observed downstream of Rogers Island.

At Rogers Island (background for OU2), there is a general downward trend in concentration over time in both Tri+ PCB and TPCB concentrations. This station is upstream of the remedial dredging effort and therefore, impacts of dredging would not be observed. The decrease over time at this station is likely attributable to continued remedial efforts performed by GE under New York State Department of Environmental Conservation (NYSDEC) at the plant sites and natural recovery of sediments located upstream of this station and the plant sites. It should be noted that the 2002 ROD anticipated that Tri+ PCB concentrations entering the project area would average 2 ng/L following source control activities in the vicinity of the GE's Hudson Falls Plant; this projected decline has been realized, with results for 95 percent of samples collected at Rogers Island during the post-dredging period being below that threshold.

For monitoring stations within the project area (Thompson Island Dam, Schuylerville, and Waterford), water column concentrations during the pre-dredging period (2004 to 2008) were elevated (relative to post-dredging) with annual geometric mean Tri+ PCB concentrations based on routine samples across the three stations typically between 8 ng/L to 14 ng/L and TPCB concentrations between 23 ng/L and 40 ng/L. During the dredging period (2009 to 2015), concentrations increased, as was expected, relative to pre-dredging levels, due to dredging related-resuspension of PCB-bearing sediment. During the post-dredging period (2016 to 2021), water column PCB concentrations decreased relative to both the pre-dredging and dredging period, with annual geometric mean Tri+ PCB concentrations less than 10 ng/L and TPCB concentrations typically less than 20 ng/L based on routine samples.

4.1.3 Routine Sampling Program PCB Concentrations: 2016 to 2021

Tables A1-2 to A1-5 present summary statistics for Tri+ PCB and TPCB concentrations for routine samples collected during the pre- and post-dredging period. Figures A1-9 to A1-13 present geometric means for Tri+ PCB and TPCB concentrations for routine samples for all three periods. Post-dredging results indicate that Bakers Falls, which is the most upstream station, had the lowest concentrations (annual geometric mean Tri+ PCB and TPCB concentrations <1 ng/L). The Rogers Island monitoring station, which is upstream of the project area but downstream of GE-related releases from its historical Hudson Falls and Fort Edward Plants, had the second lowest post-dredging concentrations (annual geometric mean Tri+ PCB and TPCB concentrations <2 ng/L). At the three monitoring stations within the project area, post-dredging water column Tri+ PCB and TPCB concentrations were elevated relative to the Bakers Falls and Rogers Island stations. The three stations within the OU2 study area (Thompson Island Dam, Schuylerville, and Waterford) had generally similar concentrations. The Schuylerville monitoring station consistently exhibited the highest annual geometric mean Tri+ PCB and TPCB concentrations of the three stations, with the exception of 2017, when Waterford had a slightly higher geometric mean Tri+ PCB concentration.

Year-to-year variability in routine sample concentrations in the Upper Hudson River can, in part, be attributed to year-to-year variation in flow given the relationship between PCB concentration and flow previously discussed in Section 4.1.1. For example, in a particularly dry year with below average flows, routine samples would be predominantly collected at flows on the far-left arm of the "V" shaped relationship shown in Figures A1-2 and A1-3. It would be expected these samples would have generally higher PCB concentrations than routine samples collected during a wetter year, when more of the routine samples would be collected at higher flows near the bottom of the "V" shaped relationship. At Schuylerville and Thompson Island Dam, where routine samples are predominantly collected during the warmer months (April to November), the lower flow years (2016, 2018, 2020; Figure A1-14) had the highest arithmetic mean concentrations (Table A1-4). This phenomenon is less apparent in the Waterford routine samples, likely in part due to routine samples being collected across all months and therefore seasonality is also exerting an influence on year-to-year variability in Waterford routine sample PCB concentrations. Natural variability in river flow and other factors controlling PCB concentrations will increase the year-to-year variability in the data and hence the number of years required to establish a meaningful time trend in the data. The need for additional years of data before meaningful time trends can be estimated in the water column data is described further in Section 4.1.6.

Figure A1-15 presents a comparison of geometric mean Tri+ PCB and TPCB concentrations for routine samples collected between May and November during the pre-dredging period and post-dredging period, plotted against long-term monitoring stations ordered from upstream to downstream. The figures demonstrate that the largest decreases in concentrations between the two periods occurred within the project area (Thompson Island Dam, Schuylerville, and Waterford). An ANCOVA regression comparing the pre- and post-dredging geometric means shown in Figure

A1-15 indicate the three stations within the project area all exhibit statistically significant declines in Tri+ PCB and TPCB concentrations. Finally, the figures also demonstrate that both pre- and post-dredging, concentrations exhibited a substantial increase as water moved through the Thompson Island Pool (the stretch of river between Rogers Island and Thompson Island Dam), with the post-dredging increase being notably smaller.

4.1.4 High-Flow Sampling Program PCB Concentrations: 2016 to 2021

Tables A1-4 and A1-5 present summary statistics for Tri+ PCB and TPCB concentrations measured in samples collected under the high-flow sampling program at Schuylerville and Waterford, respectively. As discussed in Section 4.1.1, above a certain flow value the relationship between PCB concentration and flow changes from a negative to a positive relationship, with higher flows associated with higher concentrations (Figures A1-2 and A1-3). The impact of flow on PCB concentration can also be observed in Figures A1-7 to A1-8, where the highest concentrations each year are typically associated with high-flow samples. It should be noted that while high-flow sampling targets flow events that exceed certain flow thresholds (see Section 2.2), high-flow sampling may begin at flows below these thresholds, particularly when sampling the early rising and late falling limbs of the storm hydrograph. For these samples, which may be collected toward the bottom of the "V" shaped relationship presented in Figures A1-2 and A1-3, PCB concentrations may not be substantially elevated relative to routine samples, particularly routine samples collected under low flow conditions. As a result, in some years annual summary statistics (such as the geometric mean) may not exhibit large differences between routine and high-flow samples; however, this is not indicative of changes to the relationship between PCB concentration and flow described above (Figure 2-2 of Attachment A for Tri+ PCB concentration versus flow plots for individual post-dredging years). A primary objective of the high-flow sampling program is to develop a robust annual relationship between PCB concentration and flow that can be used to understand PCB concentrations under high-flow conditions and provide meaningful estimates of PCB loads at Waterford. The non-linear relationship between PCB concentration and flow highlights the need to collect multiple high-flow samples across the range of observed flows and across multiple high-flow events to properly characterize the concentration-flow relationship.

4.1.5 Progress Towards Achieving the Water Column ARARs

Pre- and post-dredging TPCB concentrations in water column samples were compared to the most protective water column ARAR described in Section 3.2.1 (14 ng/L TPCB) (Table A1-6). Using all data collected during the pre-dredging period (2004 to 2008), the percentage of samples less than 14 ng/L at Thompson Island Dam, Schuylerville, and Waterford prior to dredging was 10 percent, 16 percent, and 18 percent, respectively. During this period, the maximum percentage of samples below 14 ng/L in a single year was 14 percent (2005), 24 percent (2006), and 33 percent (2008) for Thompson Island Dam, Schuylerville, and Waterford, respectively. For the post-dredging period (2016 to 2021), the percentage of samples less than 14 ng/L at Thompson Island Dam, Schuylerville, and Waterford was 76 percent, 44 percent, and 57 percent, respectively.

During the post-dredging period, the maximum percentage of samples below 14 ng/L in a single year for Thompson Island Dam, Schuylerville, and Waterford was 93 percent (2018), 71 percent (2017), and 75 percent (2021), respectively.

Recognizing that river flow influences water column PCB concentrations, and that the high-flow sampling program was only implemented at some locations and not consistently across the pre- and post-dredging periods, an assessment of the progress toward achieving the most protective water column ARAR using routine data only was also performed. When combining all the pre-dredging years (2004 to 2008), the total percentage of routine samples less than 14 ng/L at Thompson Island Dam, Schuylerville, and Waterford was 10 percent, 16 percent, and 16 percent, respectively. The maximum percentage of routine samples below 14 ng/L in a single year during the pre-dredging period for Thompson Island Dam, Schuylerville, and Waterford was 4 percent (2005), 24 percent (2006), and 23 percent (2007), respectively. When combining all the post-dredging years (2016 to 2021), the total percentage of routine samples less than 14 ng/L at Thompson Island Dam, Schuylerville, and Waterford was 76 percent, 40 percent, and 61 percent, respectively. The maximum percentage of routine samples below 14 ng/L in a single year during the post-dredging period for Thompson Island Dam, Schuylerville, and Waterford was 93 percent (2018), 76 percent (2017), and 82 percent (2021), respectively.

The substantial increase in the number of samples with TPCB concentrations less than 14 ng/L, achieving the most protective water column ARAR, between the pre-dredging and post-dredging period, whether using all data or just routine samples, indicates that positive progress is being made at the Site.

4.1.6 Data Requirements for Estimating Accurate Time Trends in Water Column Data

Figure A1-16 presents the results of the moving window analysis described in Section 3.2.3. At Thompson Island Dam and Schuylerville, for a given consecutive six-year grouping of pre-dredging data (EPA has six years of post-dredging data), the estimated time trend can vary approximately ±50 percent of the long-term time trend (based on the years 1998 to 2008). At Thompson Island Dam, only four of the 11 short-term trends based on six or seven years of data fell within the 95 percent confidence interval of the long-term time trend. However, for time trends calculated using eight or more years of data, all of the estimated time trends fell within the 95 percent confidence interval of the long-term time trend at Thompson Island Dam. At Schuylerville, only four of the 11 short-term time trends calculated using either six or seven years of consecutive data fell within the 95 percent confidence interval of the long-term time trend. Notably, for time trends calculated using eight or more years of data, seven of the nine short-term time trends fell within the 95 percent confidence interval of the long-term time trend at Schuylerville, with at least nine years of data needed for all short-term time trends to fall within the 95 percent confidence interval of the long-term trend. These results suggest that at least two more years of data (i.e., at least eight years of data compared with the six years of post-dredging data currently available), are necessary to assess trends in the water column data that accurately reflect the true, long-term trend.

To confirm that it is appropriate to apply the moving window evaluation to the post-dredging data, an evaluation of the annual log-transformed mean-centered standard deviation between the pre-dredging and post-dredging datasets was conducted and a statistical test of the homogeneity of variance between the two dredging periods was performed. Table A1-7 presents the annual standard deviation of the log-transformed mean-centered Tri+ PCB dataset for Thompson Island Dam and Schuylerville. The similarity in the standard deviation for the pre- and post-dredging years is evident and supports the conclusion that variability in the data collected within the two dredging periods is similar. The Levene Test for homogeneity of variances between groups indicates that the variances between the pre-dredging and post-dredging datasets at both sites were not significantly different (for TID, F-value = 0.5662, p-value = 0.4522; for Schuylerville, F-value = 2.615, p-value = 0.1065). These results indicate that this analysis is applicable to the post-dredging period. The results of the moving window analysis presented here align with those presented in the Second Five-Year Review Comment Response (EPA, 2019b) and support the conclusion that eight or more years of routine water data are needed before meaningful time trends can be calculated.

Given the intra-annual and inter-annual fluctuations in concentration data which result in part from variability in covariates like flow and temperature, long-term datasets are required to reliably estimate time trends. Insufficient or short-term datasets can result in unreliable and inaccurate estimates and may result in premature conclusions on recovery and when certain goals will be achieved. Furthermore, because a first-order rate of decline equation is used to estimate time trends, time trends based on a small number of years are sensitive to the starting and ending concentrations. This explains why the greatest variability is observed for shorter time periods and the variability exponentially decreases as the window size increases and more data is available for estimating the trend. Using data from the Great Lakes region, Gewurtz et al. (2011) analyzed datasets from different contaminant monitoring programs and demonstrated that more than 10 years of data appear optimal for use in estimating time trends and this amount of data was less sensitive to the starting and ending concentrations. Furthermore, the authors found that shorter-term datasets could provide inaccurate and premature trend estimates that exhibit decreasing, increasing, or no significant trends depending on the starting and ending concentrations.

4.2 Evaluation of PCB Load to the Lower Hudson River

4.2.1 PCB Loads at Waterford for the Pre- and Post-Dredging Periods

As stated previously in Section 1.2.3, one of the RAOs for the OU2 remedy is to minimize the long-term downstream transport of PCBs to the Lower Hudson River. As previously discussed in Section 3.3, the Waterford water column station is used to monitor the PCB concentration and load transported through the Upper Hudson River to the Lower Hudson River.

Table A1-8 presents the annual water column Tri+ PCB load at Waterford monitoring station for the pre-dredging and post-dredging periods. Note that the water column PCB load for 2004 was not estimated, since BMP water column sampling started in June of that year. During the pre-dredging period, annual Tri+ PCB loads ranged from 94 kilograms (kg) in 2007 to 150 kg in 2006. During the post-dredging period, annual Tri+ PCB loads ranged from 34 kg in 2020 to 101 kg in 2019. As expected, annual loads in part reflect the magnitude of flows within a year – 2019 was an above average flow year (median flow of 11,500 cfs), while 2016 was a below average flow year (median flow of 4,885 cfs). The long-term (2000 to 2021) median flow at the Waterford monitoring station is 8,000 cfs (Figure A1-17). Figure A1-18 presents summary information on flows at the Fort Edward gauge station for the years 2016 to 2021. Although not used to calculate load at Waterford, flows at Fort Edward represent flows in the upper sections of the river and are important in understanding the hydrodynamics within the Upper Hudson River. While year-to-year variability exists in both pre- and post-dredging period PCB loads, results indicate that PCBs loads during the post-dredging period are lower than the pre-dredging period, consistent with observed changes in water column PCB concentration between the two periods as described in Section 4.1.3. By removing variability in the water column Tri+ PCB loads introduced by flow and seasonality, the results indicate FTN Tri+ PCB loads to the Lower Hudson River are decreasing (Figure A1-19). However, even accounting for the variability introduced by flow and temperature, additional years of data are necessary to confidently estimate the trend in Tri+ PCB loads at the Waterford station.

The load estimation program (LOADEST) used to generate annual PCB loads can also generate estimates of PCB load on a daily to sub-daily timestep, which allows for an evaluation of the load of PCBs transported during high-flow events. Using the concept of the half-load discharge value ($Q_{1/2}$), or the flow value above which 50 percent of the load over a given period of time is transported (Vogel et al., 2003), results indicated that 50 percent of the PCB load between 2016 and 2021 was transported over only 17 percent of the days. This highlights the significant impact that high-flow events have on the overall load to the Lower Hudson River, as substantial PCB transport occurs during these events. Complete details on the estimation of PCB loads during high-flow events are presented in the *Tri+ PCB Load Calculation Technical Memorandum* [Attachment A].

5 CONCLUSIONS

Major conclusions of this appendix are as follows:

Water Column PCB Concentrations (2016 to 2021):

- For monitoring stations upstream of the OU2 project area (Bakers Falls and Rogers Island), annual geometric mean Tri+ PCB concentrations were less than or equal to 1 ng/L, and annual geometric mean TPCB concentrations were less than or equal to 1.5 ng/L. Ninety-five percent of samples collected at Rogers Island had water column Tri+ PCB concentrations less than 2 ng/L, which was the anticipated Tri+ PCB concentration entering the OU2 study area following source control activities in the vicinity of GE's Hudson Falls Plant. This value is discussed in the 2002 ROD.

- For monitoring stations within the OU2 study area (Thompson Island Dam, Schuylerville, and Waterford), post-dredging water column PCB concentrations decreased relative to both the pre-dredging and dredging period, with annual geometric mean Tri+ PCB concentrations less than 10 ng/L, and TPCB concentrations typically less than 20 ng/L.

- The percentage of post-dredging samples with TPCB concentrations less than 14 ng/L (the most protective water column ARAR) at Thompson Island Dam, Schuylerville, and Waterford was 76 percent, 44 percent, and 57 percent, respectively. These values are substantially improved compared to the pre-dredging period, when only 10 percent, 16 percent, and 18 percent of samples were less than 14 ng/L at Thompson Island Dam, Schuylerville, and Waterford stations, respectively. The maximum percentage of samples below 14 ng/L in a single year for Thompson Island Dam, Schuylerville, and Waterford was 93 percent (2018), 71 percent (2017), and 75 percent (2021), respectively.

- Statistical analysis indicates the current six years of water column data are insufficient to accurately determine the rate at which water column concentrations are recovering or not. Using 11 consecutive years of pre-dredging data from Thompson Island Dam and Schuylerville, it was determined that eight or more years of data are needed to estimate time trends that reflect the long-term time trends in PCB concentration. When using only six years of data (the current number of years of post-dredging data), time trends exhibit substantial variability (as measured by deviation from the long-term time trend), with trend estimates falling well outside the 95 percent confidence interval of the long-term time trend. Thus, to determine a meaningful and accurate time trend in routine water column PCB concentrations, at least eight years of routine water column data are needed. The results of this analysis are consistent with results from the Second Five-Year Review Comment Response (EPA, 2019b) using pre-dredging fish tissue data.

Water Column PCB Loads to the Lower Hudson River (2016 to 2021):

- Annual Tri+ PCB loads ranged from 34 kg in 2020 to 101 kg in 2019. As expected, annual loads in part reflect the magnitude of flows within the year, with annual loads higher in years with higher median flows. PCBs loads during the post-dredging period are lower than the pre-dredging period, consistent with observed changes in water column PCB concentrations. Finally, by removing variability in the water column Tri+ PCB loads introduced by flow and seasonality, the preliminary results indicate that Tri+ PCB loads to the Lower Hudson River are decreasing.

6 ABBREVIATIONS AND ACRONYMS

ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
ARAR	Applicable or Relevant and Appropriate Requirement
BMP	Baseline Monitoring Program
CCC	Criteria Continuous Concentration
cfs	Cubic Feet Per Second
DSR	Data Summary Report
EDI	Equal Discharge Increment
EPA	United States Environmental Protection Agency
FTN	Flow-Temperature-Normalized
FWQC	Federal Water Quality Criterion
GE	General Electric
kg	Kilogram
LOADEST	USGS Load Estimator Program
MADIS	Multiple Aliquot Depth Integrating Sampler
MCL	Maximum Contaminant Level
mGBM	Modified Green Bay Method
ng/L	Nanogram Per Liter
NYSDEC	New York State Department of Environmental Conservation
OM&M	Operation, Maintenance and Monitoring
OU	Operable Unit
PCB	Polychlorinated Biphenyls
QA/QC	Quality Assurance/Quality Control
RAMP	Remedial Action Monitoring Program
RAO	Remedial Action Objectives
RM	River Mile
ROD	Record of Decision
RS	River Section

Site	Hudson River PCBs Superfund Site
TPCBs	Total Polychlorinated Biphenyls
Tri+ PCBs	Sum of all measured PCB congeners with three or more chlorine atoms per molecule.
USGS	United States Geological Survey

7 REFERENCES

EPA (United States Environmental Protection Agency), 1997. Phase 2 Report – Review Copy, Further Site Characterization and Analysis, Volume 2C – Data Interpretation and Evaluation Report Hudson River PCBs Reassessment RI/FS. Prepared for United States Environmental Protection Agency Region 2 and United States Army Corps of Engineers by TAMs Consultants, the Cadmus Group and Gradient. February.

_________. 1999. Phase 2 Report Review Copy - Further Site Characterization and Analysis, Volume 2D - Baseline Modeling Report, Hudson River PCBs Reassessment RI/FS. Prepared for United States Environmental Protection Agency Region 2 and United States Army Corps of Engineers Kansas City District by Limno-Tech, Menzie-Cura & Associates and Tetra-Tech. May.

_________. 2002. Hudson River PCBs Site New York. Record of Decision. February.

_________. 2004. Engineering Performance Standards: Technical Basis and Implementation of the Residuals Standard, Volume 3 of 5. Prepared for the United States Army Corps of Engineers by Malcolm Pirnie and TAMS Consultants. April.

_________. 2010. Hudson River PCBs Site. Revised Engineering Performance Standards for Phase 2 Dredging. Prepared for the United States Army Corps of Engineers by the Louis Berger Group. December.

_________. 2010a. Hudson River PCBs Site EPA Phase 1 Evaluation Report. Prepared for the United States Environmental Protection Agency Region 2 and United States Army Corps of Engineers by the Louis Berger Group. March.

_________. 2012. First Five-Year Review Report for Hudson River PCBs Superfund Site. June 1.

_________. 2019a. Final Second Five-Year Review Report for the Hudson River PCBs Superfund Site. April.

_________. 2019b. Final Second Five-Year Review Report Comment Response for the Hudson River PCBs Superfund Site, Appendix B Deferral Statement - Supporting Technical Information. Prepared for the United States Environmental Protection Agency by Louis Berger US. April.

General Electric (GE). 2004. Quality Assurance Project Plan for the Hudson River PCBs Site. Baseline Monitoring Program. Prepared for General Electric by Quantitative Environmental Analysis and Environmental Standards. May.

_________. 2005. Hudson River PCBs Site Baseline Monitoring Report Data Summary Report for 2004. Prepared for General Electric by Anchor QEA and Environmental Standards, April.

_________. 2006a. Hudson River PCBs Site Phase 1 Remedial Action Monitoring Program Quality Assurance Project Plan. Prepared for General Electric by QEA, Environmental Standards, Arcadis and TRC. December.

_________. 2006b. Hudson River PCBs Site Baseline Monitoring Report Data Summary Report for 2005. Prepared for General Electric by Anchor QEA and Environmental Standards. March.

_________. 2007. Hudson River PCBs Site Baseline Monitoring Report Data Summary Report for 2006. Prepared for General Electric by Anchor QEA and Environmental Standards. March.

_________. 2008. Hudson River PCBs Site Baseline Monitoring Report Data Summary Report for 2007. Prepared for General Electric by Anchor QEA and Environmental Standards. March.

_________. 2009a. Hudson River PCBs Site Phase 1 Remedial Action Monitoring Program Quality Assurance Project Plan. Prepared for General Electric by Anchor QEA, Environmental Standards, and Arcadis. May.

_________. 2009b. Hudson River PCBs Site Baseline Monitoring Report Data Summary Report for 2008. Prepared for General Electric by Anchor QEA and Environmental Standards. March.

_________. 2010. 2009 Data Summary Report, Hudson River Water and Fish, Hudson River PCBs Superfund Site. Prepared for General Electric by Anchor QEA and Environmental Standards. May.

_________. 2011a. Hudson River PCBs Site 2011 Remedial Action Monitoring Program Quality Assurance Project Plan. Prepared for General Electric by Anchor QEA and Environmental Standards. April.

_________. 2011b. 2010 Data Summary Report Hudson River Water and Fish, Hudson River PCBs Superfund Site, Prepared for General Electric by Anchor QEA and Environmental Standards. April.

_________. 2012a. Hudson River PCBs Site Phase 2 Remedial Action Monitoring Program Quality Assurance Project Plan. Prepared for General Electric by Anchor QEA and Environmental Standards. May.

_________. 2012b. 2011 Data Summary Report Hudson River Water and Fish, Hudson River PCBs Superfund Site, Prepared for General Electric by Anchor QEA. April.

_________. 2013. 2012 Data Summary Report Hudson River PCBs Superfund Site. Prepared for General Electric by Anchor QEA and Environmental Standards. March.

_________. 2014. 2013 Data Summary Report, Hudson River PCBs Superfund Site. Prepared for General Electric by Anchor QEA and Environmental Standards. March.

_________. 2015. 2014 Data Summary Report, Hudson River PCBs Superfund Site. Prepared for General Electric by Anchor QEA and Environmental Standards. March.

_________. 2016a. 2015 Surface Water Data Summary Report, Hudson River PCBs Superfund Site. Prepared for General Electric by Anchor QEA and Environmental Standards. March.

_________. 2016b. Water Column Tri+ PCB Correlations Refined with 2015 Data. Remedial Action Monitoring Program CAM No. 11. Prepared for General Electric by Anchor QEA. January.

_________. 2017. Hudson River PCBs Superfund Site, 2016 Water and Fish Data Summary Report. Prepared for General Electric by Anchor QEA and Environmental Standards. March.

_________. 2019a. 2017 Water and Fish Data Summary Report. Prepared for General Electric by Anchor QEA and Environmental Standards. August.

_________. 2019b. 2018 Water and Fish Data Summary Report. Prepared for General Electric by Anchor QEA and Environmental Standards. November.

_________. 2020. 2019 Water and Fish Data Summary Report. Prepared for General Electric by Anchor QEA and Environmental Standards. March.

_________. 2021. 2020 Water and Fish Data Summary Report. Prepared for General Electric by Anchor QEA and Environmental Standards. March.

_________. 2022. 2021 Water, Fish and Sediment Data Summary Report. Prepared for General Electric by Anchor QEA and Environmental Standards. September.

Gewurtz S.B., S.M. Backus S.P. Bhavsar, D.J. McGoldrick, S.R. de Solla, and E.W. Murphy. 2011. "Contaminant biomonitoring programs in the Great Lakes region: Review of approaches and critical factors." *Environmental Reviews*. 2011; 19:162-184. https://doi.org/10.1139/a11-005

Hirsch, R.M., and L.A. De Cicco. 2015. "User guide to Exploration and Graphics for RivEr Trends (EGRET) and dataRetrieval: R packages for hydrologic data" (Report No. 4-A10), *Techniques and Methods*. https://doi.org/10.3133/tm4A10.

Runkel, R.L., C.G. Crawford, and T.A. Cohn. 2004. Load estimator (LOADEST): a FORTRAN program for estimating constituent loads in streams and rivers (Report No. 4-A5), *Techniques and Methods*. https://doi.org/10.3133/tm4A5.

APPENDIX 1

Table and Figures

Tables

Table A1-1
Summary Sample Counts for 2016 to 2021 Routine and High Flow Sampling Programs

	Count of Routine PCB Water Column Samples																	
	Pre-Dredging					Dredging Period							Post-Dredging Period					
Station Name	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Bakers Falls	29	44	49	43	28	10	10	7	12	8	9	9	8	7	7	6	7	8
Rogers Island	28	39	46	40	36	41	35	31	7	11	7	11	8	7	7	6	7	8
Thompson Island Dam	26	35	35	31	37	279	54	201	24	25	6	13	38	29	28	20	27	31
Schuylerville	28	43	50	42	39	238	54	201	227	57	134	171	42	29	28	24	27	31
Waterford	28	42	50	43	40	237	56	197	226	221	208	186	49	43	43	43	53	45

	Count of High Flow PCB Water Column Samples																	
	Pre-Dredging					Dredging Period							Post-Dredging Period					
Station Name	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021
Bakers Falls	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Rogers Island	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
Thompson Island Dam	-	-	-	-	-	-	43	19	-	-	-	-	-	-	-	-	-	-
Schuylerville	-	-	-	-	-	-	47	19	-	2	-	-	-	26	14	36	3	10
Waterford	-	23	29	34	36	6	43	19	-	2	11	-	3	21	20	51	7	10

Notes:

1. Pre-dredging period sample counts include pseudo-time of travel study samples. Dredging period sample counts do not include samples collected for purposes other than far-field monitoring of PCBs (e.g., non-compliance and diagnostic sampling). Therefore, counts may differ slightly from counts presented in the annual Data Summary Reports.
2. Parent-duplicate pairs are counted as one sample.

Table A1-2
Summary Statistics for Routine PCB Concentration at the Bakers Falls and Rogers Island Monitoring Stations

		BAKERS FALLS ROUTINE SAMPLE SUMMARY STATISTICS									
		Count of Samples		Tri+ PCB (ng/L)				Total PCB (ng/L)			
Dredging Period	Year	Detects	Non-Detect	Minimum	Maximum	Mean	Geometric Mean	Minimum	Maximum	Mean	Geometric Mean
Pre-Dredging	2004	29	0	0.01	6.1	0.51	0.13	0.02	6.9	1.8	0.98
	2005	44	0	0.01	0.67	0.12	0.074	0.01	3.5	1.2	0.81
	2006	49	0	0.004	0.38	0.11	0.066	0.01	1.8	0.74	0.55
	2007	43	0	0.10	0.5	0.15	0.11	0.12	2.2	1.1	0.97
	2008	28	0	0.01	1.4	0.26	0.14	0.02	2.9	1.1	0.90
Post-Dredging	2016	8	0	0.32	1.1	0.66	0.62	0.44	1.4	0.83	0.78
	2017	7	0	0.25	1.7	0.65	0.52	0.27	1.9	0.71	0.56
	2018	7	0	0.06	0.68	0.35	0.28	0.06	0.78	0.41	0.32
	2019	6	0	0.08	0.22	0.15	0.14	0.09	0.26	0.16	0.15
	2020	7	0	0.06	0.43	0.17	0.14	0.06	0.46	0.19	0.15
	2021	8	0	0.01	0.17	0.09	0.06	0.01	0.17	0.09	0.07

		ROGERS ISLAND ROUTINE SAMPLE SUMMARY STATISTICS									
		Count of Samples		Tri+ PCB (ng/L)				Total PCB (ng/L)			
Dredging Period	Year	Detects	Non-Detect	Minimum	Maximum	Mean	Geometric Mean	Minimum	Maximum	Mean	Geometric Mean
Pre-Dredging	2004	28	0	0.35	6.6	2.3	1.8	0.87	8.4	4.2	3.7
	2005	39	0	0.07	3.9	1.7	1.3	0.87	6.0	3.4	2.9
	2006	46	0	0.15	7.4	1.4	1.1	0.97	9.8	2.4	2.1
	2007	40	0	0.44	4.6	1.6	1.3	1.1	6.7	3.1	2.8
	2008	36	0	0.48	27	3.3	2.0	0.84	28	4.8	3.6
Post-Dredging	2016	8	0	0.18	1.8	0.80	0.67	0.40	3.7	1.2	0.97
	2017	7	0	0.38	0.81	0.55	0.54	0.51	0.94	0.68	0.67
	2018	7	0	0.34	3.7	1.4	1.0	0.68	4.3	1.8	1.5
	2019	6	0	0.52	1.8	1.1	0.99	0.85	2.3	1.5	1.4
	2020	7	0	0.23	1.8	0.97	0.74	0.39	2.5	1.4	1.1
	2021	8	0	0.18	1.1	0.50	0.43	0.38	1.5	0.81	0.74

Notes:
1. Non-detect sample results are omitted from statistical analyses.
2. Parent and duplicate samples are only counted once, and the average result is used in the statistical analyses.
3. For duplicate samples with only one detect between the parent and duplicate, the sample is counted as detected and the detected result is used in the statistical analyses.

