Portland Harbor Superfund Site

Sampling Plan for Pre-Remedial Design, Baseline and Long-Term Monitoring

June 6, 2017

REVISED WORKING DRAFT

Table of Contents

Tables

Figures

Appendices

LIST OF ACRONYMS

ARAR	applicable or relevant and appropriate requirement
ATSDR	Agency for Toxic Substances and Disease Registry
BERA	baseline ecological risk assessment
BMP	best management practice
cfs	cubic feet per second
cm	centimeter
COC	contaminant of concern
cPAH	carcinogenic polycyclic aromatic hydrocarbon
CRD	Columbia River datum
DDD	dichlorodiphenyldichloroethane
DDE	dichlorodiphenyldichloroethene
DDT	dichlorodiphenyltrichloroethane
DDx	DDT+DDD+DDE
ENR	enhanced natural recovery
EPA	United States Environmental Protection Agency
FMD	future maintenance dredge
FS	feasibility study
ft^2	square feet
g	grams
HI	hazard index
HQ	hazard quotient
IC	institutional control
m^2	square meters
mm	millimeter
MNR	monitored natural recovery
NAPL	non-aqueous phase liquid
NRWQC	National Recommended Water Quality Criteria
ODHS	Oregon Department of Human Services
PAH	polycyclic aromatic hydrocarbon
PCB	polychlorinated biphenyl
PBDE	polybrominated biphenyl ether
PTW	principal threat waste
RAL	remedial action level
RAO	remedial action objective
RI	remedial investigation
RM	river mile
ROD	Record of Decision
SEA	Stiplin Environmental Associates
Site	Portland Harbor Superfund Site
SMA	sediment management area
SPMD	semipermeable membrane device
SVOC	semi-volatile organic compound
TBT	tributyltin

TOC	total organic carbon
TPH	total petroleum hydrocarbons
USFWS	U.S. Fish and Wildlife Service
USGS	U.S. Geological Survey
WQS	water quality standard

1 Introduction

On January 3, 2017, the United States Environmental Protection Agency (EPA) issued the Record of Decision (ROD), which presents the Selected Remedy for the in-river portion of the Portland Harbor Superfund Site (Site) in Multnomah County, Oregon. Monitoring requirements for the Selected Remedy include baseline monitoring and remedial design data collection before construction; short-term monitoring during construction; and long-term monitoring of caps, dredge areas, and monitored natural recovery (MNR) areas after construction to evaluate long-term effectiveness and ensure the remedies function as designed.

1.1 Purpose and Scope

This document provides the objectives and sampling approach for preliminary remedial design characterization and baseline and long-term monitoring at the Site. The preliminary design characterization includes sampling of surface and subsurface sediments for the purpose of establishing the preliminary remedial footprint for dredging and capping, including establishing the preliminary depth of contamination. The baseline and long-term monitoring approach includes monitoring of surface sediment, biota tissue, surface water, and sediment traps for the purpose of establishing baseline conditions and monitoring the effectiveness of the remedy toward achieving the remedial action objectives (RAOs) established for the Site. Due to the different objectives and sampling approach, the preliminary remedial design characterization and baseline and long-term monitoring programs are envisioned as separate sampling efforts although coordination between the two efforts may be required. This document also includes the sampling approach for anadromous and migratory fish tissue monitoring that will be used in conjunction with resident fish tissue monitoring to support the evaluation of the tribal fish consumption exposure scenario through the 5-year review process. Specific elements addressed in this document are outlined below.

[redacted]

The sampling guidelines presented in this document represent an initial sampling approach to monitor various media in Portland Harbor relative to RAOs, media specific cleanup and target levels and other metrics. By necessity, this sampling approach will be modified based on the results of the baseline and long-term monitoring programs to ensure that progress is monitored in an effective manner.

Preliminary Remedial Design Characterization

Preliminary remedial design characterization will focus on delineating the horizontal and vertical extent of contamination associated with sediment management areas identified in the ROD, evaluating the effectiveness of MNR in areas targeted for enhanced natural recovery (ENR) through placement of a 12-inch thick sand cover, and initiating the remedial design process. The objectives, sampling boundaries, and sampling approach for preliminary remedial design characterization are provided in Section 3.

Baseline Monitoring

Baseline monitoring will be conducted to identify existing conditions at the Site for bedded surface sediment, surface water, sediment traps, and biota tissue (i.e., smallmouth bass, carp, crayfish, clams, and osprey eggs) samples. These data will serve as the basis for comparison to data collected during the long-term monitoring program. The objectives, monitoring boundaries, and sampling approach for baseline monitoring are provided in Section 4.

Long-Term Monitoring

The long-term monitoring program will compare Site conditions against baseline conditions and cleanup levels to assess the performance of the remedial action. Long-term monitoring will include bedded surface sediment, surface water, sediment trap, and biota tissue samples from upstream, within, and/or downstream of the Site. The sampling design will be consistent with the baseline data collection effort to evaluate trends and remedy effectiveness over the long term. The objectives, monitoring boundaries, and sampling approach for long-term monitoring are provided in Section 5.

Anadromous and Migratory Fish Tissue Monitoring

Monitoring of anadromous and migratory fish tissue (i.e., lamprey, salmon, and sturgeon) will be conducted to evaluate the tribal fish consumption exposure scenario as part of the 5-year review process. Consistent with the remedial investigation (RI), exposures to tribal fishers will be evaluated assuming direct contact with contaminants in sediment and via fish consumption. The objectives, monitoring methods and temporal approach for anadromous and migratory fish tissue monitoring are provided in Section 6.

This document is to be used as a guide during the performing parties' development of characterization and monitoring plans. The performing parties are expected to prepare investigation-specific planning documents, including applicable work plans, field sampling and analysis plans, standard operating procedures, quality assurance project plans and health and safety plans. These plans will include the necessary details required to implement the sampling programs and will be prepared in accordance with applicable guidance to describe the sampling rationale, data quality objectives, sampling procedures, analytical methods, and data analysis approach. EPA and partners will review the performing parties' planning documents to ensure they meet the requirements for preliminary remedial design characterization, and baseline and long-term monitoring.

1.2 Activities Not Included in Sampling Plan

This document focuses on preliminary design characterization, and baseline and long-term monitoring and does not include detailed design level characterization and remedy performance monitoring. Detailed design level data represent the data necessary to design and implement the active remedial technologies within sediment management areas (SMAs). Additionally, the performance monitoring of the constructed components of the remedy are not included in this document. Specifically, this document does not include data collection requirements associated with upland source control monitoring; river bank, porewater, and groundwater characterization and monitoring; detailed remedial design; construction and post-construction performance monitoring; and other activities necessary to implement the Selected Remedy. Data collection requirements for these and other activities not discussed in this plan will be defined in the performing parties' planning documents reviewed by EPA

and partners, as applicable, or in later phases of remedy implementation. **Table 1** provides examples of data collection and monitoring activities that are not included in this document.

2 Remedial Goal Decision Criteria

This section describes the RAOs and identifies decision criteria that will be used for monitoring the progress of the remedial action and determining if the RAOs have been achieved.

2.1 Remedial Action Objectives

The nine RAOs included in the ROD consist of media-specific goals for protecting human health and the environment. Achieving the RAOs relies on effectively implementing the Selected Remedy to meet the cleanup levels discussed in Section 2.2. The nine RAOs are provided below.

Human Health RAOs

- RAO 1 – Sediment: Reduce cancer and non-cancer risks to people from incidental ingestion of and dermal contact with contaminants of concern (COCs) in sediment and beaches to exposure levels that are acceptable for fishing, occupational, recreational, and ceremonial uses
- RAO 2 – Biota: Reduce cancer and non-cancer risks to acceptable exposure levels (direct and indirect) for human consumption of COCs in fish and shellfish
- RAO 3 – Surface Water: Reduce cancer and non-cancer risks to people from direct contact (ingestion, inhalation, and dermal contact) with COCs in surface water to exposure levels that are acceptable for fishing, occupational, recreational, and potential drinking water supply
- RAO 4 – Groundwater: Reduce migration of COCs in groundwater to sediment and surface water such that levels are acceptable in sediment and surface water for human exposure

Ecological RAOs

- RAO 5 – Sediment: Reduce risk to benthic organisms from ingestion of and direct contact with COCs in sediment to acceptable exposure levels
- RAO 6 – Biota (Predators): Reduce risks to ecological receptors that consume COCs in prey to acceptable exposure levels
- RAO 7 – Surface Water: Reduce risks to ecological receptors from ingestion of and direct contact with COCs in surface water to acceptable exposure levels
- RAO 8 – Groundwater: Reduce migration of COCs in groundwater to sediment and surface water such that levels are acceptable in sediment and surface water for ecological exposure

Human Health and Ecological RAO

- RAO 9 – River Banks: Reduce migration of COCs in river banks to sediment and surface water such that levels are acceptable in sediment and surface water for human health and ecological exposures

Data collected during construction and long-term monitoring will be used to evaluate effectiveness of the Selected Remedy to meet the RAOs

2.2 Cleanup Levels

Cleanup levels are the long-term contaminant concentrations that need to be achieved by the Selected Remedy to meet RAOs. Site-specific cleanup levels were developed for each RAO for the following media: sediment (including beaches) and river bank soil, surface water, and groundwater. **ROD Table 17**, which summarizes the media-specific cleanup levels, is included in this document.

2.2.1 Human Health Risk-Based Cleanup Levels

Human health risk-based cleanup levels for sediment were calculated based on direct contact with beach and in-river sediment (RAO 1) and sediment concentrations to be protective of indirect exposure through consumption of fish and shellfish (RAO 2). The risk-based cleanup levels for RAOs 1 and 2 represent the lowest value in each medium (beach or in-river sediment) to be protective of all potential receptors and are applicable to sediment and river bank soils.

2.2.2 Human Health Fish Tissue Targets

Human health risk-based fish tissue targets are not cleanup levels but will be used to inform fish advisories, assess effectiveness of best management practices (BMPs), and monitor the progress of the Selected Remedy to reduce COCs in biota (RAO 2 and 6). Consistent with footnote (6) in **ROD Table 17** of the ROD, additional data will be collected during baseline and long-term monitoring to determine background fish tissue concentrations for polychlorinated biphenyls (PCBs) and other fish tissue COCs as appropriate.

2.2.3 Ecological Risk-Based Cleanup Levels

Ecological risk-based cleanup levels to meet RAOs 5 through 8 were developed for sediment, surface water, and groundwater/pore water. To be protective of all species, the lowest value for each COC was selected as the risk-based cleanup level for RAOs 5 and 6. COC-specific water concentrations from the baseline ecological risk assessment (BERA) that are protective of ecological receptors were selected as risk-based cleanup levels for RAOs 7 and 8, except for the manganese cleanup level, which was developed after the BERA.

2.2.4 ARAR-Based Cleanup Levels

As required by the Comprehensive Environmental Response, Compensation, and Liability Act, applicable or relevant and appropriate requirement (ARAR)-based cleanup levels were developed. The ARAR-based cleanup levels for RAOs 3 and 4 were selected from the lower of the federal National Recommended Water Quality Criteria (NRWQC) (organism + water) and Oregon water quality standards (WQSs) (organism + water), maximum contaminant levels, and non-zero maximum contaminant level goals. The cleanup levels for RAOs 7 and 8 are based on the lower of the NRWQC (chronic aquatic life) and Oregon WQS (chronic aquatic life) only when risk-based values are not available or are greater than ARARs. ARARs-based numbers are used for tributyltin for RAO 7 and for arsenic, chromium, and DDx (dichlorodiphenyltrichloroethane [DDT], dichlorodiphenyldichloroethane [DDD], and dichlorodiphenyldichloroethylene [DDE]) for RAO 8.

2.2.5 Background-Based Cleanup Levels

Sediment samples collected from areas not influenced by Site releases were used to develop background concentrations during the RI and feasibility study (FS). As shown on **ROD Table 17**, sediment and river bank soil cleanup levels for arsenic, PCBs, carcinogenic polycyclic aromatic hydrocarbons (cPAHs), and selected dioxins/furans were selected based on background concentrations. Insufficient

data are currently available to calculate defensible background concentrations for other media. Additional monitoring of surface water and biota tissue in the Upriver Reach is required to evaluate background concentrations for these media.

3 Preliminary Remedial Design SMA Delineation

This section describes the approach for performing preliminary remedial design characterization for delineating SMAs identified in the ROD for the Selected Remedy. As described in Section 1.1, preliminary remedial design characterization is focused on delineating the horizontal and vertical extent of contamination. More detailed design level data will be needed to complete design. SMAs identified in the ROD were delineated based on the remedial action levels (RALs) and/or principal threat waste (PTW) thresholds (whichever is the lower value) for the Selected Remedy and represent areas of sediment contamination where capping and/or dredging will be performed to reduce risk. Site-wide RALs, navigation channel RALs, and PTW thresholds are listed in the attached **ROD Table 21**. Navigation channel RALs only apply to the navigation channel, and Site-wide RALs apply to all other areas of the Site. Within the navigation channel, PTW thresholds are lower than the navigation channel RAL levels and thus could be applicable. The approximate lateral extent of the SMA for the Selected Remedy is shown in **ROD Figure 30**, included in this document. The SMA lateral extent was estimated based on sediment COC concentrations and is not contingent on property boundaries. The sampling approach for pre-remedial design characterization, baseline monitoring, long-term monitoring, and anadromous and migratory fish tissue monitoring is summarized on **Table 2**. A summary of the monitoring associated with each media, including the data use, reach, spatial scale, sample design, sample type, and sample parameters, is provided in **Table 3**.

3.1 Problem Definition

The SMA boundary for the Selected Remedy, as shown on **ROD Figure 30**, was estimated using RI data. Additional data are needed to further refine and delineate the current lateral and vertical extent of cap and/or dredge areas in accordance with the Selected Remedy. Obtaining new sediment data is needed to accurately delineate SMA boundaries and enable the entire Portland Harbor Site to move closer to completing the remedial design.

3.2 Pre-Remedial Design Sampling Objectives

Data shall be collected to delineate lateral and vertical extent of sediment exceeding the Site-wide RALs, navigation channel RALs, and/or PTW thresholds, as listed in the attached **ROD Table 21**. Contaminated sediment shall be delineated to the depth of the applicable RAL concentrations or PTW concentrations, whichever value is lower.

The objectives of pre-remedial design sampling include the following:

- Delineate the horizontal and vertical extent of contaminated sediment requiring remediation in accordance with the requirements specified in the ROD
- Collect preliminary sediment data that can be used to support remedial design activities
- Collect preliminary data for the purposes of applying the decision tree (**ROD Figure 28**) and supporting remedial design
- Evaluate the effectiveness of MNR in areas targeted for ENR in the ROD
- Characterize nature and extent of contaminated materials intended to be left in place

3.3 Pre-Remedial Design Sampling Boundaries

Current surface and subsurface sediment data are needed to delineate SMAs for each in-water region of the Site (Navigation Channel, Future Maintenance Dredge Areas, Intermediate Region, and Shallow Region). The characterization approach should be consistent with the Selected Remedy within each region, as described in Sections 3.3.1 through 3.3.4 below.

In each region, the horizontal and vertical extent of sediment contamination exceeding the RALs and/or PTW thresholds (whichever is the lower value) will be determined to identify the remedial footprint and the depth of contamination (DoC) required to be dredged. Characterization of sediment will require identification of PTW and be of sufficient depth to characterize material to be left in place to support cap design if required. Dredging and/or capping in each region shall be performed in accordance with the technology application decision tree provided in attached **ROD Figure 28**.

3.3.1 Navigation Channel

The Selected Remedy in the federally authorized navigation channel is restricted to dredging in order to avoid constructing a remedy (cap or residual layer) within the authorized dredge depth. Contaminated sediment will be dredged to the depth of the Navigation Channel RAL concentrations or PTW threshold concentrations specified in **ROD Table 21**, whichever value is lower.

3.3.2 Future Maintenance Dredge Areas

Future maintenance dredge (FMD) areas are those locations in the river that are periodically dredged to allow continued vessel activity such as shipping and docking. As presented in the technology application decision tree (attached **ROD Figure 28**), contaminated sediment shall be dredged to the depth of the Site-wide RAL concentrations, the depth of PTW, or the depth required to allow placement of a cap or backfill sufficient to be effective over the long term, whichever is deeper.

3.3.3 Intermediate Region

The intermediate region is defined as outside the horizontal limits of the navigation channel and FMD areas to the riverbed elevation of approximately -2 feet Columbia River Datum (CRD). In the intermediate region, contaminated sediment will be dredged to the depth required by the technology application decision tree (**ROD Figure 28**) to achieve Site-wide RALs and remove PTW or to allow placement of cap or backfill material sufficient to be effective over the long term.