Table A1-3
Summary Statistics for Routine PCB Concentration at the Thompson Island Dam Monitoring Station

		THOMPSON ISLAND DAM ROUTINE SAMPLE SUMMARY STATISTICS									
		Count of Samples		Tri+ PCB (ng/L)				Total PCB (ng/L)			
Dredging Period	Year	Detects	Non-Detect	Minimum	Maximum	Mean	Geometric Mean	Minimum	Maximum	Mean	Geometric Mean
Pre-Dredging	2004	26	0	3.8	27	14	13	9.8	84	44	40
	2005	35	0	1.1	26	11	8.8	5.2	78	33	27
	2006	35	0	1.9	43	11	9.0	11	95	34	30
	2007	31	0	2.9	30	12	10	9.2	120	40	34
	2008	37	0	1.8	84	13	9.9	9.6	140	36	30
Post-Dredging	2016	38	0	1.5	37	6.1	4.8	2.4	52	12	10
	2017	29	0	0.80	24	5.2	3.9	1.6	33	8.3	6.5
	2018	28	0	1.6	18	5.7	5.1	3.7	28	11	10
	2019	20	0	2.1	9.2	5.7	5.3	4.9	31	15	13
	2020	27	0	2.8	12	7.1	6.7	6.6	50	16	15
	2021	31	0	0.95	16	4.6	3.8	2.8	37	11	8.9

Notes:
1. Non-detect sample results are omitted from statistical analyses.
2. Parent and duplicate samples are only counted once, and the average result is used in the statistical analyses.
3. For duplicate samples with only one detect between the parent and duplicate, the sample is counted as detected and the detected result is used in the statistical analyses.

Table A1-4
Summary Statistics for Routine and High Flow PCB Concentration at the Schuylerville Monitoring Station

		SCHUYLERVILLE ROUTINE SAMPLE SUMMARY STATISTICS									
		Count of Samples		Tri+ PCB (ng/L)				Total PCB (ng/L)			
Dredging Period	Year	Detects	Non-Detect	Minimum	Maximum	Mean	Geometric Mean	Minimum	Maximum	Mean	Geometric Mean
Pre-Dredging	2004	28	0	1.9	28	16	14	6.4	84	46	39
	2005	43	0	0.34	34	13	8.9	0.85	81	33	24
	2006	50	0	0.76	37	12	8.6	2.4	82	32	25
	2007	42	0	2.1	32	12	11	8.3	122	36	30
	2008	39	0	1.9	48	14	11	6.0	75	34	29
Post-Dredging	2016	39	3	2.6	26	7.9	6.9	3.4	62	17	15
	2017	29	0	2.0	17	6.8	6.2	4.1	36	12	11
	2018	28	0	3.0	14	8.5	8.0	5.6	38	17	16
	2019	24	0	3.0	12	8.3	7.8	7.0	32	19	17
	2020	27	0	4.5	16	11	10	9.7	33	23	22
	2021	31	0	2.9	15	7.3	6.6	7.3	29	15	14

		SCHUYLERVILLE HIGH FLOW SAMPLE SUMMARY STATISTICS									
		Count of Samples		Tri+ PCB (ng/L)				Total PCB (ng/L)			
Dredging Period	Year	Detects	Non-Detect	Minimum	Maximum	Mean	Geometric Mean	Minimum	Maximum	Mean	Geometric Mean
Pre-Dredging	2004	-	-	-	-	-	-	-	-	-	-
	2005	-	-	-	-	-	-	-	-	-	-
	2006	-	-	-	-	-	-	-	-	-	-
	2007	-	-	-	-	-	-	-	-	-	-
	2008	-	-	-	-	-	-	-	-	-	-
Post-Dredging	2016	-	-	-	-	-	-	-	-	-	-
	2017	26	0	2.2	28	9.6	7.8	3.0	38	13	11
	2018	14	0	1.8	22	9.2	6.9	2.8	32	14	11
	2019	36	0	2.4	66	17	12	4.2	120	29	19
	2020	3	0	2.3	21	9.2	5.9	4.0	30	14	9.5
	2021	10	0	3.4	15	8.6	7.5	5.6	22	13	12

Notes:
1. No high flow samples were collect at Schuylerville during the pre-dredging period or in 2016.
2. Non-detect sample results are omitted from statistical analyses.
3. Parent and duplicate samples are only counted once, and the average result is used in the statistical analyses.
4. For duplicate samples with only one detect between the parent and duplicate, the sample is counted as detected and the detected result is used in the statistical analyses.

Table A1-5
Summary Statistics for Routine and High Flow PCB Concentration at the Waterford Monitoring Station

		WATERFORD ROUTINE SAMPLE SUMMARY STATISTICS									
		Count of Samples		Tri+ PCB (ng/L)				Total PCB (ng/L)			
Dredging Period	Year	Detects	Non-Detect	Minimum	Maximum	Mean	Geometric Mean	Minimum	Maximum	Mean	Geometric Mean
Pre-Dredging	2004	28	0	4.5	28	15	14	12	65	36	33
	2005	42	0	1.9	30	14	12	2	67	30	25
	2006	50	0	1.5	45	12	9.3	6.1	79	26	23
	2007	43	0	2.8	27	10	9.2	6.3	62	26	23
	2008	40	0	2.8	37	12	9.9	5.6	69	26	23
Post-Dredging	2016	44	5	2.6	16	7.3	6.8	3.2	36	15	13
	2017	43	0	1.9	13	7.3	6.4	2.7	19	11	10
	2018	43	0	2.2	13	6.7	5.9	3.6	24	12	10
	2019	43	0	1.8	13	6.3	5.4	3.1	22	12	9.8
	2020	53	0	2.0	11	6.1	5.2	3.7	26	12	10
	2021	45	0	1.3	10	5.2	4.5	3.6	21	10	9.4

		WATERFORD HIGH FLOW SAMPLE SUMMARY STATISTICS									
		Count of Samples		Tri+ PCB (ng/L)				Total PCB (ng/L)			
Dredging Period	Year	Detects	Non-Detect	Minimum	Maximum	Mean	Geometric Mean	Minimum	Maximum	Mean	Geometric Mean
Pre-Dredging	2004	-	-	-	-	-	-	-	-	-	-
	2005	23	0	4.4	94	35	28	10	120	50	43
	2006	29	0	7	150	35	26	9.7	260	51	37
	2007	34	0	4.8	44	19	16	8.9	68	28	24
	2008	36	0	3.4	126	17	12	6	150	21	16
Post-Dredging	2016	3	0	9.5	65	34	26	12	81	44	33
	2017	21	0	2.9	29	14	12	3.7	37	18	15
	2018	20	0	2.2	19	8.8	7.2	3.6	28	13	11
	2019	51	0	1.9	140	18	11	3.8	370	32	17
	2020	7	0	3.9	40	12	8.1	6.7	56	17	12
	2021	10	0	6.0	39	13	11	8.8	92	22	17

Notes:
1. No high flow samples were collect at Waterford during 2004.
2. Non-detect sample results are omitted from statistical analyses.
3. Parent and duplicate samples are only counted once, and the average result is used in the statistical analyses.
4. For duplicate samples with only one detect between the parent and duplicate, the sample is counted as detected and the detected result is used in the statistical analyses.

Table A1-6

Percentage of All Samples and Routine Samples with Total PCB Concentrations < 14 ng/L

	Percentage of All Samples <14 ng/L TPCB		
Year	Thomspson Island Dam	Schuylerville	Waterford
2004	3.8	7.1	7.1
2005	14	19	9
2006	11	24	13
2007	9.7	12	22
2008	11	13	33
Pre-Dredging Period	10	16	18
2016	76	44	55
2017	90	71	53
2018	93	36	63
2019	60	33	49
2020	52	10	55
2021	81	59	75
Post-Dredging Period	76	44	57

	Percentage of Routine Samples <14 ng/L TPCB		
Year	Thomspson Island Dam	Schuylerville	Waterford
2004	3.8	7.1	7.1
2005	14	19	12
2006	11	24	18
2007	9.7	12	23
2008	11	13	15
Pre-Dredging Period	10	16	16
2016	76	44	57
2017	90	76	53
2018	93	29	65
2019	60	25	56
2020	52	4	53
2021	81	58	82
Post-Dredging Period	76	40	61

Notes:

1. 14 ng/L TPCB represents the criteria continuous concentration (CCC) federal water quality criterion (FWQC) for protection of aquatic life in freshwater.
2. Non-detect values were excluded from the calculations.
3. Parent-Duplicate pairs are counted as one sample.

Table A1-7

Annual Variability in Routine Water Column Tri+ PCB Data Collected Between 1998 and 2008 and 2016 to 2021

	Thompson Island Dam				**Schuylerville**		
Year	Dataset Time Period	Number of Routine Samples Included in Analysis	Standard Deviation	Year	Dataset Time Period	Number of Routine Samples Included in Analysis	Standard Deviation
1998	Pre-Dredging Period	27	0.46	1998	Pre-Dredging Period	28	0.35
1999		30	0.60	1999		30	0.59
2000		28	0.33	2000		29	0.28
2001		31	0.31	2001		33	0.23
2002		27	0.43	2002		30	0.38
2003		6	0.80	2003		6	0.56
2004		26	0.46	2004		26	0.55
2005		30	0.65	2005		31	0.91
2006		31	0.66	2006		31	0.53
2007		31	0.53	2007		31	0.36
2008		30	0.71	2008		30	0.50
2016	Post-Dredging Period	31	0.57	2016	Post-Dredging Period	31	0.47
2017		28	0.69	2017		28	0.42
2018		28	0.48	2018		28	0.40
2019		20	0.42	2019		24	0.37
2020		27	0.37	2020		27	0.34
2021		28	0.60	2021		28	0.45

Notes:

1. Due to differences in sample collection frequency, pre- and post-dredging data included in this table are restricted to routine samples collected in the months of May to November. For data collected before 2004, data also restricted to samples collected when flows were less than 10,000 cfs at Ft. Edward gauging station (USGS #01327750).
2. Standard deviation is calculated using data that is first log-transformed, then mean centered on an annual basis. This is the same data handling procedure used for the Levene Test.

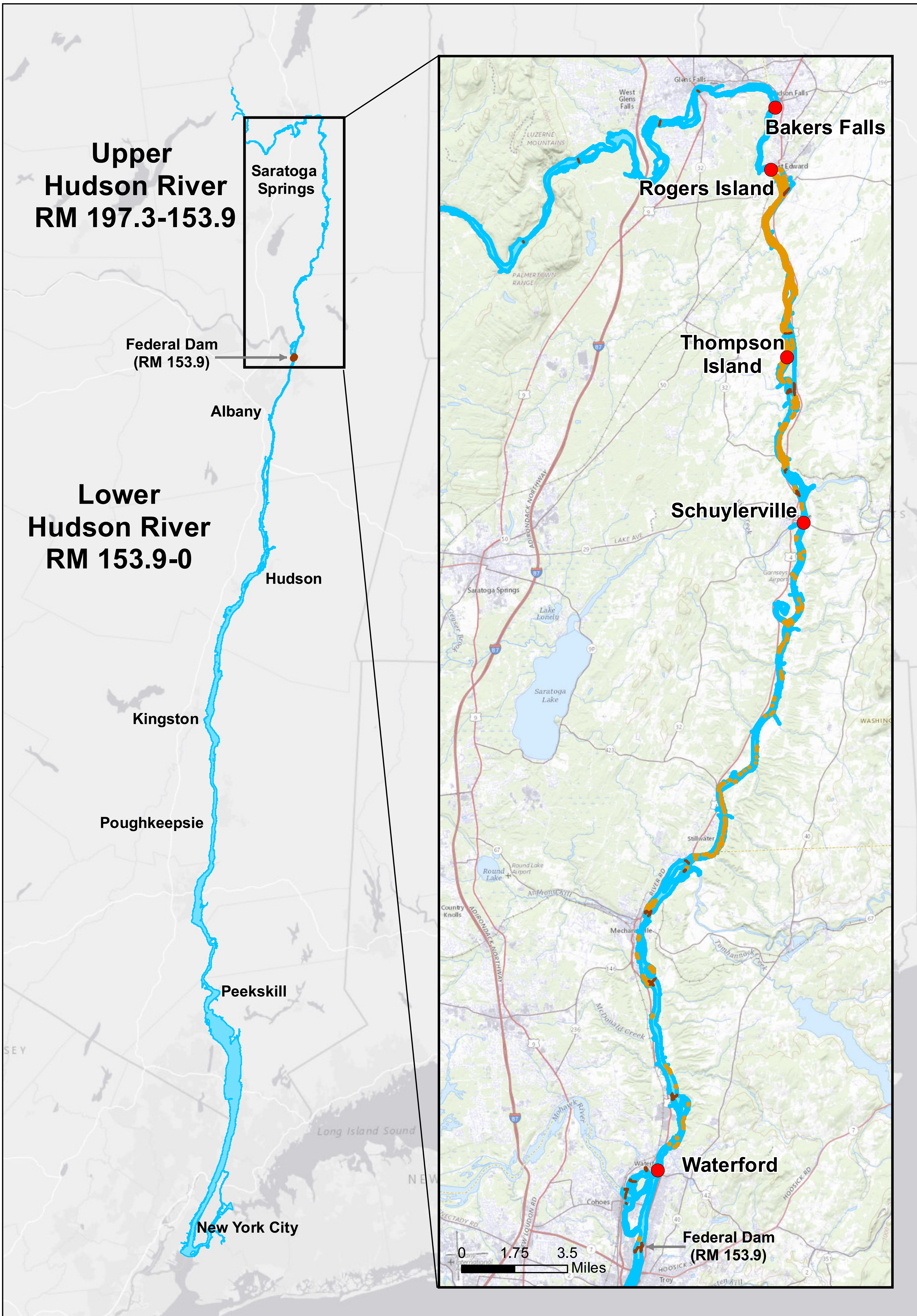
Table A1-8

Pre-Dredging and Post-Dredging Estimated Tri+ PCB load at Waterford

		Waterford Annual Tri+ PCB Load (kg)
Dredging Period	**Year**	**LOADEST Site-Specific Equation (15-Minute Flow)**
Pre-Dredging	2004	Sampling started in June 2004; no load calculated.
	2005	**130** (120, 150)
	2006	**150** (130, 170)
	2007	**94** (86, 100)
	2008	**130** (120, 150)
Post-Dredging	2016	**39** (36, 42)
	2017	**68** (63, 72)
	2018	**45** (43, 47)
	2019	**101** (90, 112)
	2020	**34** (32, 36)
	2021	**46** (44, 49)

Notes:
1. 95% confidence interval displayed in parentheses for LOADEST results.
2. All values reported to 2 significant digits.
3. For detailed description of the development of the LOADEST site-specific rating curve equation, see *Load Calculation Technical Memorandum* (Attachment A).

Figures



Legend

- Water Monitoring Location
- Dredging Area
- Hudson River
- Dam and Lock Locations

Coordinate System:
NAD 1983 StatePlane New York East FIPS 3101 Feet

July 2024

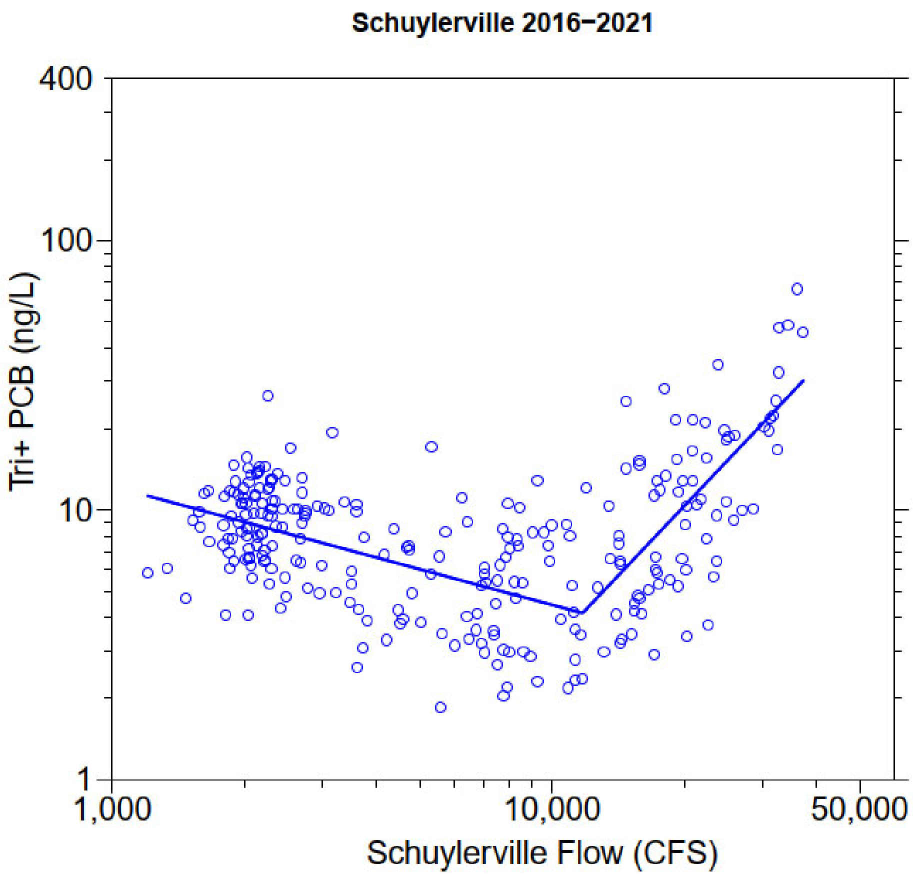
Figure A1-1
Hudson River Site Map with Water Column Monitoring Stations

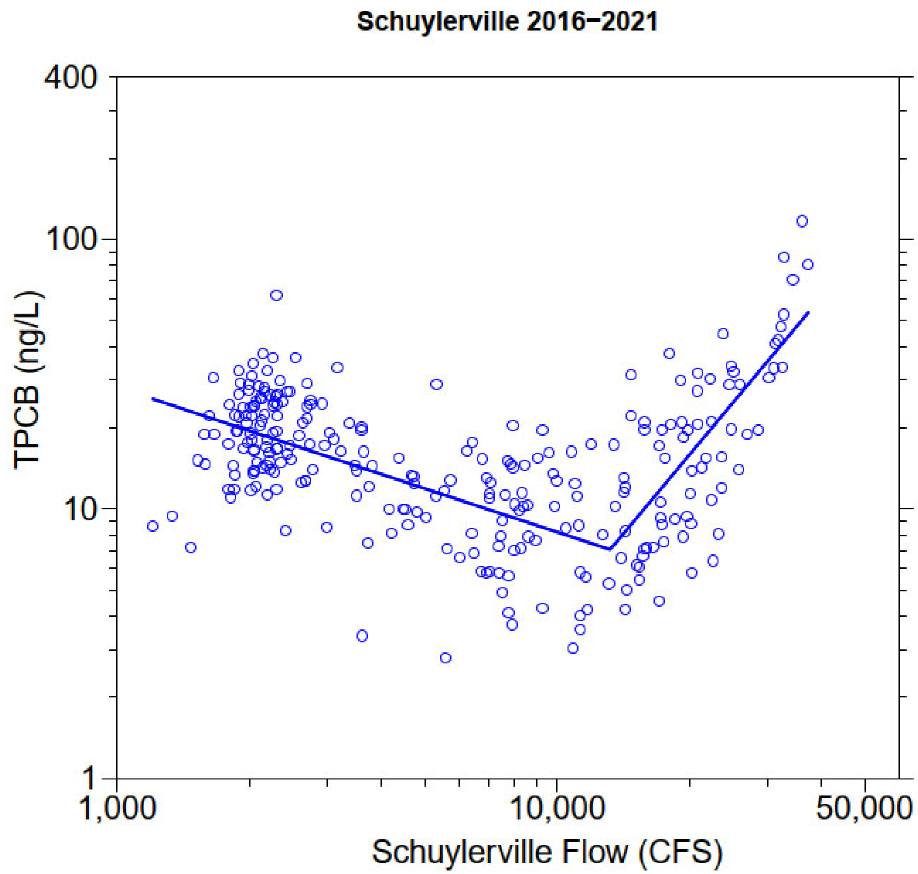
wsp

N

0 10 20 Miles

Date Saved: 2022/05/12
C:\Users\kgarvey\OneDrive - WSP O365\Desktop\UpperHudson Overview5-12-22 map.mxd;





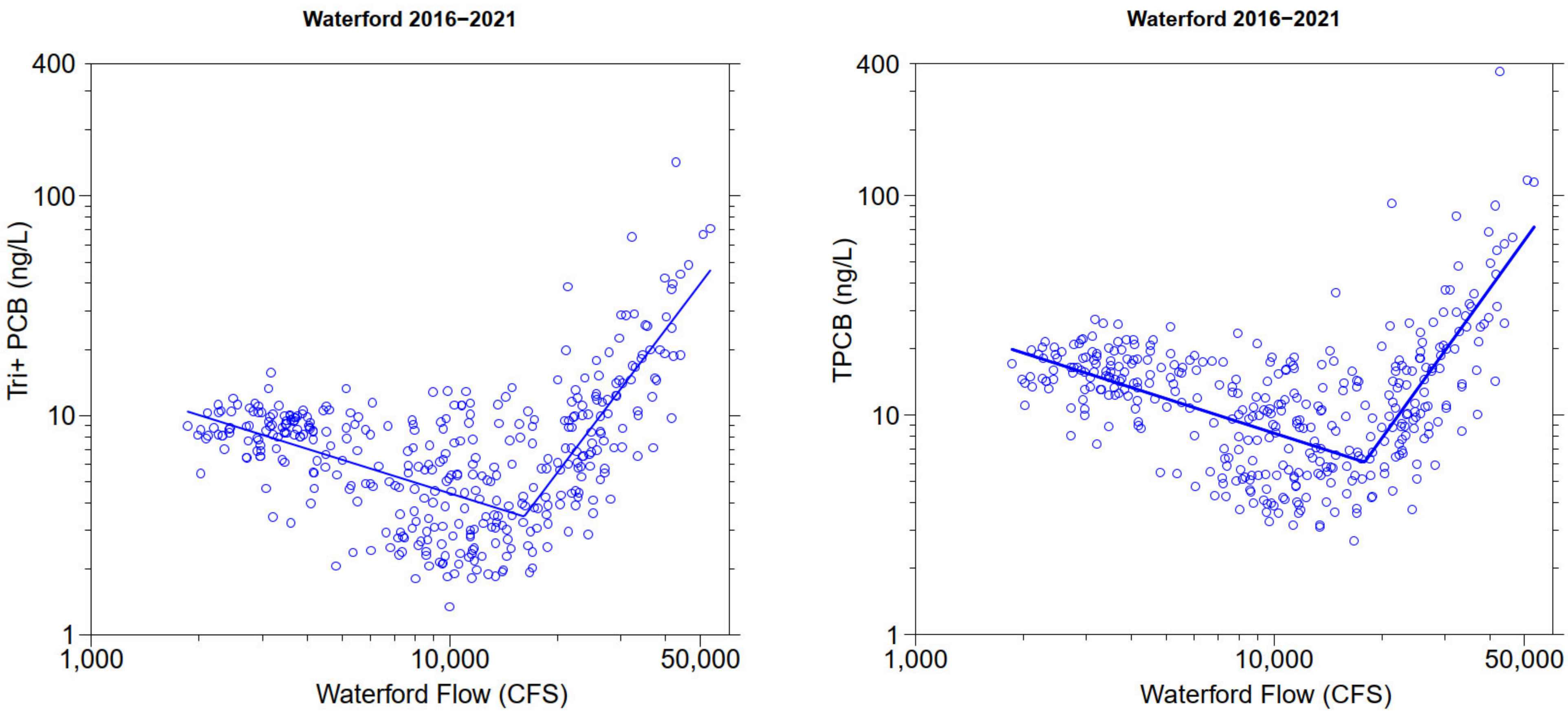
Notes:

- The regression line represents the segmented regression fit to the relationship between PCB concentration and river flow.

Relationship Between Post-Dredging PCB Concentration and River Flow at the Schuylerville Monitoring Station

Figure A1-2

July 2024



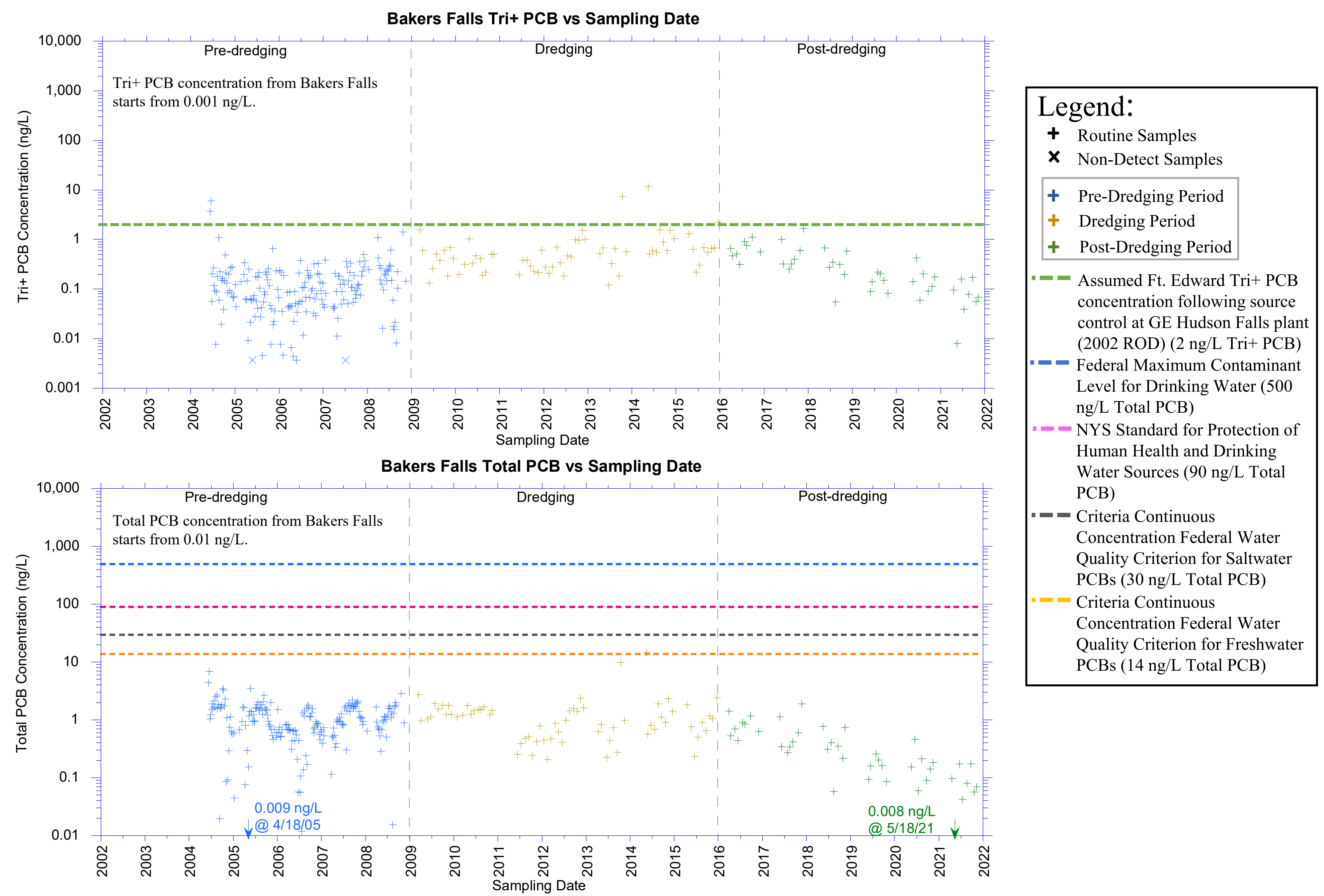
Notes:

- The regression line represents the segmented regression fit to the relationship between PCB concentration and river flow.

Relationship Between Post-Dredging PCB Concentration and River Flow at the Waterford Monitoring Station

Figure A1-3

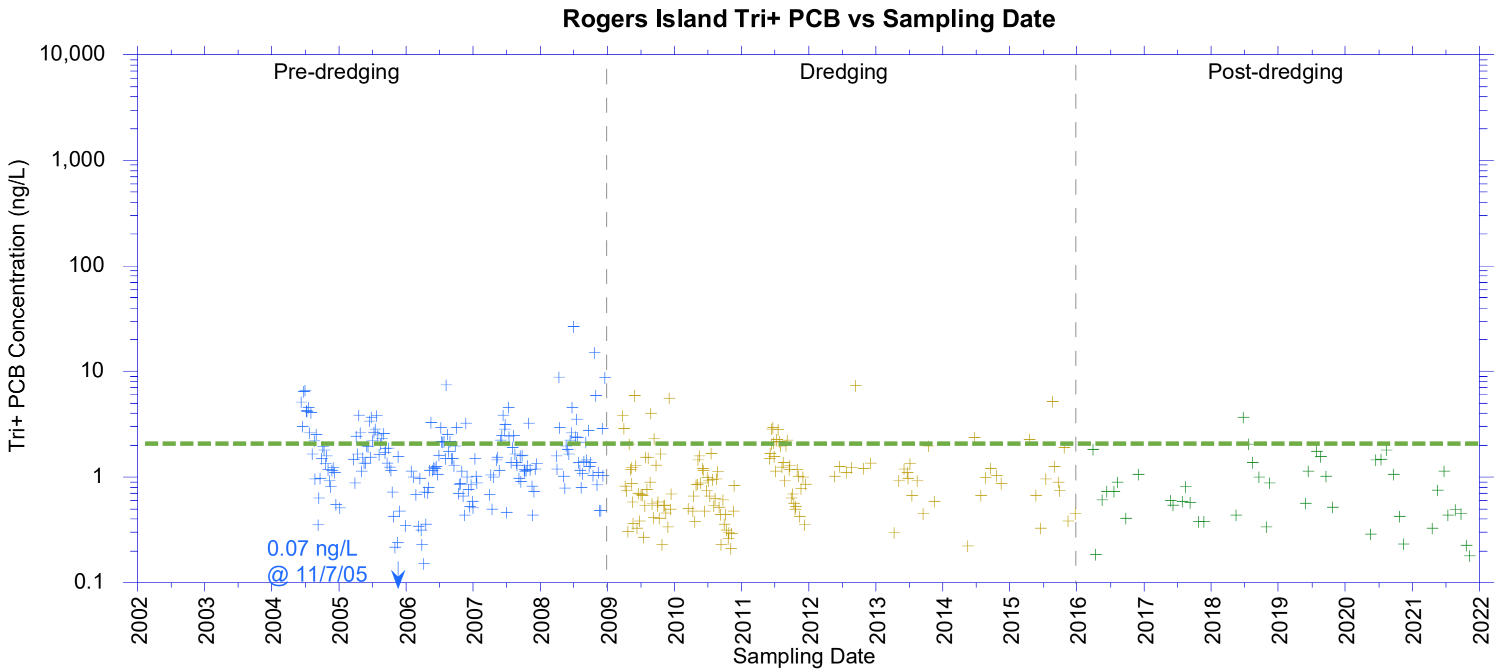
July 2024

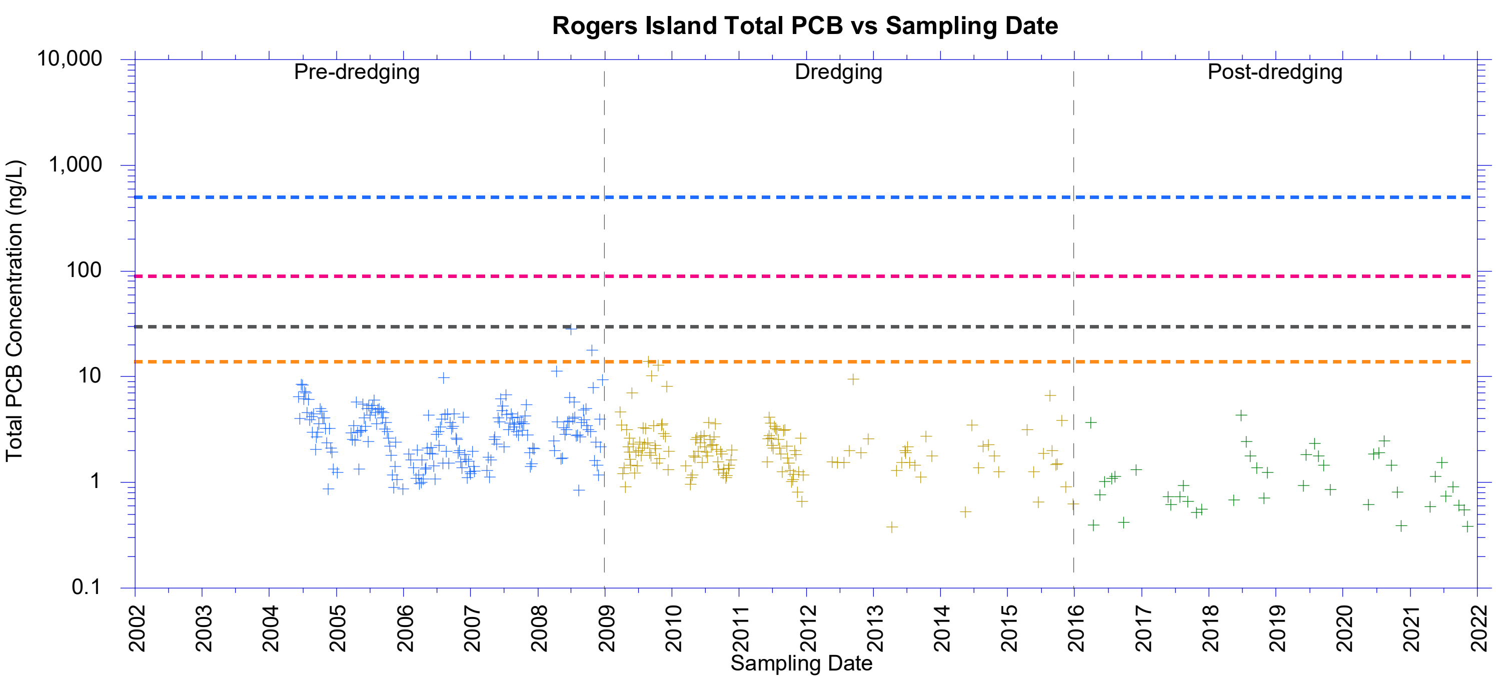


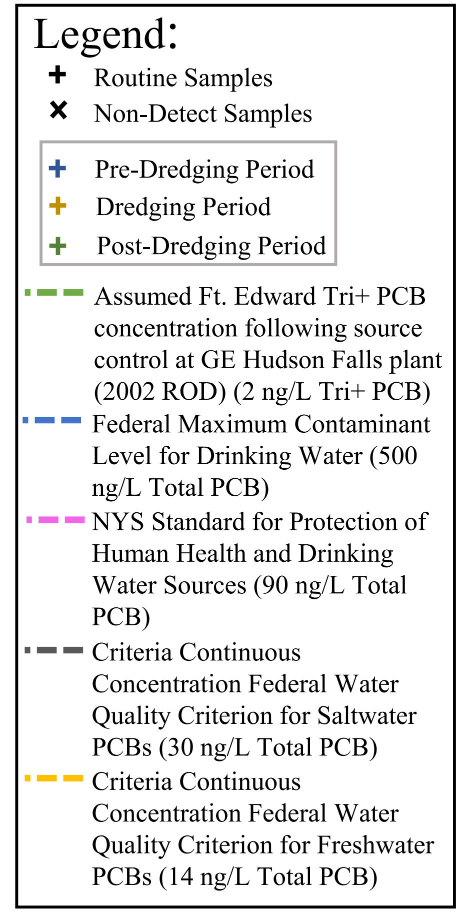
PCB Concentration at the Bakers Falls Monitoring Station from 2004 to 2021

Figure A1-4

July 2024



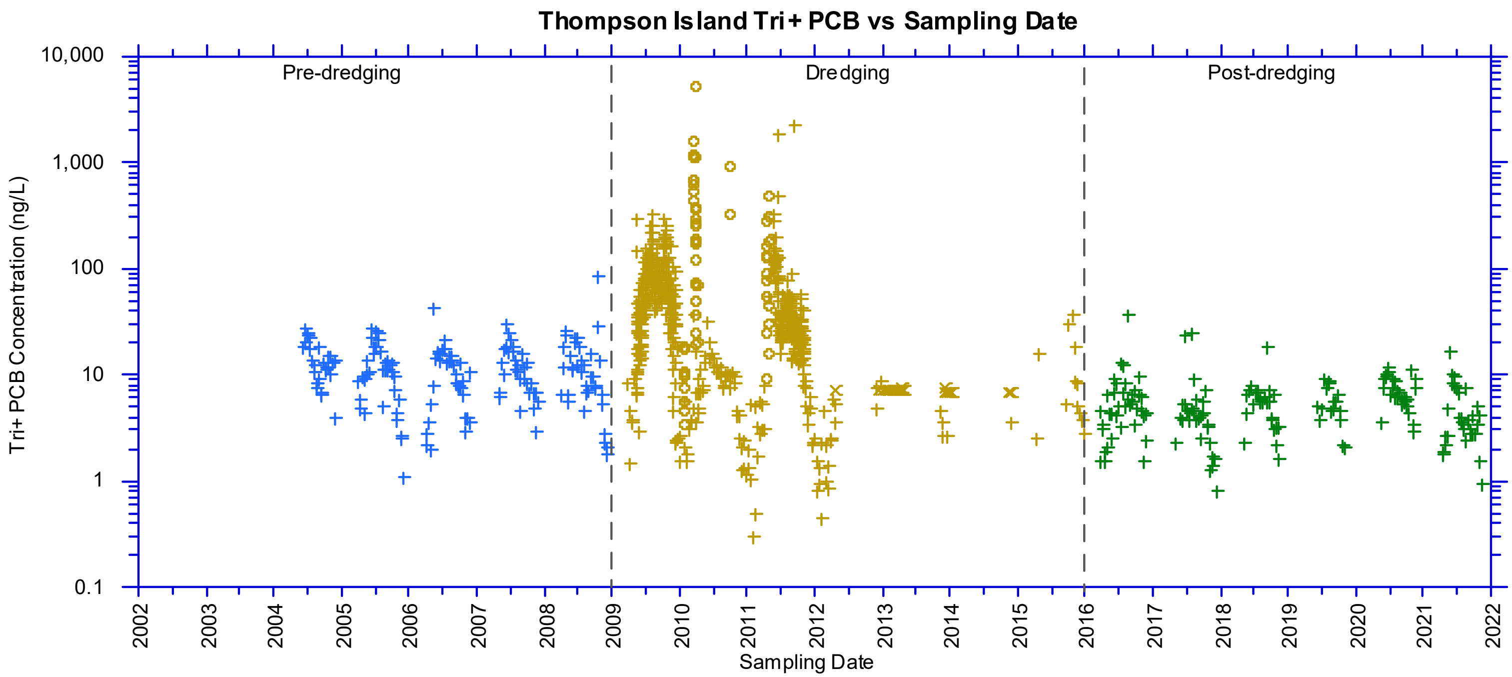


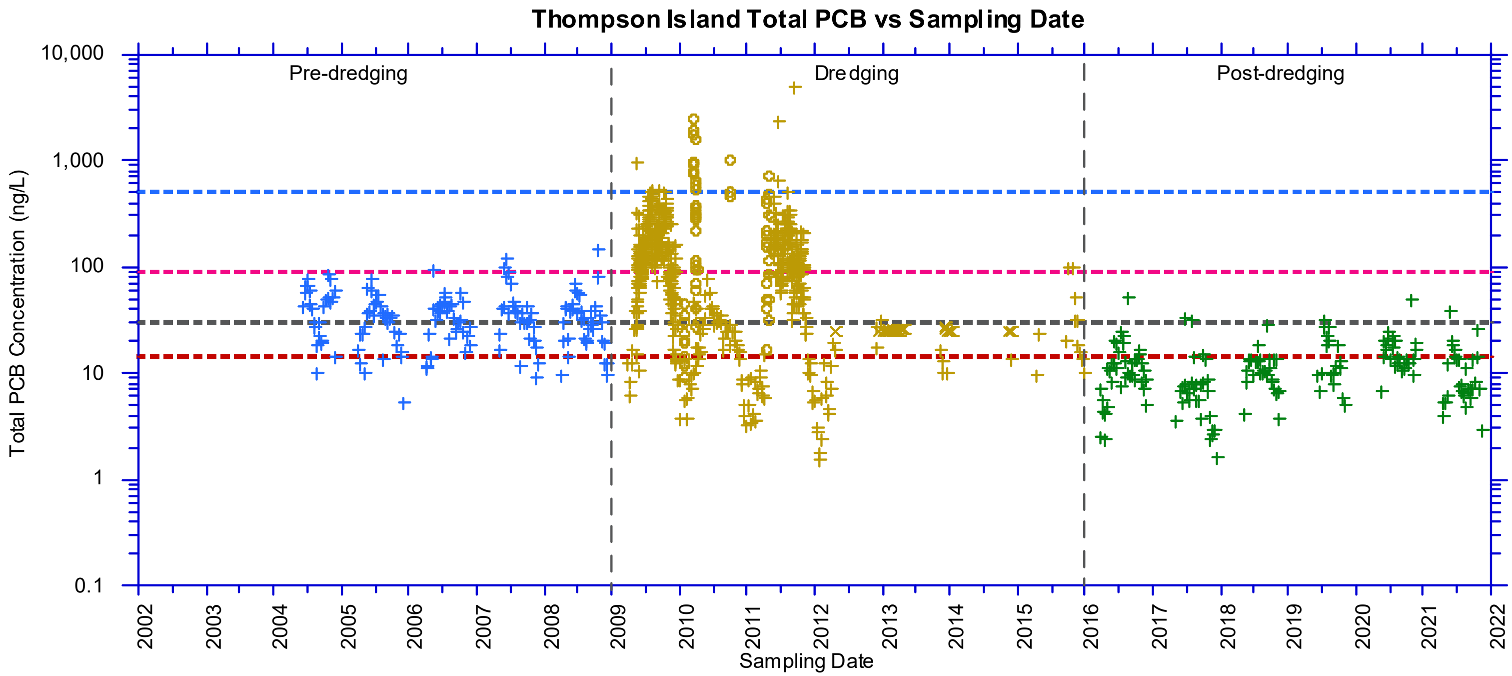


PCB Concentration at the Rogers Island Monitoring Station from 2004 to 2021

Figure A1-5

July 2024





Legend:

- ✚ Routine Samples
- O High Flow Samples
- ✕ Non-Detect Samples

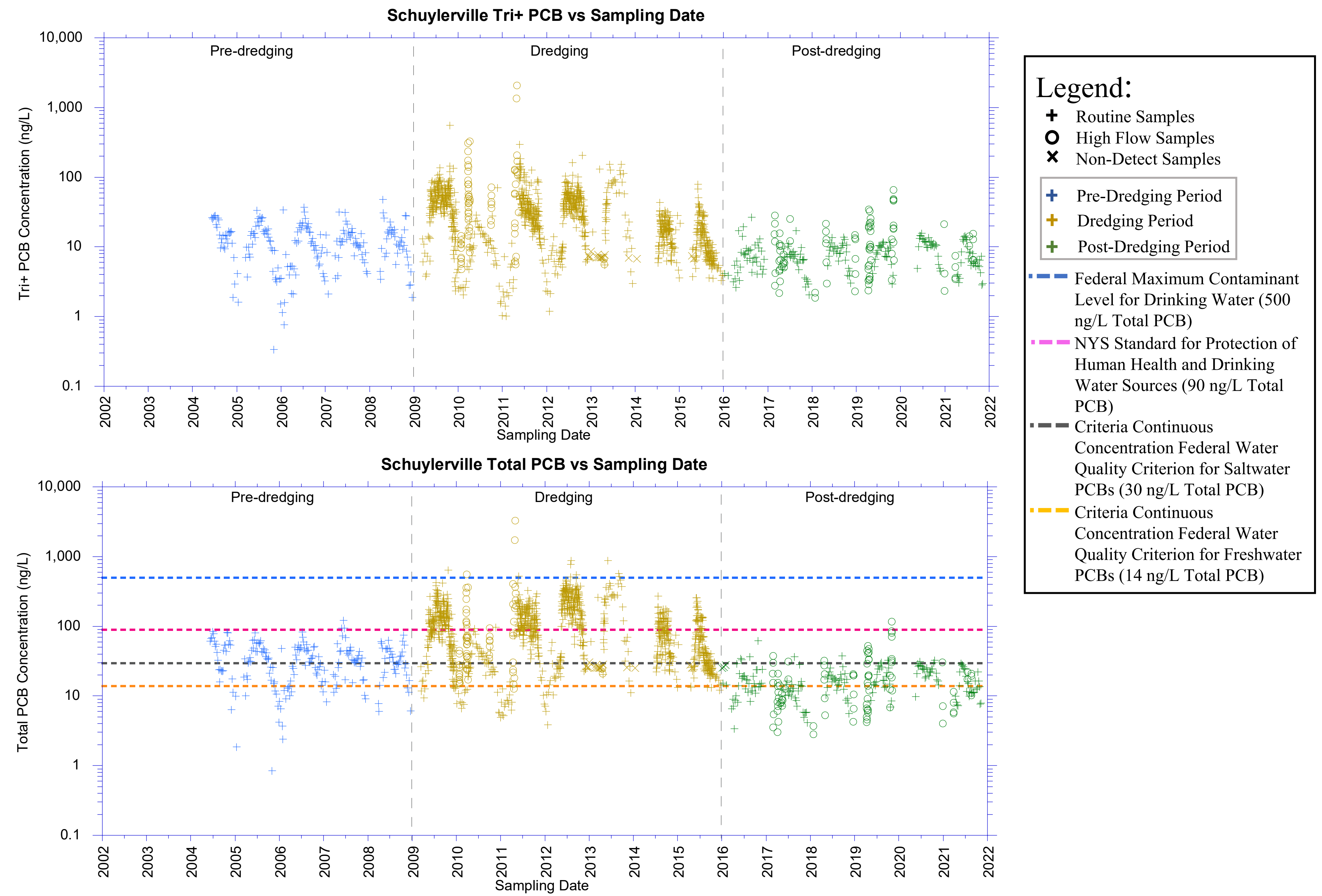
- ✚ Pre-Dredging Period
- ✚ Dredging Period
- ✚ Post-Dredging Period

- Federal Maximum Contaminant Level for Drinking Water (500 ng/L Total PCB)
- NYS Standard for Protection of Human Health and Drinking Water Sources (90 ng/L Total PCB)
- Criteria Continuous Concentration Federal Water Quality Criterion for Saltwater PCBs (30 ng/L Total PCB)
- Criteria Continuous Concentration Federal Water Quality Criterion for Freshwater PCBs (14 ng/L Total PCB)

PCB Concentration at the Thompson Island Dam Monitoring Station from 2004 to 2021

Figure A1-6

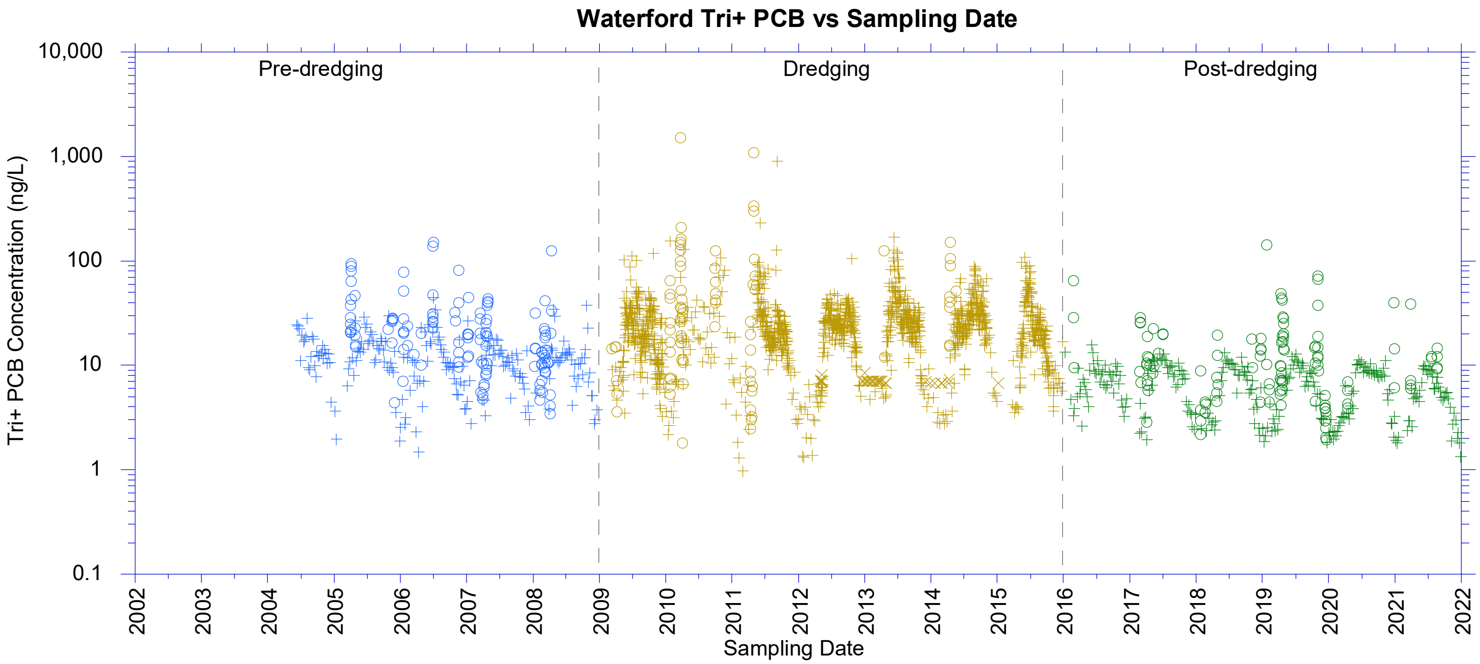
July 2024

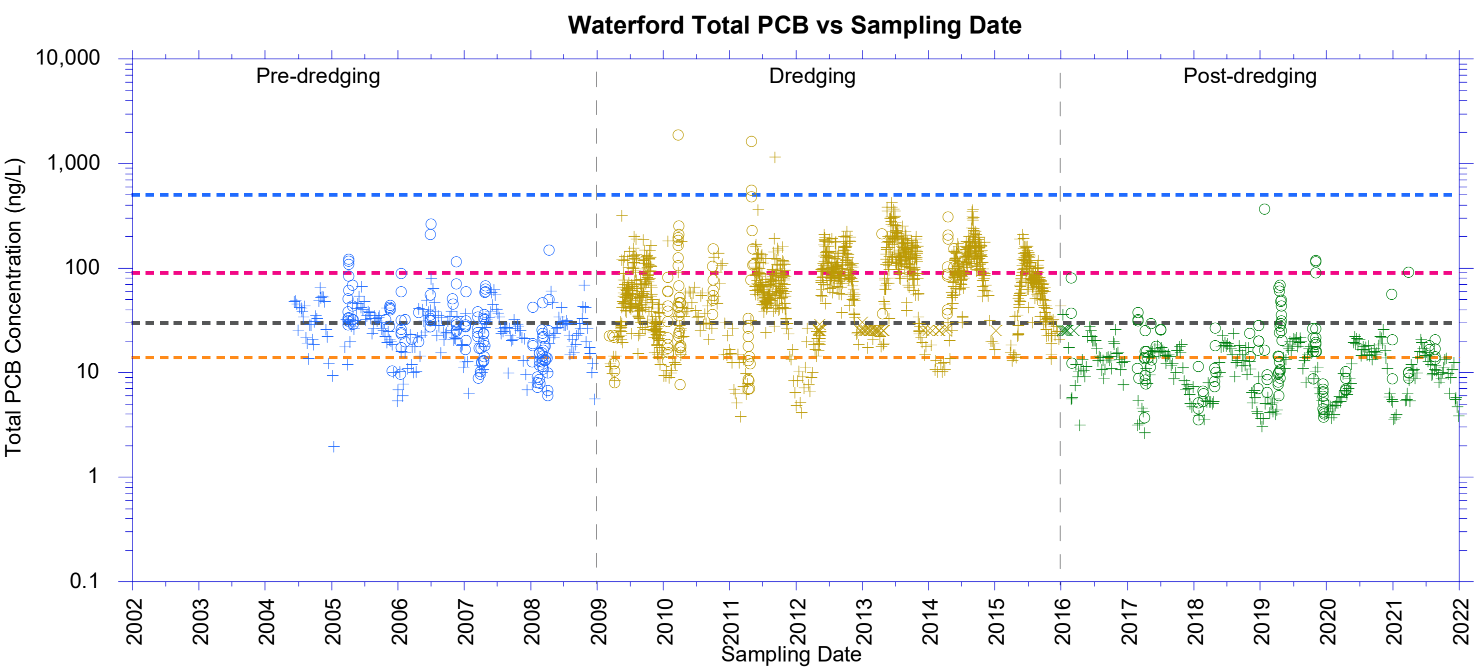


PCB Concentration at the Schuylerville Monitoring Station from 2004 to 2021

Figure A1-7

July 2024





Legend:

- Routine Samples
- High Flow Samples
- Non-Detect Samples

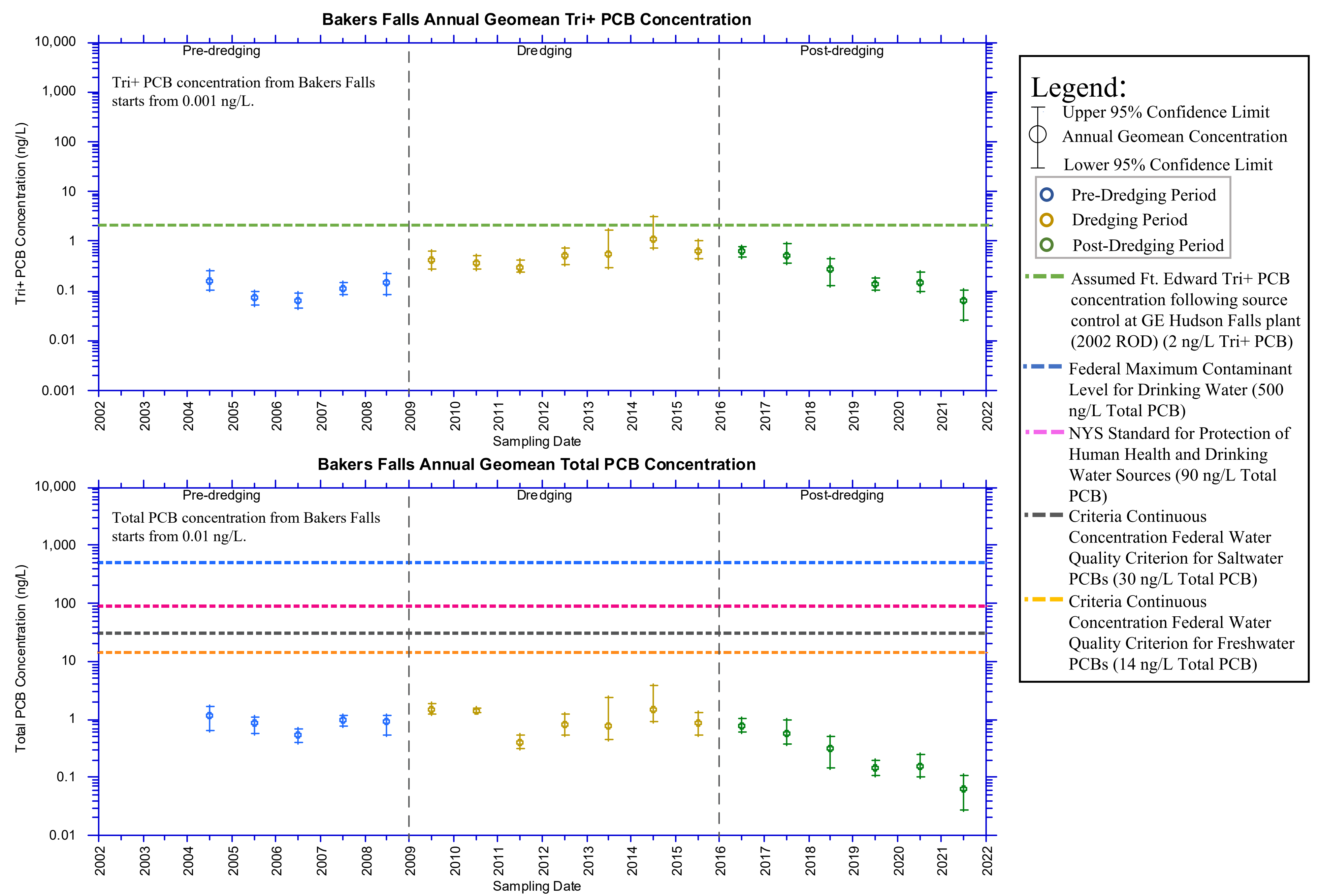
- Pre-Dredging Period
- Dredging Period
- Post-Dredging Period

- Federal Maximum Contaminant Level for Drinking Water (500 ng/L Total PCB)
- NYS Standard for Protection of Human Health and Drinking Water Sources (90 ng/L Total PCB)
- Criteria Continuous Concentration Federal Water Quality Criterion for Saltwater PCBs (30 ng/L Total PCB)
- Criteria Continuous Concentration Federal Water Quality Criterion for Freshwater PCBs (14 ng/L Total PCB)