3.3.4 Shallow Region

The shallow region is defined as shoreward of the riverbed elevation of approximately -2 feet CRD. Contaminated sediment in this area will be dredged to the depth required to remove all non-aqueous phase liquid (NAPL) or PTW that cannot be reliably contained unless it is present below the feasible depth limit of excavation technology in which case it will be dredged to a depth to accommodate a cap and then capped. Where NAPL/not reliably contained (NRC) PTW is not present but the depth of excavation to achieve RAL concentrations is greater than 5 feet, the area will be dredged to 5 feet with placement of a cap and backfilled to grade.

3.4 Preliminary Remedial Design Data Collection

Preliminary remedial design data collection will focus on sediment core sampling to delineate the lateral and vertical extent of contamination. **Table 2** summarizes the preliminary remedial design data collection activities and laboratory analyses for characterization of surface and subsurface sediment.

As shown on **Figure 1**, sediment cores will be advanced within the SMA footprints and spaced on a 150- by 150-foot grid. This grid spacing was selected to be consistent with the SMA characterization performed by the RM11E group. Additionally, a 150-foot buffer zone of sediment core "step out" locations will be advanced to delineate the extent of COCs above the applicable RAL and PTW thresholds. If COCs are above the applicable RAL and PTW thresholds at a step out location, then an additional core will be advanced no farther than 150 feet from the core location(s) with COCs above the RAL and PTW thresholds. This process will continue until the lateral extent of sediment above the applicable RAL and PTW thresholds is delineated and the vertical extent of COCs above the applicable RAL and PTW thresholds is delineated to the depth of feasible dredge limits or to characterize material to be capped consistent with the Selected Remedy.

Delineating the vertical extent of COCs may require characterization of subsurface sediment contamination to depths of up to 20 feet below mudline. The final core depth and sampling intervals will be determined based on the expected depth of contamination using existing data, the CSM (e.g., contaminant release mechanism and hydrodynamic setting), and the anticipated technology to be applied. Sediment will be collected with a Vibracore or similar method that allows penetration to the required depth, collects sediment samples over 30–centimeters (cm) (1-foot) sampling intervals, recovers 80 percent of the sampling interval, and establishes the vertical extent of contamination. The vertical extent of contamination will be considered delineated when two consecutive 30-cm sampling intervals are below the applicable RAL or PTW threshold. Specific details regarding the sampling program (e.g., sample location, depth, step out core installation procedures, archiving procedures) shall be presented in project plans and subject to EPA approval. Sediment samples will be analyzed for the parameters listed in **Tables 2 and 3**. It should be noted that the preliminary remedial design characterization described above may need to be supplemented with SMA-specific data collection and evaluations for the purpose of developing dredge prisms and/or sediment cap design.

4 Baseline Monitoring

Baseline monitoring will be used to establish baseline conditions prior to the commencement of remedial activities and evaluate the trajectory of natural recovery processes over time.

4.1 Problem Definition

The establishment of baseline conditions is required for evaluating progress toward achieving the RAOs for the Site during long-term monitoring.

Much of the RI data was collected approximately 10 years ago. Since that time, natural recovery processes have occurred and a series of upstream and source control cleanup actions have taken place. Updated contaminant concentrations are needed in surface sediment, biota tissue, surface water, and sediment being transported through the system to establish baseline conditions and for comparison to long-term monitoring results. In addition, updated biota tissue data are needed to support the development of updated fish consumption advisories.

Baseline data are needed within the Site, and for the Upriver and Downtown reaches to support the evaluation of remedy effectiveness. Background-based sediment cleanup levels were established for some COCs in the ROD. Surface sediment data collected in the Upriver and Downtown reaches will be used to evaluate recontamination potential and support the evaluation of remedy effectiveness. Background based surface water cleanup levels and fish tissue target levels were not established in the ROD due to the lack of data available to statistically estimate background concentrations for these media. Monitoring of surface water in the Upriver Reach is required to support the evaluation of whether surface water cleanup levels based on ARARs are achievable within a reasonable timeframe. Similarly, monitoring of biota tissue in the Upriver Reach is required to determine whether the fish tissue target levels are achievable within a reasonable timeframe.

4.2 Baseline Monitoring Objectives

The objectives of baseline monitoring include the following:

- Establish baseline conditions for evaluating remedy performance toward achieving RAOs
- Provide baseline data for evaluating the trend and effectiveness of natural recovery in MNR areas
- Provide baseline data for evaluating the effectiveness of source control measures
- Provide baseline data for evaluating downstream migration of Site COCs
- Evaluate contaminant transport into and from the Downtown Reach for evaluating remedy performance
- Support development of appropriate and effective institutional controls (ICs) for fish consumption
- Develop estimates of background biota tissue and surface water to support the evaluation of remedy performance

- Support performance of a statistically valid equivalency evaluation and comparisons at the specified scales

4.3 Baseline Monitoring Boundaries

The boundaries of the study will extend from RM 28.4 (just upstream of Willamette Falls) to RM 1.9 just upstream of the confluence of the Willamette and Columbia Rivers. Sampling efforts are proposed in four areas:

- Site (RM 1.9 to 11.8). This area encompasses the Superfund site and includes the areas specified for active remediation or natural recovery.
- Downtown Reach (RM 11.8 to RM 16.6). This reach, directly upstream from the site, includes a series of ongoing and planned sediment and upland source remediation activities. It has a higher level of contamination than the Upriver Reach. and immediate proximity to the site.
- Upriver Reach (RM 16.6 to 28.4). This area is upstream of the Downtown Reach and extends to just upstream of Willamette Falls. The contaminant concentrations in this area more closely represent those areas not impacted by point sources undergoing remediation.
- Downstream Reach (RM 0 – 1.9). Downstream monitoring will focus on monitoring of surface water and sediment traps at RM 1.9 and the upper portion of Multnomah Channel and the collection of osprey eggs from nests downstream of the Site.

4.4 Baseline Monitoring Data Collection

This section describes data collection activities and analytical requirements for characterization of surface sediment, surface water, sediment traps, fish and shellfish tissue, and other biota. A summary of the baseline monitoring approach, including analytes, is shown on **Table 2**. Sample and/or transect locations for bedded sediment, surface water, and sediment traps are shown on **Figures 2, 3, and 4**, respectively. Sample locations for biota tissue are shown on **Figure 5.**

4.4.1 Surface Sediment

Surface sediment monitoring will be performed using sampling procedures consistent with those utilized during the RI (Integral 2004). Surface sediments will be collected from the 0- to 30- cm depth interval. The 0- to 30-cm surface sediment interval, or upper 12 inches, was used in the RI to represent an appropriate interval for monitoring surface sediments based on the expected depth physical and biological processes in the sediment bed. In the absence of site-specific information regarding the depth of the biotic zone, the depth of the biotic zone is expected to range between 15 and 35 cm in lotic environments (USEPA 2016). In addition, the 0- 30-cm depth interval is consistent with the majority of surface sediment data collected at the Site during the RI and used to support the baseline human health and ecological risk assessments. Each sediment sample submitted for analysis shall be comprised of a three-point composite to increase the statistical power of the baseline and long-term monitoring program.

Surface sediments will be collected with a Ponar grab Eckman dredge, power grab or similar sampling device and analyzed for sediment COCs specified in **ROD Table 17**. The required number of surface sediment samples was developed based on an evaluation of expected declines in surface sediment concentrations associated with the selected remedy, the variability in the existing surface sediment data set, and the spatial scale of the monitoring program (see **Appendix A**). Surface sediment transects will

be used to facilitate sampling locations. The transects are 0.2 miles apart between RM 1.9 to 11.8 and shown on **Figure 2.**

Eight samples will be collected from each transect, with the exception of Swan Island Lagoon where four samples will be collected from each transect. Two surface sediment samples each will be collected from the International Slip and Fire Boat Cove. For each transect in the Willamette River, three samples will be collected from the left bank sediment shelf, three samples will be collected from the right bank sediment shelf, and two samples will be collected from within the navigation channel. Sample locations depicted on **Figure 2** are for illustrative purposes only; specific sampling locations will be identified by the performing parties in the applicable sampling and work plans.

4.4.2 Surface Water

Consistent with RI sampling methods (Integral 2006), baseline surface water sampling will use hydrophobic polyaromatic resin (XAD) high volume samplers and a peristaltic pump for grab samples. As shown on **Figure 3** and **Table 2**, surface water samples will be collected at five transects:

- RM 16.5 between Upriver and Downtown reaches
- RM 11.8 between the Downtown Reach and the Site
- RM 6 near the mid-point of the Site
- RM 1.9 between the Site and the Downstream Reach
- Upper portion of Multnomah Channel

Each transect will have three sample locations: east-channel, mid-channel, and west-channel. At each sample location, single-point near bottom samples and single-point near surface samples will be collected using XAD high volume samplers and a peristaltic pump for grab samples.

Surface water samples will be monitored quarterly over a 12-month period, for a total of four rounds of samples during this period. One of the four events shall be representative of early fall low flow conditions, and one of the four sampling events shall be representative of winter high flow conditions. The sampling is not designed to capture extreme events but rather seasonal low flow (less than 10,000 cubic feet per second [cfs]) and high flow (greater than 50,000 cfs) conditions that occur on an annual basis.

Surface water will be analyzed for the surface water COCs specified in **ROD Table 17**. Surface water grab samples collected with a peristaltic pump will be analyzed for conventional analytes, total and dissolved metals, and semi-volatile organic compounds (SVOCs). XAD high-volume samples will be analyzed by high-resolution gas chromatography/high-resolution mass spectrometry for PCB congeners, polychlorinated dibenzo-p-dioxin/furans, organochlorine pesticides, phthalate esters, and polycyclic aromatic hydrocarbons (PAHs). Samples will include total and dissolved constituents. Detection limits must be sufficient to demonstrate compliance with the surface water cleanup levels specified in **ROD Table 17**.

Additionally, semipermeable membrane device (SPMD) samplers will be deployed at the RM 11.8, RM 6, and RM 1.9 transects during one quarterly monitoring event to provide time-integrated, freely dissolved water COC concentrations. These data will provide a second line of evidence and depict a longer term

COC concentration (instead of a single point in time). At the RM 11.8-, 6-, and 1.9-transects, SPMD samplers will be deployed in triplicate with the sediment trap samplers (two locations per transect as described below) for a total of 6 SMPDs per transect or 18 SPMDs total. The samplers will be deployed for approximately one month during the low-flow quarterly monitoring period. The SPMDs will be analyzed for pesticides, PCBs, PAHs, and dioxin/furans. Following each sampling event, SPMD monitoring results will be evaluated to determine whether the SPMD monitoring program should be adjusted.

4.4.3 Sediment Traps

Sediment traps will be deployed on an annual basis and monitored quarterly, using sampling procedures consistent with the RI (Anchor QEA 2006), to characterize baseline conditions of sediment movement into, within, and out of the Site. As shown on **Figure 4** and **Table 2**, sediment traps will be deployed at five transects:

- RM 16.5 between Upriver and Downtown reaches
- RM 11.8 between the Downtown Reach and the Site
- RM 6 near the mid-point of the Site
- RM 1.9 between the Site and the Downstream Reach
- Upper portion of Multnomah Channel

These transect locations are the same as those identified for surface water sampling. Each transect will have two sediment trap sampling locations: east-channel and west-channel.

Sediment traps will be deployed for 3-month intervals to collect quarterly data over a 12-month period. Collected sediment will be analyzed for sediment COCs specified in **ROD Table 17**, total organic carbon (TOC), and grain size.

Sediment traps will be constructed using polyvinyl chloride pipe or similar material and have sufficient sample capacity for the required laboratory analyses. The sediment trap inlet will be perpendicular to the sediment floor and approximately 3 feet above the mudline.

4.4.4 Biota Tissue

Biota tissue will be monitored using sampling procedures consistent with those utilized during the various tissue monitoring events conducted at the Portland Harbor Site as part of the RI, including Round 1, Round 3B, and the 2011/2012 RI (Striplin Environmental Associates [SEA] 2002a; SEA 2002b; Oregon Department of Human Services [ODHS], EPA, and Agency for Toxic Substances and Disease Registry [ATSDR] 2003; Integral 2007a; Integral 2007b; Windward Environmental 2007; U.S. Geological Survey [USGS] 2009; and GSI Environmental 2011). To be consistent with previous tissue sampling efforts, sampling will take place between August 1 and October 31.

Biota tissue monitoring will be used to evaluate progress toward achieving the target tissue levels specified in **ROD Table 17**. This will include PCB congeners, dioxins and furans, metals, PAHs, organochlorine pesticides, phenols, PBDEs, and SVOCs, as shown in **Table 3.** Because fish readily metabolize PAHs, only shellfish (clams and crayfish) will be analyzed for PAHs. Species to be monitored include smallmouth bass, carp, crayfish, clams, and osprey. These species were selected to monitor

progress toward achieving RAOs 2 and 6 on both a Site-wide and SDU/RM scale. These species were selected based on the results of the human health and ecological risk assessments and due to their likelihood of being consumed, relative abundance, ease of capture, contaminant levels, and biological characteristics (including home range).

4.4.4.1 Smallmouth Bass

Smallmouth bass (*Micropterus dolomieu*) is a key species for human health risk assessment and can act as a surrogate for other species. Methods similar to smallmouth bass sampling conducted by in 2011 and 2012 For smallmouth bass tissue, samples will be collected using a standard rod and reel or equivalent method. Fish will have a target size of 225 to 355 millimeters (mm) and may be of either sex. To obtain data that can be used to evaluate progress toward achieving RAOs 2 and 6, fillet and offal (the remainder of the carcass after filleting) will be analyzed. From those data, a whole body (reconstituted) value will be derived.

Smallmouth bass samples will be collected from the Upriver and Downtown reaches and within the Site, (see **Table 2**). Within the Site, individual smallmouth bass will be collected on a river mile basis from each river bank. Within the Downtown Reach, 15 fish will be collected from each side of the river bank. Within the Upriver Reach, fish would optimally be collected from locations throughout the reach to meet the needed number of samples. Each fish will comprise one sample (no compositing). As shown in **Table 2**, the number of samples from each reach should include:

- Upriver Reach – 20 samples collected throughout reach
- Downtown Reach – 30 samples (15 from each side)
- Site – 100 samples (15 samples from each side every 3.3 miles and 10 from Swan Island Lagoon)

Sample locations at the Site may correspond to those collected during the RI.

4.4.4.2 Carp

Carp are bottom feeder fish with a high lipid content. Carp represent a key species for the human health risk assessment and were found to contain the highest levels of PCBs of any species tested at the Site (USEPA 2016). Carp represent a second trophic level of fish to evaluate remedy effectiveness, attainment of RAOs, and the carp contaminant data will be used to inform the meal recommendations in a fish advisory. Carp will be collected using boat electrofishing, backpack electrofishing or angling. To be consistent with the RI data set, the target total length range for carp is 508 to 677 mm. To obtain data that can be used to evaluate progress toward achieving RAOs 2 and 6, fillet and offal (the remainder of the carcass after filleting) will be analyzed.

Since it is believed that carp have a relatively large home range, carp will be collected every 3.3 miles within the Site from both sides of the river (including Swan Island Lagoon) and from throughout the Upriver and Downtown reaches. Consistent with the RI sampling methods, five 3-fish composite carp samples will be collected from each reach. As shown in **Table 2**, the number of samples from each reach should include:

- Upriver Reach – 5 samples (Five 3-fish composites throughout reach)
- Downtown Reach – 5 samples (Five 3-fish composite samples throughout reach)

- Site – 15 samples (Five 3-fish composite samples every 3.3 miles, including from Swan Island Lagoon)

4.4.4.3 Crayfish

Crayfish were evaluated in the baseline human health risk assessment and are likely consumed by humans within Portland Harbor. Although there is no evidence that crayfish are currently harvested commercially, information from Oregon Department of Fish and Wildlife indicated that an average of 4,300 pounds of crayfish were harvested commercially from the portion of the Willamette River within Multnomah County each of the five years from 1997 to 2001 (Kennedy/Jenks Consultants 2013). Crayfish are also consumed by numerous fish wildlife species within Portland Harbor (Windward, 2013). Crayfish will be monitored to assess progress toward achieving RAOs 2 (human health crayfish consumption) and RAO 6 (ecological prey consumption).

Crayfish will be collected on a composite basis using crayfish traps. A minimum of 300 grams (g) will be targeted for each sample (approximately 12 individual crayfish per composite) with a minimum size of 100 mm for each individual. Crayfish will be collected only within the Site. Samples will be collected from approximately the same locations as they were sampled during the RI. Target sample locations are shown on **Figure 5**.