PCB Concentration at the Waterford Monitoring Station from 2004 to 2021

Figure A1-8

July 2024

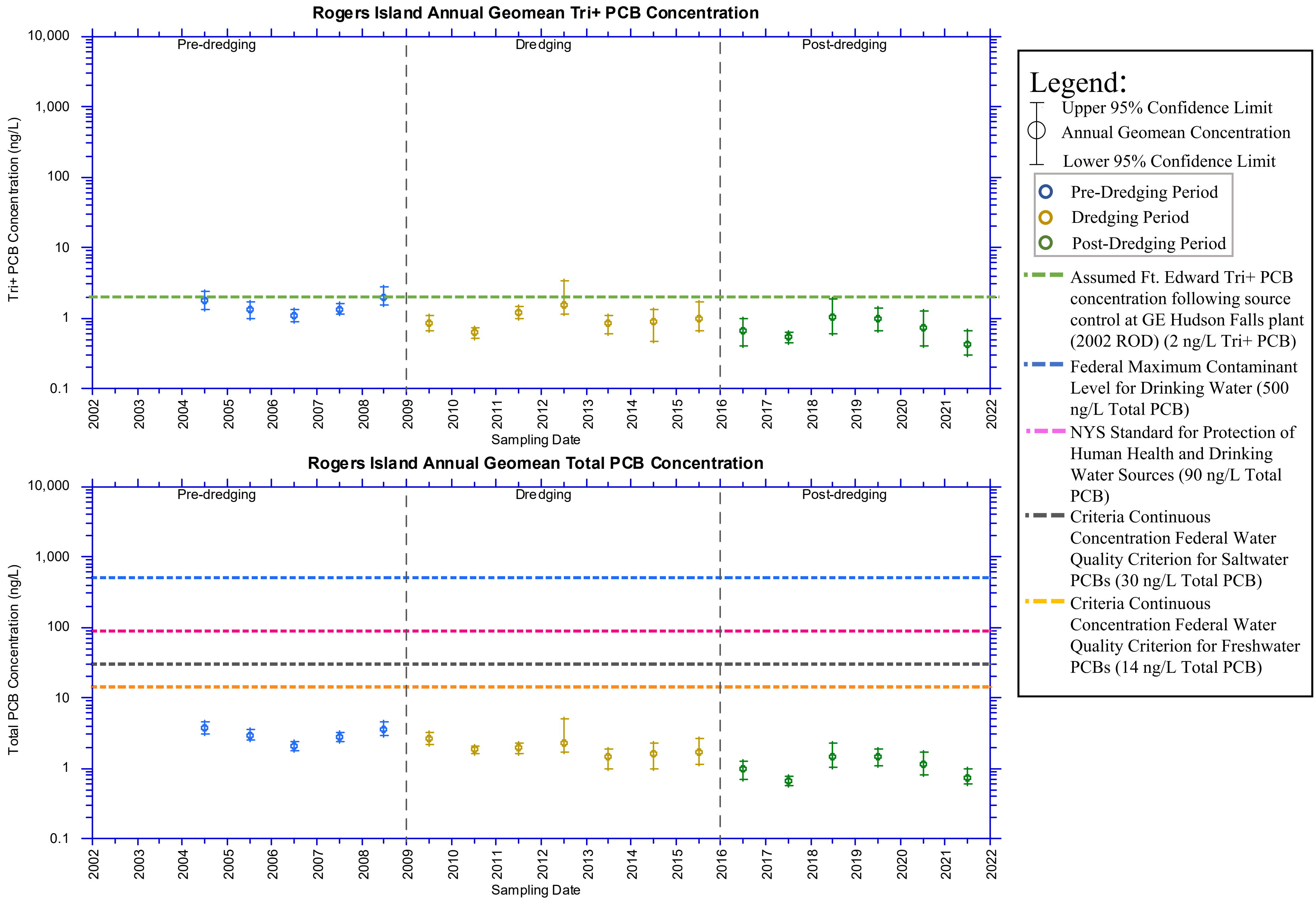


Hudson River PCBs Superfund Site | wsp

Annual Geometric Mean PCB Concentration for Routine Samples Collected at the Bakers Falls Monitoring Station from 2004 to 2021

Figure A1-9

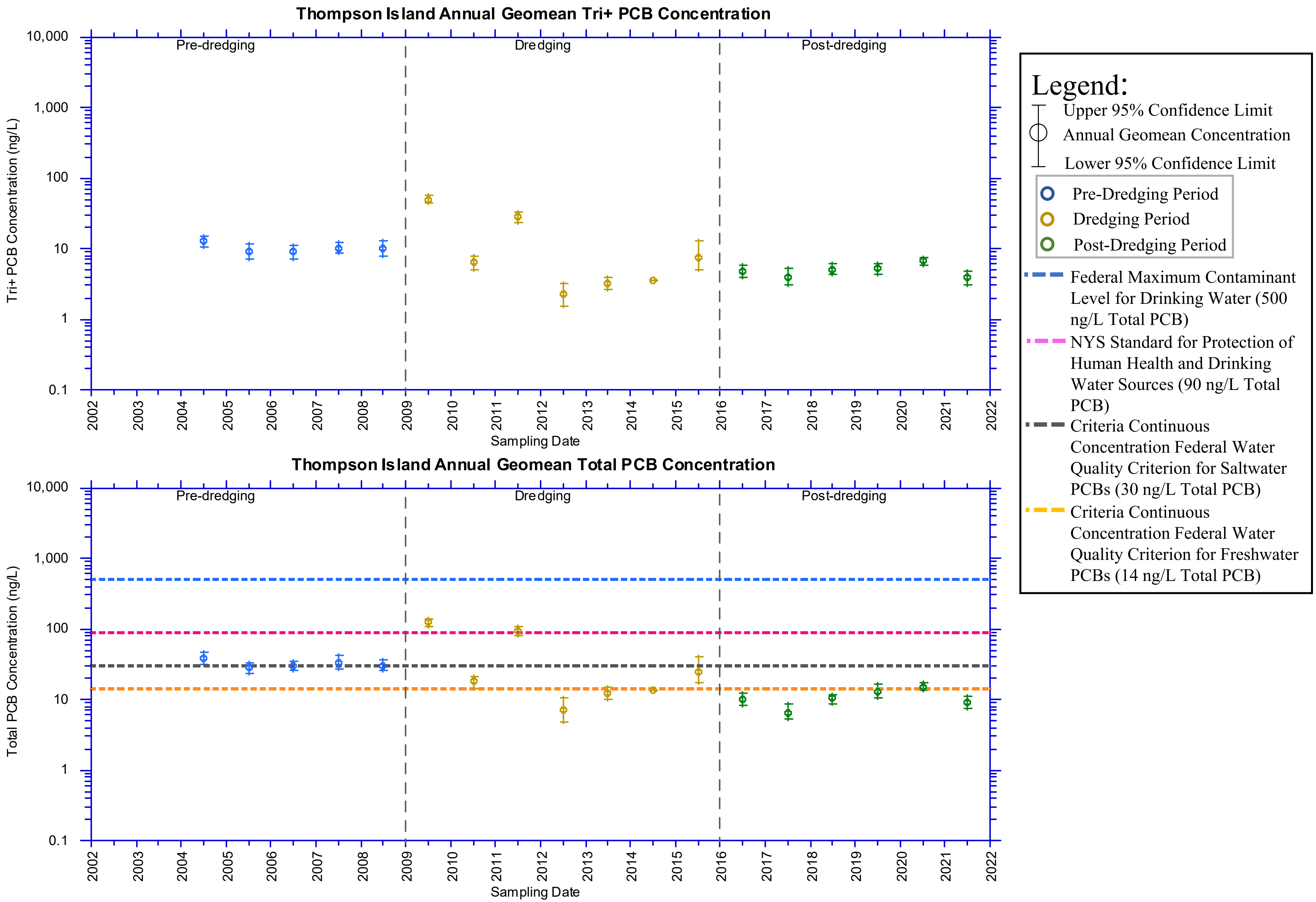
July 2024



Annual Geometric Mean PCB Concentration for Routine Samples Collected at the Rogers Island Monitoring Station from 2004 to 2021

Figure A1-10

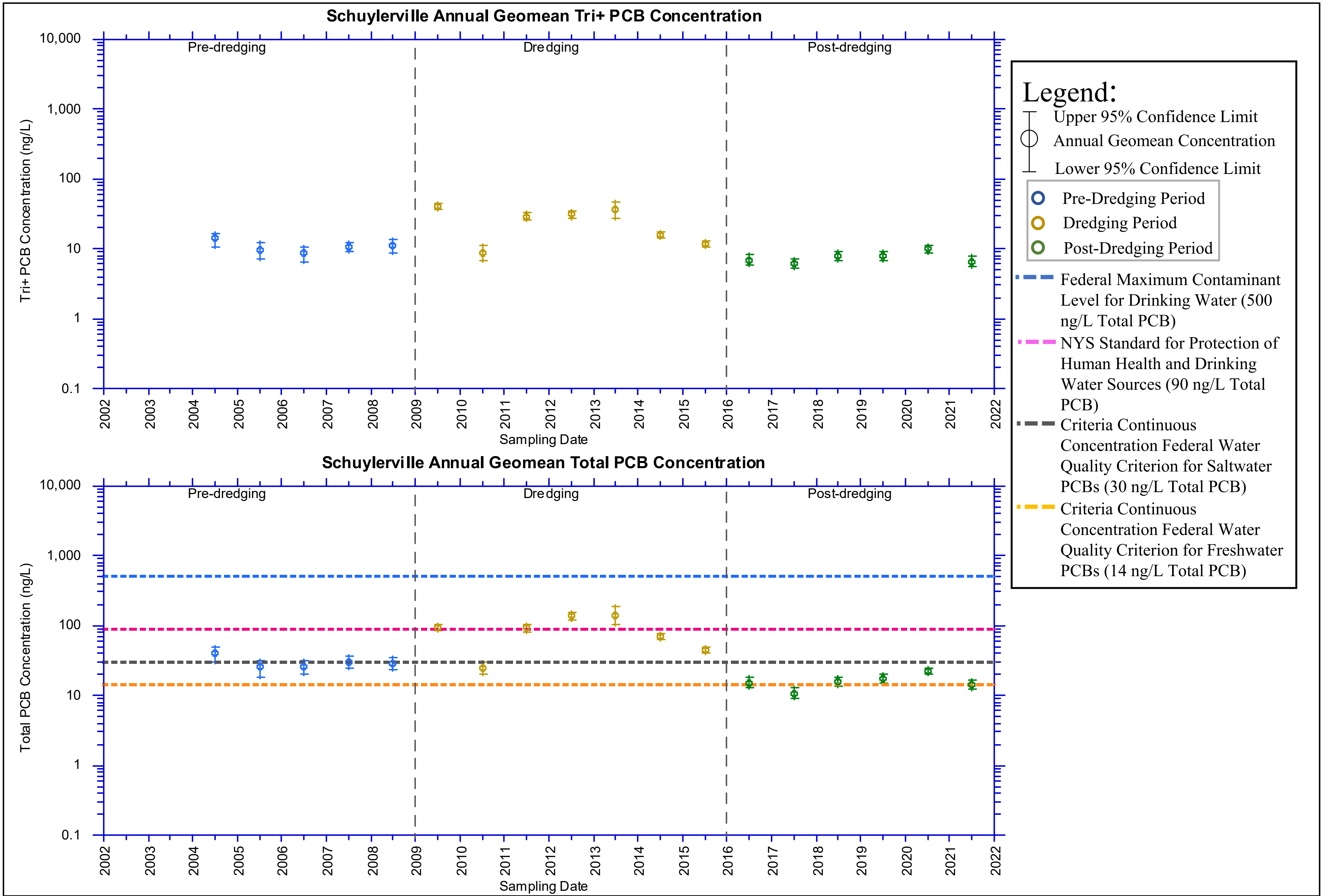
July 2024



Annual Geometric Mean PCB Concentration for Routine Samples Collected at the Thompson Island Dam Monitoring Station from 2004 to 2021

Figure A1-11

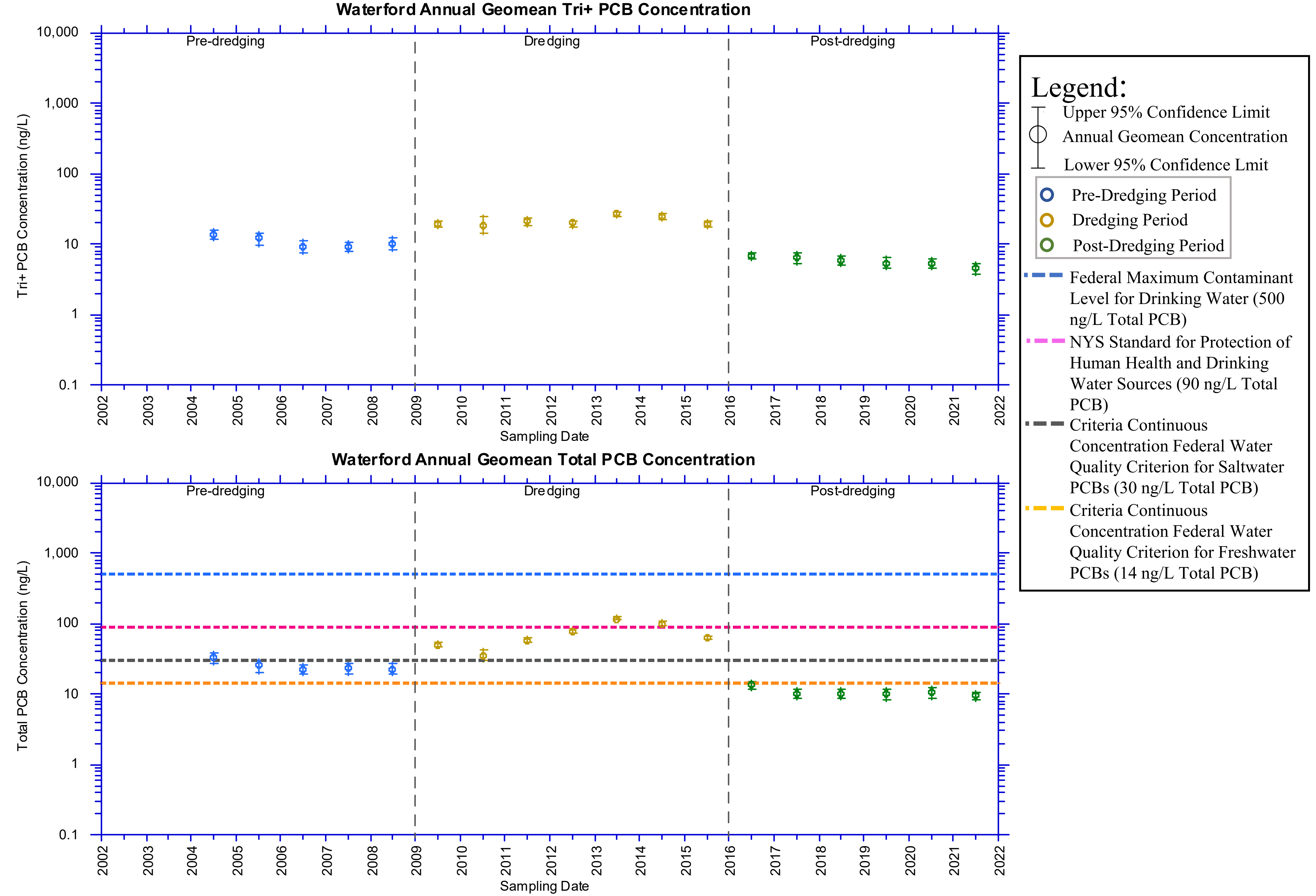
July 2024



Annual Geometric Mean PCB Concentration for Routine Samples Collected at the Schuylerville Monitoring Station from 2004 to 2021

Figure A1-12

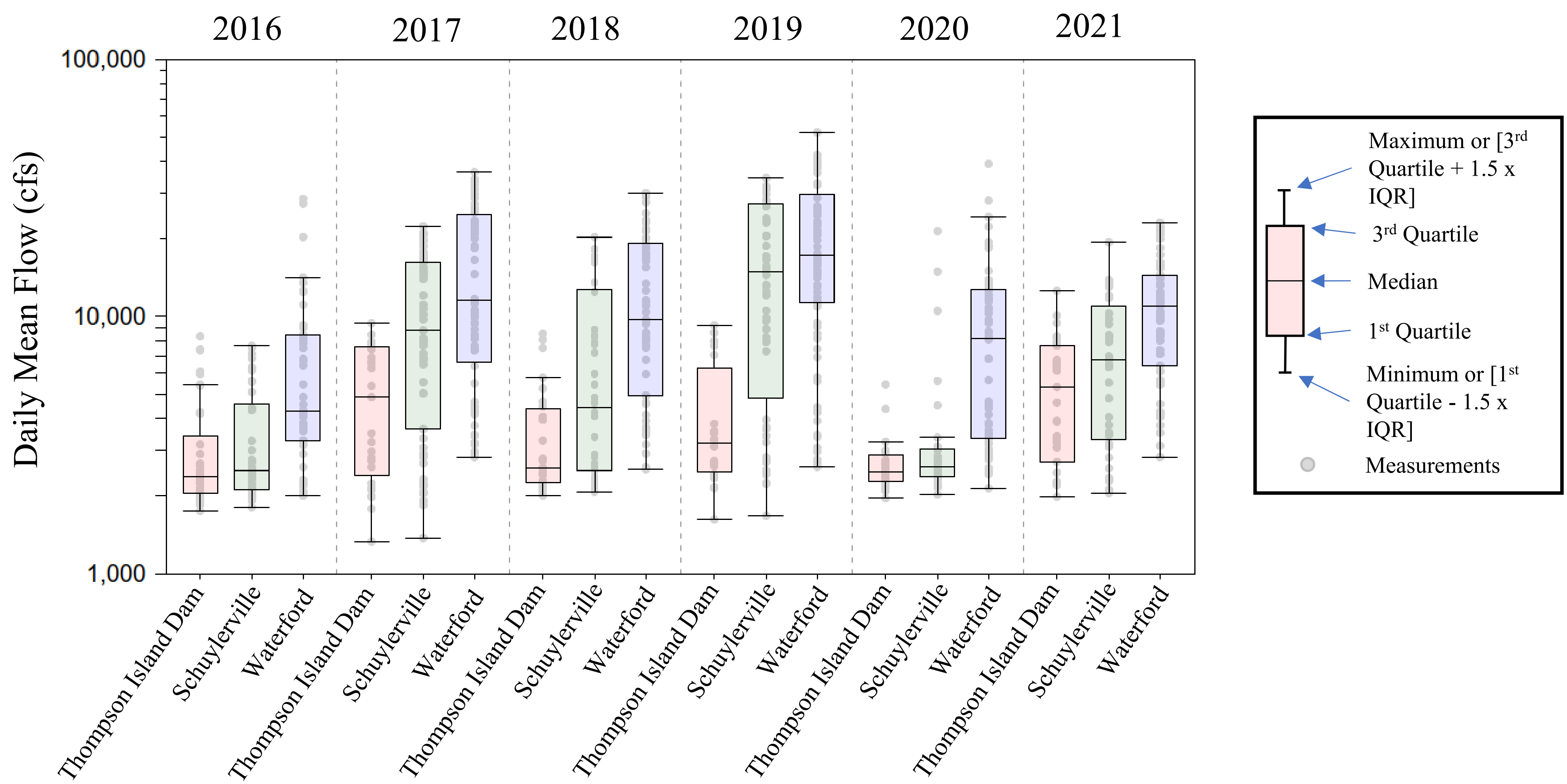
July 2024



Annual Geometric Mean PCB Concentration for Routine Samples Collected at the Waterford Monitoring Station from 2004 to 2021

Figure A1-13

July 2024



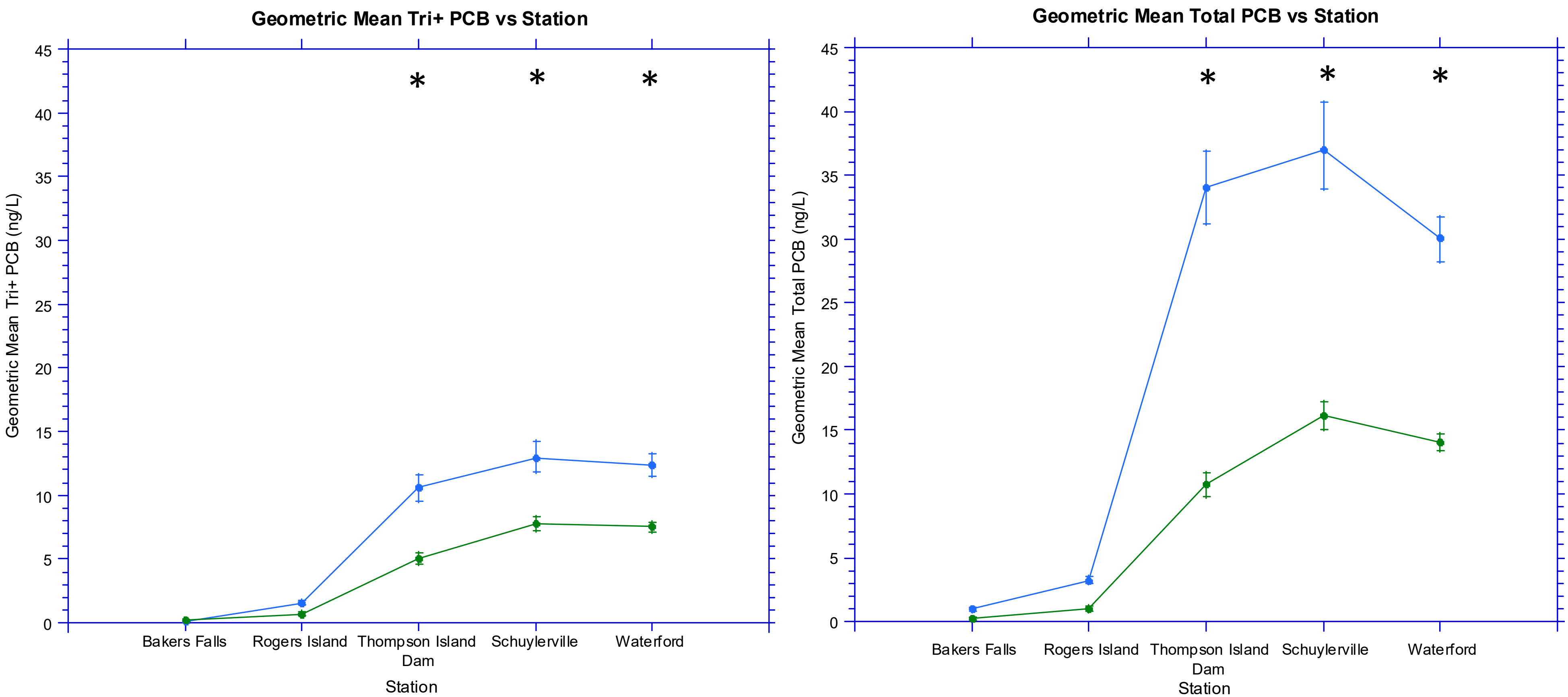
Notes:

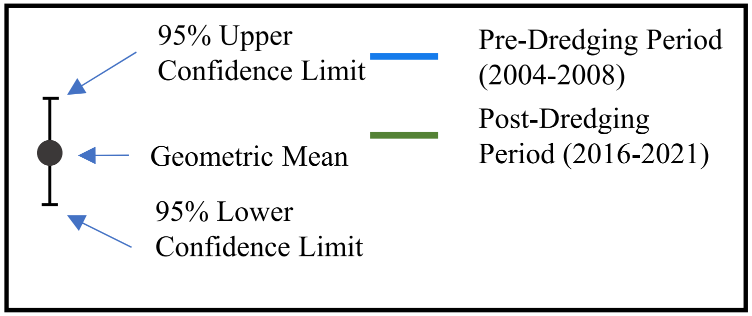
1. Thompson Island Dam flow represents daily mean flows measured at the Ft. Edward USGS gauge station (#01327750), multiplied by a flow multiplier to account for increased drainage area relative to the Ft. Edward USGS gauge station (1.030919). No high flow samples are collected at Thompson Island Dam.
2. Schuylerville flow represents daily mean flows measured at the Ft. Edward USGS gauge station (#01327750), multiplied by a flow multiplier to account for increased drainage area relative to the Ft. Edward USGS gauge station (1.067093).
3. Waterford flow represents daily mean flows measured at the Waterford USGS gauge station (#01335754).
4. Data plotted includes both routine samples and high flow samples.

Daily Mean Flows on Sampling Days in 2016 to 2021 at the Thompson Island Dam, Schuylerville and Waterford Monitoring Stations

Figure A1-14

July 2024





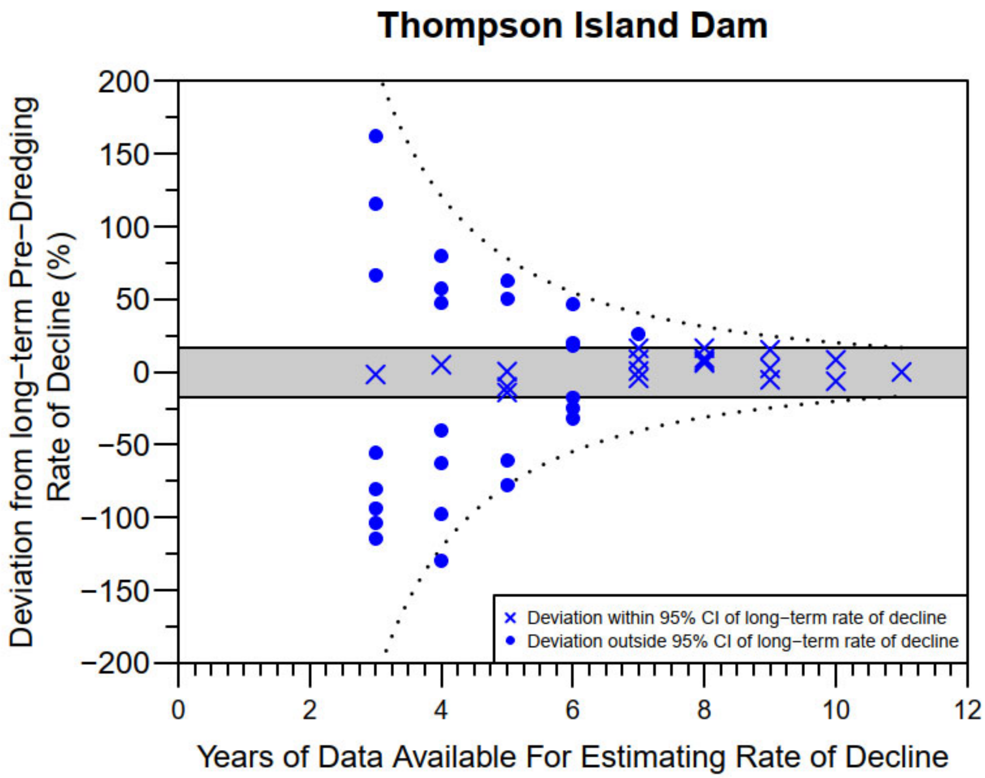
Note:

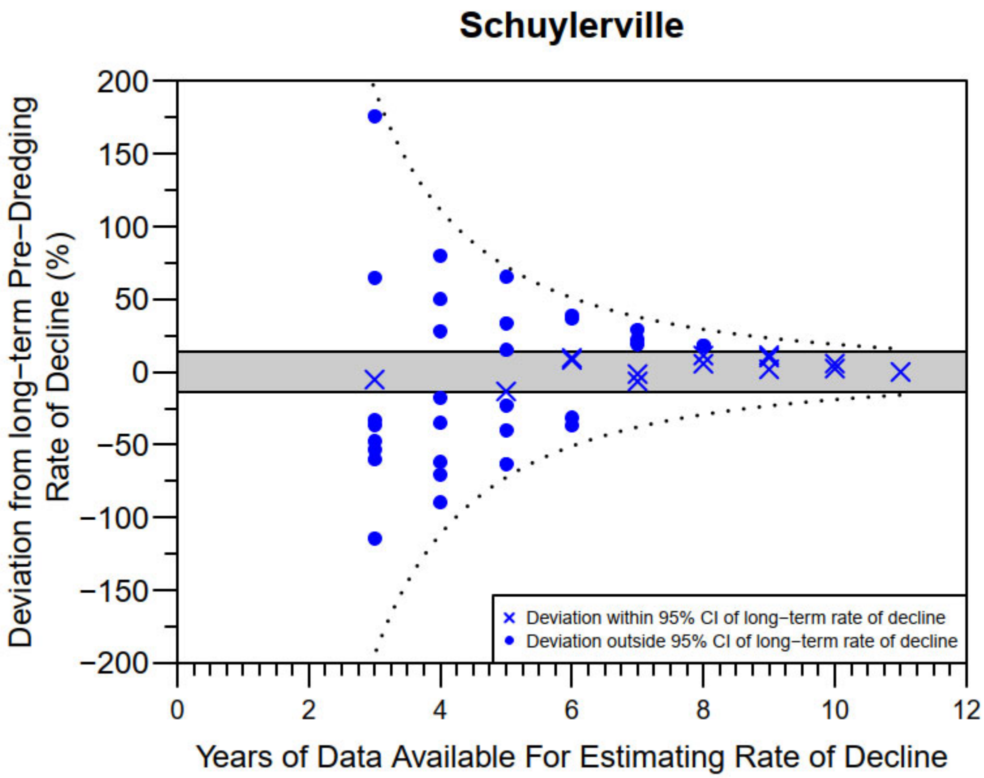
- Due to differences in sample collection frequency, pre- and post-dredging data included in this figure are restricted to samples collected in the months of May to November.
- Distance between points on x-axis is not to scale.
- Stars along top of each plot indicate whether the geometric mean concentration between the pre- and post-dredging period were significantly different at the $\alpha = 0.05$ level based on an ANCOVA analysis. See Section 3.2.1 for additional details of model formulation.