As shown on **Table 2**, composite crayfish samples will be collected within the Site reach only and should include:

- Fourteen west side of Site
- Thirteen east side of Site
- Two from Swan Island Lagoon

Composite crayfish samples will be analyzed for all fish tissue COCs presented in **ROD Table 17**.

4.4.4.4 Clams

Clams are an important component of the benthic invertebrate community and are consumed by humans (Kennedy/Jenks Consultants, 2013) and many fish and wildlife species (Windward Environmental, 2013). The widespread distribution and abundance of the Asiatic clam (*Corbicula fluminea*) in the Lower Willamette River makes it a useful species for environmental monitoring and investigations of environmental quality. Clams will be monitored to assess progress toward achieving RAOs 2 (human health clam consumption) and RAO 6 (ecological prey consumption). Clam tissue may also be compared to tissue residue values used in the baseline ecological risk assessment to evaluate progress toward achieving RAO 5 (reduce benthic risk).

Clam tissue samples will be collected only within the Site reach. Samples may be collected at approximately the same locations as the RI Round 1 and Round 3 sample locations, as shown on **Figure 5**. At each sampling station, clams will be collected from an area of up to 3,500 square meters (m^2) (approximately 37,600 square feet [ft^2]). Clam stations will be determined in consultation with EPA during a reconnaissance survey of potential clam habitat. All clams collected at a station will be composited into one tissue sample. A minimum of 50 g of soft tissue will be targeted for each station. Samples will be depurated prior to analysis. As shown on **Table 2**, composite clam samples are to be collected within the Site reach only and should include:

- Nineteen west side of Site
- Seventeen east side of Site
- Four from Swan Island Lagoon

Composite clam samples will be analyzed for all fish tissue COCs presented in **ROD Table 17**.

4.4.5 Osprey Eggs

The osprey is of interest because it occupies a unique ecological niche and has been observed nesting and foraging throughout the Willamette River and its tributaries. Osprey feed almost exclusively on fish and have relatively small home ranges while nesting. Osprey are present from March until September, with several breeding pairs nesting in or near the Site. The collection of osprey eggs is needed for predator/prey evaluation to assess if the remedy can achieve RAO 6.

As was done during the RI, osprey eggs likely will be collected by USGS and the U.S. Fish and Wildlife Service (USFWS). The number of osprey eggs to be collected for baseline monitoring includes:

- Upriver Reach – Five osprey eggs
- Site – Five osprey eggs
- Downstream of Site – Five osprey eggs

Actual sample locations will be determined following reconnaissance surveys of osprey nesting sites. To replicate sampling conducted during the RI, eggs representative of the Upriver Reach may be collected along the Willamette River near Salem. Per the sampling approach summarized on **Table 2**, osprey egg sample collection will be consistent with the RI sampling scheme. Osprey egg samples will be analyzed for pesticides, PBDEs, PCBs and dioxin/furans as shown on **Table 3**.

5 Long-Term Monitoring

Long-term monitoring will be used to monitor the progress of the remedy toward achieving the RAOs established in the ROD through comparison to cleanup levels and target tissue levels provided in **ROD Table 17**. Long-term sediment and fish tissue monitoring data will be evaluated statistically to estimate the decay rate and determine whether COC reductions are occurring as expected. Due to the smaller sample numbers, trends in surface water, sediment trap, and osprey egg data will be evaluated on a qualitative or semi-quantitative basis to evaluate long-term remedy effectiveness.

5.1 Problem Definition

Long-term monitoring is required to evaluate progress toward achieving RAOs established for the Site, the effectiveness of MNR, and the effectiveness of source control measures and to support the 5-year review process.

5.2 Long-Term Monitoring Objectives

The objectives of long-term monitoring include the following:

- Evaluate remedy performance toward achieving RAOs on a Site-wide and reach-specific basis
- Evaluate the effectiveness of MNR
- Evaluate the effectiveness of source control measures
- Evaluate the effectiveness of remedial measures in the Downtown Reach
- Support the 5-year review process
- Support fish consumption advisory program

5.3 Long-Term Monitoring Boundaries

Sample locations and monitoring boundaries selected for baseline monitoring will be utilized for long-term monitoring (see Section 4.3).

5.4 Long-Term Monitoring Data Collection

Long-term monitoring will focus on the characterization of surface sediment, surface water, sediment trap samples, fish and shellfish tissue, and other biota (i.e., osprey eggs). To the extent possible, data should be collected in a manner that is consistent with the baseline sampling effort and prior long-term sampling events, once performed, to ensure valid comparisons over time. A summary of the long-term monitoring approach, including analytes, is shown on **Table 2**.

The first long-term monitoring event will commence 2 years after the start of the initial baseline monitoring event to provide a second data set for Site-wide natural recovery monitoring prior to significant construction. The second long-term monitoring event will commence 5 years after the initial baseline monitoring event to provide a third data set needed for the 5-year review evaluation. Subsequent long-term monitoring events will continue every 5 years thereafter during construction. Following construction, samples will again be taken during the final year of construction, 2 years after; and 5 years after completion, before resuming sampling at 5 year intervals. Each long-term monitoring event will duplicate the activities conducted under the baseline monitoring event.

5.4.1 Surface Sediment

Long-term monitoring of surface sediment will duplicate the surface sediment sampling and analysis efforts described for baseline monitoring such that samples will be collected along transects advanced every 0.2 mile from RM 1.9 to 11.8. Sample types, numbers, and analysis will be the same as described in Section 4.4.1. Sample location transects shown on **Figure 2** will be shifted 0.1 mile upstream for the first long-term monitoring event and then shifted back downstream 0.1 mile during the next long-term monitoring event. The shifting the sample location transects upstream or downstream 0.1 mile for each long-term monitoring event will more fully sample the Site and eliminate potential sampling bias from the initial sampling locations.

As described in Section 4.4.1, surface sediments will be collected from the 0- 30-cm depth interval and analyzed for sediment COCs specified in **ROD Table 17**. The required number of surface sediment samples was developed based on an evaluation of expected declines in surface sediment concentrations associated with the selected remedy, the variability in the existing surface sediment data set, and the spatial scale of the monitoring program (see Appendix A). Long-term monitoring data will be evaluated on a Site-wide, segment (3.3 miles), and river mile basis. A statistical evaluation will be performed to estimate the decline in COC sediment concentrations and to estimate the uncertainty in these estimates based the statistical power of the analysis relative to the sample numbers. The data will also be used to support the equivalency analysis described in Section 2.4.

5.4.2 Surface Water

Long-term monitoring of surface water will duplicate the surface water sampling and analysis efforts described for baseline monitoring. Sample types, numbers, locations and analysis will be the same as described in Section 4.4.2. As described in Section 4.4.2, surface water will be collected using XAD high volume samplers and grab samplers to monitor reductions of surface water concentrations and progress toward achieving RAOs 3 and 7 on a Site-wide and transect-specific basis. As further described in Section 4.4.2, SPMD samplers will be deployed at the RM 11.8, RM 6, and RM 1.9 transects during one quarter of each annual surface water monitoring event to provide time-integrated, freely dissolved water COC concentrations.

Similar to the baseline monitoring program, surface water samples will be monitored quarterly over a 12-month period, for a total of four rounds of samples during this period. One of the four events shall be representative of early fall low flow conditions, and one of the four sampling events shall be representative of winter high flow conditions. The sampling is not designed to capture extreme events but rather seasonal low flow (less than 10,000 cfs) and high flow (greater than 50,000 cfs) conditions that occur on an annual basis. In addition to regular surface water monitoring, an additional round (or rounds) of surface water sampling may be required in the event of an abnormally high flow event (e.g., 100-year flood event).

Surface water will be analyzed for the surface water COCs specified in **ROD Table 17**. Due to the limited number of surface water sample locations and the temporal component of the surface water monitoring program, surface water data collected during the long-term monitoring program will be used to evaluate progress toward achieving RAOs 3 and 4 on a qualitative or semi-quantitative basis.

5.4.3 Sediment Traps

Long-term monitoring of sediment traps will duplicate the sediment trap deployment and monitoring efforts described for baseline monitoring. Sample types, numbers, locations, and analysis will be the same as described in Section 4.4.3. As described in Section 4.4.3, sediment traps will be monitored quarterly over a 12-month period during each long-term monitoring event to assess changes in contaminated sediment load associated with implementation of the Portland Harbor Remedy on a Site-wide and transect-specific basis.

Due to the limited number of sediment trap locations and the temporal component of the sediment trap monitoring program, sediment trap data collected during the long-term monitoring program will be evaluated over time to assess spatial and temporal trends indicative of remedy effectiveness at decreasing contaminant load within and exiting the site.

5.4.4 Fish and Shellfish Tissue

Long-term fish tissue monitoring will duplicate the sampling and analysis efforts described for baseline monitoring. Sample types, numbers, locations, and analysis will be the same as described in Section 4.4.4. Fish and shellfish tissue will be collected to monitor reductions in tissue concentrations and progress toward achieving RAOs 2 and 6 on a Site-wide and reach-specific basis. Fish and shellfish tissue will be analyzed for the fish tissue COCs specified in **ROD Table 17**. This will include PCB congeners, dioxins and furans, metals, PAHs, organochlorine pesticides, phenols, PBDEs, and SVOCs, as shown in **Table 3.** Because fish readily metabolize PAHs, only shellfish (clams and crayfish) will be analyzed for PAHs. A statistical evaluation will be performed to estimate the decline in COC fish and shellfish tissue concentrations and to estimate the uncertainty in these estimates based on the statistical power of the analysis relative to the sample numbers.

5.4.5 Osprey Eggs

Long-term monitoring of osprey eggs will duplicate the sampling and analysis efforts described for baseline monitoring. Sample types, numbers, locations, and analysis will be the same as described in Section 4.4.5. Due to the limited number of osprey eggs collected during each long-term monitoring event, changes in osprey egg concentrations will be evaluated on a qualitative or semi-quantitative basis in the context of historical osprey egg data collected from the region.

5.5 Data Evaluation

Data collected during the long-term monitoring program will be evaluated on a Site-wide, segment (3.3 miles), and river mile basis as appropriate based on sampling media to assess progress towards attaining RAOs through a comparison to cleanup and target tissue levels specified in the Portland Harbor ROD.

Evaluation of long-term monitoring will be evaluated against baseline monitoring data using statistical procedures. These statistical procedures will take on the traditional null hypothesis framework that considers the statistical power of the analysis based on sample numbers. In addition, it is expected that long-term monitoring data for all media will be fit to first order decay functions to estimate post-remedial recovery rates.

Since areas upstream of the Site can have sediment concentrations that exceed risk-based cleanup levels and can be a source of COC contamination associated with incoming sediments, site data will also be compared with sediment, surface water, and biota data collected from areas upstream of the Site

using an equivalence analysis (McDonald and Erickson 1994). This type of evaluation assesses whether the Site data are statistically equivalent to upstream data. The purpose of the equivalency evaluation is to determine whether Site concentrations are approaching Upriver Reach or Downtown Reach levels and evaluate whether the remedy is functioning as intended.

After the Selected Remedy is implemented (and natural recovery occurs), Site and upstream concentrations are expected to converge and be similar (equivalent). This convergence will be evaluated by estimating the ratio of the Site geometric mean to the Upriver and Downtown reaches geometric mean. Equivalence will be established when the upper 90% confidence limit for this ratio is no more than 1.5. Data from the Upriver and Downtown reaches will be collected synoptically with Site data and evaluated against Site data on a Site-wide, segment (3.3 miles), and river mile basis. The equivalency evaluation will consider the statistical power of the analysis based on the sample numbers.

6 Anadromous and Migratory Fish Tissue Monitoring

Monitoring of anadromous and migratory fish tissue will be used in conjunction with resident fish tissue monitoring to support the evaluation of the tribal fish consumption exposure scenario. Monitoring of anadromous and migratory fish tissue will be conducted consistent with the previous multi-agency monitoring effort (ODHS, EPA, and ATSDR 2003).

6.1 Problem Definition

Lamprey, salmon, and sturgeon have traditionally represented a substantial portion of the fish diet of tribal members. Therefore, monitoring of these species is required to evaluate the tribal fish consumption exposure scenario as part of the 5-year review process. Consistent with the RI, exposures to tribal fishers will be evaluated, assuming direct contact with contaminants in sediment and via fish consumption.

Anadromous and migratory fish spend only a portion of their life cycles within the Site. Adult salmon use the Site primarily during migration, whereas sturgeon may spend more of their lives within the Site. Lamprey ammocoetes drift downstream after hatching during high flow events to areas of low velocity and fine substrates where they burrow, grow, and live as filter feeders for 3 to 7 years. Metamorphosis to the juvenile stage occurs over several months before the lamprey migrate to the ocean where they reside for 1 to 3 years (USFWS 2008). As a result, contaminant concentrations in these species may not solely reflect sediment concentrations within the Portland Harbor Site. However, tribal consumption rates of anadromous species account for approximately 50 percent of total intake, as determined during the RI. Therefore, fish consumption will be evaluated during the 5-year review process, assuming a multi-species diet consisting of anadromous and resident fish species on a Site-wide basis for consistency with the RI.

6.2 Monitoring Objectives

The objectives of anadromous and migratory fish tissue monitoring include the following:

- Monitor tissue levels in anadromous and migratory fish to develop a comprehensive characterization of harvested fish for the fish consumption advisory effort.
- Provide information to the community about the risks associated with consuming these species. Fish advisories will distinguish between anadromous species (e.g., lamprey, salmon, and sturgeon) and resident species (e.g., smallmouth bass and carp).
- Support the 5-year review process for assessment of remedy performance and protectiveness of the remedy.

6.3 Monitoring Methods

Previous sampling of anadromous and migratory fish was conducted under a cooperative effort between the Oregon Department of Human Services, Agency for Toxic Substances and Disease Registry, Oregon Department of Fish and Wildlife, the City of Portland, Tribes, and EPA. To the extent feasible, the previous sample locations and methods will be duplicated, as described below.

6.3.1 Chinook Salmon

Consistent with previous sampling, adult Chinook salmon (*Oncorhynchus tshawytscha*) will be collected from the Clackamas fish hatchery. Sample collection will target adult Chinook salmon greater than 24 inches in length to be consistent with fishing regulations. A total of 15 individual fish will be collected. Fillet with skin samples will be analyzed for pesticides, PBDEs, PCBs and dioxin/furans as shown on **Table 3**.

6.3.2 Pacific Lamprey

Consistent with previous sampling, adult Pacific lamprey (*Entosphenus tridentatus*) will be collected at Willamette Falls, in coordination with tribal fishing activities. A total of 10 individual fish will be collected. The target length for the lamprey individuals is 15 to 25 inches. Whole body samples will be analyzed for pesticides, PBDEs, PCBs and dioxin/furans as shown on **Table 3**.

6.3.3 White Sturgeon

Consistent with previous sampling, adult white sturgeon (*Acipenser transmontanus*) will be collected from the Site using a standard rod and reel or equivalent method. Five individual fish will be collected between RMs 3.5 and 9.2. The target length for fish collected should range from 42 to 60 inches to be consistent with the sturgeon collection effort conducted for the RI. Fillet samples without skin will be analyzed for pesticides, PBDEs, PCBs and dioxin/furans as shown on **Table 3**.

6.4 Temporal Approach

Monitoring of anadromous and migratory fish tissue will be conducted post-construction to evaluate if fish tissue concentrations have changed relative to data collected during the RI. Special consideration regarding the timing of the anadromous and migratory fish monitoring program will be based on the migratory habits of these species. The frequency for subsequent long-term monitoring of anadromous and migratory fish tissue will be determined following review of the post-construction and RI fish tissue data.

7 References

Anchor QEA. 2006. Portland Harbor RI/FS Round 3A Field Sampling Plan - Sediment Traps. AE06-0003. Prepared for the Lower Willamette Group, Portland, OR. Anchor Environmental, L.L.C., Seattle, WA. August, 2006.

GSI Environmental, Inc. 2011. Sampling and Analysis Plan, Portland Harbor 2011 Baseline Smallmouth Bass Tissue Study Prepared for U.S. Environmental Protection Agency, U.S. Army Corp of Engineers and City of Portland. September 2011.

Integral. 2004. Round 2 Field Sampling Plan, Sediment Sampling and Benthic Toxicity Testing. Prepared for the Lower Willamette Group, Portland, OR. Integral Consulting, Inc., Mercer Island, WA. June 21, 2004.

Integral. 2006. Portland Harbor RI/FS Round 3A Surface Water Field Sampling Plan. IC060012. Prepared for the Lower Willamette Group, Portland, OR. Integral Consulting Inc., Mercer Island, WA. September 18, 2006.

Integral. 2007a. Portland Harbor RI/FS Round 3B Field Sampling Plan for Fish and Invertebrate Tissue and Collocated Surface Sediment. IC07-0019. Prepared for the Lower Willamette Group, Portland, OR. Integral Consulting Inc., Mercer Island, WA. October 22, 2007.

Integral. 2007b. Portland Harbor RI/FS Standard Operating Procedure for Homogenization of Sturgeon Tissue for the Portland Harbor RI/FS. Draft. IC07-0015. Prepared for the Lower Willamette Group. Integral Consulting Inc., Mercer Island, WA. June 18, 2007.

Kennedy/Jenks Consultants. 2013. Baseline Human Health Risk Assessment. Presented in Appendix F of the Portland Harbor Remedial Investigation Report. Produced for the Lower Willamette Group and United States Environmental Protection Agency. March 28.

McDonald, L. L and W. P. Erickson. 1994. Testing for bioequivalence in field studies: Has a disturbed site been adequately reclaimed? In book: Statistics in Ecology and Environmental Monitoring, Publisher: University of Otago Press, Dunedin, New Zealand, Editors: D.J. Fletcher, B.F.J. Manly, pp.183-197. January.

ODHS, EPA, and ATSDR. 2003. Salmon, Sturgeon, and Lamprey Tissue Investigation, Portland Harbor Site. WLTASE03. Prepared for Oregon Department of Health Services, Portland, OR, U.S. Environmental Protection Agency Region 10, Seattle, WA, and U.S. Department of Health and Human Services, Public Health Service, Agency for Toxic Substances and Disease Registry, Atlanta, GA. 2003.

SEA. 2002a. Fish Tissue Homogenization and Shipping SOP Round 1 Portland Harbor RI/FS. Prepared for Lower Willamette Group, Portland, OR. Striplin Environmental Associates, Inc., Olympia, WA. August 28, 2002.

SEA. 2002b. Field Sampling Plan, Round 1A Lower Willamette River ISA RI/FS. Prepared for the Lower Willamette Group. Striplin Environmental Associates, Inc., Olympia, WA. 2002.

USEPA. 2016. Portland Harbor RI/FS, Final Remedial Investigation Report. February 8, 2016.

USFWS. 2008. Fact Sheet, Pacific Lamprey. January.

USGS. 2009. Field Sampling Report for the Collection of Eggs and Determination of Productivity of Osprey Nesting Within the Portland Harbor Superfund Site and Vicinity. WLRASE08. U.S. Geological Survey. October 30, 2009.

Windward Environmental. 2007. Portland Harbor RI/FS Field Sampling Plan: Round 3 Sampling for Pre-Breeding White Sturgeon (*Acipenser Transmontanus*) Tissue. WE-06-0004. Prepared for the Lower Willamette Group. Windward Environmental LLC, Seattle, WA. February 16, 2007.

Windward Environmental. 2013. Final Baseline Ecological Risk Assessment. Presented in Appendix G of the Portland Harbor Remedial Investigation Report. Produced for the Lower Willamette Group and United States Environmental Protection Agency. December 16.

Tables

ROD Table 17. Summary of Cleanup Levels or Targets by Media

	Surface Water (1)			Groundwater (2)			River Bank Soil/Sediment (3)			Fish Tissue (4)		
Contaminant	**Unit**	**Conc.**	**Basis**	**Unit**	**Conc.**	**Basis**	**Unit**	**Conc.**	**Basis**	**Unit**	**Conc.**	**Basis**
Aldrin	µg/L	0.00000077	A				µg/kg	2	R	µg/kg	0.06	R
Arsenic	µg/L	0.018	A	µg/L	0.018	A	mg/kg	3	B	mg/kg	0.001	R
Benzene				µg/L	0.44	A						
BEHP	µg/L	0.2	A				µg/kg	135	R	µg/kg	72	R
Cadmium				µg/L	0.091	A/R(5)	mg/kg	0.51	R			
Chlordanes	µg/L	0.000081	A				µg/kg	1.4	R	µg/kg	3	R
Chlorobenzene				µg/L	64	R						
Chromium	µg/L	100	A	µg/L	11	A						
Copper	µg/L	2.74	A	µg/L	2.74	A/R	mg/kg	359	R			
Cyanide				µg/L	4	A						
DDx	µg/L	0.01	R	µg/L	0.001	A	µg/kg	6.1	R	µg/kg	3	R
DDD	µg/L	0.000031	A	µg/L	0.000031	A	µg/kg	114	R			
DDE	µg/L	0.000018	A	µg/L	0.000018	A	µg/kg	226	R			
DDT	µg/L	0.000022	A	µg/L	0.000022	A	µg/kg	246	R			
1,1-Dichloroethene				µg/L	7	A						
cis-1,2-Dichloroethene				µg/L	9.9	A						
Dieldrin							µg/kg	0.07	R	µg/kg	0.06	R
2,4-Dichlorophenoxyacetic acid				µg/L	70	A						
Ethylbenzene	µg/L	7.3	R	µg/L	7.3	R						
Hexachlorobenzene	µg/L	0.000029	A				µg/kg			µg/kg	0.6	R
Lindane							µg/kg	5	R			
Lead				µg/L	0.54	A/R	mg/kg	196	R			
Manganese				µg/L	430	R						
MCPP	µg/L	16	R									
Mercury							mg/kg	0.085	R	mg/kg	0.031	A
Pentachlorophenol	µg/L	0.03	A	µg/L	0.03	A				µg/kg	2.5	R
Perchlorate				µg/L	15	A						
PBDEs										µg/kg	26	R
PCBs	µg/L	0.0000064	A	µg/L	0.014	A/R	µg/kg	9	B	µg/kg	0.25 (6)	R
PAHs							µg/kg	23000				
cPAHs (BaP eq)	µg/L	0.00012	A	µg/L	0.00012	A	µg/kg	12 (7)	B	µg/kg	7.1	R
Acenaphthene				µg/L	23	R						
Acenaphthylene												
Anthracene				µg/L	0.73	R						
Benzo(a)anthracene	µg/L	0.0012	A	µg/L	0.0012	A						
Benzo(a)pyrene	µg/L	0.00012	A	µg/L	0.00012	A						
Benzo(b)fluoranthene	µg/L	0.0012	A	µg/L	0.0012	A						
Benzo(g,h,i)perylene												
Benzo(k)fluoranthene	µg/L	0.0013	A	µg/L	0.0013	A						
Chrysene	µg/L	0.0013	A	µg/L	0.0013	A						
Dibenz(a,h)anthracene	µg/L	0.00012	A	µg/L	0.00012	A						
Fluoranthene												
Fluorene												
Indeno(1,2,3-c,d)pyrene	µg/L	0.0012	A	µg/L	0.0012	A						
2-Methylnaphthalene												
Naphthalene	µg/L	12	R									
Phenanthrene												
Pyrene												
Dioxins/Furans (2,3,7,8-TCDD eq)	µg/L	0.0000000005	A									
1,2,3,4,7,8-HxCDF							µg/kg	0.0004	B	µg/kg	0.00008	R
1,2,3,7,8-PeCDD							µg/kg	0.0002	B	µg/kg	0.000008	R
2,3,4,7,8-PeCDF							µg/kg	0.0003	B	µg/kg	0.00003	R
2,3,7,8-TCDF							µg/kg	0.00040658	R	µg/kg	0.00008	R
2,3,7,8-TCDD							µg/kg	0.0002	B	µg/kg	0.000008	R
Tetrachloroethene				µg/L	0.24	A						
Toluene				µg/L	9.8	R						
TPH-Diesel							mg/kg	91	R			
TPH-Diesel (C10-C12 Aliphatic)				µg/L	2.6	R						
Tributyltin	µg/L	0.063	A				µg/kg	3080	R			
Trichloroethene				µg/L	0.6	A						
2,4,5-Trichlorophenol				µg/L	50	A						
Vanadium				µg/L	20	R						
Vinyl Chloride				µg/L	0.022	A						
Xylenes				µg/L	13	R						
Zinc	µg/L	36.5	R	µg/L	36.5	R	mg/kg	459	R			

Notes:

(1) Surface Water Cleanup Levels - RAOs 3 and 7

(2) Groundwater Cleanup Levels - RAOs 4 and 8

(3) Sediment Cleanup Levels - RAOs 1 and 5

(4) Fish Tissue Targets - RAOs 2 and 6

(5) A/R indicates that the ARARs-based number and the risk-based number are the same.

(6) The tissue target is a risk-based number and does not represent background levels. Additional data will be collected to determine background fish tissue concentrations for PCBs during design and construction of the Selected Remedy.

(7) The cleanup level for cPAHs of 12 µg/kg is based on direct contact with sediment and is applicable to nearshore sediment. The cleanup level applicable to sediments in the navigation channel is 3,950 µg/kg and is based on human consumption of clams.

Abbreviations:

A- ARAR-based number

ARAR - applicable or relevant and appropriate requirement

B - Background-based number

BEHP - bis(2-ethylhexyl)phthalate

BaP eq - benzo(a)pyrene equivalent

C - carbon

ROD Table 17. Summary of Cleanup Levels or Targets by Media

Abbreviations (continued):
Conc - concentration
cPAH - carcinogenic polycyclic aromatic hydrocarbon
DDD - dichlorodiphenyldichloroethane
DDE - dichlorodiphenyldichloroethene
DDT - dichlorodiphenyltrichloroethane
DDx - DDD + DDE + DDT
HxCDF - 1,2,3,7,8,9-hexachlorodibenzofuran
MCPP - 2-(4-chloro-2-methylphenoxy)propanoic acid
mg/kg - milligram per kilogram
PAH - polycyclic aromatic hydrocarbon
PBDE - polybrominated diphenyl ether
PCB - polychlorinated biphenyl
PeCDD - pentachlorodibenzo-p-dioxin
PeCDF - pentachlorodibenzofuran
R - risk-based number
RAO - remedial action objective
TCDD - 2,3,7,8-tetrachlorodibenzo-p-dioxin
TCDF - tetrachlorodibenzofurans
TPH - total petroleum hydrocarbons
µg/kg - microgram per kilogram
µg/L - microgram per liter

ROD Table 21. Sediment RALs and PTW Thresholds for Selected Remedy

Contaminants	Site Wide RALs[1]	PTW Thresholds [2]	Navigation Channel RALs
Focused COCs			
PCBs	75	200	1,000
Total PAHs [4]	13,000	NA	170,000
2,3,7,8-TCDD	0.0006	0.01	0.002
1,2,3,7,8-PeCDD	0.0008	0.01	0.003
2,3,4,7,8-PeCDF	0.2	0.2	1
DDx	160	7,050	650
Additional Contaminants			
2,3,7,8-TCDF	NA	0.6	NA
1,2,3,4,6,7,8-HxCDF	NA	0.04	NA
cPAHs (BaP Eq)	NA	106,000	NA
Chlorobenzene	NA	>320	NA
Naphthalene	NA	>140,000	NA

Notes:
1 – Site wide includes all areas of the Site except the navigation channel. FMD areas are subject to these RALs.
2 – PTW thresholds are based on highly toxic PTW values (10^{-3} risk) except chlorobenzene and naphthalene, which are threshold values for not reliably contained PTW.

Abbreviations:
BaP Eq – benzo(a)pyrene equivalent
cPAH –carcinogenic polycyclic aromatic hydrocarbon
COC – Contaminant of concern
DDx – dichlorodiphenyldichloroethane + dichlorodiphenyldichloroethene + dichlorodiphenyltrichloroethane
FMD – future maintenance dredge
HxCDF - hexachlorodibenzofuran
NA – not applicable
PAH – polycyclic aromatic hydrocarbon
PCB – polychlorinated biphenyl
PeCDD – pentachlorodibenzo-p-dioxin
PeCDF – pentachlorodibenzofuran
PTW – principal threat waste
RAL – remedial action level
TCDD – tetrachlorodibenzo-p-dioxin
TCDF – tetrachlorodibenzofuran
µg/kg – microgram per kilogram
> – greater than

Table 1. Data Collection Activities Not Included in Document

Work Phase for Data Collection	Data Collection Examples	Regulatory and Programmatic Responsibility
Upland Source Control Program	Groundwater monitoring to design and evaluate the effectiveness of groundwater-based source control measures (e.g., COC concentration, groundwater discharge rate, contaminant flux)	Oregon DEQ Source Control
	Stormwater monitoring to design and evaluate the effectiveness of stormwater-based source control measures	
	River bank characterization (e.g., COC concentration, erodibility assessment)	
	Downtown Reach effectiveness characterization	
Design Phase Data Collection	Groundwater characterization to support cap design (e.g., COC concentrations in groundwater and/or pore water, groundwater discharge rate, hydrogeologic characteristics)	EPA Remedial Design
	Detailed characterization of sediment and river bank COC concentrations	
	Waste characterization for disposal purposes	
	Site bathymetry	
	Geotechnical and physical characterization of sediment to support cap and dredge design (e.g., water content, bulk density, Atterberg limits, grain size distribution, shear strength)	
	Elutriate testing	
	Debris and infrastructure surveys	
	Shoreline and bulkhead stability	
	Characterization of hydrodynamic conditions (e.g., currents, waves, tides), including seasonal variations for flood rise and flood storage	
Construction Phase Performance/BMP Short-Term Monitoring	Water quality monitoring to measure the effectiveness of remediation BMPs	EPA Construction Monitoring
	Post remediation bathymetry surveys	
	Post-remediation sediment confirmation sampling	
Post-Construction Performance Monitoring	Monitoring to assess whether onsite habitat mitigation measures are functioning as intended	EPA Performance Monitoring
	Monitoring to assess whether compensatory mitigation measures are functioning as intended	
	Physical inspection of sediment caps to monitor cap integrity (e.g., diver and bathymetric surveys)	
	Porewater monitoring to assess whether the cap is effectively isolating contamination	
	Benthic recolonization monitoring	

Table 2. Sampling Approach Summary

Phase	Media	Sampling Method	Parameters	Depth	Sample Detail	Locations (by Reach)	Total Station Locations	Frequency	Rationale Summary
Pre-Remedial Design SMA Delineation	Surface and Subsurface Sediment	Vibracore or equivalent	Focused COCs and additional contaminants listed in ROD Table 21, NAPL identification where appropriate, TOC, grain size, and additional geotechnical parameters, as appropriate.	Parties will install sediment cores to establish the depth of contamination exceeding the applicable RALs and PTW thresholds. Sediment cores will include characterization of material beneath the contaminated sediment to confirm that the RALs and PTW thresholds were achieved. Sediment samples shall be a composite of each 30-cm (1-foot) interval to be submitted for analyses. Depth of contamination shall be bound by two consecutive composite samples of 30-cm layers with focused COCs less than RALs and PTW thresholds. Maximum core depth based on clean sediment materials, bedrock, or the limits of technology. Where the selected remedy is capping, sediment cores will be installed to a depth sufficient to characterize the material to be capped.	Laterally, sediment cores should be advanced every 150 feet on a grid pattern. Step out sediment cores shall be advanced as necessary beyond the minimum core array to delineate remedial footprints. Surface sediment samples will be collected from 0 to 30 cm below mudline. Subsurface samples within the cores will be collected based on a maximum 30-cm sampling interval for defining exceedances of RAL and PTW thresholds and characterizing the material to be left behind.	See Figure 1 More density at specific locations may be needed for remedial design and will be specified by performing parties	Estimated 1,080 to 1,470 core locations Actual number of sediment samples to be determined	One event prior to remedial design. Additional characterization may be required during remedial design.	• RI/FS sediment cores are insufficient for design of dredging/capping due to the limited number of core samples available and the age of the data. • Accurate delineation of the extent of sediment contamination exceeding RAL and PTW thresholds on an SMA basis is necessary to develop preliminary remedial footprints. • Grid spacing based on existing SMA characterization performed by RM11E group. • Analytical parameters are based on focused COCs for which RALs have been established.