Pre-Dredging (2004 to 2008) and Post-Dredging (2016 to 2021) Annual Geometric Mean PCB Concentration for Routine Samples by Station

Figure A1-15

July 2024





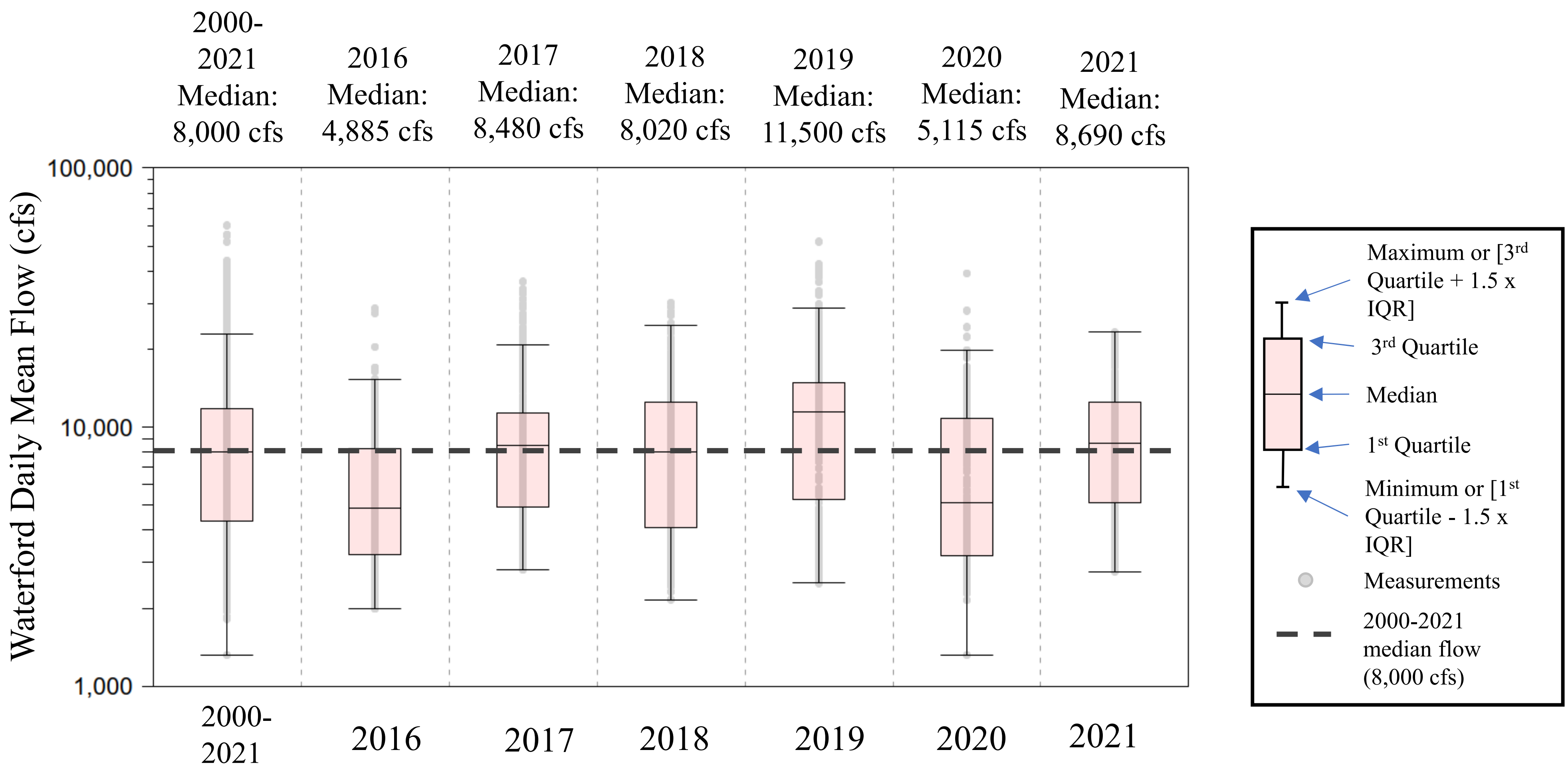
Notes:

- Shaded region represented 95% confidence bands on the 11-year rate of decline.
- Deviation from long-term pre-dredging rate of decline was calculated as the relative change of the short-term average rate of decline to the 11-year rate of decline.
- As an example, the symbols at the six-year interval on the X- axis represent the rates calculated for the following intervals: 1998 to 2003, 1999 to 2004, 2000 to 2005, 2001 to 2006, 2002 to 2007 and 2003 to 2008, resulting in six separate estimates of the rate of decline, represented by the seven points on the graph at 6 years.
- Dotted lines are empirical lines to show the approximate decline in variance with increasing number of years for averaging.
- Note that a positive deviation of 100% is equal to a decay rate that is twice as fast as the 11-year rate, whereas a negative deviation of 100% is equal to a decay rate of 0%/year (a flat line trend).
- The data used in this figure represents the 11-year period 1998 to 2008.

Variability in the Pre-Dredging Rate of Decline When Using Different Temporal Lengths of Routine Water Column Data between 1998 to 2008

Figure A1-16

July 2024

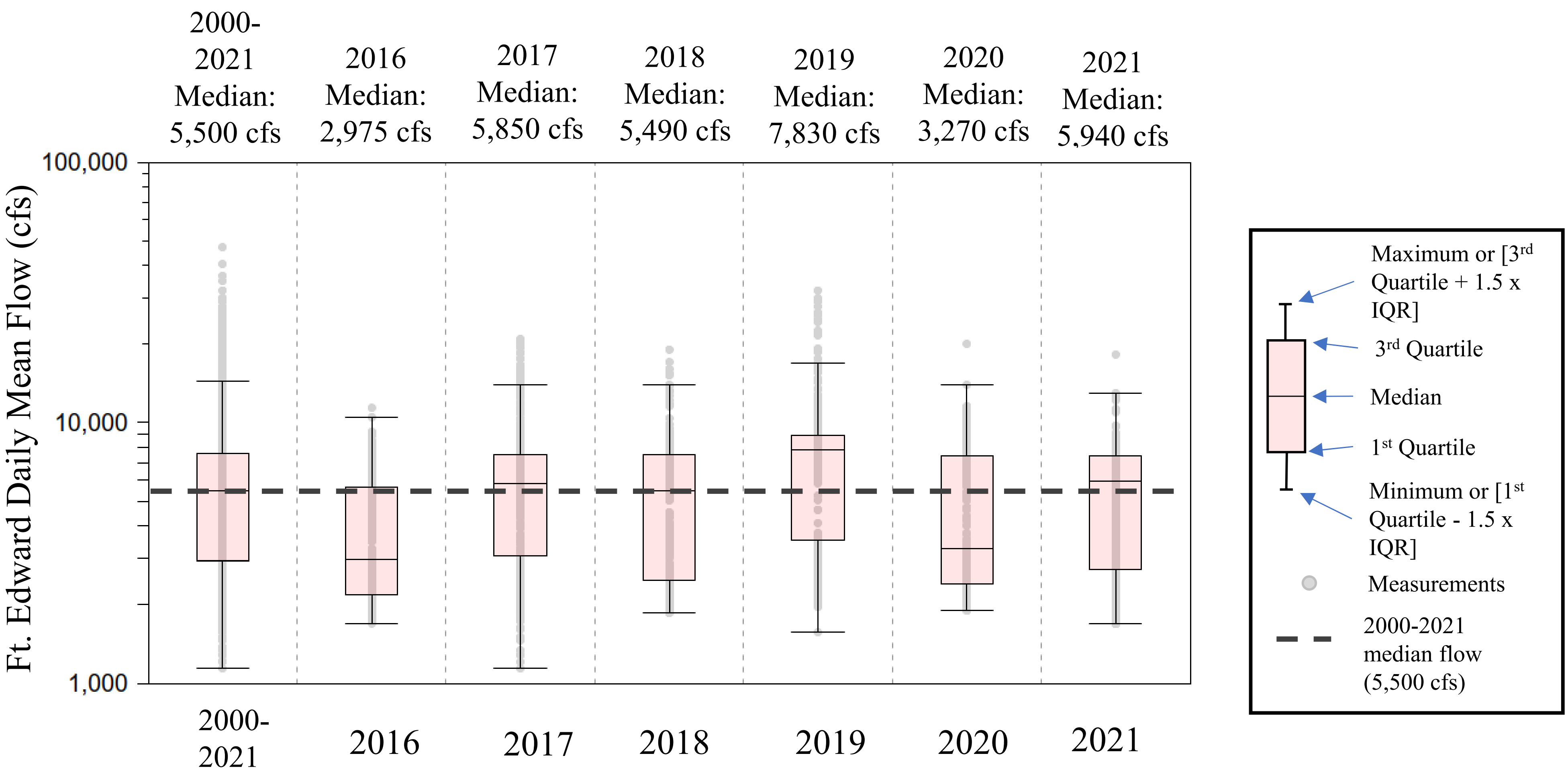


Note: Daily mean flows measured at the Waterford USGS gauge station (#01335754).

Long-term (2000 to 2021) Daily Mean Flows and Annual 2016 to 2021 Daily Mean Flow Statistics at the Waterford Monitoring Station

Figure A1-17

July 2024

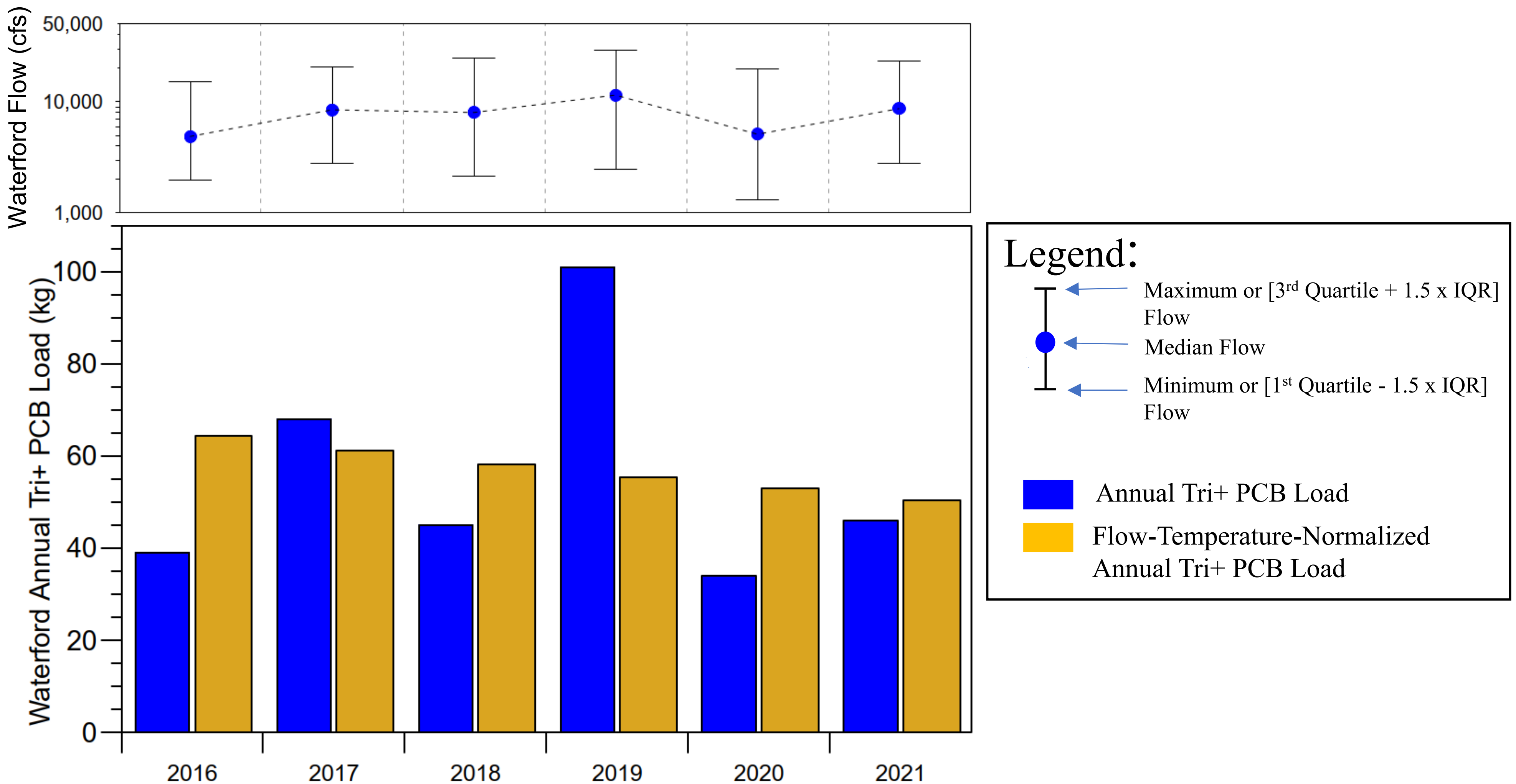


Note: Daily mean flows measured at the Ft. Edward USGS gauge station (#01327750).

Long-term (2000 to 2021) Daily Mean Flows and Annual 2016 to 2021 Daily Mean Flow Statistics at the Ft. Edward Monitoring Station

Figure A1-18

July 2024



Note: Annual PCB loads are estimated using the USGS LOADEST load estimation program (Runkel et al., 2004). Flow-Temperature-Normalized (FTN) PCB loads adjusts annual loads to remove the influence of year-to-year variability in flow and seasonality such that the FTN PCB loads reflect changes in PCB concentration only.

Annual Tri+ PCB Loads at the Waterford Monitoring Station from 2016 to 2021

Figure A1-19

July 2024

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

APPENDIX 1

ATTACHMENT A

TRI+ PCB LOAD CALCULATION METHOD MEMORANDUM

Prepared by
WSP USA Solutions Inc.

July 2024

Tri+ PCB Load Calculation Method Memorandum
Hudson River PCBs Superfund Site

TABLE OF CONTENTS

Tri+ PCB Load Calculation Method Memorandum
Hudson River PCBs Superfund Site

LIST OF TABLES

Tri+ PCB Load Calculation Method Memorandum

Hudson River PCBs Superfund Site

LIST OF FIGURES

1 INTRODUCTION

1.1 Purpose of this Memorandum

The purpose of this memorandum is to present the analyses that support the development of the method used by United States Environmental Protection Agency (EPA) during the Operations, Monitoring and Maintenance (OM&M) program to estimate annual PCB loads at Waterford.

1.2 Memorandum Organization

This memorandum is organized into the following sections:

- Section 1 "Introduction" presents the purpose and organization of this memo.
- Section 2 "Load Estimation Programs and Rating Curve Model Development" presents an overview of load estimation in riverine systems, describes the load estimation programs used in the analyses, and describes the rating curve models used to estimate PCB load at Waterford.
- Section 3 "Waterford Tri+ PCB Load Estimates and Model Performance for the 2016 to 2021 Post-Dredging Period" presents estimates of PCB loads for 2016 to 2021 and presents model performance metrics for different load estimation methods.
- Section 4 "Importance of High-Flow Events for Transport of PCBs at Waterford" discusses the importance of sampling high-flow events with regard to estimating loads at Waterford.
- Section 5 "Methodology for Calculating Tri+ PCB Loads at Waterford During the OM&M Program" presents the preferred method for estimating Tri+ PCB loads at Waterford and describes the data requirements.
- Section 6 "Conclusions" summarizes memorandum findings.
- Section 7 defines acronyms and abbreviations used.
- Section 8 provides the complete references for documents cited in this memorandum.

2 LOAD ESTIMATION PROGRAMS AND RATING CURVE MODEL DEVELOPMENT

2.1 Overview of Load Estimation Methods for Riverine Systems

In riverine systems, the constituent load is defined as the mass of a constituent (e.g., Tri+ PCBs[1]) transported past a given location over a specific period of time (e.g., a year). If the constituent concentration and flow are measured continuously throughout the period of interest, the constituent load for the time period can be calculated by multiplying concurrent measurements of concentration and flow and then summing the product over the period of interest. However, it is not always feasible or necessary to continuously measure constituent concentration over the period of interest, due to logistical, budgetary considerations, or overall data quality objectives. Instead, it is common for sampling programs to collect discrete samples throughout the period of interest (e.g., weekly, or monthly sampling). Assuming that flow is measured continuously over the period of interest (as would be the case if a flow gauge is present at or near the measurement location), a method is required to estimate the constituent concentration for time periods between discrete measurements to calculate the constituent load for the period of interest.

Various mathematical and statistical methods have been developed to estimate constituent load for a period of interest, including interpolation methods, averaging methods, ratio estimator methods, and rating curve methods (Richards, 1998; Lee et al., 2016). The performance of a particular method is influenced by the characteristics of the riverine system as well as the sampling frequency and sampling strategy used (e.g., stratified sampling or fixed interval sampling). For example, a relatively simple interpolation approach may be sufficient if the time between discrete samples is short (i.e., high-frequency sampling), but a rating curve approach may be preferred if there is a well-defined relationship between flow and constituent concentration.

PCB loads in the Upper Hudson River were previously estimated using AutoBeale (EPA, 2019). AutoBeale is a FORTRAN-based program that uses a time-stratified Beale Ratio Estimator method for load estimation (Richards, 1998). While the Beale Ratio Estimator is relatively straightforward mathematically and has been successfully applied to load estimation in the Great Lakes region, a key assumption is that there is a positive linear relationship between load and flow for a given time stratum (Richards, 1998). The long-term monitoring data for the Upper Hudson River obtained through 2021 indicate a non-linear relationship between load and flow, suggesting that a rating curve-based estimation method may be more appropriate than a ratio estimator-based method (see Figure 2-1). The goal of the analyses presented in this memorandum was to develop a rating curve model that can be used to provide robust load estimates for the Upper Hudson River. A review of

[1] 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 10 chlorine atoms per molecule, each with its own set of chemical properties.

available modeling frameworks in the literature indicated that the LOADEST program (Runkel et al., 2004), developed by the U.S. Geological Survey (USGS), will meet the objective of estimating annual loads to the Lower Hudson River and their associated uncertainties. LOADEST is a well-established load estimation method and is used by the USGS National Water Quality Network (NWQN) for estimating constituent load at sites across the United States. WRTDS (Weighted Regression on Time, Discharge, and Season), a more recent USGS load estimation program, which is being incorporated into the NWQN program alongside LOADEST (Lee et al., 2020) was also evaluated.

2.2 Load Estimation Programs Used in Analyses

The LOADEST load estimation software was originally developed as a FORTRAN program to estimate constituent loads by the rating curve method (Runkel et al., 2004). The FORTRAN version of the software is available without cost through the USGS website (http://water.usgs.gov/software/loadest/). More recently, a version of LOADEST was developed to run within the R® programming environment (available via the "rloadest" R® package). LOADEST has several desirable features, including: 1) robust bias-correction of log-transformed values for predicted concentrations and loads, 2) ability to accept sub-daily (e.g., 15-minute) data as inputs for calibrating the rating curve model and the ability to output predicted loads and concentrations at sub-daily timesteps, and 3) flexibility in formulating the model used to predict loads and concentrations. In addition to the user being able to develop their own rating curve model, LOADEST provides 11 built-in rating curve models, as well as the ability to automatically select the "best" model out of the available built-in models. Interested readers are referred to Cohn (2005), Cohn et al. (1989), and Gilroy et al. (1990) for additional details on the statistical development of LOADEST.

The USGS also developed a different load estimation program, written in R®, called WRTDS with an extension called WRTDS-Kalman that is based on the rating curve method (Hirsch et al., 2010; Zhang and Hirsch, 2019). Although both WRTDS and WRTDS-Kalman (hereafter abbreviated as WRTDS-K) estimate loads, WRTDS was originally developed for calculating trends in constituent loads through time when constituent datasets of approximately 20 years or more are available, while WRTDS-K was developed more specifically to predict load for defined time periods (e.g., month, season, or year) (Zhang and Hirsch, 2019). WRTDS and WRTDS-K are included as part of the publicly available "EGRET" R® package (Hirsch and De Cicco, 2015), and share some of the desirable features of LOADEST. However, an important difference between LOADEST and WRTDS/WRTDS-K is that WRTDS/WRTDS-K can only accept daily mean flow values for calibrating the rating curve model and predictions. Therefore, WRTDS/WRTDS-K only produces loads and concentrations at daily timesteps. The ability to generate sub-daily estimates of load is important for systems like the Upper Hudson River, where concentrations vary non-linearly with flow and flows can vary substantially within a day due to dam releases and storm events. The underlying statistical and model fitting procedures also differ from LOADEST. Readers are referred to Hirsch et al. (2010), Hirsch and De Cicco (2015), and Zhang and Hirsch (2019) for details on implementing WRTDS and for additional information on the statistical methods used by WRTDS and WRTDS-K to predict loads.

It is important to note that the actual annual Tri+ PCB load at Waterford each year is unknown since measurements of both flow and Tri+ PCB concentration at every instant within a year do not exist. Therefore, it is not possible to assess the accuracy of the annual load estimates determined from different load estimation methods relative to the actual annual load. By comparing estimates produced by different load estimation methods (that use different rating curve models and underlying statistical procedures) and evaluating model performance metrics, we can gain some insight into whether a load estimate is reasonable. If different methods produce similar annual load estimates, this indicates the estimates are robust to differences in the estimation methods and provides a measure of confidence that the estimates are reasonable, based on the available data. If different methods produce substantially different annual load estimates, this will reduce confidence in the load estimates and suggest one or more estimates of the load may not be reasonable. Similarly, if model performance metrics indicate that the model is meeting the assumptions of the statistical procedures used to develop the estimate, this would also provide confidence that the load estimates are reasonable. While this memo is not intended to be a comprehensive comparison of load estimation programs, different load estimation methods are employed to provide a measure of context for whether the estimated loads are reasonable.

2.3 Development of the LOADEST Rating Curve Model for Estimating Annual Loads at Waterford

As mentioned earlier, LOADEST allows the user to develop a site-specific rating curve model for estimating loads. The rating curve method for load estimation involves developing a statistical model (referred to as the rating curve model, which is commonly a log linear regression model) between the constituent whose load is to be estimated (i.e., Tri+ PCB) and river flow, while also including other measured variables that may improve the predictive ability of the model (Cohn et al., 1992). This type of load estimation method is most appropriate when strong relationships exist between concentration and flow (and possibly other auxiliary variables) (Marsh and Waters, 2009). Previous studies on the Upper Hudson River have demonstrated a strong relationship between Tri+ PCB concentrations and river flow, as well as strong seasonality in Tri+ PCB concentrations (EPA, 1997; EPA, 1999). Thus, the development of a site-specific rating curve model for Waterford must include flow and seasonality to explain the variance in the observed concentrations of Tri+ PCB.

2.3.1 Tri+ PCB Concentration vs. Flow

Earlier studies observed a non-linear relationship between Tri+ PCB concentration and flow in the Upper Hudson River (EPA, 1999). Under low flow conditions, the relationship between concentration and flow is negative, with increasing flow values resulting in decreasing concentrations; however, above a certain flow threshold, the relationship becomes positive, with increasing flow values resulting in increasing concentration. This non-linear relationship between concentration and flow is likely reflective of a dilution-dominated flow regime at relatively low flows and a resuspension-dominated flow regime at higher flows.

Figure 2-2 presents the relationship between Tri+ PCB concentration and flow for the post-dredging years 2016 to 2021. All years exhibit a "V"-shaped, piecewise relationship between Tri+ PCB concentration and flow, indicative of the two distinct flow regimes described above. The consistent piecewise relationship between Tri+ PCB concentration and flow across the six years of data supports the use of a two-flow regime rating curve model for predicting concentrations and load in the Upper Hudson River.

Segmented regression was utilized to model the non-linear piecewise relationship between concentration and flow. Segmented regression uses an indicator variable and a change point value to fit two different slopes to the observations, depending on whether the observed flow is above or below a certain flow threshold value, referred to as the change point. For the analyses in this memorandum, the flow change point for an individual year is determined using a change point detection algorithm implemented in the R® package "segmented" and is not defined *a priori* (Muggeo, 2008). Figure 2-2 demonstrates the fit of the segmented regression model to measured data and Table 2-1 presents the estimated change point of the fitted segmented regression model for each year as well as all years combined. The change point represents the critical river flow needed to erode and transport sediments, and thus depends on the hydrodynamics and sediment physical characteristics of the Upper Hudson River. The empirically derived change points for the various years vary within a factor two, with the variability likely associated with year-to-year differences in PCB concentrations, watershed supply of sediments, variability on the rising and falling limb of storm hydrographs, suspended sediment concentrations, storm magnitudes, and the number of data points available to constrain the estimate.

2.3.2 Seasonality in Tri+ PCB Concentrations

Seasonality in water column PCB concentrations has been observed in earlier studies of PCBs with higher concentrations during the warmer summer months and lower concentrations during the colder fall and winter months under routine flow conditions (EPA, 1997; EPA, 1999). The observed seasonality is likely a response to increased biological activity in the near-surface sediment during the warmer months, which enhances the release of sediment-bound PCBs to the water column (Larsson, 1983; Larsson et al., 1990). Figure 2-3 presents the seasonality in Tri+ PCB concentration for the post-dredging years 2016 to 2021 along with the average daily water temperature measured at the Albany USGS gauge station (Station ID# 01359139). Temperature measured at the Albany USGS gauge station is used as an auxiliary variable in the rating curve model to account for the observed seasonality in Tri+ PCB concentration. Multiple factors, including temperature, interact to produce the seasonality observed in PCB concentration. While recognizing that temperature alone does not fully explain all of the variability we see in the seasonality in PCB concentration, temperature is well-suited for use as a proxy for seasonality in a rating curve model. First, temperature correlates well with the timing of seasonal increases and decreases in PCB concentrations, which is important when modeling seasonal fluctuations. Second, temperature data is measured at the time of sample collection and is also measured by the USGS at the same temporal frequency as river flow. This allows for temperature to be used to both develop the model and predict concentrations between sampling events.

2.3.3 Rating Curve Models for Estimating Annual Loads

This section presents the LOADEST site-specific rating curve model that incorporates both segmented regression and temperature discussed in Sections 2.3.1 and 2.3.2. Then the USGS-provided LOADEST Model "6" rating curve model, which differs from the site-specific rating curve model in the treatment of flow and seasonality, is introduced. Model "6" uses a quadratic term and cosine and sine terms for modeling flow and seasonality, respectively. Model "6" is included in the analysis for comparison with the site-specific rating curve model. Lastly, the rating curve model in the WRTDS/WRTDS-K program is introduced.

The site-specific rating curve model that includes both segmented flow and water temperature can be written in linear regression form using logarithmic terms as:

$$\ln(C) = b_o + b_1 \ln(Q) + \ b_2\big(\ln(Q) - \ CP_Q\big)I_Q + b_3 Temp + \varepsilon \qquad (Eq. 1)$$

Where *C* is the measured concentrations of the constituent of interest (Tri+ PCB), b_i are regression coefficients estimated by the model, *Q* is river flow, CP_Q is the change point flow value estimated using the "segmented" R package, I_Q is an indicator variable that equals 1 when the measured flow is greater than CP_Q and zero when flow is less than CP_Q, *Temp* is water temperature, and ε is the regression model error term.

The LOADEST Model "6" mentioned earlier has the following form (Runkel et al., 2004; Cohn et al., 1992):

$$\ln(C) = b_o + b_1 \ln(Q) + \ b_2(\ln(Q))^2 + b_3 \sin(2\pi t) + b_4 \cos(2\pi t) + \varepsilon \qquad (Eq. 2)$$

where *t* represents the fractional year, and the other terms are defined as above.

LOADEST uses the same equation form to develop a model for estimating concentrations and loads. In the case of the rating curve model for loads, $\ln(C)$ is replaced with $\ln(L)$, where L is measured instantaneous load ($L = Q*C$). For each year, a regression model was developed using all Tri+ PCB concentrations (or Tri+ PCB loads) measured within the year as the response variable and river flow and water temperature as predictor variables. LOADEST can accept either instantaneous river flow and water temperature measured at 15-minute intervals or daily mean river flow and temperature which are the averages of the 15-minute interval measurements. River flow is measured at the Waterford USGS gauge station (Station ID# 01335754) and water temperature is measured at the time of sample collection, or if not available, the closest-in-time water temperature at the Albany USGS gauge station (Station ID# 01359139) is used. When

LOADEST is run using daily mean flow and temperature values, same-day measurements of Tri+ PCB concentrations are averaged to produce a mean Tri+ PCB concentration for that day. Due to averaging of concentrations and the use of daily mean flow (versus 15-minute interval flow), LOADEST results using daily mean values will produce different load estimates compared with model results that use 15-minute interval flow and temperature data without averaging same-day samples.

LOADEST uses an adjusted maximum likelihood estimation (AMLE) procedure to estimate the coefficients of the rating curve models. The AMLE estimation procedure converges to maximum likelihood estimation (MLE) when the datasets contain no censored (i.e., non-detect) data, as is the case here. The calibrated model can then be used to predict Tri+ PCB concentration and load at 15-minute intervals (or other time intervals if requested) in log-units. The predictions in log-units are then back-transformed to original units and a minimum variance unbiased estimator (MVUE) procedure is used to correct for bias in back-transformed predictions. LOADEST returns predicted loads in units of mass/day along with the standard error of the model fit, standard error of prediction, and the 95 percent prediction interval. Although LOADEST returns loads in units of mass/day, the user can request load estimates for a longer time period (e.g., calendar year) and LOADEST will adjust the standard errors and confidence intervals to reflect that the estimate represents a summation of daily load estimates.

The WRTDS and WRTDS-K rating curve models are based on a rating curve model built into LOADEST (Model "7"), which is similar to Model "6" but removes the squared term for flow and includes time as a separate variable:

$$\ln(C) = b_o + b_1 \ln(Q) + b_2 \sin(2\pi t) + b_3 \cos(2\pi t) + b_4 t + \varepsilon \qquad (Eq.\,3)\ \mathbf{ln(C)}\ \boldsymbol{b_o b_1}\ \mathbf{ln(Q)}\ \boldsymbol{b_2}\ \mathbf{sin(2\pi t)}\ \boldsymbol{b_3}\ \mathbf{cos(2\pi t)}\ \boldsymbol{b_4}$$

where the terms are defined as above.

Unlike LOADEST, which fits the rating curve model once using all input data within a given year, WRTDS develops a grid of regularly spaced nodes in flow-time space and applies the rating curve model at each node. When applying the rating curve model at a specific node, only observations that are within a user-defined "window" of the node are included in the model fit. The "window", centered at each node with user-defined half-lengths in flow and time, identifies observations to be included when fitting the model at the specific node. Different weights are applied to the selected observations based on the distance in flow–time space between the observations and the node, with observations closer to the node having more weight (i.e., influence) on the model fit compared with points further away. The model fitted at a specific node is then used to predict the concentration at that node. Note that the fitted model weighting at each node is unique because of different observations included during model fitting, as selected by the moving time and flow window, and the node-specific weights applied to the observations. Concentrations between nodes can be estimated by bi-linear interpolation of nearby node concentrations (i.e., the predicted value is a linear interpolation of node values in the flow-time space). Because of these procedures, WRTDS requires the user to choose values for the parameters that control the size of the "window" (Hirsch and De Cicco, 2015). The developers of WRTDS provide some guidelines for selecting values for these parameters, but each site (and possibly each year) may require determining the appropriate parameter values. Different combinations of parameter values can produce different load estimates, resulting in a trade-off between the potential for more reasonable load estimates with increased methodological complexity. By comparison, LOADEST is more straightforward during model setup (i.e., a single regression model is used), and avoids the increase in complexity that can result from having too many "knobs to tweak".