Table 2. Sampling Approach Summary

Phase	Media	Sampling Method	Parameters	Depth	Sample Detail	Locations (by Reach)	Total Station Locations	Frequency	Rationale Summary
Baseline and Long-Term Monitoring	Surface Sediment	Ponar, Eckman dredge, power grab, or equivalent	Full COC set based on ROD Table 17, plus TOC and grain size at Year 0[1]. With the exception of PCBs, pesticides, dioxins/furans and PAHs, the required analytical parameter list will be re-assessed based on previous monitoring results to determine the sampling requirements for subsequent long-term monitoring events.	Surface: 0–30 cm	Site reach transects advanced every 0.2 mile, starting at RM 1.9 of the Willamette River and ending at RM 11.8. Eight sediment samples per transect (three nearshore east, two navigation channel, and three nearshore west), except for Swan Island Lagoon; four samples per transect will be collected within Swan Island Lagoon. Downtown and Upriver reaches samples (60 each) are non-transect and located to ensure a representative spatial distribution and the presence of bedded sediment. Each sediment sample submitted for analysis shall be comprised of a three-point composite.	Site: 50 transects x 8 samples per transect = 400 samples Swan Island Lagoon: 5 transects x 4 samples per transect = 20 samples International Slip and Fire Boat Cove: 2 samples per area = 4 samples Site total = 424 Downtown Reach = 60 samples Upriver Reach = 60 samples	Total = 544 station locations	Years 0[1], 2, 5, 10, and every 5 years thereafter until cleanup levels are achieved. Year 2 provides a second data set prior to significant construction for natural recovery monitoring Site-wide and a third point for 5-year review evaluation. See Section 5.4. Following construction, samples will again be taken during the final year of construction, 2 years after; and 5 years after completion, before resuming sampling at 5 year intervals.	• Downtown Reach and Upriver Reach sampling included to support evaluation of remedy effectiveness by determining the equivalency of two sample populations (upstream vs. Site). • Downtown Reach and Upriver Reach will facilitate identification of ongoing sources of contamination upstream of the Site. • Site sediment data will be used to evaluate the long-term effectiveness of the remedy.

Table 2. Sampling Approach Summary

Phase	Media	Sampling Method	Parameters	Depth	Sample Detail	Locations (by Reach)	Total Station Locations	Frequency	Rationale Summary
Baseline and Long-Term Monitoring	Surface Water	XAD high volume sampling, and grab samples using a peristaltic pump, water pump, or equivalent SPMD samplers will be deployed at RM 11.8-, 6-, and 1.9-transects	Full COC set based on ROD Table 17, plus TOC and DOC at Year 0. With the exception of PCBs, pesticides, dioxins/furans and PAHs, the required analytical parameter list will be re-assessed based on previous monitoring results to determine the sampling requirements for subsequent long-term monitoring events.	Near bottom and near surface water column samples will be collected at each sample location.	Sample transects located at RM 1.9, RM 6, RM 11.8, RM 16.5, and Multnomah Channel. Three samples will be collected at each transect location. Transects will be monitored quarterly, including during fall low flow conditions and winter high flow conditions. At the RM 11.8-, 6-, and 1.9-transects, two SPMD samplers will be co-located with sediment trap samplers for comparison purposes. SPMD samplers will be deployed for approximately one month during the low-flow quarterly monitoring period. The SPMDs will be analyzed for pesticides, PCBs, PAHs, and dioxin/furans.	Boundary of Downstream and Site (RM 1.9) Multnomah Channel Mid-Site (RM 6) Boundary of Site and Downtown Reach (RM 11.8) Boundary of Downtown Reach and Upriver Reach (RM 16.5)	Three surface water sample locations per transect. Each sample location will include a near surface and a near bottom sample. Samples will include peristaltic pump and XAD high volume samplers (30 point samples total). Semipermeable membrane device (SPMD) samplers will be deployed at the RM 11.8, RM 6, and RM 1.9 transects to provide a second line of evidence that depicts a longer term COC concentration during one quarterly monitoring event. Two SPMD samples will be deployed in triplicate at the three transects (6 SMPDs per transect or 18 SPMDs total). Total = 48 samples	Year 0 - Quarterly sample events within the 1-year period. Repeated for long-term monitoring on Years 2, 5, and 10. See Section 5.4. Following construction, samples will again be taken during the final year of construction, 2 years after; and 5 years after completion, before resuming sampling at 5 year intervals.	• Surface water sampling is consistent with RI sampling scheme • Water column data will be used in conjunction with USGS flow estimates to develop COC loading estimate to evaluate changes in loading to the Site from upstream, within the Site, and from the Site to the Columbia River. • Used to compare to cleanup levels in the ROD. • Upriver and Downtown Reach transects will be used to determine if surface water quality criteria are achievable. • Evaluate changes in surface water concentrations on a Site-wide and transect-specific basis. • Downtown Reach and Upriver Reach transects will be used to support evaluation of remedy effectiveness by determining the equivalency of two sample populations (Upriver Reach and Downtown Reach vs. Site). • SPMD samples provide time-integrated, dissolved water COC concentration data for comparison to the point sampling data.

Table 2. Sampling Approach Summary

Phase	Media	Sampling Method	Parameters	Depth	Sample Detail	Locations (by Reach)	Total Station Locations	Frequency	Rationale Summary
Baseline and Long-Term Monitoring	Sediment Traps	Sediment trap capable of collecting enough sediment for sample analysis	Full COC set based on ROD Table 17, plus TOC and grain size at Year 0. With the exception of PCBs, pesticides, dioxins/furans and PAHs, the required analytical parameter list will be re-assessed based on previous monitoring results to determine the sampling requirements for subsequent long-term monitoring events.	The top of the sediment trap should be located no higher than 3 feet above the mudline.	Sediment traps shall be co-located with surface water transects. Sediment traps should be deployed for 3-month intervals to collect quarterly data over a 12-month period.	Collocated with surface water locations	5 transects with 2 sediment traps per transect Total = 10 samples	Year 0 - Deployed for 1 year. Monitored quarterly within the 1- year deployment period. Repeated for long-term monitoring on Year 2, 5, and 10. See Section 5.4. Following construction, samples will again be taken during the final year of construction, 2 years after; and 5 years after completion, before resuming sampling at 5 year intervals.	• Sediment trap data will support ROD metric for evaluation of contaminant flux and to understand the quality of the sediment moving through the Site. • Secondary line of evidence for MNR and recontamination. • Supports a semi-quantitative evaluation for increase/decrease trend. • Used to evaluate changes in sediment load associated with construction on a Site-wide basis.

Table 2. Sampling Approach Summary

Phase	Media	Sampling Method	Parameters	Depth	Sample Detail	Locations (by Reach)	Total Station Locations	Frequency	Rationale Summary
Baseline and Long-Term Monitoring	Biota	Rod/reel, trot or hand line, or similar method for fish; bass tissue processed via fillet and whole body minus filet (i.e., offal) Carp will be collected with electrofishing and/or angling	Fish tissue COCs listed on Table 17 on Year 0 and subsequent sampling events.	Not applicable	**Smallmouth Bass** Individual fish will be collected. Target approximately five fish per river mile per side based on being able to statistically determine a 10% decay rate. 15 fish every 3.3 miles per side Target size: 225–355 mm, either sex	Upriver Reach – 20 samples collected throughout reach Downtown Reach – 30 samples (15 from each side) Site – 100 samples (15 samples from each side every 3.3 miles and 10 from Swan Island Lagoon)	Total = 150 fish	Years 0, 2, 5, 10... Year 2 provides a second data set prior to significant construction for natural recovery monitoring Site-wide and a third point for 5-year review evaluation. See Section 5.4. Following construction, samples will again be taken during the final year of construction, 2 years after; and 5 years after completion, before resuming sampling at 5 year intervals. Aug 1-Oct 31 fish collection window.	• Stratification of sample locations based on exposure areas evaluated in the human health and ecological risk assessments. • Smallmouth bass is a key species for human health assessment and a surrogate for other resident species (e.g., crappie). • Carp is a key species for human health assessment and a surrogate for other resident species (e.g., bullhead). • Upriver and Downtown reaches are expected to be less variable, allowing for a reduction in the number of samples. • Fillet tissue data needed for human health risk evaluation. • Reconstructed whole body tissue data need for ecological risk evaluation, including comparison to tissue toxicity reference values and evaluation of wildlife.
					Carp Composite fish will be collected. Samples will be filleted, and separate composites will be prepared of the fillets and the remaining fish body. Target size: 508-677 mm, either sex	Upriver Reach – Five 3-fish composites throughout reach Downtown Reach – Five 3-fish composite samples throughout reach Site – 15 samples (five 3-fish composite samples every 3.3 miles, including from Swan Island Lagoon)	Total = 25 composite samples (three fish per composite)		

Table 2. Sampling Approach Summary

Phase	Media	Sampling Method	Parameters	Depth	Sample Detail	Locations (by Reach)	Total Station Locations	Frequency	Rationale Summary
Baseline and Long-Term Monitoring	Biota	Benthic sledge	Fish tissue COCs listed on Table 17 on Year 0 and subsequent sampling events, except for osprey eggs which will be analyzed for pesticides, PBDEs, PCBs and dioxin/furans.	Not applicable	**Clams** Reoccupy RI sampling locations. Clam station locations determined to be identified during Site reconnaissance in consultation with EPA. All clams collected at a station will be composited to generate sufficient sample mass for chemical analysis.	Site – Reoccupy RI Stations (19 west side, 17 east side, and 4 from Swan Island Lagoon	Total = 40 composite samples	Years 0, 2, 5, 10... Year 2 provides a second data set prior to significant construction for natural recovery monitoring Site-wide and a third point for 5-year review evaluation. See Section 5.4. Following construction, samples will again be taken during the final year of construction, 2 years after; and 5 years after completion, before resuming sampling at 5 year intervals.	• Evaluated on a river mile basis in the risk assessment. • Benthic invertebrate with small home range. • Clam tissue provides the basis for human health risk-based cPAH cleanup levels.
		Crayfish traps			**Crayfish** Reoccupy RI sampling locations. Crayfish station locations determined to be identified during Site reconnaissance in consultation with EPA. All clams collected at a station will be composited to generate sufficient sample mass for chemical analysis. Target size = greater than 100 mm.	Site – Reoccupy RI Stations (14 west side, 13 east side, and 2 from Swan Island Lagoon)	Total = 29 composite samples		• Benthic invertebrate with limited home range. • Evaluated in human health risk assessment on a station-specific basis. • Crayfish are consumed by humans and a variety of wildlife species.
		Single egg collected from selected nests			**Osprey** Site reconnaissance required to identify nests for egg collection.	Upriver – 5 eggs Site – 5 eggs Downstream – 5 eggs	Total = 15 eggs		• Sample collection consistent with RI sampling scheme • Needed for RAO 6 predator/prey evaluation

Table 2. Sampling Approach Summary

Phase	Media	Sampling Method	Parameters	Depth	Sample Detail	Locations (by Reach)	Total Station Locations	Frequency	Rationale Summary
Baseline and Long-Term Monitoring	Anadromous and Migratory Fish	Rod/reel for sturgeon. Salmon will be collected from the Clackamas fish hatchery using appropriate methods. Lamprey will be collected from Willamette Falls with tribal coordination.	Pesticides, PBDEs, PCBs and dioxin/furans	Not applicable	**Chinook Salmon** – 15 fish collected from the Clackamas fish hatchery. Fillet with skin samples Target size = greater than 24 inches **Pacific Lamprey** – 10 fish collected from Willamette Falls for whole body sample Target size = 15 – 25 inches **White Sturgeon** – 5 fish collected between RM 3.5 and 9.2 for fillet without skin samples Target size = 42 – 60 inches	Chinook Salmon: Clackamas Fish Hatchery Pacific Lamprey: Willamette Falls White Sturgeon: Site (RM 3.5–RM 9.2)	Total = 30 fish	Conducted post-construction to evaluate if fish tissue concentrations have changed relative to data collected during the RI. The frequency for subsequent long-term monitoring of anadromous and migratory fish tissue will be determined following review of the post-construction and RI fish tissue data.	• Consistent with 2003 RI sampling by ODHS • Used in the human health risk assessment to evaluate effectiveness of the remedy based on the tribal fish consumption scenario.

Note(s)
1 – Year 0 is considered the baseline sampling event.

Table 3. Monitoring Plan Sample Summary

Media	Data Use	Reach	Spatial Scale	Sample Design	Sample Type	Parameters
Surface and Subsurface Sediment	Preliminary remedial design sampling to delineate the horizontal and vertical extent of contamination.	Site (RM 1.9–11.8).	SMA basis.	Sediment cores on a 150-foot interval. Step-out cores as necessary. Depth of contamination determined by two consecutive sample intervals below applicable RAL/PTW threshold.	Sediment sample composited over a 30-cm sampling interval.	Pesticides, PAHs, PCBs, and dioxins/furans.
Surface Sediment	Baseline and long-term monitoring. Evaluate progress toward achieving RAOs 1, 2, 5, and 6.	Site (RM 1.9–11.8).	Site-wide, segment (3.3 miles, one side) and river mile (1 mile, each side) basis.	Eight surface sediment (0- to 30-cm) transects every 0.2 mile. Three nearshore each side and two in navigation channel.	Three sample surface sediment composites; 0- to 30-cm depth interval.	Metals, pesticides, SVOCs, PAHs, PCBs, dioxins/furans, total petroleum hydrocarbons (TPH), and tributyltin (TBT).
Surface Sediment	Baseline and long-term monitoring. Supports equivalency analysis.	Upriver (RM 16.5–28.4) and Downtown Reach (RM 11.8–16.5).	Reach-specific basis.	30 surface sediment (0- to 30-cm) within Downtown Reach, and 20 surface sediment samples within the Upriver Reach evenly distributed.	Three sample surface sediment composites; 0- to 30-cm depth interval.	Metals, pesticides, SVOCs, PAHs, PCBs, dioxins/furans, TPH, and TBT.

Table 3. Monitoring Plan Sample Summary

Media	Data Use	Reach	Spatial Scale	Sample Design	Sample Type	Parameters
Surface Water	Baseline and long-term monitoring. Evaluate progress toward achieving RAOs 3 and 7. Estimate contaminant load entering and exiting the Site.	Upriver (RM 16.5), Downtown (RM 11.8), and Site (RM 6, RM 1.9, and Multnomah Channel).	Not applicable.	Five transects (RM 16.5, RM 11.8, RM 6, RM 1.9, and Multnomah Chanel). Near bottom and near surface samples collected from each side and the center of the channel. Four quarters of monitoring. SPMD samples will be deployed at the RM 11.8, RM 6, and RM 1.9 transects during one quarterly monitoring event per year .	XAD high volume samples and peristaltic pump grab samples. Total and dissolved fractions. Two SPMD samples will be deployed in triplicate at the RM 11.8, RM 6, and RM 1.9 (6 SMPDs per transect; 18 total).	Metals, pesticides, PAHs, PCBs, SVOCs, dioxins/furans, and TBT for XAD and grab samples. Pesticides, PCB, PAHs and dioxins/furans for SPMD samples.

Table 3. Monitoring Plan Sample Summary

Media	Data Use	Reach	Spatial Scale	Sample Design	Sample Type	Parameters
Sediment Traps	Baseline and long-term monitoring. Estimate contaminant load entering and exiting the Site.	Upriver (RM 16.5), Downtown (RM 11.8), and Site (RM 6, RM 1.9, and Multnomah Channel).	Not applicable.	Sediment traps to be collocated with surface water transects. One sediment trap to be installed on each side of the channel. Sediment traps to be deployed for 1 year and monitored quarterly.	Sediment will be collected from each trap on a quarterly basis if sufficient sediment has accumulated.	Metals, pesticides, SVOCs, PAHs, PCBs, dioxins/furans, TPH, and TBT.
Smallmouth Bass Tissue	Baseline and long-term monitoring. Evaluate progress toward achieving RAOs 2 and 6.	Site (RM 1.9–11.8).	Site-wide, segment (3.3 miles, one side) and river mile (1 mile, one side) basis.	Five individual fish collected on a river mile basis, each side. Fish will be collected using hand lines. Target size range is 225–355 mm.	Individual fish. Fillet and whole body minus fillet (offal).	Metals, pesticides, polybrominated biphenyl ethers (PBDEs), phenols, SVOCs, PCBs, and dioxins/furans.

Table 3. Monitoring Plan Sample Summary

Media	Data Use	Reach	Spatial Scale	Sample Design	Sample Type	Parameters
Smallmouth Bass Tissue	Baseline and long-term monitoring. Supports equivalency analysis.	Upriver (RM 16.5–28.4) and Downtown Reach (RM 11.8–16.5).	Reach-specific basis.	Ten individual fish samples from each reach. Within the Downtown Reach, 5 individual fish will be collected from each side of the river. Fish will be collected using hand lines. Target size range is 225–355 mm.	Individual fish. Fillet and whole body minus fillet (offal).	Metals, pesticides, PBDEs, phenols, SVOCs, PCBs, and dioxins/furans.
Carp Tissue	Baseline and long-term monitoring. Evaluate progress toward achieving RAOs 2 and 6.	Site (RM 1.9–11.8).	Segment basis (3.3 miles).	Five fish tissue composites collected within each segment. Fish will be collected using hand lines. Target size range is 508–677 mm.	Five fish composite. Fillet and whole body minus fillet (offal).	Metals, pesticides, PBDEs, phenols, SVOCs, PCBs, and dioxins/furans.
Carp Tissue	Baseline and long-term monitoring. Supports equivalency analysis.	Upriver (RM 16.5–28.4) and Downtown Reach (RM 11.8–16.5).	Reach-specific basis.	Five fish tissue composites collected within each reach. Fish will be collected using hand lines. Target size range is 508–677 mm.	Five fish composite. Fillet and whole body minus fillet (offal).	Metals, pesticides, PBDEs, phenols, SVOCs, PCBs, and dioxins/furans.

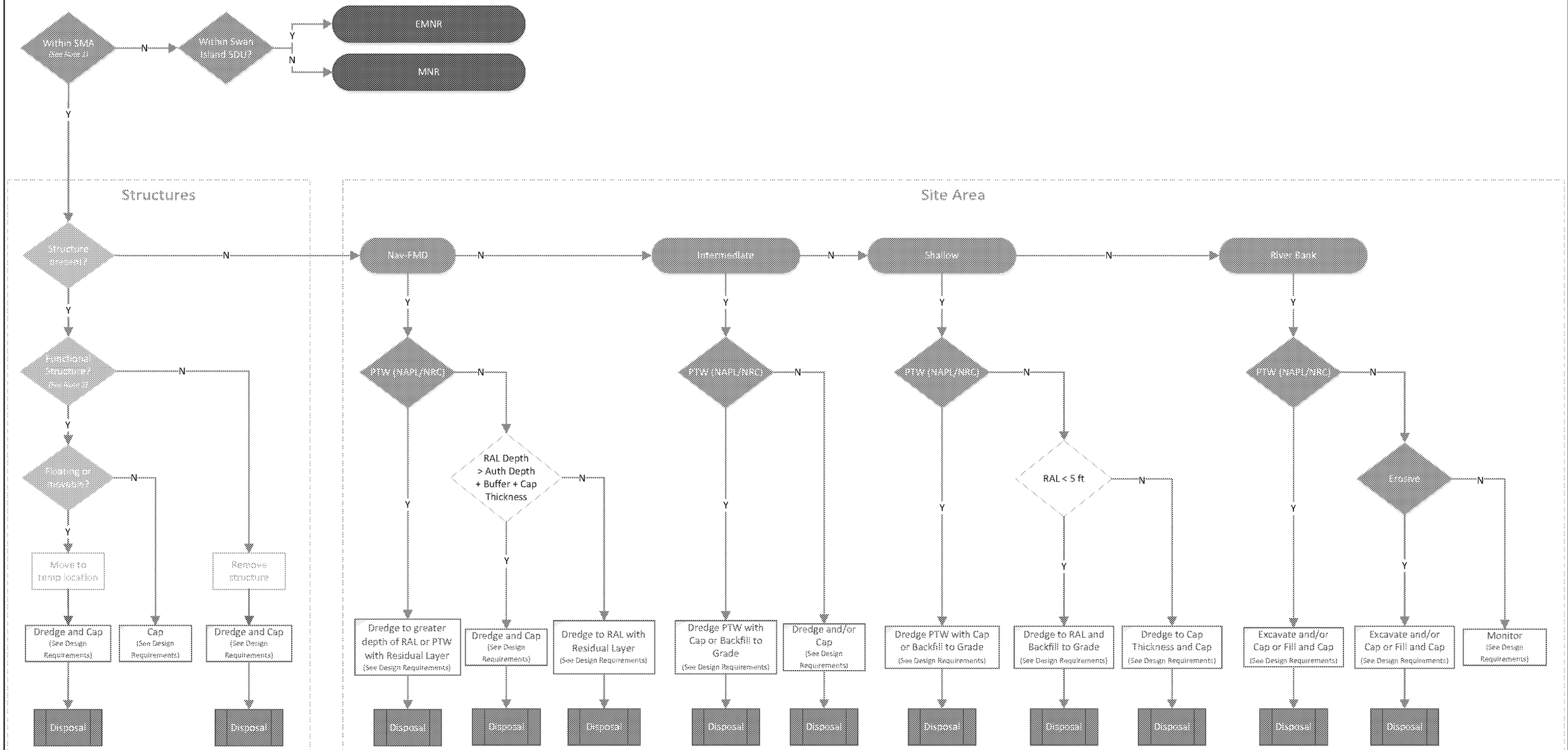
Table 3. Monitoring Plan Sample Summary

Media	Data Use	Reach	Spatial Scale	Sample Design	Sample Type	Parameters
Clam Tissue	Baseline and long-term monitoring. Evaluate progress toward achieving RAOs 2 and 6.	Site (RM 1.9–11.8).	Site-wide basis.	Reoccupy RI sample locations; 40 clam tissue composite sample locations will be targeted. Clam tissue samples will be collected using a benthic sledge. Target sample mass is 50 g from each location.	Depurated tissue composite samples without shell. The number of clams within each composite will be based on collecting sufficient tissue mass for chemical analysis.	Metals, pesticides, PBDEs, phenols, SVOCs, PAHs, PCBs, and dioxins/furans.
Crayfish Tissue	Baseline and long-term monitoring. Evaluate progress toward achieving RAOs 2 and 6.	Site (RM 1.9–11.8).	Site-wide basis.	Reoccupy RI sample locations; 29 crayfish tissue composite sample locations will be targeted. Crayfish will be collected using crayfish traps. Target size range is greater than 100 mm.	Whole body crayfish samples will be collected. The number of crayfish within each composite will be based on collecting sufficient tissue mass for chemical analysis.	Metals, pesticides, PBDEs, phenols, SVOCs, PAHs, PCBs, and dioxins/furans.
Osprey Tissue	Baseline and long-term monitoring. Evaluate progress toward achieving RAO 6.	Site (RM 1.9–11.8).	Site-wide basis.	One egg from 5 nests between RM 1.9 and 11.8. Attempt to reoccupy RI sample locations.	Whole egg.	Pesticides, PBDEs, PCBs, and dioxins/furans.

Table 3. Monitoring Plan Sample Summary

Media	Data Use	Reach	Spatial Scale	Sample Design	Sample Type	Parameters
Osprey Tissue	Baseline and long-term monitoring. Supports equivalency analysis.	Upstream and downtown stream areas.	Regional.	One egg from five nests upstream and five nests downstream of the Site.	Whole egg.	Pesticides, PBDEs, PCBs, and dioxins/furans.
Adult Pacific Lamprey Tissue	Long-term monitoring. Supports 5-year review process.	Upstream (Willamette Falls).	Regional	Collect 10 individuals from Willamette Falls; target size range is 15–25 inches.	Whole body tissue sample.	Pesticides, PBDEs, PCBs, and dioxins/furans.
Adult Chinook Salmon Tissue	Long-term monitoring. Supports 5-year review process.	Upstream (Clackamas River Fish Hatchery).	Regional.	Collect 15 individuals from Clackamas Fish Hatchery. Target size range is greater than 24 inches.	Fillet with skin samples will be analyzed.	Pesticides, PBDEs, PCBs, and dioxins/furans.
Pre-Breeding White Sturgeon Tissue	Long-term monitoring. Supports 5-year review process.	Site (RM 3.5–9.2).	Regional.	Collect five individuals. Target size range is 42–60 inches.	Fillet with skin samples will be analyzed.	Pesticides, PBDEs, PCBs, and dioxins/furans.

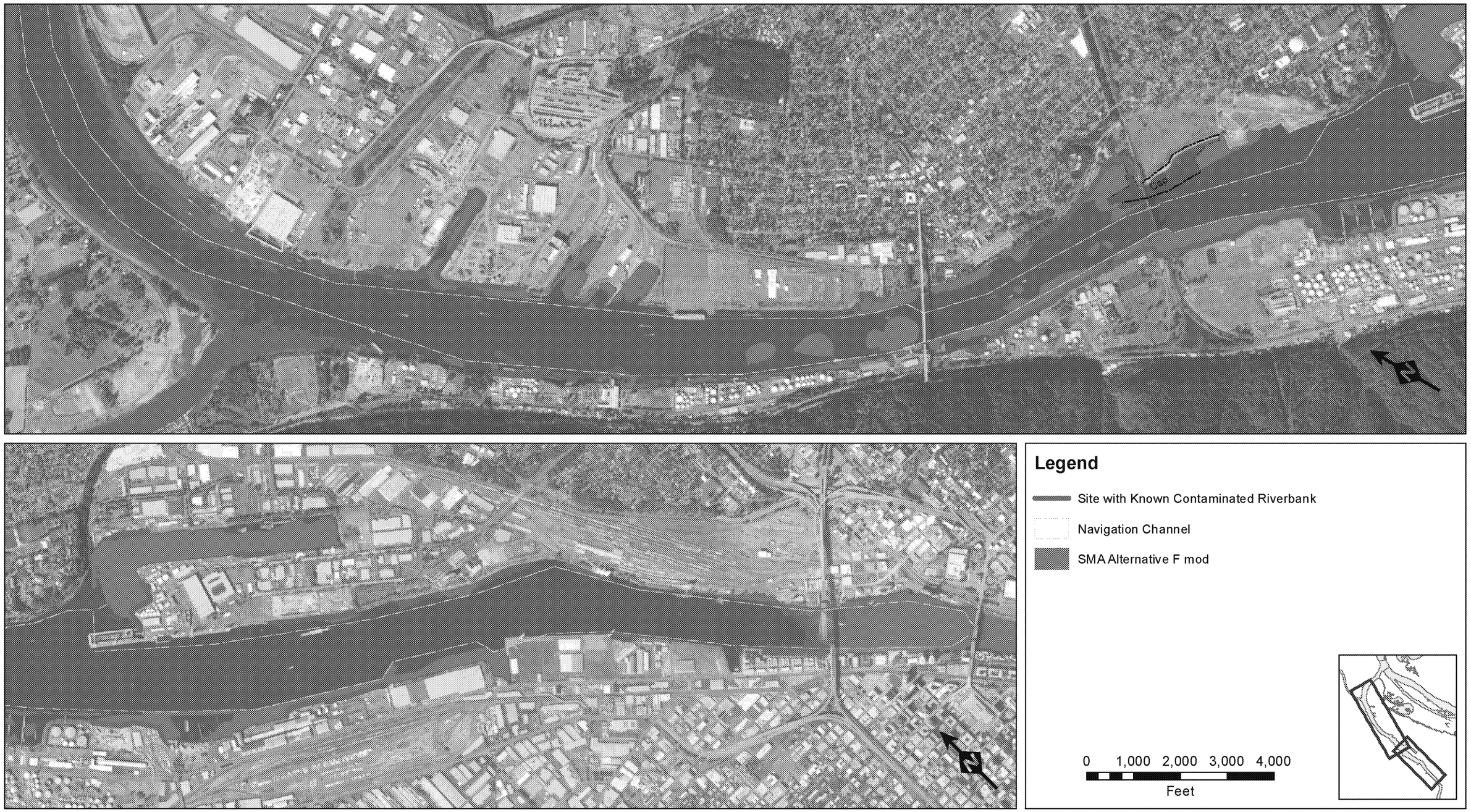
Figures



Notes:
(1) Contamination is defined in three dimensions.
(2) Currently operating or used to stabilize bank. Service life > 50 yrs.

ROD Figure 28. Technology Application Decision Tree

Portland Harbor Superfund Site



Source Credits: Source: Esri, DigitalGlobe, GeoEye, Earthstar Geographics, CNES/Airbus DS, USDA, USGS, AeroGRID, IGN, and the GIS User Community

ROD Figure 30. Sediment Mangement Areas, Selected Remedy

Portland Harbor Superfund Site

Source: Portland Harbor Record of Decision, Figure 30

Path: E:_Projects\Portland Harbor\GIS\MapDocuments\Remedial Design\Figure-4-1_Selected-Remedy-Site-Wide.mxd, Created by: MLF

Date: 4/13/2017

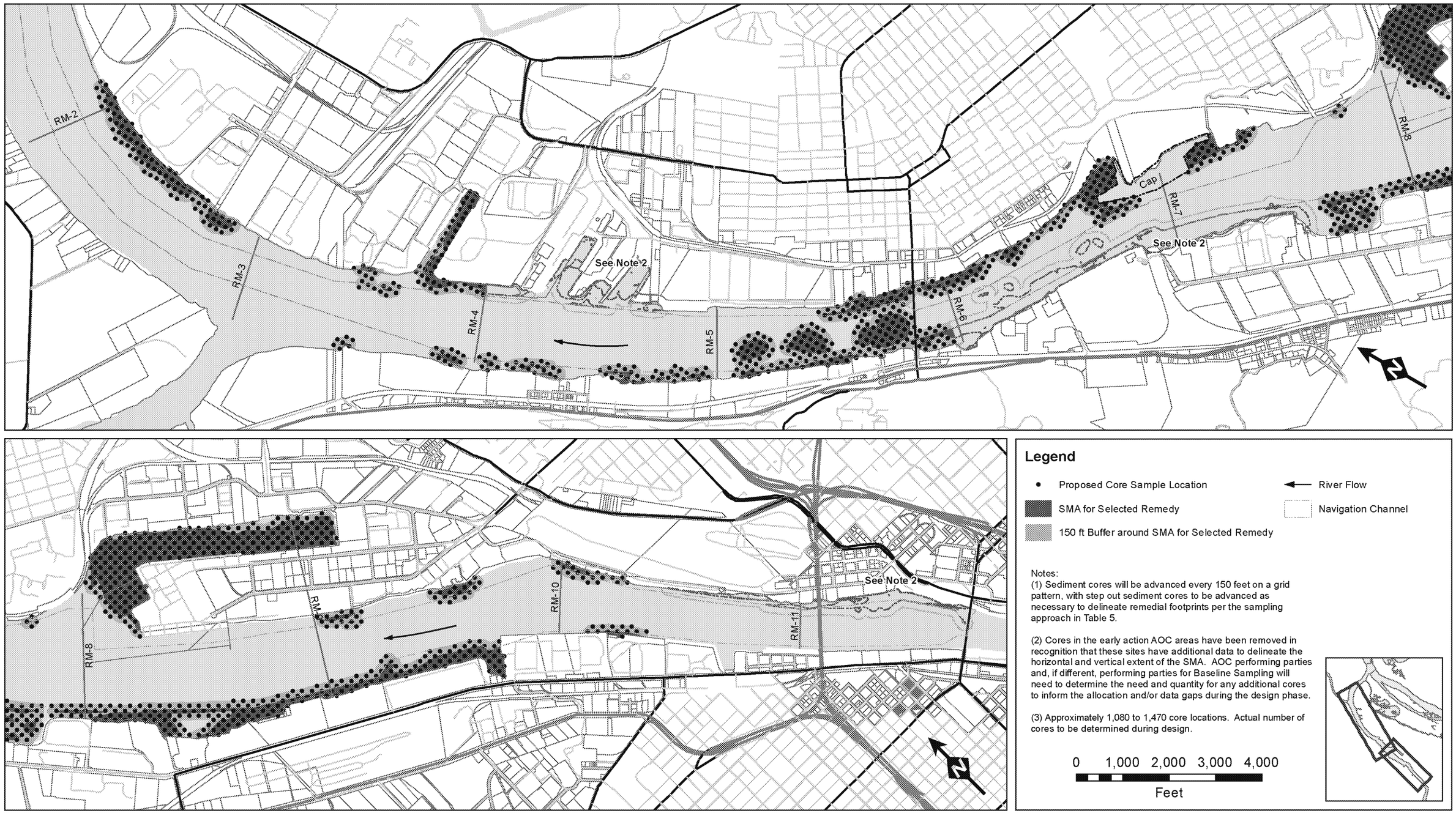


Figure 1. Conceptual Sediment Core Locations for Pre-Remedial Design Characterization

Portland Harbor Superfund Site

Path: E:_Projects\Portland Harbor\GIS\MapDocuments\Remedial Design Phase\Figure-1_Proposed-Core-Sample-Locations-Site-Wide.mxd, Created by: MLF