The WRTDS-K is an extension of WRTDS and can be run after running WRTDS. WRTDS-K performs two primary procedures that distinguish it from WRTDS: it replaces the predicted concentrations with the observed concentrations on the measurement days, and it addresses correlated model residuals on the days between observations using the observations, WRTDS model residuals, and a Kalman filter in a Monte Carlo framework (Zhang and Hirsch, 2019). Briefly, WRTDS-K assumes that the WRTDS model residuals are serially correlated. In other words, if WRTDS overpredicts two consecutive observed values, WRTDS-K assumes that the WRTDS model predictions for the time between the two observations are also overpredicted. WRTDS-K accounts for the serial correlation in model residuals and produces an updated predicted value for each time interval. This process is conducted within a Monte Carlo simulation framework to acquire a distribution of WRTDS-K model predictions for each time interval rather than the single predicted value by WRTDS. The final predictions produced by WRTDS-K is the average of the predictions based on the Monte Carlo simulations. Complete details on the methods used by WRTDS-K are provided in Zhang and Hirsch (2019).

As described above, WRTDS and WRTDS-K produce different annual load estimates. As WRTDS-K was specifically developed for estimating loads for a defined period of time (rather than for estimating time trends in loads, as was the case for WRTDS), we will primarily focus on the results of WRTDS-K but include results for WRTDS for completeness. The "EGRET" R® package provides functions to estimate bootstrap-based uncertainty for loads estimated by WRTDS-K; however, the package does not include functions to estimate uncertainty for the WRTDS annual loads. Since WRTDS/WRTDS-K only accept daily mean flow values as inputs to the program, daily mean river flow measured at the Waterford USGS gauge station (Station ID# 01335754) was used and same-day measurements of Tri+ PCB concentrations were averaged prior to running the program. Default values for the parameters controlling the weights were used for load estimates, and the minimum number of samples for the WRTDS regression at each node was set to 20. Analysis of the impact of different parameter values (e.g., window sizes, minimum number of samples) on WRTDS/WRTD-K load estimates was beyond the scope of this memorandum.

3 WATERFORD TRI+ PCB LOAD ESTIMATES AND MODEL PERFORMANCE FOR THE 2016 TO 2021 POST-DREDGING PERIOD

3.1 Data Handling for Tri+ PCB Load Estimation at Waterford

Water column PCB concentration data from the Waterford monitoring station were obtained from GE's annual Data Summary Reports. Tri+ PCB concentrations used to calculate the Tri+ PCB load in 2016 were quantified using the modified Green Bay Method (mGBM), with the exception of two samples collected early in 2016 that used an Aroclor-based method. Beginning in 2017, Tri+ PCB concentrations were quantified using the EPA congener-based method 1668C. Parent and field duplicate samples were combined based on the following criteria:

- If concentrations in both the parent and duplicate samples were detected, then the values were averaged.
- If one sample was detected and the other was reported as non-detect, then the detected value was used.
- If both samples were reported as non-detect, then the maximum of the two reporting limit (RL) values was used.

Summation of individual congeners (for samples analyzed using Method 1668) or peaks (for samples analyzed using the mGBM) for estimation of Tri+ PCB concentrations was based on detected values and values flagged with a “J” qualifier (i.e., estimated). For samples analyzed using the Aroclor-based method, the calculation of Tri+ PCB concentrations was calculated using a site-specific regression equation (GE, 2016). Aroclor-based calculations also included concentrations based on both detected values and values flagged with a “J” qualifier. In the case of parent-duplicate pairs, calculation of Tri+ PCB values was done prior to averaging.

Flow data were obtained from the Waterford USGS gauging station (#01335754). The closest 15-minute flow value was assigned to each sample based on the sample date and time; however, if no 15-minute flow value was within 6 hours of a sample date and time, or if the sample did not have a reported sample time, the daily mean flow value for the sample date was used.

Water temperature data for each sample were obtained from water quality measurements collected by GE concurrent with the collection of the water samples. If no water temperature value was available in GE’s water quality database for a given sample, water temperature data from the Albany USGS gauge station (#01359139) were assigned to the sample, following similar procedures to the flow data with respect to date and time. The closest 15-minute water temperature value was assigned to each sample based on the sample date and time; however, if no 15-minute temperature value was within 6 hours of a sample date-time, or if the sample did not have a sample time, the daily mean temperature value for the sample date was used. If there was no daily mean temperature value for a sample date, then the long-term average (2002-2022) daily mean Albany water temperature on the month and day of the sample collection date was used.

3.2 Tri+ PCB Load Estimates for the 2016 to 2021 Post-Dredging Period

Models were developed for each individual calendar year (January 1st to December 31st) from 2016 to 2021 using Tri+ PCB concentration and flow and water temperature measurements within each respective year. Model predictions were then made for all time intervals within the year, with the time interval depending on whether daily mean or 15-minute interval river flow and water temperature data were used.

Figure 3-1 presents time series of model predicted concentration and observed concentrations for 2016 to 2021, and Table 3-1 and Figure 3-2 present the estimated annual Tri+ PCB load at Waterford for the post-dredging years 2016 to 2021 using the two LOADEST models and WRTDS/WRTDS-K. For completeness, we have included annual loads calculated using the AutoBeale software, as this was the method used to present loads in the Second Five-Year Review (EPA, 2019). However, the remainder of the memorandum will focus on the LOADEST and WRTDS/WRTDS-K models. The two LOADEST models (the site-specific model and Model “6”) produced load estimates within 7 percent of each other when comparing estimates based on the same data time interval (i.e., 15-minute, and daily mean values). Estimates were also similar when comparing LOADEST estimates based on daily mean values to WRTDS-K estimates, with annual differences within 10 percent. WRTDS consistently produced the highest load estimates, with annual estimates between 7 percent and 23 percent higher than estimates based on the LOADEST models using daily mean values. Overall, the similarity in load estimates across different methods provides confidence that the annual load estimates are reasonable.

A comparison of load estimates based on 15-minute versus daily mean values using the LOADEST program indicates the estimates were essentially the same, except for 2019 (where the estimate based on daily mean values was 11 percent higher for the site-specific model and 8 percent higher for Model “6”). Similarly, WRTDS-K estimates were very similar to LOADEST estimates based on the 15-minute interval data except for 2019, where the load estimate for WRTDS-K was 14 percent and 19 percent higher than the LOADEST site-specific and Model “6”, respectively. As noted previously, WRTDS consistently produced slightly higher load estimates of all models. The higher loads estimated by WRTDS (and WRTDS-K to an extent) are likely in part related to the use of daily mean flow values (as discussed below).

Of the six years analyzed, 2019 exhibited the largest difference between the LOADEST and WRTDS/WRTDS-K models. It should be noted that 2019 was the highest flow year of all years analyzed, and it included a sample with the highest concentration measured between 2016 and 2021 (142 ng/L on 1/25/2019 with an associated 15-minute flow value of 42,700 cubic feet per second [cfs] and a daily mean flow value of 39,300 cfs). Other sample days within the post-dredging period had similar (or higher) flows than 1/25/2019 but Tri+ PCB concentrations were substantially lower (as an example, on 11/02/2019 15-minute flows were recorded as 53,200 cfs with a Tri+ PCB concentration of 71 ng/L). A review of daily loads for 2019 based on the LOADEST site-specific model using 15-minute values and WRTDS-K (which uses daily mean values), indicates the difference in the daily load estimate for 1/25/2019 between these two models can account for most of the differences in the predicted annual load between the two models (daily load estimated on 1/25/2019 was 13 kg using WRTDS-K and 1.5 kg using the LOADEST site-specific model with 15-minute values). The higher estimated daily load on 1/25/2019 using WRTDS-K reflects how this model uses the high Tri+ PCB value. That is, the model assumes the Tri+ PCB concentration over the entire 24-hour period was 142 ng/L. Note that, the 15-minute flow-based LOADEST site-specific model underpredicted the concentration for the time interval when the anomalously high sample was collected (predicted value of 20 ng/L vs. observed value of 142 ng/L), so it likely underestimated the total load for the day. No other water samples were collected at Waterford (or any other monitoring station) on 1/25/2019 to provide information on why this particular sample had a very high PCB concentration, or for how long PCB concentrations were elevated.

The use of 15-minute flow data is preferred as it more accurately reflects flow variation during storm events and does not require the averaging of samples collected within the same day. Representation of sub-daily data, especially during storm events is the reason why the LOADEST program is preferable to the WRTDS/WRTDS-K program for estimating loads at Waterford. Overall, the LOADEST site-specific model produced essentially the same estimates as the built-in LOADEST rating curve model (Model "6"). Additionally, LOADEST models using daily mean values produced broadly similar load estimates to WRTDS and WRTDS-K. This general agreement provides support that load estimates presented here can be considered reasonable estimates of annual Tri+ PCB loads for the range of flow values observed between 2016 and 2021.

3.3 Performance Metrics for Load Estimation Models

In general, the evaluation of the performance of a model on a dataset requires the determination of how well the predictions made by the model match the observed data. This section describes several metrics to evaluate model performance. Evaluations of these metrics to assess model performance are given later in the section. The first performance matrix is a "leave-one-out" cross validation (LOOCV) for all load estimation methods using daily mean data. LOOCV involves iteratively removing a single observation from the dataset, fitting the model to the remaining observed data, then using the resulting model to predict the concentration of the observation that was removed. LOOCV provides a measure of how well a model predicts unobserved values. The root mean squared error for the LOOCV analysis (RMSECV) was used to represent this performance metric where:

$$RMSECV = \sqrt{\frac{\sum_{k=1}^{N}\left(\widehat{L_{k-cv}} - L_k\right)^2}{N}} \qquad (Eq.4)$$

Where for the case of LOOCV, $\widehat{L_{k-cv}}$ is the predicted load for the *k*th observation that was removed, L_k is the observed load for the *k*th observation and *N* is the total number of observations used in the analysis. The LOOCV analysis was based on load estimates using daily mean values in order to compare LOADEST results to WRTDS/WRTDS-K results. In general, when comparing model RMSECV results, the model that best fits the data is the one with the lowest RMSECV. The RMSECV results for the four models on a daily load basis are shown in Table 3-2.

Separately, four additional model performance metrics were developed using the LOADEST site-specific model and Model "6" with 15-minute data because the higher resolution dataset more accurately reflects the fluctuations in storm events flows. The model performance metrics were not calculated for WRTDS/WRTDS-K, since these methods only accept daily mean values, comparing results to the LOADEST results based on 15 minutes would not be meaningful. Further, since WRTDS-K uses observed values when available in place of predictions, these metrics are not applicable to WRTDS-K.

The four model performance metrics for the two LOADEST models are presented in Tables 3-3 through 3-6, including the root mean square error of estimation (RMSEE), partial load ratio (PLR), the percent bias (Bp), and the Nash-Sutcliffe Efficiency Index (E). The RMSEE is calculated using:

$$RMSEE = \sqrt{\frac{\sum_{k=1}^{N}\left(\widehat{L_{k-e}} - L_k\right)^2}{N}} \qquad (Eq. 5)$$

Where $\widehat{L_{k-e}}$ is estimated from a model that includes all observations (i.e., no data is left out), and L_k is the same as defined for Equation 4 above, there is an important difference between RMSEE and RMSCV. While Equation 4 for RMSECV and Equation 4 for RMSEE look similar, it is important to note that difference as follows:

- When calculating the RMSECV as presented in Table 3-2, each $\widehat{L_{k-cv}}$ is estimated from a model developed with one observation (L_k) left out. Thus, *N* models are fit when calculating the RMSECV (with each of the *N* models having a different observation removed).
- When calculating the RMSEE, for the 15-minute interval model results, as presented in Table 3-6, $\widehat{L_{k-e}}$ is estimated from a model that includes all observations (i.e*.,* no data is left out). Thus, all $\widehat{L_{k-e}}$ are derived from a single model fit.

Therefore, while the RMSECV measures how well the model predicts unobserved values, the RMSEE on the other hand measures the average deviation of the model estimates for all predicted *N* measurements from the observed *N* values.

PLR compares the total observed load summed over all observed time intervals (15-minute intervals in this case) with predicted loads over the same time intervals. PLR is defined as (Runkel et al., 2004):

$$PLR = \frac{\sum_{k=1}^{N} \widehat{L_{k-e}}}{\sum_{k=1}^{N} L_k} \qquad (Eq. 6)$$

where the $\widehat{L_{k-e}}$ term is as defined for RMSEE. A PLR value of 1 indicates a perfect estimation of load on the sampled time intervals, while a value less than 1 indicates an underprediction, and a value greater than 1 indicates an overprediction.

Bp is a measure of the bias in the total load estimated for the sampled time intervals, as a percent of the measured total load for the sampled time intervals. Bp is defined as:

$$Bp = 100 * \left(\frac{\sum_{k=1}^{N} \left(\hat{L}_{k-e} - L_k \right)}{\sum_{k=1}^{N} L_k} \right) \qquad (Eq. 7)$$

where the terms are defined as in Equation 5. A Bp value of zero indicates no bias in predictions, while a value less than zero indicates a low bias and a value greater than one indicates a high bias.

E compares the sum of the squared difference between the estimated and measured load for the sampled time intervals to the sum of squared difference between the estimated load and the average measured load on the sampled time intervals (Runkel et al., 2004). E is defined as:

$$E = 1 - \frac{\sum_{k=1}^{N} \left(L_k - \hat{L}_{k-e} \right)^2}{\sum_{k=1}^{N} (L_k - \bar{L})^2} \qquad (Eq. 8)$$

where $\bar{L}$ is the average 15-minute measured load over the period analyzed (a calendar year in our analysis) and the other terms are defined as in Equation 5. An E value of one equals a perfect match between observed and predicted loads. A value of zero indicates the use of the average measured load is equally as good as using predicted values. A value less than zero indicates using the average measured load is better than using the predicted values.

Table 3-2 provides the results of the RMSECV for the LOOCV analysis for both the LOADEST models and WRTDS/WRTDS-K models on a daily mean basis. The results of the LOOCV analysis indicate that the LOADEST models and WRTDS/WRTDS-K methods produced similar RMSECV values for load. The WRTDS/WRTDS-K methods performed best in one of the five years (2019), while the LOADEST site-specific model performed best in three of the six years (2016, 2018, and 2020). The LOADEST Model "6" performed best in the remaining two years (2017 and 2021). While no single model performed best in all of the years analyzed, the results based on this statistical measure show that the LOADEST site-specific model performed best over the largest number of years.

A review of PLR, Bp, E, and RMSEE values using 15-minute interval data for the LOADEST site-specific model and Model "6" provided in Tables 3-3 to 3-6 indicates that, overall, the LOADEST site-specific model performed better than Model "6" for the 2016 to 2021 time period. as noted previously, the WRTDS/WRTDS-K models could not be tested in this manner because they require, among other things, daily average data and WRTDS-K replaces predicted values with observed values. The site-specific model metrics were consistently better in four of the five years. The notable exception was 2017, where Model "6" performed slightly better for all four metrics. In general, the models performed poorest in 2019, with a Bp of -9.4 percent and -14.6 percent for the site-specific model and Model "6", respectively, suggesting an underpredicted load value. This is consistent with our discussion of 2019 results in Section 3.1. For the LOADEST site-specific model, the remaining four years have absolute percent bias less than 6 percent, with three years (2017, 2018, and 2020) having an absolute percent bias of less than 2 percent.

The performance of the LOADEST site-specific model was further analyzed using model diagnostic plots (see Figure 3-3). Hirsch (2014) demonstrated that load estimates may be biased if linear regression-based rating curve models, such as those used here, do not conform to assumptions of linear regression analysis (particularly normally distributed model residuals with constant variance). To assess this, three plots were created for each year: measured load vs. predicted load, model residuals vs. predicted load, and a quantile-quantile plot of model residuals. The first plot provides an assessment of how well predicted and observed values agree. The second plot provides information on the assumption of homoscedasticity (constant variance over the range of predicted values) in linear regression models. The last plot assesses whether the residuals are normally distributed. The plots are shown in log-units, which are the units used in fitting the model. The plots in Figure 3-3 indicate that the LOADEST site-specific model performs well overall and broadly conforms to the assumptions of linear regression analysis. An exception is in 2019, where one sample (1/25/2019, discussed in Section 3.1) is not fit well by the model and appears to be an outlier for this dataset. However, given a lack of additional data during the event to further constrain and improve the model prediction during this day, there is no basis to remove it from the dataset. Exclusive of this point, the rest of the data for 2019 generally conform to the assumptions of linear regression analysis.

4 IMPORTANCE OF HIGH-FLOW EVENTS FOR TRANSPORT OF PCBS AT WATERFORD

Previous studies conducted in PCB-contaminated rivers, including the Upper Hudson River, have found that high-flow events can be associated with a significant portion of the total PCB load for a given year (EPA, 1997; EPA, 1999). This section uses the Site-specific model developed within the LOADEST estimation program to provide estimates of the amount of Tri+ PCB mass transported during high-flow events. To estimate Tri+ PCB load during high-flow events, the LOADEST site-specific model described previously was calibrated with mean daily flow values and used to predict daily Tri+ PCB loads between 2016 and 2021. High-flow loads were defined as Tri+ PCB loads occurring on days when daily mean flow was greater than 22,500 cfs measured at the Waterford USGS gauge station. This flow value was used in both the Baseline Monitoring Program (BMP) and the Remedial Action Monitoring Program (RAMP) associated with the dredging conducted in the Upper Hudson River between 2009 and 2015 to trigger more frequent sampling during storm events (GE, 2012). Additionally, the half-load discharge value ($Q_{1/2}$), or the flow value above which 50 percent of the load over a given period of time is transported (Vogel et al., 2003), was calculated as another metric to assess the importance of high flows on Tri+ PCB transport in the Upper Hudson River. The $Q_{1/2}$ was determined by first ranking the flow from lowest to highest for the given time period, then estimating the fraction of accumulative daily load vs total load sequentially. The $Q_{1/2}$ represents the flow corresponding to the point where the accumulative daily load equals half of the total load for the given time period. The percentage of flow days in excess of $Q_{1/2}$ (see Figure 4-1) in the given time period (the number of days where flow was higher than $Q_{1/2}$ divided by total days in the time period, e.g., 365) can highlight the importance of high flows in the transport of PCB load.

Table 4-1 provides summary statistics for Tri+ PCB load estimates at Waterford based on the LOADEST site-specific model, including load estimate totals for the days when daily mean flows exceeded 22,500 cfs at the Waterford USGS gauge station. The percentage of high-flow days varied from less than 1 percent of all days within a year for 2016, 2020 and 2021 to 10 percent in 2019 (see Table 4-1). The Tri+ PCB load transported on days with flows greater than 22,500 cfs ranged from 1 kg (2 percent of total load) in 2021 to 65 kg (58 percent of total load) in 2019. Interestingly, the amount of Tri+ PCB load transported under low flow conditions (<22,500 cfs) was fairly consistent across all years, ranging between 34 kg and 47 kg. These results indicate that even during a relatively low flow year with few high-flow days (i.e., 2016, 2020, and 2021), the few high-flow days can transport on the order of 15 percent of the annual load. On the other hand, in a high-flow year like 2019, 10 percent of the days can deliver almost 60 percent of the total annual load. Overall, from 2016 to 2021, high-flow days (equal to or greater than 22,500 cfs) occurred just 3 percent of the time but delivered 30 percent of the total load for the five-year period (see Table 4-1).

To provide additional context regarding the importance of high-flow days on Tri+ PCB transport, the $Q_{1/2}$ for each year and for all years combined was calculated. To better visualize $Q_{1/2}$, Figure 4-1 presents the cumulative Tri+ PCB load fraction between 2016 and 2021 derived from LOADEST vs. daily mean river flow. Also presented in Figure 4-1 is the exceedance fraction of measured flows for 2016 to 2021. The cumulative Tri+ PCB load fraction provides information on the amount of Tri+ PCB mass transported at flows greater than a particular flow value, while the exceedance fraction of measured flows provides information on how many times a particular flow value was exceeded in a given period of time. Table 4-1 provides $Q_{1/2}$ values for each year as well as all years combined. $Q_{1/2}$ values range from 7,890 cfs in 2016 to 26,800 cfs in 2019, reflecting in part whether a year was a low flow year or a high-flow year. In 2016, a relatively low flow year, 50 percent of the load was transported over 27 percent of the days within the year, while in 2019, a high-flow year, the $Q_{1/2}$ value indicates that 50 percent of the load was transported on only 5 percent of the days within the year. Combining all years together, $Q_{1/2}$ was calculated as 13,700 cfs. Based on this value, 50 percent (or approximately 170 kg) of the cumulative Tri+ PCB load between 2016 and 2021 was transported over only 17 percent of the days, or about 62 days annually.

The results of these analyses indicate that relatively infrequent high-flow events can transport a substantial amount of Tri+ PCBs. The rating curve model with the 15-minute flow interval data provides the best basis to integrate total load during these events and estimate their contribution to the total load for the entire year. In selecting a model for load calculation, it is important that the model be able to properly represent these short-term measurements in estimating annual load, supporting the use of the LOADEST-based models.

5 METHODOLOGY FOR CALCULATING TRI+ PCB LOADS AT WATERFORD DURING THE OM&M PROGRAM

This section outlines the recommended modeling methodology to be used to estimate Upper Hudson River PCB loads under the OM&M program and briefly describes the basis for these recommendations, based on the analyses described previously.

5.1 Preferred Method for Estimating Annual Loads at Waterford

Although both the LOADEST models and WRTDS/WRTDS-K produced similar annual load estimates, the primary drawback for WRTDS/WRTDS-K is its inability to accept sub-daily input data. As discussed in Sections 3 and 4, properly characterizing short-term variation during high-flow events is critical for development of an accurate rating curve model and estimation of annual loads. Since sub-daily sampling is needed and will occur during storm events to properly characterize these dynamic events, the selected calculation method should be able to accept sub-daily measurements of flow and concentration as input data. Therefore, LOADEST is preferable to WRTDS/WRTDS-K. An additional advantage of LOADEST over the other approach discussed in this memorandum is the ability for the user to develop a site-specific rating curve model that explicitly includes physical processes known to influence water column PCB concentrations. Based on review of the model performance metrics for the two LOADEST models, the site-specific LOADEST rating curve model that includes segmented regression and temperature (Equation 1) generally performed better over the years analyzed compared with Model “6”. The site-specific LOADEST model is the recommended method for estimating Tri+ PCB loads moving forward.

The “rloadest” package will be used to estimate annual loads, although the FORTRAN version is also available for comparison if necessary. The WRTDS/WRTDS-K program is also implemented in R® and is relatively easy to run. Load estimates should be generated by both methods as an additional check on whether estimated loads are reasonable. Finally, the study of methods for estimating constituent loads in rivers is an area of ongoing research at the USGS and elsewhere, and if in the future another load estimation method is determined to provide more robust estimates of load (compared to LOADEST), the use of LOADEST may be re-evaluated at that time. Finally, as additional years of data are collected, it may become appropriate to utilize WRTDS/WRTDS-K’s procedures for estimating long-term time trends in PCB load.

5.2 Data Requirements for Estimating Annual Loads at Waterford

Data requirements for estimating loads are as follows: 1) measurements of Tri+ PCB concentration and associated flow and water temperature at the time of measurement spanning a range of time and flow across the period of interest (for model calibration), and 2) flow and water temperature data for the entire year where annual loads are to be estimated, preferably on a 15-minute interval basis (for prediction). Both 15-minute and daily mean flow and water temperature are currently available via the USGS gauge station at Waterford (Station ID# 01335754) and the USGS gauge station at Albany (Station ID# 01359139), respectively. USGS 15-minute flow and water temperature data will be the preferred input data to estimate annual loads. Prior to running LOADEST, flow and water temperature data will be reviewed for spurious values (e.g., zero or unrealistic flow and temperature values). If spurious or missing values are identified, these values will be replaced with either estimated values using linear interpolation of temporally proximal data values or daily mean values, depending on the amount of spurious data. If for some reason 15-minute data are not available for a given year, daily mean flow data and water temperature will be used instead. If daily mean flow and temperature data are used to for load estimation, same-day Tri+ PCB measurements should be averaged to produce daily mean Tri+ PCB concentrations prior to load estimation.

6 CONCLUSIONS

A goal of the OM&M program is to monitor and track Tri+ PCB loads transported from the Upper Hudson to the Lower Hudson. There are three essential requirements to produce accurate estimates of annual Tri+ PCB loads including:

- the characterization of high-flow events by collection of samples to represent the transport under these conditions,
- the collection of routine samples to characterize the relationship between flow and concentration at low flow conditions, and
- the development of an appropriate rating curve model to represent and integrate these results.

This memorandum evaluated and described the rating curve model that will be used to estimate Tri+ PCB at the Waterford monitoring station during the OM&M program. The model consists of two components: a segmented regression component that accounts for the non-linear relationship between Tri+ PCB concentration and flow, and a water temperature component that accounts for the seasonality in Tri+ PCB concentration. This rating curve model is implemented using the USGS LOADEST software in the R® programming environment. Using data from 2016 to 2021, this rating curve model was compared to another common LOADEST model (Model "6") and also to a separate USGS-produced load estimation program (WRTDS/WRTDS-K) using multiple model performance metrics. While all methods and models presented produced similar annual load estimates, a key differentiator between LOADEST and WRTDS/WRTDS-K is the ability of LOADEST to accept sub-daily input data. This is important because 1) sub-daily sampling is essential to characterize short-term load variation associated with short-term flow variation during high-flow events, and 2) 15-minute flow data more accurately reflect flow variability, particularly during storm events when flow (and corresponding Tri+ PCB concentration) can vary widely over relatively short time periods. Based on these considerations, the LOADEST rating curve program was selected for use in the OM&M monitoring program since it was able to use both forms of sub-daily data while yielding model performance statistics comparable to other models, when compared on a daily mean basis. Analysis of the reliability of LOADEST Tri+ PCB load estimates will continue moving forward. If in the future, analyses suggest the reliability of load estimates produced by LOADEST has decreased or another method consistently produces more reliable results, the use of LOADEST as the primary load estimation software may be re-evaluated.

Regardless of whether a year is characterized as a high-flow or low-flow year, substantial Tri+ PCB mass is transported during the highest flow days within a year. Estimation of the half-load discharge value ($Q_{1/2}$) for 2016 to 2021 indicates that 50 percent of the Tri+ PCB load was transported on only 17 percent of the days. To further emphasize this point, days with flows in excess of 22,500 cfs represent only 3 percent of the total period but were responsible for 30 percent of the total load. Therefore, it is critical that the water column sampling program continues to target periods of high flows in addition to the routine sampling in order to properly account for the large amount of Tri+ PCB transport during high-flow periods and provide the data needed to properly calibrate the rating curve model.

7 ABBREVIATIONS AND ACRONYMS

AMLE	Adjusted Maximum Likelihood Estimation
BMP	Baseline Monitoring Program
Bp	Percent bias
cfs	cubic feet per second
DQO	Data quality objective
E	Nash-Sutcliffe efficiency index
EPA	United States Environmental Protection Agency
GE	General Electric
kg	Kilogram
LOADEST	USGS Load Estimator Program
LOOCV	Leave-one-out cross validation
MLE	Maximum Likelihood Estimation
MVUE	Minimum Variance Unbiased Estimator
NWQN	National Water Quality Network
ng/L	Nanogram per Liter
OM&M	Operation, Maintenance & Monitoring
PCB	Polychlorinated Biphenyl
PLR	Partial load ratio
$Q_{1/2}$	Half-load discharge value
RAMP	Remedial Action Monitoring Program
RMSECV	Root mean squared error for the LOOCV analysis
RMSEE	Root mean square error of estimation
USGS	United States Geological Survey
WRTDS	Weighted Regression on Time, Discharge and Season
WRTDS-K	Weighted Regression on Time, Discharge and Season – Kalman

8 REFERENCES

Cohn, T.A. 2005. "Estimating contaminant loads in rivers: An application of adjusted maximum likelihood to type 1 censored data." In *Water Resources Research* 41. https://doi.org/10.1029/2004WR003833.

Cohn, T.A., D.L. Caulder, E.J. Gilroy, L.D. Zynjuk, and R.M. Summers. 1992. "The validity of a simple statistical model for estimating fluvial constituent loads: An Empirical study involving nutrient loads entering Chesapeake Bay." In *Water Resources Research* 28, 2353–2363. https://doi.org/10.1029/92WR01008.

Cohn, T.A., L.L Delong, E.J. Gilroy, R.M. Hirsch, and D.K. Wells. 1989. "Estimating constituent loads." In *Water Resources Research* 25, 937–942. https://doi.org/10.1029/WR025i005p00937.

EPA (United States Environmental Protection Agency). 1997. Phase 2 Report – Review Copy, Further Site Characterization and Analysis, Volume 2C – Data Interpretation and Evaluation Report, Hudson River PCBs Reassessment RI/FS. Prepared for the United States Environmental Protection Agency Region 2 and United States Army Corps of Engineers by TAMs Consultants, the Cadmus Group and Gradient. February.

_______. 1999. Phase 2 Report Review Copy - Further Site Characterization and Analysis, Volume 2D - Baseline Modeling Report, Hudson River PCBs Reassessment RI/FS. Prepared for United States Environmental Protection Agency Region 2 and United States Army Corps of Engineers Kansas City District by Limno-Tech, Menzie-Cura & Associates and Tetra-Tech. May.

_______. 2010. Attachment E to Statement of Work, Hudson River PCBs Site, Operations, Maintenance and Monitoring Scope for Phase 2 of the Remedial Action. December.

_______. 2019. Final Second Five-Year Review Report for the Hudson River PCBs Superfund Site, Appendix 1 Evaluation of Water Column PCB Concentrations and Loadings. Prepared by Louis Berger and Limno Tech. April.

General Electric (GE). 2012. Hudson River PCBs Site Phase 2 Remedial Action Monitoring Program Quality Assurance Project Plan. Prepared for General Electric by Anchor QEA and Environmental Standards. May.

_______. 2016. Water Column Tri+ PCB Correlations Refined with 2015 Data. Remedial Action Monitoring Program CAM No. 11. Prepared for General Electric by Anchor QEA. January.

Gilroy, E.J., R.M. Hirsch, and T.A. Cohn, T.A. 1990. "Mean square error of regression-based constituent transport estimates." *Water Resources Research* 26, 2069–2077. https://doi.org/10.1029/WR026i009p02069.

Hirsch, R.M. 2014. "Large biases in regression-based constituent flux estimates: Causes and diagnostic tools." *JAWRA Journal of the American Water Resources Association*, 50(6), 1401-1424. https://doi.org/10.1111/jawr.12195.

Hirsch, R.M., and L.A. De Cicco. 2015. "User guide to Exploration and Graphics for RivEr Trends (EGRET) and dataRetrieval: R packages for hydrologic data" (Report No. 4-A10), *Techniques and Methods*. https://doi.org/10.3133/tm4A10.

Hirsch, R.M., D.L. Moyer, and S.A. Archfield. 2010. "Weighted Regressions on Time, Discharge, and Season (WRTDS), with an Application to Chesapeake Bay River Inputs." *JAWRA Journal of the American Water Resources Association* 46, 857–880. https://doi.org/10.1111/j.1752-1688.2010.00482.x.

Larsson, P. 1983. "Transport of 14C-labelled PCB compounds from sediment to water and from water to air in laboratory model systems." *Water Research* 17, 1317–1326. https://doi.org/10.1016/0043-1354(83)90258-0.

Larsson, P., L. Okla, S.O. Ryding, and B. Westoo. 1990. "Contaminated sediment as a source of PCBs in a river system." *Canadian Journal of Fisheries and Aquatic Sciences* 47, 746-754. https://doi.org/10.1139/f90-085.

Lee, C.J., R.M. Hirsch, G.E. Schwarz, D.J. Holtschlag, S.D. Preston, C.G. Crawford, and A.V. Vecchia. 2016. "An evaluation of methods for estimating decadal stream loads." *Journal of Hydrology* 542, 185–203. https://doi.org/10.1016/j.jhydrol.2016.08.059.

Lee, C.J., J.C. Murphy, C.G. Crawford, and J.R. Deacon. 2020. "Methods for computing water-quality loads at sites in the U.S. Geological Survey National Water Quality Network" (ver. 1.2, November 2020) (USGS Numbered Series No. 2017–1120), Open-File Report. U.S. Geological Survey, Reston, VA. https://doi.org/10.3133/ofr20171120.

Marsh, N., and D. Waters. 2009. Comparison of load estimation methods and their associated error. 18th World IMACS Congress and MODSIM09 International Congress on Modelling and Simulation.

Muggeo, V.M.R. 2008. "Segmented: an R package to fit regression models with broken-line relationships." *R News* 8/1, 20–25.

Richards, P.R. 1998. Estimation of Pollutant Loads in Rivers and Streams: A Guidance Document for NPS Programs. United States Environmental Protection Agency Region VIII Grant X998397-01-0, Water Quality Laboratory, Heidelberg University, Tiffin, OH.

Runkel, R.L., C.G. Crawford, and T.A. Cohn. 2004. Load estimator (LOADEST): a FORTRAN program for estimating constituent loads in streams and rivers (Report No. 4-A5), *Techniques and Methods*. https://doi.org/10.3133/tm4A5.

Vogel, R.M., J.R. Stedinger, and R.P. Hooper. 2003. "Discharge indices for water quality loads." *Water Resources Research* 39. https://doi.org/10.1029/2002WR001872.

Zhang, Q., and R.M. Hirsch. 2019. "River Water-Quality Concentration and Flux Estimation Can be Improved by Accounting for Serial Correlation Through an Autoregressive Model." *Water Resources Research* 55, 9705–9723. https://doi.org/10.1029/2019WR025338.

ATTACHMENT A

Table and Figures

Tables

Table 2-1

Estimated Flow Change Point for Tri+ PCB and Flow using Segmented Regression Included in the "segmented" R® Package

Station	Analyte	Year(s)	Change Point (cfs) (95% Confidence Interval)
Waterford	Tri+ PCB	2016	**13,400** (10,800, 16,700)
Waterford	Tri+ PCB	2017	**16,900** (12,200, 23,400)
Waterford	Tri+ PCB	2018	**14,600** (12,900, 16,600)
Waterford	Tri+ PCB	2019	**19,400** (17,200, 22,000)
Waterford	Tri+ PCB	2020	**11,900** (10,500, 13,500)
Waterford	Tri+ PCB	2021	**12,800** (10,500, 15,600)
Waterford	Tri+ PCB	All Years	**16,000** (14,900, 17,300)

Notes:

1. Change point based on segmented regression model between concentration and flow.

Table 3-1
Estimated Tri+ PCB Load for Years 2016 to 2021

	Annual Load (Kg)						
Year	LOADEST Site-Specific Model (15-Minute Flow)	LOADEST Model "6" (15-Minute Flow)	LOADEST Site-Specific Model (Daily Mean Flow)	LOADEST Model "6" (Daily Mean Flow)	WRTDS-K Daily (Daily Mean Flow)	WRTDS (Daily Mean Flow)	AutoBeale (Daily Mean Flow)
2016	**39** (36, 42)	**40** (37, 43)	**40** (34, 46)	**41** (35, 48)	**44** (41, 47)	**49**	**45** (30, 50)
2017	**68** (63, 72)	**67** (62, 73)	**67** (60, 75)	**66** (58, 74)	**68** (66, 71)	**73**	**87** (73, 101)
2018	**45** (43, 47)	**48** (46, 51)	**46** (42, 51)	**48** (44, 53)	**46** (44, 48)	**53**	**48** (43, 52)
2019	**101** (90, 112)	**97** (87, 108)	**112** (93, 130)	**105** (88, 122)	**115** (112, 119)	**123**	**139** (97, 181)
2020	**34** (32, 36)	**34** (32, 36)	**35** (32, 38)	**34** (31, 38)	**36** (35, 37)	**37**	**37** (29, 45)
2021	**46** (44, 49)	**46** (43, 48)	**47** (42, 52)	**46** (41, 50)	**43** (42, 45)	**47**	**44** (39, 49)

Notes:
1.95% Prediction Intervals are displayed in parentheses. Note that the WRTDS does not output prediction intervals for WRTDS load estimates.
2.LOADEST Site-Specific Model is the rating curve model that uses segmented regression and water temperature to account for flow and seasonality, respectively. See Section 2 for additional details.
3.All models were developed using only Tri+ PCB measurements which fell within each calendar year, and the load was the summation of all model predictions for that year based on the data resolution (i.e., daily mean or 15-minute interval).

Table 3-2
Root Mean Squared Error of Cross Validation (RSMECV) Values Determined from "Leave-one-out Cross Validation" (LOOCV) Results for LOADEST and WRTDS/WRTDS-K m\Methods with Daily Mean Data

	LOOCV RMSE (RMSECV) (kg/day)			
Year	**LOADEST Site-Specific Model**	**LOADEST With Model "6"**	**WRTDS-K**	**WRTDS**
2016	0.318	0.392	0.458	0.483
2017	0.385	0.346	0.361	0.437
2018	0.089	0.113	0.131	0.156
2019	1.41	1.41	1.38	1.42
2020	0.097	0.258	0.281	0.300
2021	0.239	0.230	0.281	0.238

Notes:

1.All results based on loads estimated using daily mean values. See Section 3.2 for additional details on LOOCV RMSE analysis.

2.LOADEST Site-Specific Model is the rating curve model that uses segmented regression and water temperature to account for flow and seasonality, respectively. See Section 2 for additional details.

3.A lower RMSECV value indicates a better model fit. Highlighted cells indicate the model that performed best for the respective year.

Table 3-3

Partial Load Ratio (PLR) for the LOADEST Site-Specific Model and "Model 6" for Years 2016 to 2021 using 15-minute Flow and Water Temperature Measurements

Year	Partial Load Ratio (PLR)	
	Site-Specific Model	Model 6
2016	0.943	0.809
2017	1.02	1.01
2018	0.983	0.949
2019	0.906	0.854
2020	0.924	0.821
2021	0.990	0.959

Notes:
1.LOADEST Site-Specific Model is the rating curve model that uses segmented regression and water temperature to account for flow and seasonality, respectively. See Section 2 for additional details.
2.A PLR value closer to one indicates a better model fit. Highlighted cells indicate which model performed better for an individual year.

Table 3-4
Percent Bias (Bp) for the LOADEST Site-Specific Model and "Model 6" for Years 2016 to 2021 using 15-minute Flow and Water Temperature Measurements

Year	Percent Bias (Bp, %)	
	Site-Specific Model	Model 6
2016	-5.69	-19.1
2017	2.13	1.23
2018	-1.74	-5.06
2019	-9.38	-14.6
2020	-7.60	-17.9
2021	-1.05	-4.08

Notes:
1.LOADEST Site-Specific Model is the rating curve model that uses segmented regression and water temperature to account for flow and seasonality, respectively. See Section 2 for additional details.
2.A Bp value closer to zero indicates a better model fit. Highlighted cells indicate which model performed better for an individual year.

Table 3-5

Nash-Sutcliffe Efficiency Index (E) for the LOADEST Site-Specific Model and "Model 6" for Years 2016 to 2021 using 15-minute Flow and Water Temperature Measurements

Year	Nash-Sutcliffe Efficiency Index (E)	
	Site-Specific Model	Model 6
2016	0.899	0.748
2017	0.322	0.382
2018	0.860	0.828
2019	0.553	0.508
2020	0.897	0.655
2021	0.345	0.407

Notes:

1.LOADEST Site-Specific Model is the rating curve model that uses segmented regression and water temperature to account for flow and seasonality, respectively. See Section 2 for additional details.

2.A higher E value indicates a better model fit. Highlighted cells indicate which model performed better for an individual year.

Table 3-6

Root Mean Squared Error of Estimation (RMSEE) for the LOADEST Site-Specific Model and "Model 6" for Years 2016 to 2021 using 15-minute Flow and Water Temperature Measurements

Year	Root Mean Squared Error (RMSEE, kg/day)	
	Site-Specific Model	Model 6
2016	0.247	0.392
2017	0.458	0.437
2018	0.124	0.137
2019	1.42	1.49
2020	0.169	0.309
2021	0.245	0.234

Notes:
1.LOADEST Site-Specific Model is the rating curve model that uses segmented regression and water temperature to account for flow and seasonality, respectively. See Section 2 for additional details.
2.A lower RMSEE value indicates a better model fit. Highlighted cells indicate which model performed better for an individual year.

Table 4-1
Summary Flow Statistics for Calendar Years 2016 to 2021, Information on the Number of High Flow Days and Amount of Tri+ PCB Load Transported during High Flow Days, and $Q_{1/2}$

Year	Minimum Daily Mean Flow (cfs)	Mean Daily Mean Flow (cfs)	Median Daily Mean Flow (cfs)	Maximum Daily Mean Flow (cfs)	Number of High Flow Days	Annual Load (kg)	Annual Load for Low Flow Days (kg)[1]	Annual Load for High Flow Days (kg)[1]	Percentage of High Flow Days (%)	Percentage of Load on High Flow Days (%)	$Q_{1/2}$ (cfs)	Percentage of Days Greater than $Q_{1/2}$ (%)
2016	1,990	6,140	4,890	28,700	2	40	34	6	0.5	14	7,850	28
2017	2,820	9,800	8,480	36,400	22	67	46	21	6	31	13,000	19
2018	2,160	9,040	8,020	30,200	11	46	38	8	3	18	11,900	28
2019	2,500	12,200	11,500	51,900	37	112	47	65	10	58	27,700	5
2020	1,320	7,340	5,120	39,200	3	35	30	5	0.8	13	9,310	34
2021	2,840	9,310	8,690	23,200	1	47	46	1	0.3	2	12,500	25
2016-2021	1,320	8,970	8,000	51,900	76	347	242	105	3	30	13,600	17

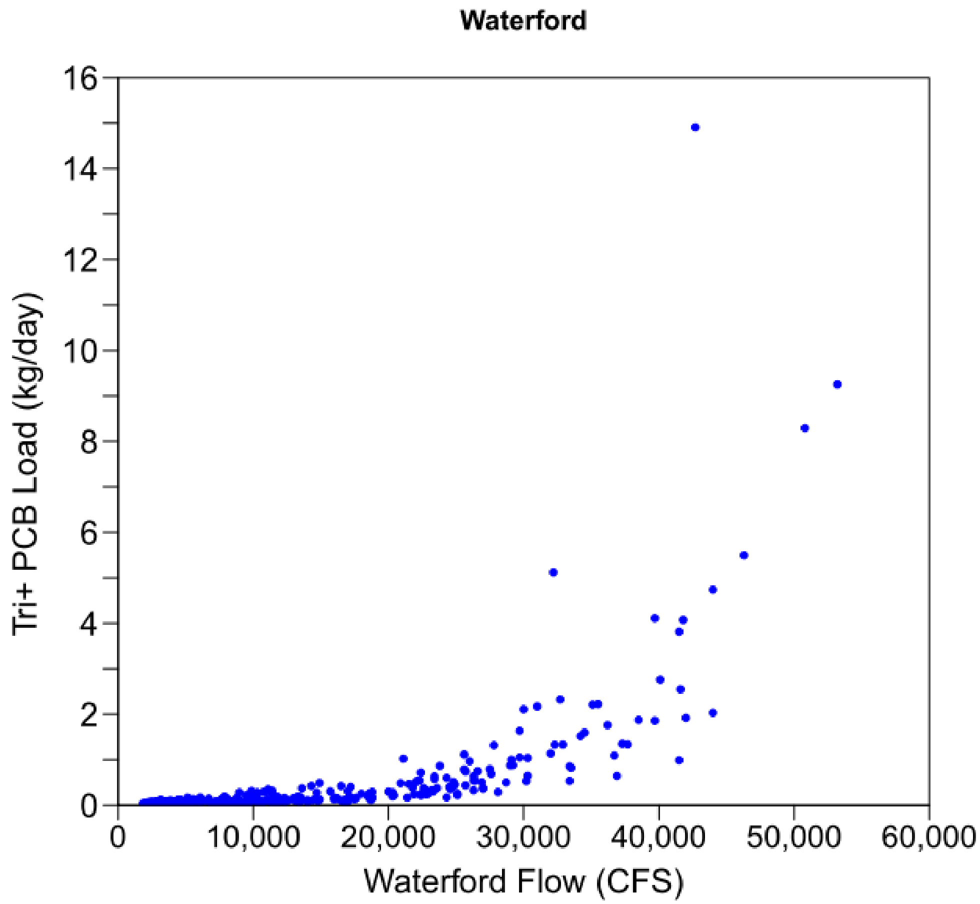
Notes:

1. Low flow days defined as days when mean daily flow <22,500 cfs. High flow days defined as days when mean daily flow was ≥ 22,500 cfs.

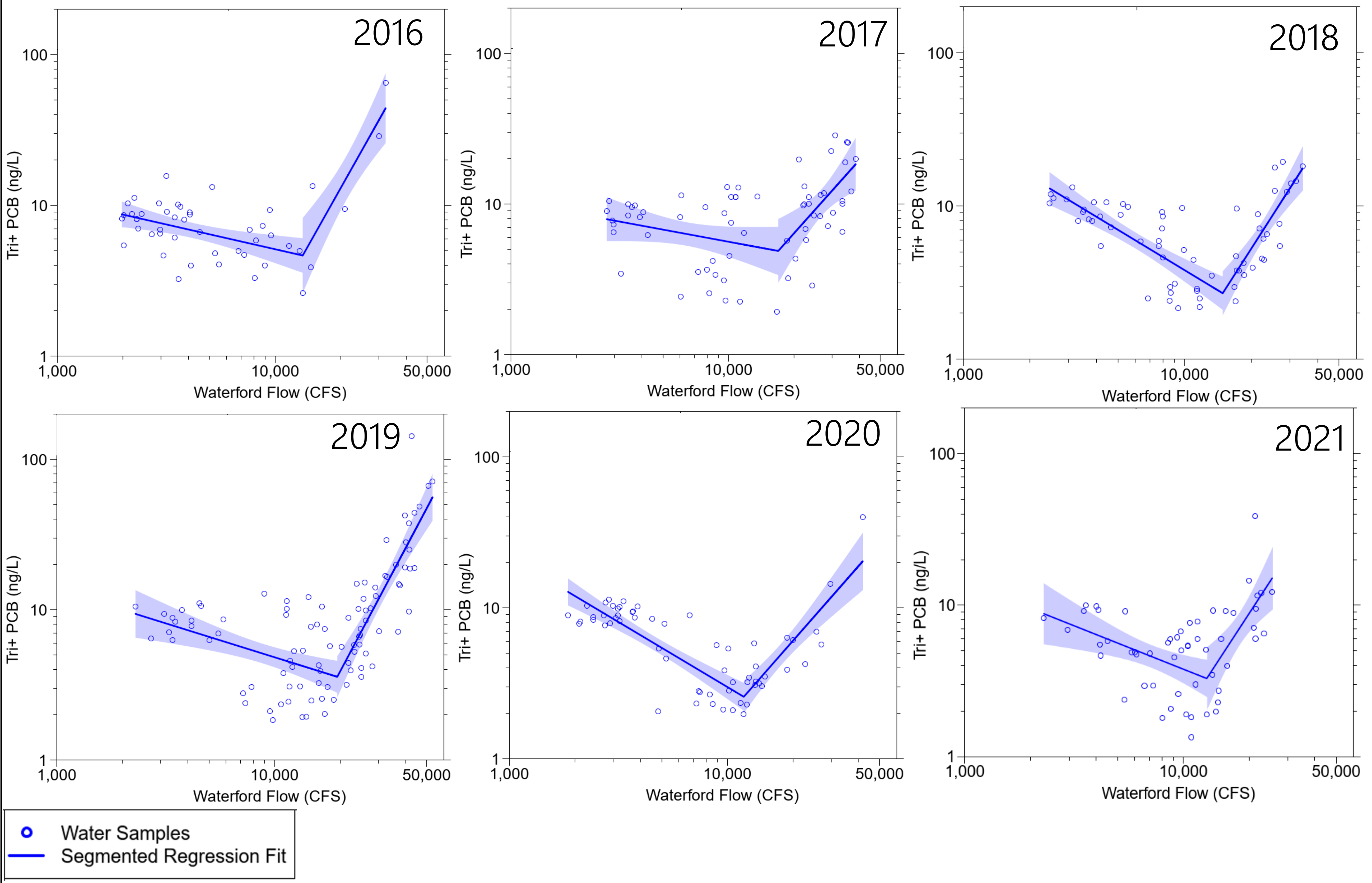
2. $Q_{1/2}$ is the flow value where half the annual Tri+ PCB load (or half the total 2016-2021 Tri+ PCB load in the case of all years combined) occurs at a greater flow value. See Section 4 for additional details.

3. For values associated with the Year 2016-2021 results, $Q_{1/2}$ and Percentage of Days Greater than $Q_{1/2}$ were recalculated using the full 6 years of daily flow and load values.

Figures

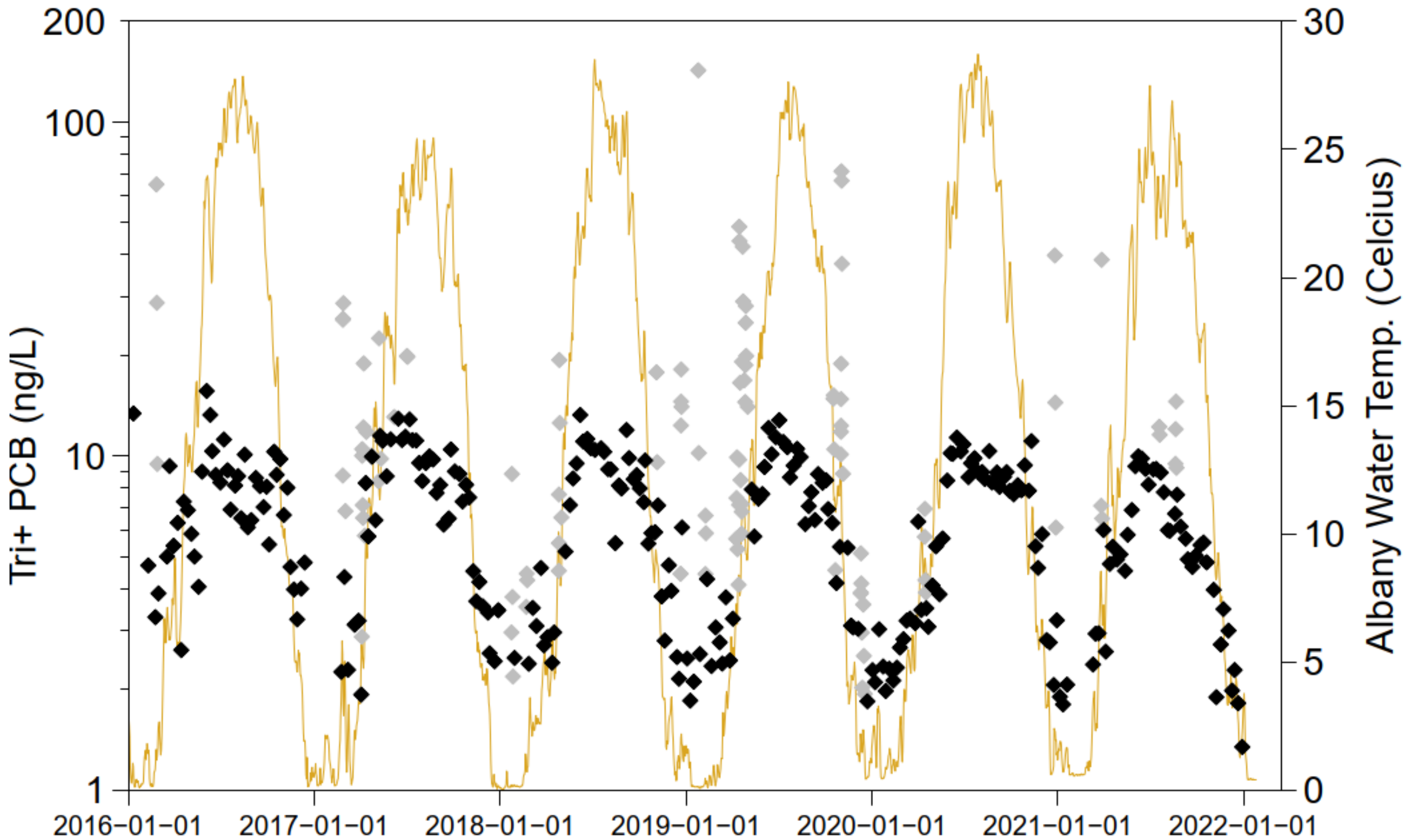


Measured Tri+ PCB Load Versus Flow for Post-Dredging Years 2016-2021 at the Waterford Monitoring Station	**Figure 2-1**
	July 2024



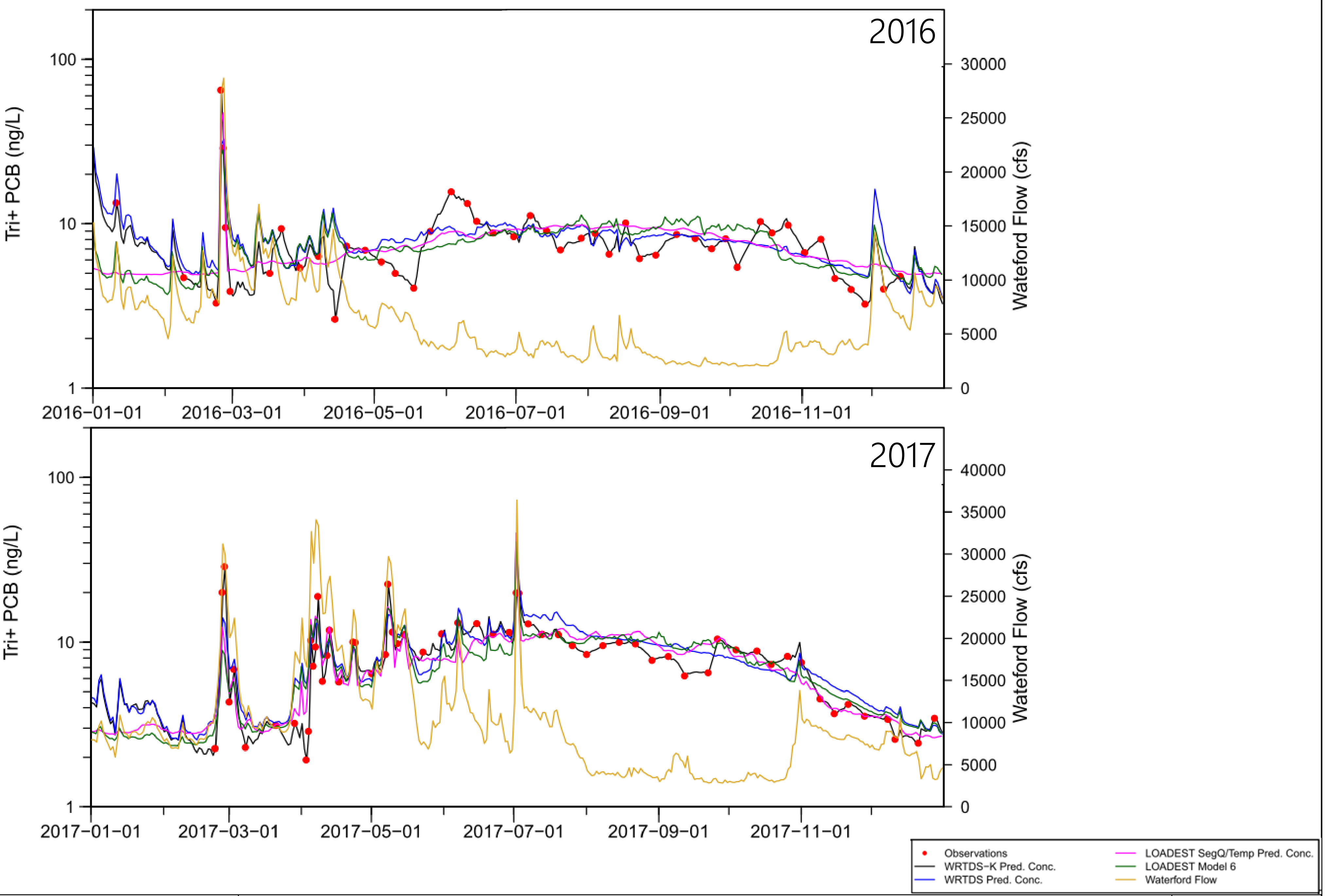
Note: Blue line represents best-fit of the segmented regression model between concentration and flow. Blue shaded area represents the 95% confidence band about the fit.

Tri+ PCB Concentration Versus Flow for Post-Dredging Years 2016-2021 at the Waterford Monitoring Station	**Figure 2-2**
	July 2024



Notes: Black diamonds are samples collected at flows less than or equal to 22,500 cfs to highlight seasonality in the water column Tri+ PCB concentrations. Water temperature data recorded at Albany USGS Station (#01359139).