Date: 5/1/2017

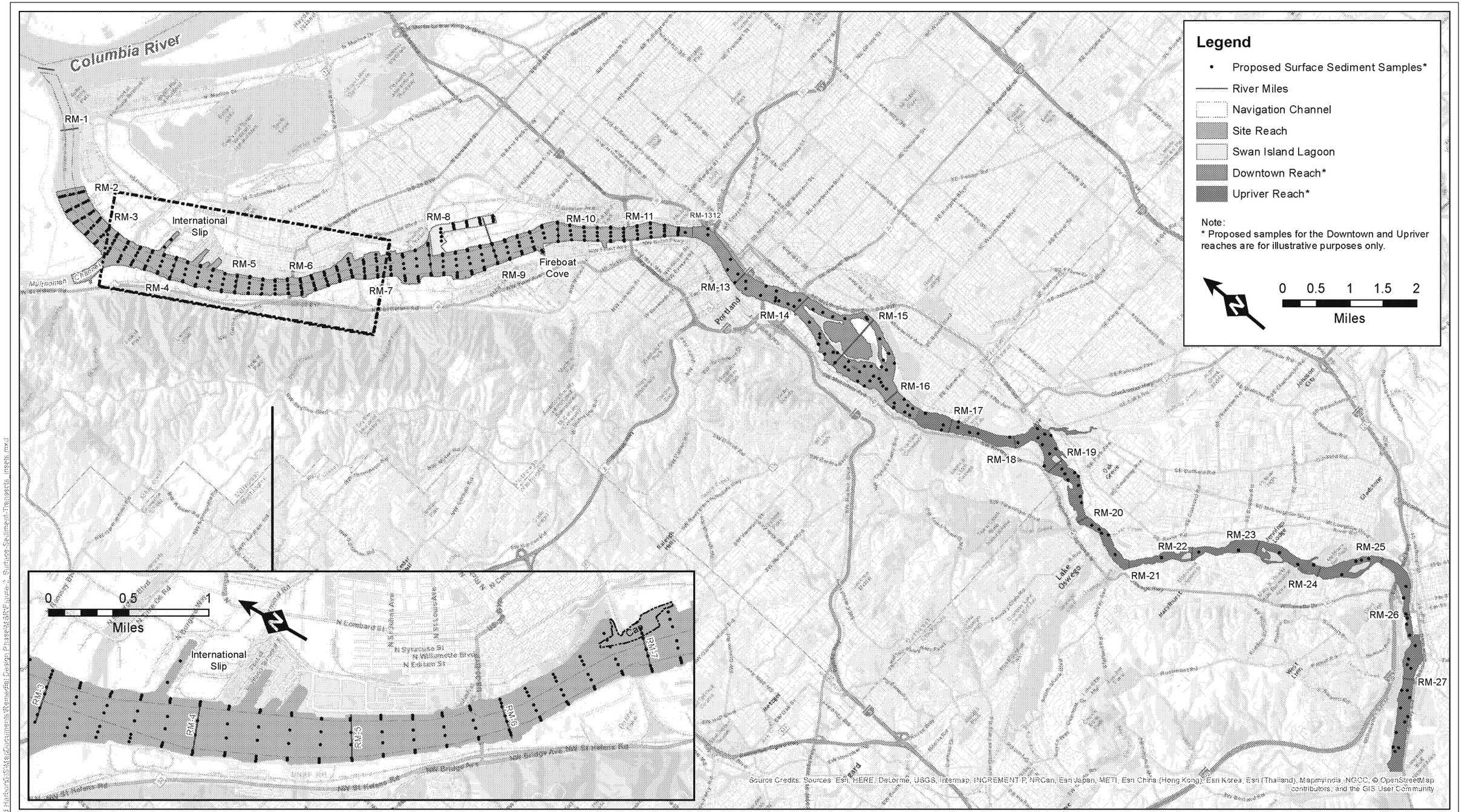


Figure 2. Surface Sediment Transect Locations for Baseline and Long-Term Monitoring

Portland Harbor Superfund Site

Path: G:\Projects\Portland Harbor\GIS\Maps\Documents\Remedial Design Phase\RDWP\Figure 1. Surface Sediment Transects.mxd
Date: 5/31/2017

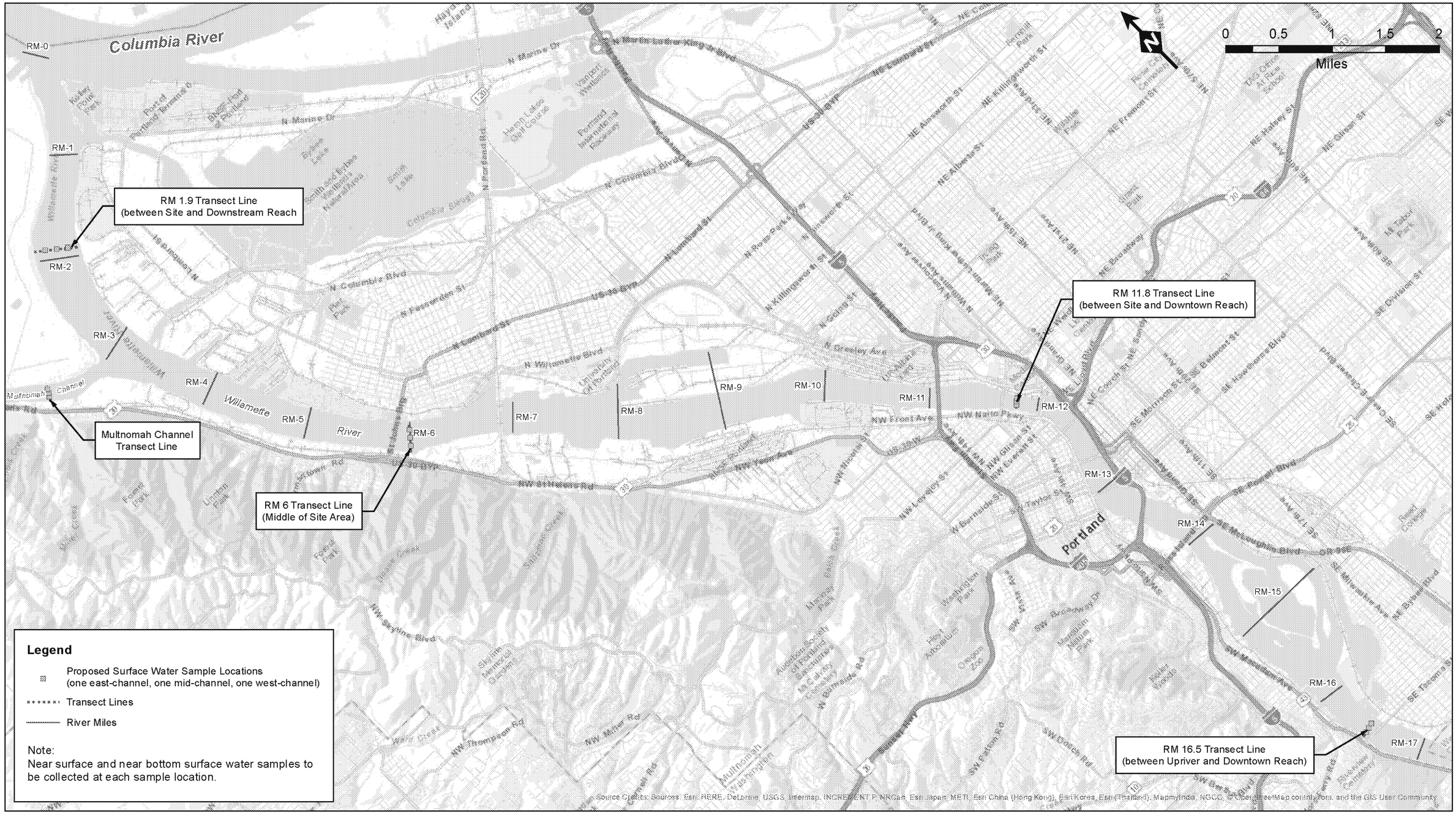


Figure 3. Surface Water Sample Transect Locations for Baseline and Long-Term Monitoring

Portland Harbor Superfund Site

Path: E:\Projects\Portland Harbor\GIS\MapDocuments\Remedial Design Phase\Figure 3_Proposed Surface Water Transects.mxd

Date: 4/19/2017

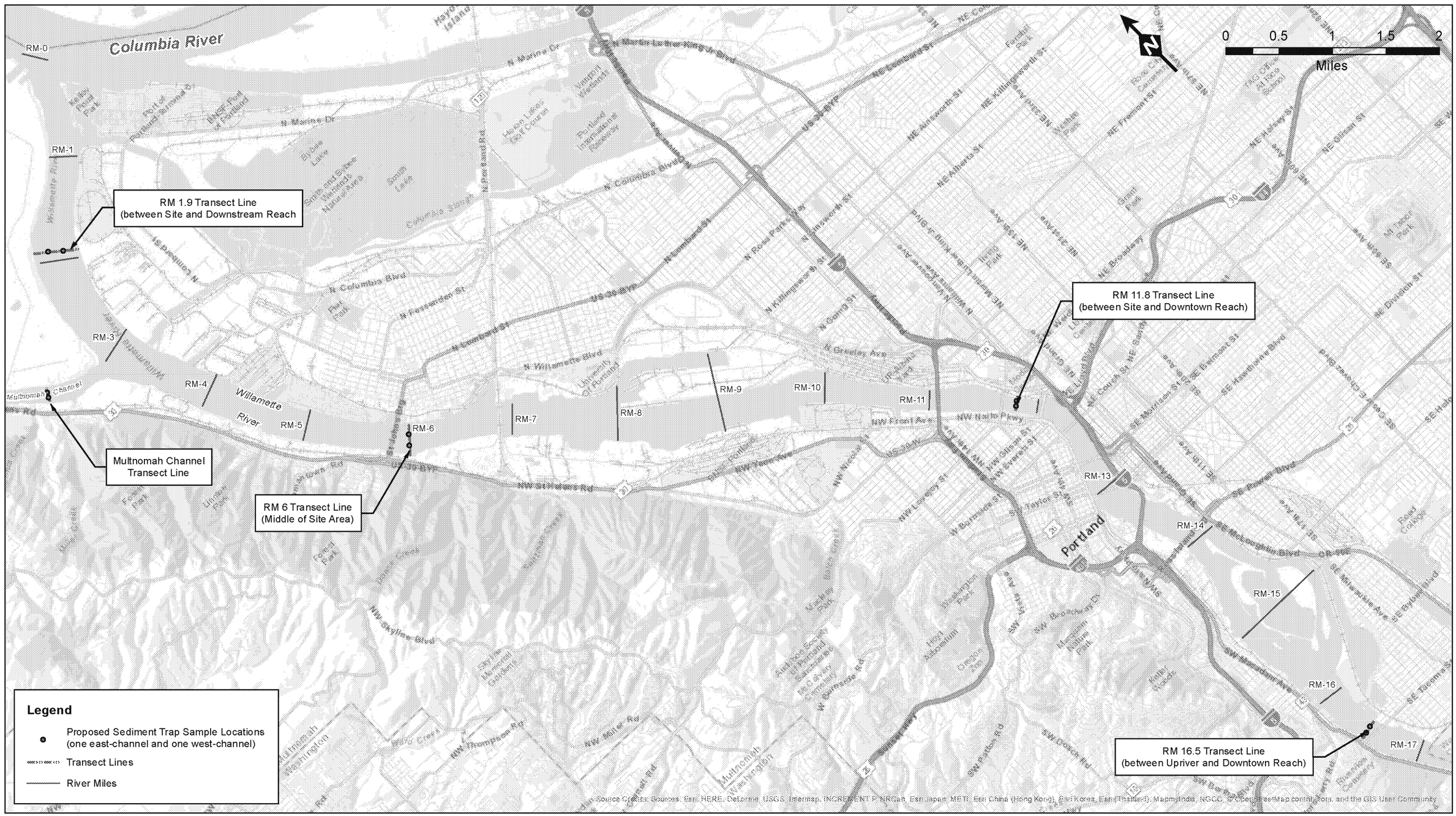
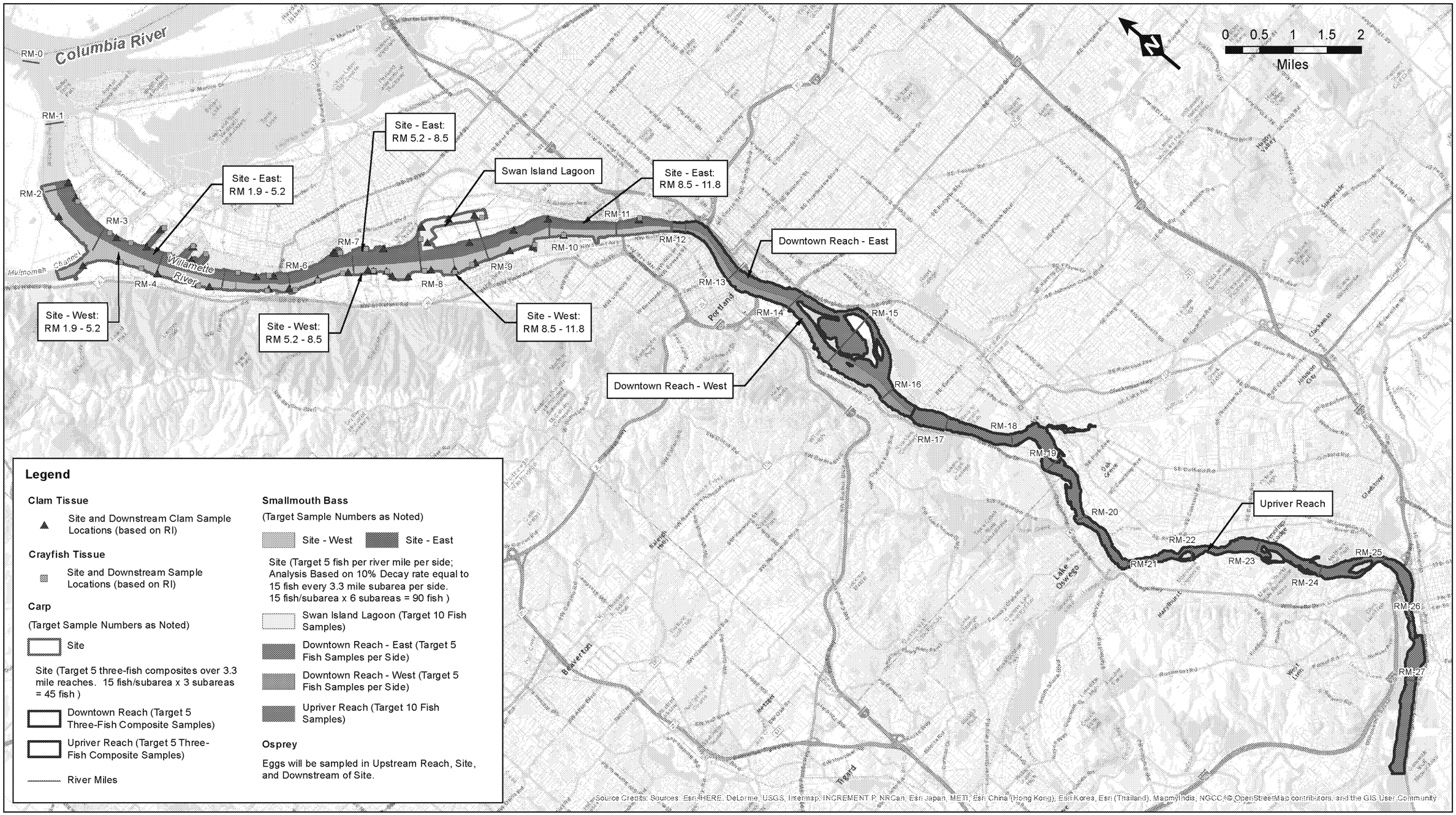


Figure 4. Sediment Trap Transect Locations for Baseline and Long-Term Monitoring

Portland Harbor Superfund Site

Date: 4/13/2017



Path: E:\Projects\Portland Harbor\GIS\Maps\Documents\Remedial Design Phase\BSP\Figure5_Biota Sampling.mxd
Date: 5/16/2017

Figure 5. Biota Sampling Locations

Portland Harbor Superfund Site

Appendix A

Statistical Basis for Sampling Approach

Appendix A

Statistical Basis for Sampling Approach

Introduction

Baseline and long term monitoring programs are intended to quantify remedial effectiveness and progress toward attaining remedial objectives. It is anticipated that statistical analyses will be used to evaluate monitoring data, and the number and frequency of measurements must be adequate to support reliable statistical estimation and hypothesis testing procedures needed for decision making. This appendix documents the basis for the number and frequency of sampling locations to meet monitoring objectives. **Attachment A1** to this appendix provides further discussion equivalency evaluations and explains the benefit of this approach for post-remediation monitoring.

The performance of statistical procedures is characterized by precision of estimates and power to reject tests of hypotheses. For unbiased sampling designs, statistical precision is the difference between a parameter estimate and the true population parameter. More precise parameter estimates have narrower confidence intervals and when estimation procedures are unbiased, precision also provides a measure of accuracy. Statistical power is the probability of correctly rejecting the null hypothesis when in fact the alternative hypothesis is true. Power and precision of statistical procedures are determined by the specific statistical procedure; the number and frequency of sampling events the natural variability of measurements and the magnitude of effects under study. The number and frequency of samples to be collected were determined to provide statistical basis to estimate confidence intervals with specified precision and to test hypotheses of temporal trends with adequate statistical power. Analyses and information necessary to select adequate sample sizes are summarized in the following sections.