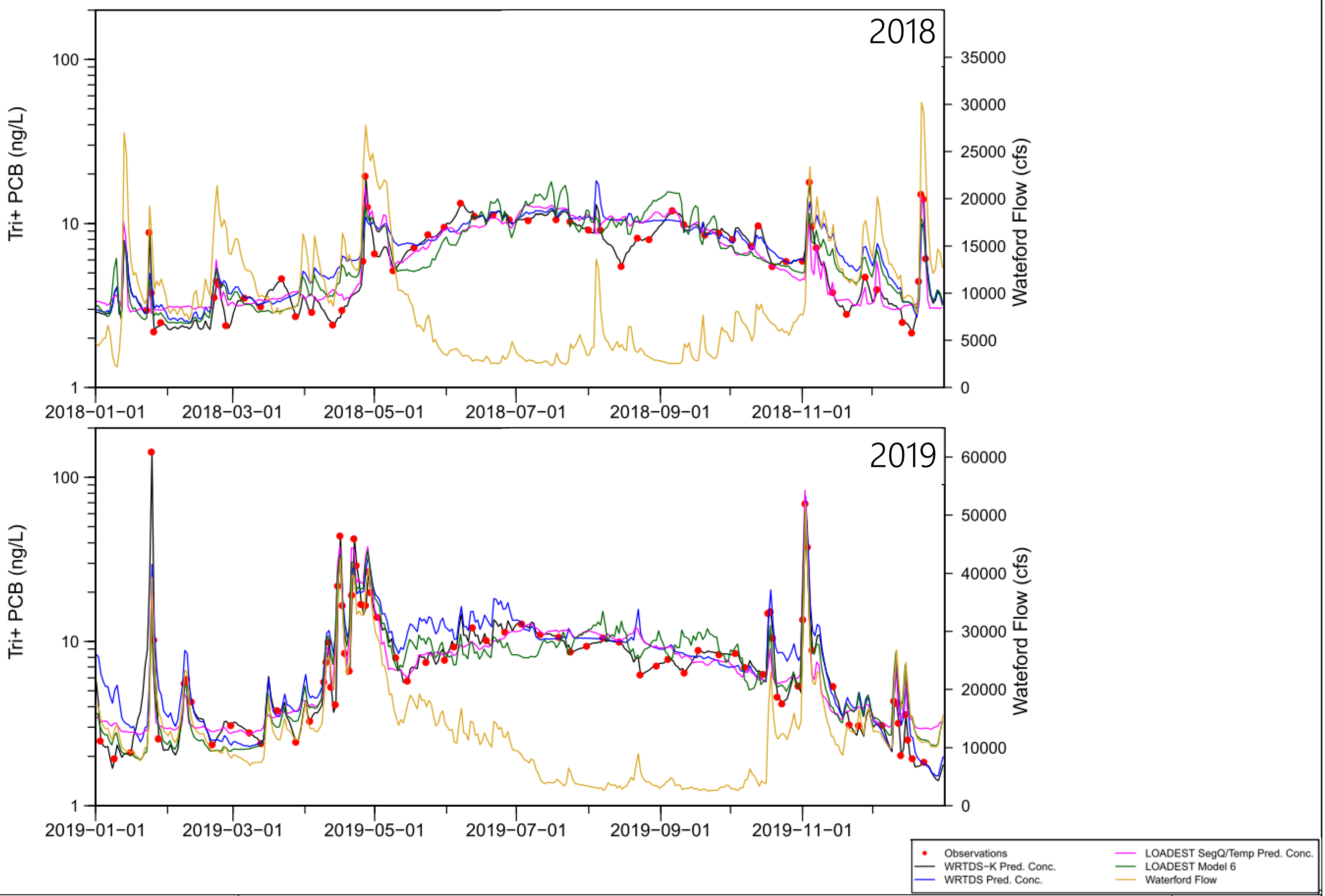
Hudson River PCBs Superfund Site / wsp	Seasonality in Tri+ PCB Concentration Measured at the Waterford Monitoring Station and Water Temperature at the Albany USGS Gauge Station	**Figure 2-3** July 2024



Comparison of Observed Tri+ PCB Concentration and Predicted Concentrations Using Different Load Estimation Models for 2016-2021

Figure 3-1

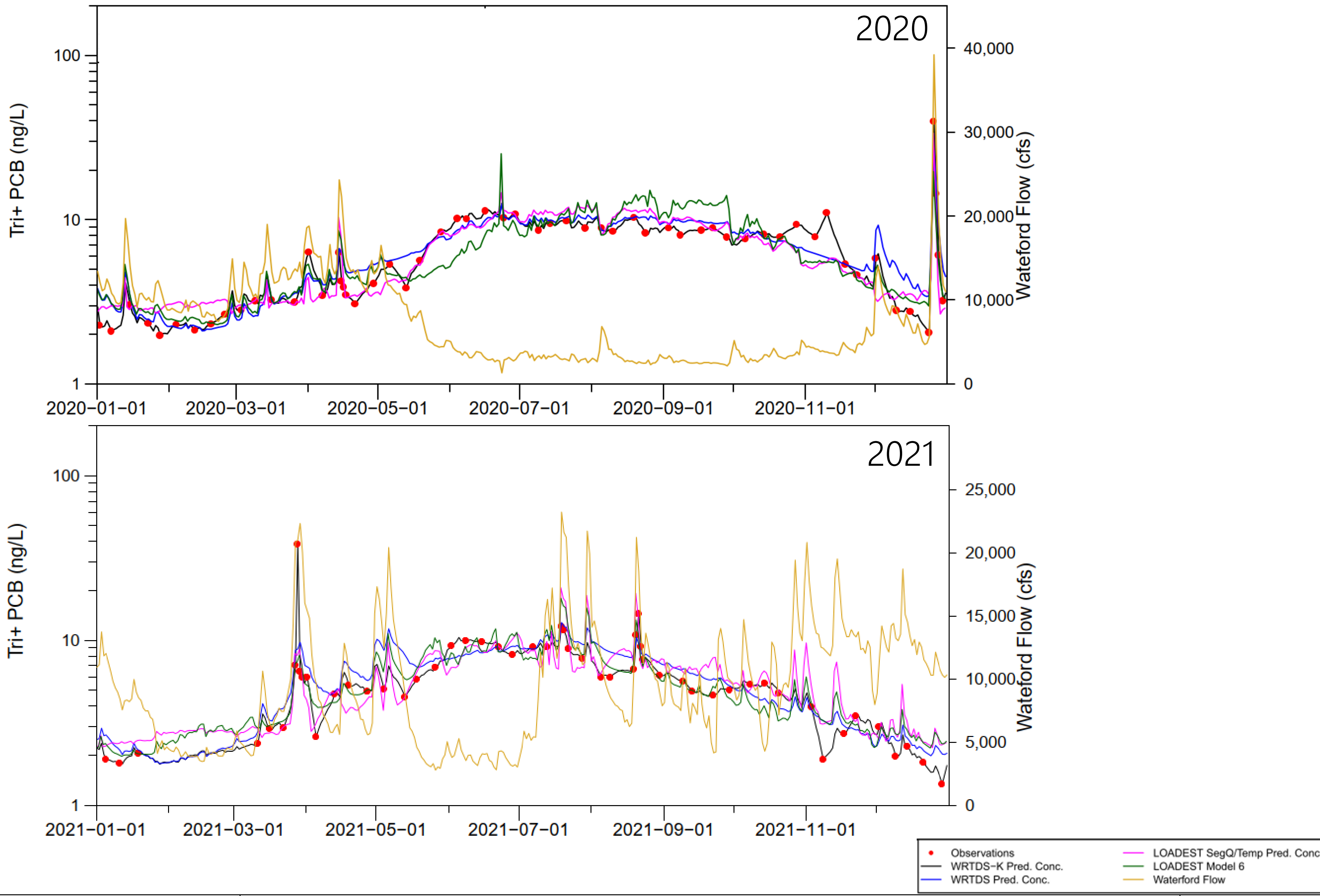
July 2024



Comparison of Observed Tri+ PCB Concentration and Predicted Concentrations Using Different Load Estimation Models for 2016-2021

Figure 3-1 (cont'd)

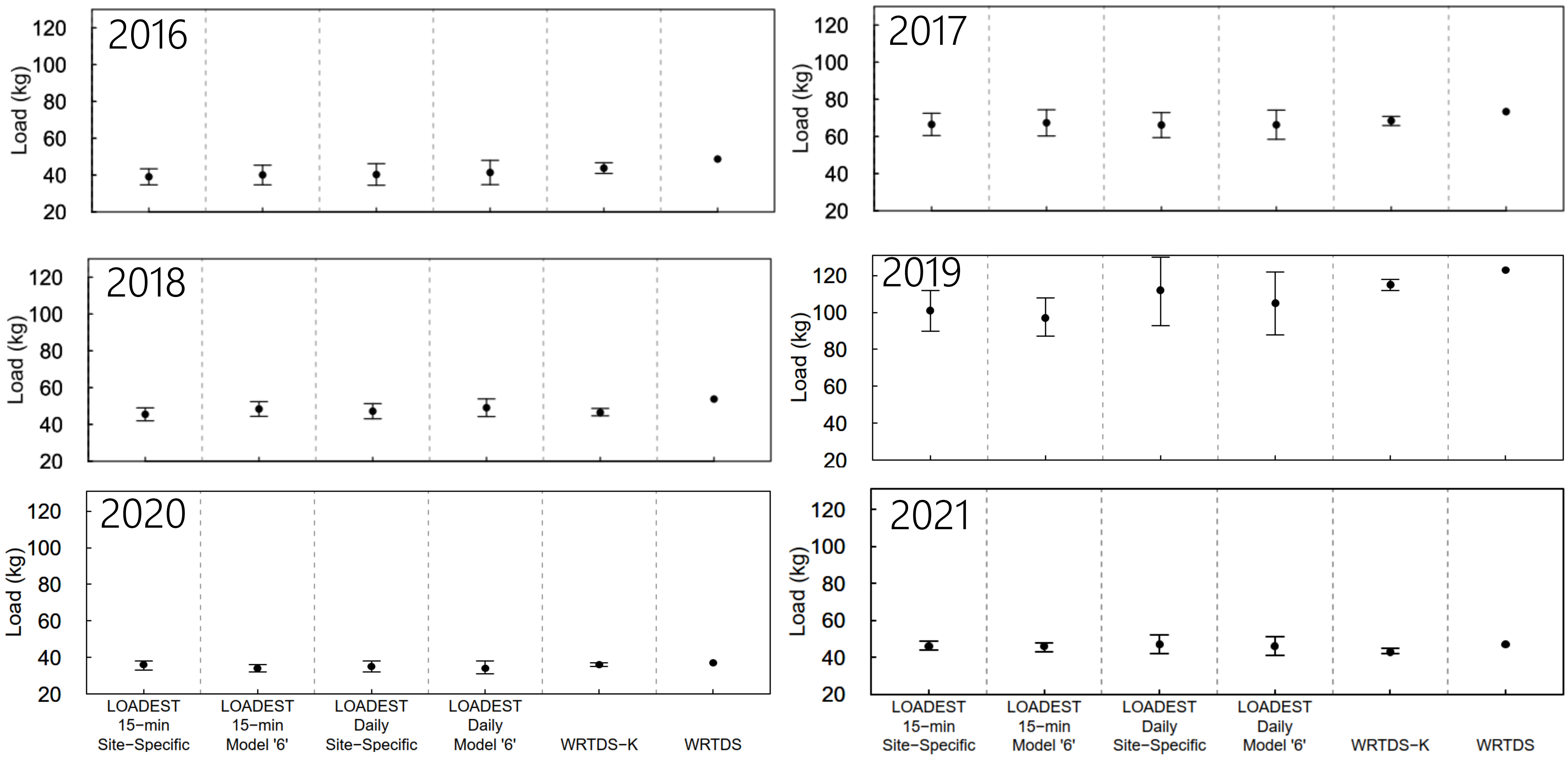
July 2024



Comparison of Observed Tri+ PCB Concentration and Predicted Concentrations Using Different Load Estimation Models for 2016-2021

Figure 3-1 (cont'd)

July 2024

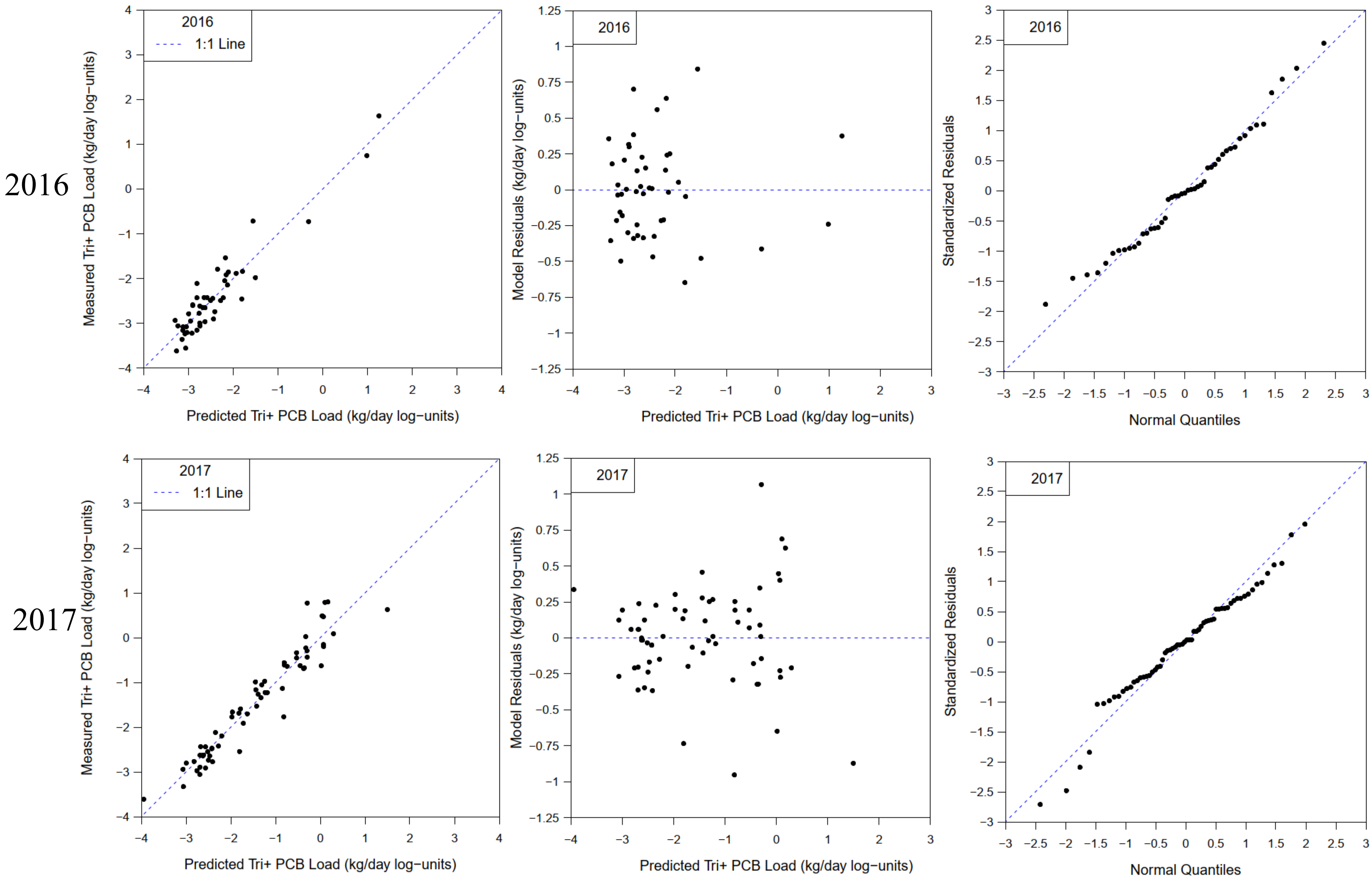


Notes: Error bars represent 95% confidence intervals on the estimated annual Tri+ PCB load.

Comparison of Estimated Annual Tri+ PCB Load Using Different Load Estimation Models for 2016-2021

Figure 3-2

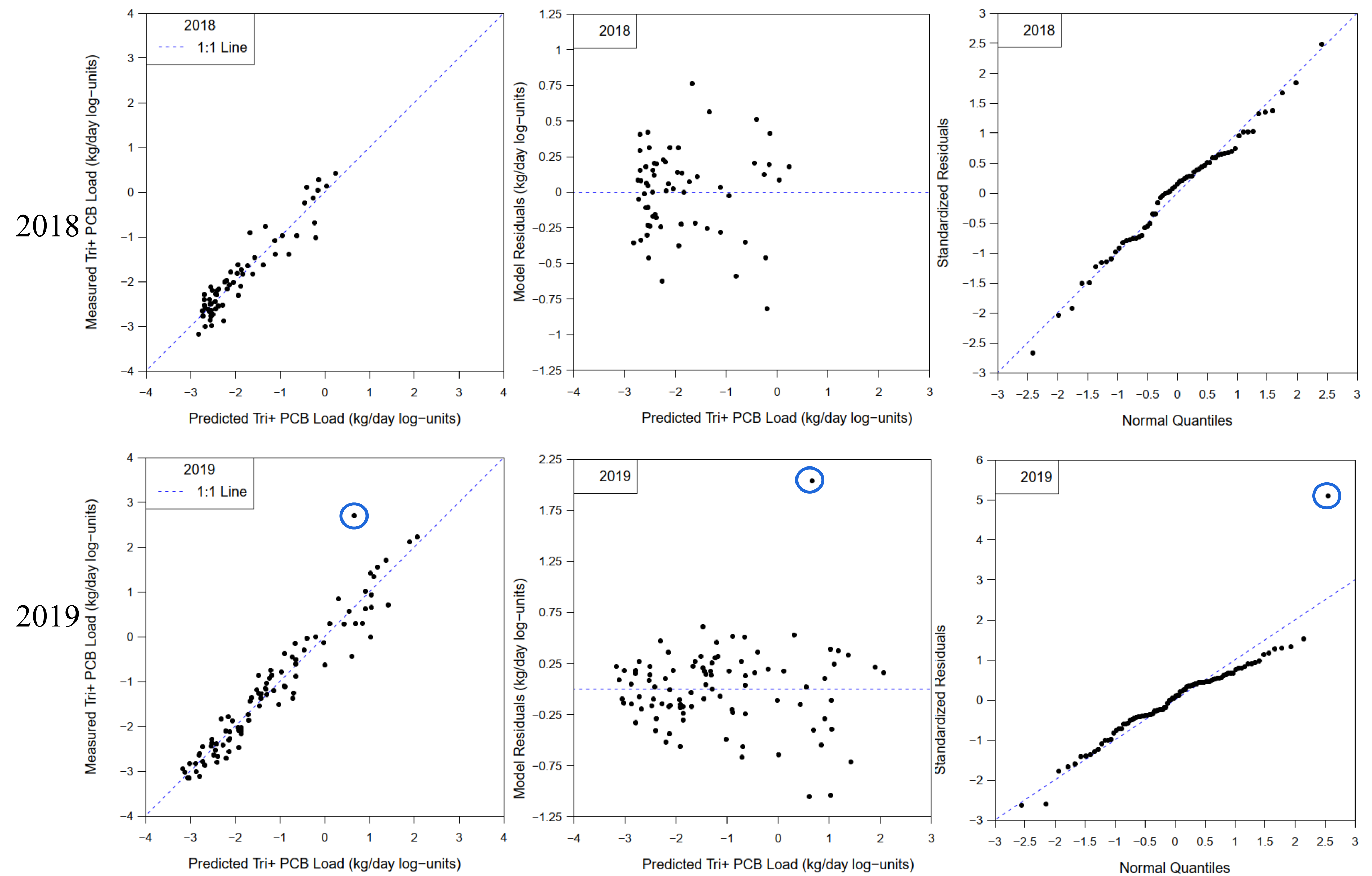
July 2024



Model Performance Diagnostic Plots for the LOADEST Site-Specific Model

Figure 3-3

July 2024



Notes: Circled point in 2019 plots is a sample with an anomalously high Tri+ PCB concentration (142 ng/L) for both the flow conditions at the time it was collected and for the post-dredging period overall. See Section 3.2 for discussion of this sample.

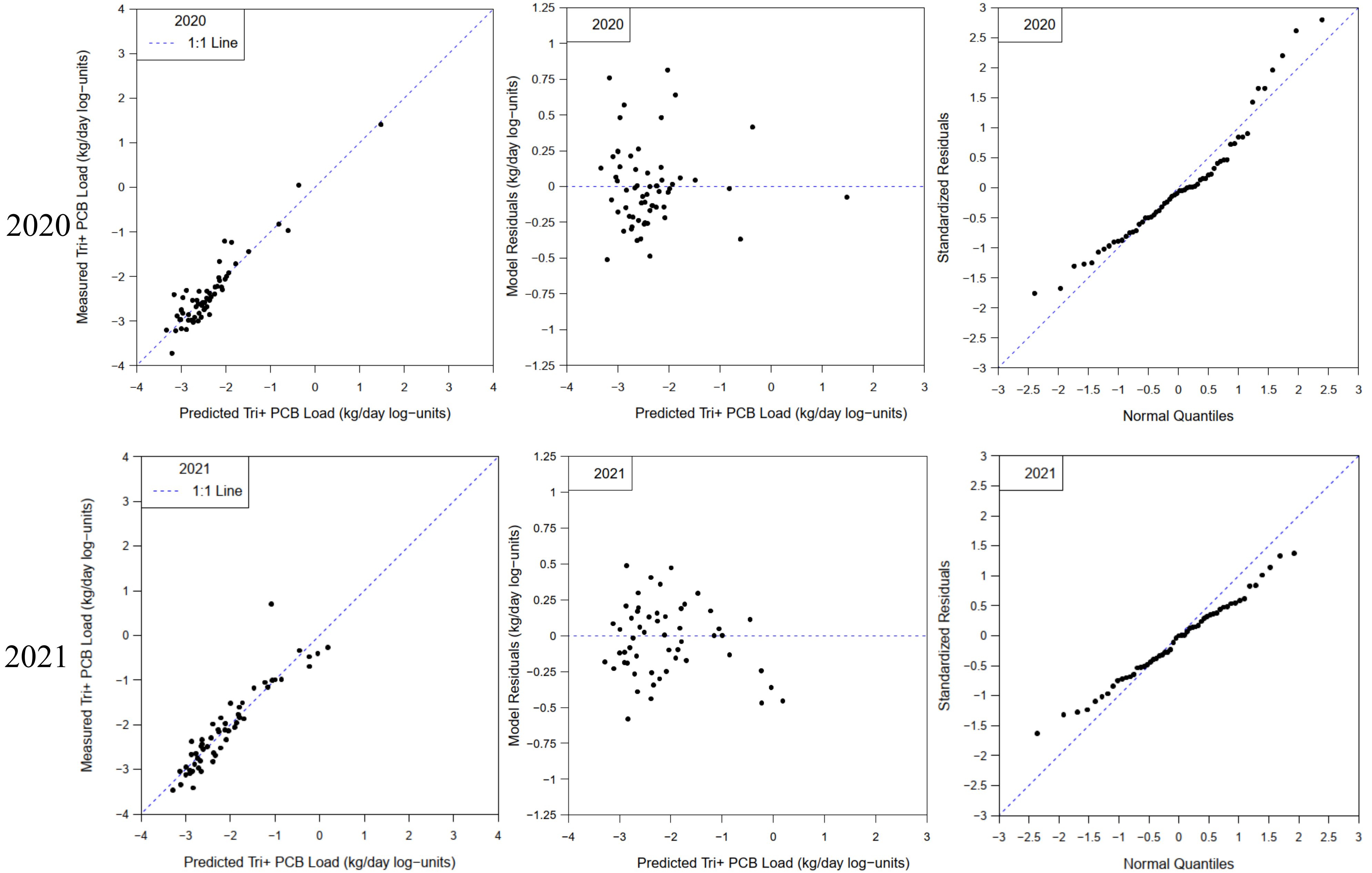
Hudson River PCBs Superfund Site

wsp

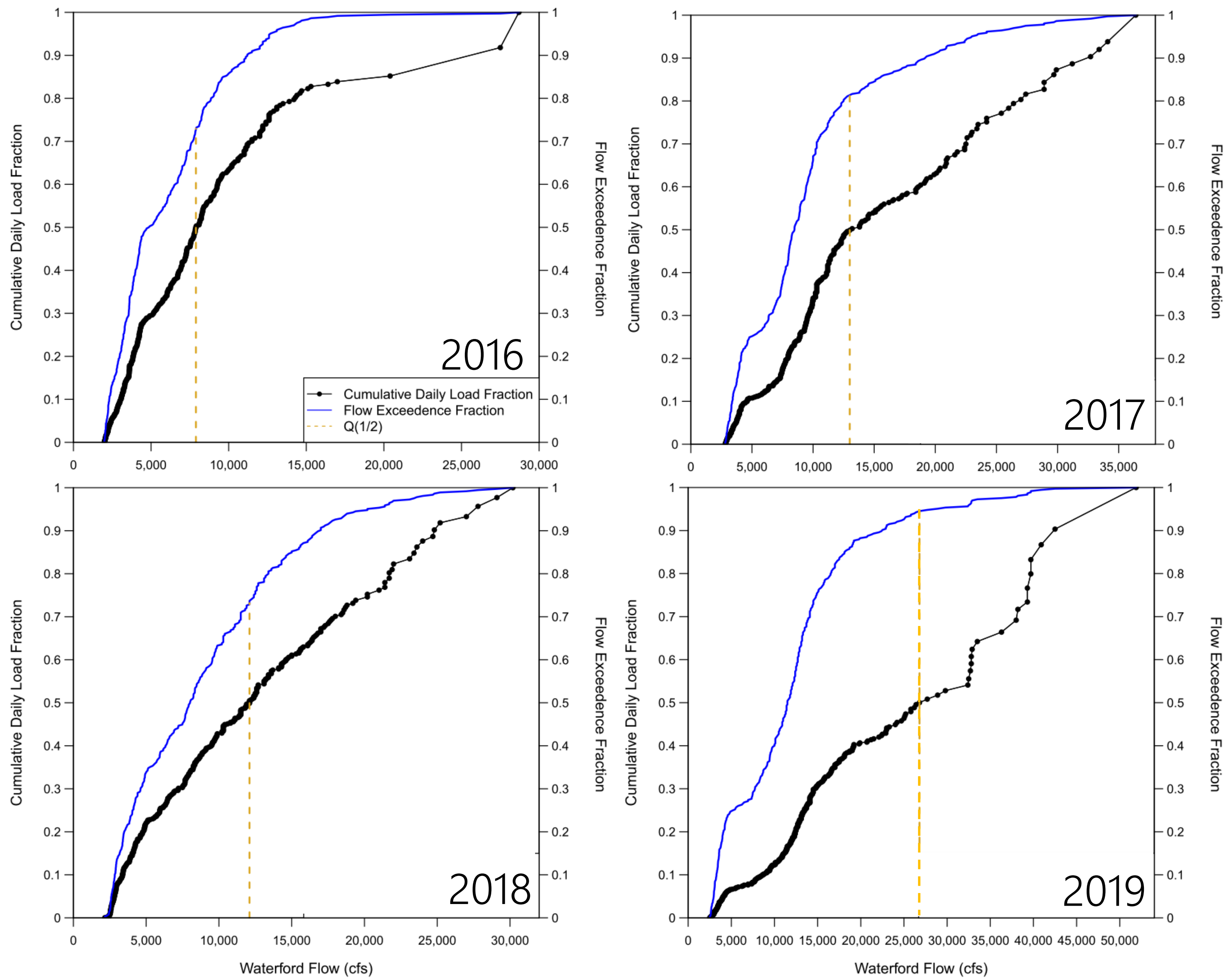
Model Performance Diagnostic Plots for the LOADEST Site-Specific Model

Figure 3-3 (cont'd)

July 2024



Hudson River PCBs Superfund Site	Model Performance Diagnostic Plots for the LOADEST Site-Specific Model	**Figure 3-3 (cont'd)** July 2024

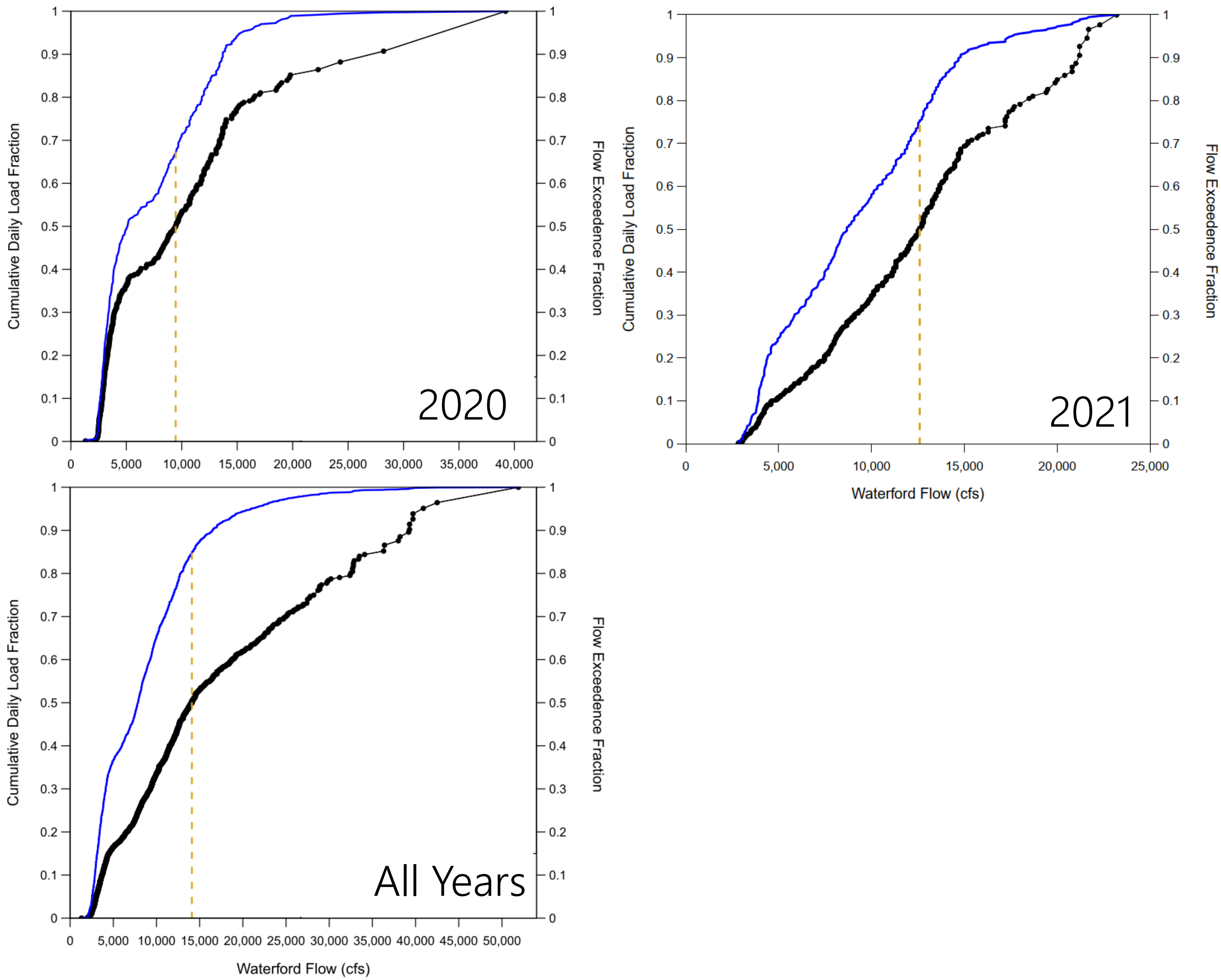


Notes: Cumulative daily load fraction is calculated by ordering daily mean flow from smallest to largest, and cumulatively summing the corresponding daily loads. The cumulative sums are then divided by the total annual load to produce a fraction. The vertical yellow dashed line indicates $Q_{1/2}$, which is the flow value where half the annual Tri+ PCB load (or half the total 2016-2021 Tri+ PCB load in the case of all years combined) occurs at a greater flow values. The intersection of the dashed line with the flow exceedance fraction curve indicates the flow exceedance fraction value for $Q_{1/2}$. See Section 4 for additional details.

Cumulative Daily Load Fraction and Flow Exceedance Fraction for Individual Years 2016-2021 and All Years Combined

Figure 4-1

July 2024



Notes: Cumulative daily load fraction is calculated by ordering daily mean flow from smallest to largest, and cumulatively summing the corresponding daily loads. The cumulative sums are then divided by the total annual load to produce a fraction. The vertical yellow dashed line indicates $Q_{1/2}$, which is the flow value where half the annual Tri+ PCB load (or half the total 2016-2021 Tri+ PCB load in the case of all years combined) occurs at a greater flow values. The intersection of the dashed line with the flow exceedance fraction curve indicates the flow exceedance fraction value for $Q_{1/2}$. See Section 4 for additional details.

Cumulative Daily Load Fraction and Flow Exceedance Fraction for Individual Years 2016-2021 and All Years Combined

Figure 4-1 (cont'd)

July 2024