Sample size determinations were conducted for fish and sediment sampling intended to support:

1) estimation of mean COC concentrations under pre-dredge baseline conditions,
2) tests of the null hypothesis of no change in COC concentrations,
3) estimation of post remedial recovery rates for fish and sediment at a range of scales, and
4) tests of hypotheses of equivalence between site and upstream reference area conditions.

Data Objectives

For each evaluation, a data objective was set so that a number of samples and an appropriate frequency of monitoring could be determined to meet power and precision objectives.

For comparison of post-remedial COC concentrations with baseline conditions the minimum number of samples was identified so that a 20% reduction in geometric mean (e.g. median of a lognormal distribution) concentration could be identified with at least 80% statistical power.

For temporal trend analysis, the number of samples per monitoring period were determined by identifying the number of samples necessary to detect an approximately 5% to 10% annualized rate of change over a 10-year period with approximately 80% statistical power.

For equivalency analysis comparing site and background concentrations[1], the number of samples per monitoring time step and spatial group were determined so that the upper 90% confidence limit for the ratio of site to upstream geometric means was less than 1.5 when site data are equivalent to background (i.e. R=1.0).

Methods (Sample Size Determination)

A statistical simulation approach was used to develop relationships between sampling designs (i.e. number, frequency and compositing scheme) and precision or power of statistical procedures. Relationships between sample size and power or precision were plotted for selected design configurations and the numbers of samples and design specifications meeting power and precision data objectives were selected. The simulation method was selected because a broad range of statistical tests and sampling designs can be evaluated easily with a single robust framework. Development of the final sampling plan requires inspection of power curves for each media and the statistical tests of interest applied to the primary chemicals of interest. Final sample sizes should meet as many of the data objectives as possible for each media and chemical. It should also be noted that some compromises may be necessary in cases where projected sample size determinations may not be feasible for all combinations of objectives, media and chemicals—a balance of practicality and rigorously meeting data objectives is necessary.

Simulation of Power and Precision

The simulation procedure for developing relationships between statistical sampling design configurations and precision or power of analysis techniques proceeds through 7 steps:

1. Obtain pilot scale data suitable to understand the statistical distribution of data likely to be collected under the sampling program
   1.1. For temporal trend analysis site RI data were used for pilot data within river miles 1.9 to 12
   1.2. For equivalency analysis, data from upstream of river mile 15
2. Develop a model describing statistical distributions based on the pilot data
3. Modify statistical distributions as needed to represent null and alternative hypothesis situations
   3.1. For temporal trends this requires adjusting the mean to represent selected temporal decay rates
   3.2. For equivalency analysis this includes specification of the on-site and off-site concentration mean and variance
4. Select random samples of data from statistical model developed in step 3
   4.1. For temporal trends this includes sampling in a way that mimics the study design, including number and frequency of samples under consideration
   4.2. For equivalency analysis both site and off-site data were assumed to have the same distribution (i.e. populations assumed to be equivalent)
5. Apply anticipated statistical procedures to the synthetic data
   5.1. Estimate equivalency statistics
   5.2. Test for temporal trends
6. Repeat steps 4 and 5 many (N=1000) times to develop a distribution of test statistics

[1] Background data are representative of the distribution of contaminant concentrations representative of limiting conditions and levels of recontamination that may be expected to recontaminate the site post remedy.

7. Post process distributions
   7.1. The number of times the null hypothesis was correctly rejected divided by 1000 is an estimate of statistical power
   7.2. The half width of confidence intervals is an estimate of the precision of the sampling design for equivalency analysis

Data for Planning

Implementation of the steps described above requires an understanding of the statistical distributions that can be anticipated in future monitoring efforts. Generally, pilot data from existing studies are used to infer the likely nature of future data to be collected. The primary parameters needed to develop the power analysis are the general shape of the distributions, symmetric or skewed histograms, and estimates of mean and variance. For analyses reported herein, sample fish and sediment chemistry data collected as part of the remedial investigation (RI) were used to develop these necessary inputs.

For fish tissue power and sample size determinations, PCB concentrations in samples collected from 2002, 2007, 2011 and 2012 were used to develop distributional assumptions. Differing distributions were developed for areas on-site (River Miles 1.9 to 12) and areas upstream of River Mile 15, representing background areas. For sediment analyses chemistry data for total PAHs, PCBs and DDx from the site and upstream of River Mile 15 were used for background evaluations and data from on-site (River Miles 1.9 to 12) were used for other evaluations unrelated to background concentrations.

All power analyses assumed a lognormal distribution with log-mean and log-variance estimated from the RI data. Although environmental data are generally right skewed, frequently they are less skewed than expected under a log-normal distribution. If actual data are less skewed than assumed in this evaluation, one can expect the precision and power of statistical analyses to be better than planned. Using a log-normal approach represents an effort to plan for the worst and hope for the best.

Pre- and Post-Remedial Action Comparisons

Immediately after the remedy is completed, there is an expectation that sample media will exhibit lower contamination concentrations indicative of short term remedial performance. It is not expected that impacted media will attain cleanup levels at this point in time, as the selected remedy generally entails a combination of active remediation and natural recovery processes to reach long term cleanup goals. In the short-term it will be desirable to test for changes in COC concentrations in media immediately after remedy completion as an indicator that short term "step" change in COC levels through implementation of the remedy have taken place as expected.

The short-term effects of the remedy will be examined by testing the null hypothesis of no change in COC concentrations prior to and immediately after remedy completion. Although several tests of change could be considered, the two independent samples Wilcoxon rank sum test is evaluated here (For details see Hollander and Wolfe, 1999, or other standard nonparametric statistics texts). The power of the Wilcoxon Rank Sum test was evaluated using the simulation framework described above where the test was applied to independent simulated samples for a range of effect sizes and numbers of samples. The test compares median (e.g. geometric mean for log-normally distributed data) COC concentrations for samples collected before and after remediation, so the post remedial concentrations were simulated with median levels expressed as a multiplier of the pre-remedial concentrations. For example, an 80%

reduction in concentration was simulated by multiplying the median parameter by a factor of 0.20. Both samples were assumed to be log-normally distributed with log-mean and log-standard deviation estimated from existing fish or sediment data from within the site. Sample sizes under consideration varied from 5 to 200 and percentage reduction varied from 0 to 80% in increments of 20%.

Figure 1 shows power curves for PCB concentrations in fish tissue for the Wilcoxon rank sum test of the null hypothesis of equal medians for pre- and post-remedial concentrations. Percentage changes investigated ranged from no-change to 80% reduction. The vertical axis in each plot is the power (i.e. probability of rejecting the null hypothesis) and the horizontal axis represents the corresponding sample size. The black line represents the case of no change and should be approximately 0.05 indicating the Type I error rate (α) of the test. When there is no actual change in concentration the tests incorrectly reject the null with at most 5% probability. Power curves increase with sample size and with the magnitude of the change in concentration. Based on the power curve presented in Figure 1, approximately 10 samples are required to detect a 20% reduction in median concentration of PCBs in fish tissue after remediation.

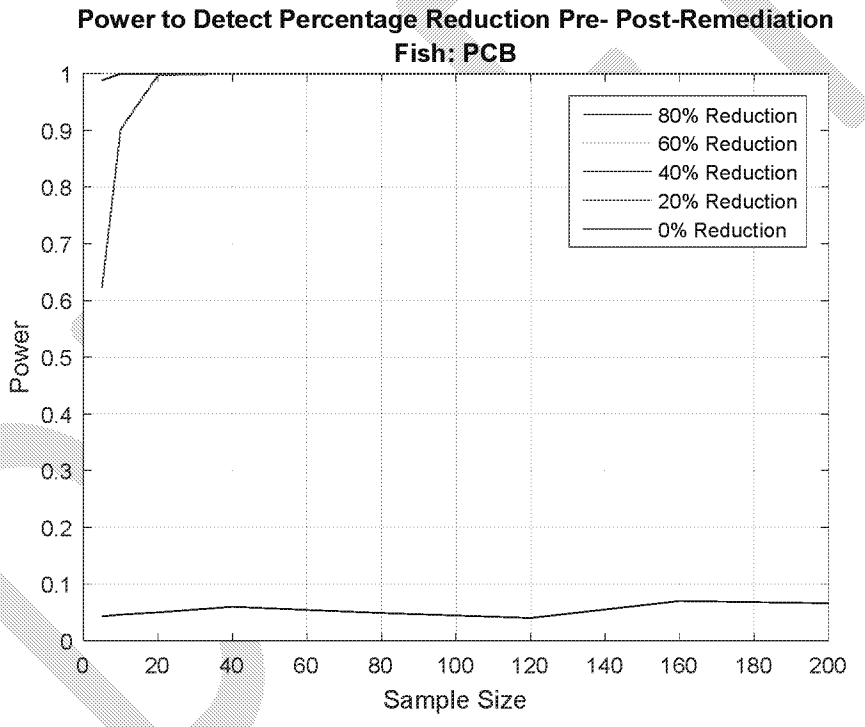


Figure 1. Power to detect a percentage reduction in concentration after completion of the remedy for total PCBs in whole body smallmouth bass tissue.

Figure 2 through Figure 4 summarize the simulated power curves for before after comparison of total PAH (PAH), total DDT, DDD and DDE (DDx) and total PCBs (PCB) in sediment. For PAH 20% reduction in concentration can be detected with at least 80% power for sample sizes of at least 30. For DDx, similar level of power would require over 200 samples, and for PCB, approximately 60 samples would be required. Generally, one would select the chemical with the worst-case scenario to determine sample size for each medium, however the DDx results would push the sampling design beyond what would be

expected to be practical, so it is anticipated that some compromise will be selected as a final sample size to be applied to spatial areas within which pre-post comparisons will be evaluated. Final selection of sample size will require consideration of the numbers needed to meet temporal decay data objectives and equivalency analyses.

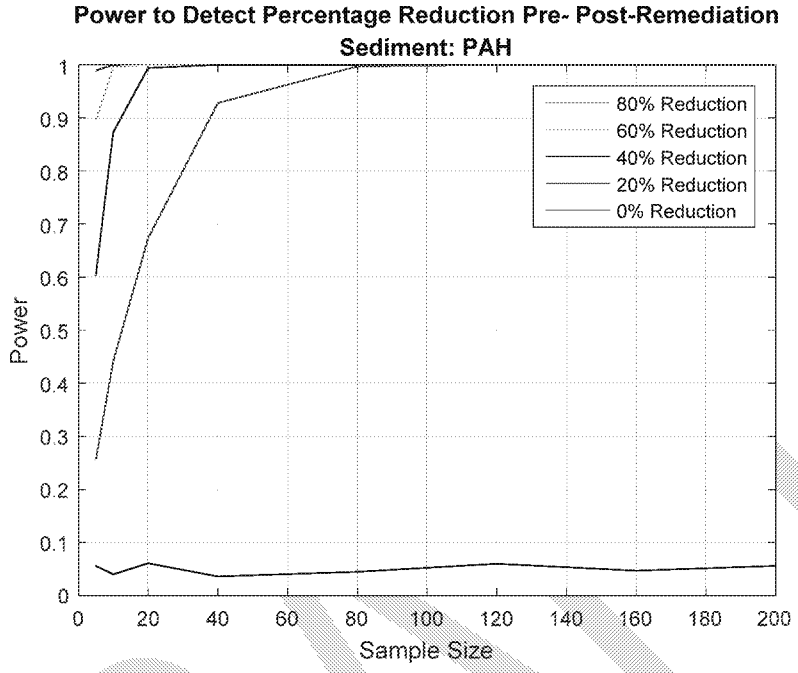


Figure 2. Power to detect a percentage reduction in concentration after completion of the remedy for total PAHs in sediment.

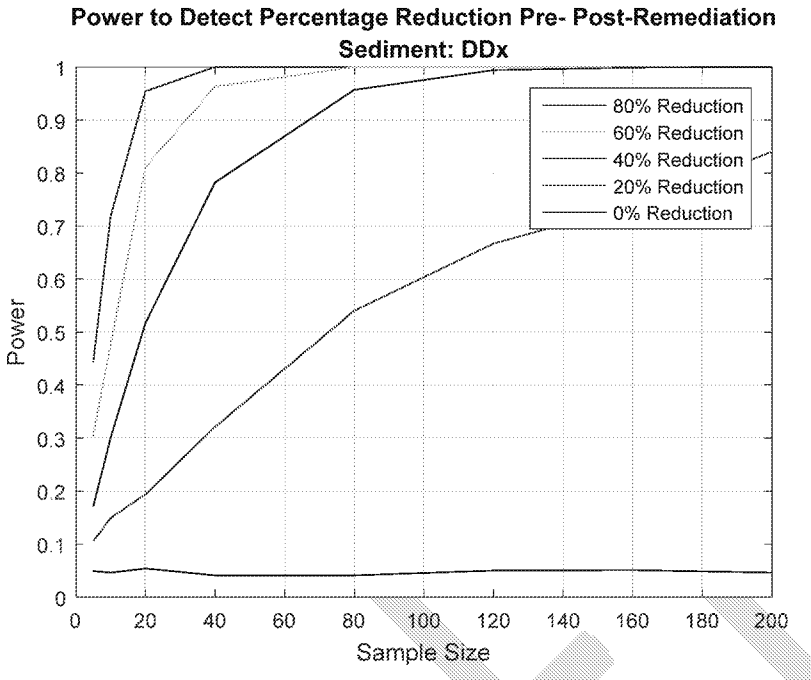


Figure 3. Power to detect a percentage reduction in concentration after completion of the remedy for total DDx in sediment.

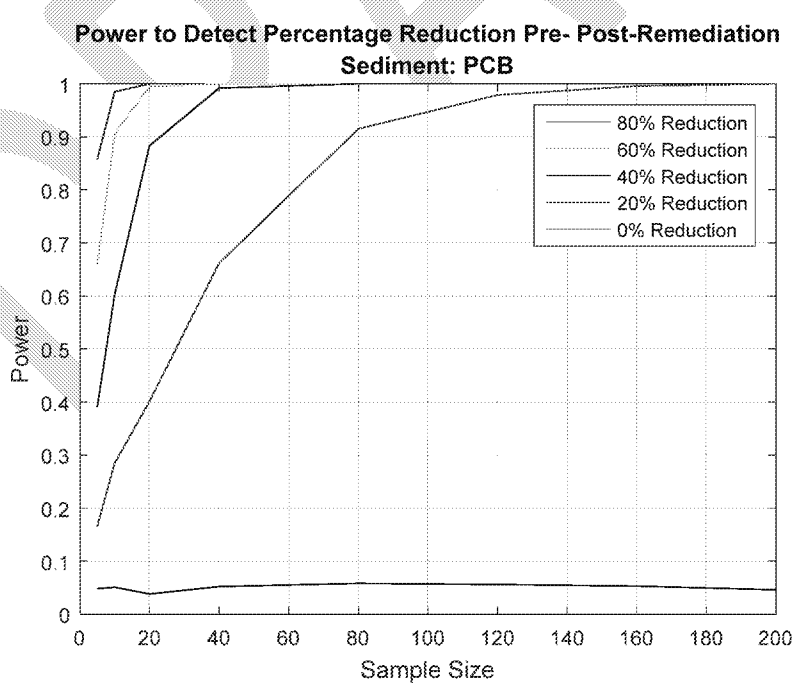


Figure 4. Power to detect a percentage reduction in concentration after completion of the remedy for total PCBs in sediment.

Temporal Trends

It is anticipated that temporal recovery rates will be estimated by fitting first order decay functions which assume concentrations are proportional to an exponential decay function:

$$C_f = C_0 e^{kt}$$

where C_f is the COC concentration in fish and k is the recovery rate which approximately represents the annualized recovery rate. For example, when k=0.05 the rate of recovery is approximately 5% per year. This decay, or recovery, rate represents the combined effects of source control, burial and dilution and other natural recovery processes. For study planning purposes, the number and frequency of samples was determined by identifying combinations of frequency and sample size that are expected to reject the null hypothesis H_0: k=0 in favor of the alternative H_a: k<0 when the actual recovery rate is in the range of 5% to 10%. Preliminary analysis of site smallmouth bass PCB data suggested that natural recovery may be functioning at approximately 10%, however this estimate should be treated cautiously due to limited temporal record. It will be better to plan for detecting a slower rate to insure adequate numbers of samples to accurately interpret long term monitoring data.

Power of statistical tests for specified combinations of number of samples and frequency of sampling was simulated using Monte Carlo techniques based on the assumption that future sample data will be distributed similarly to the RI data with the exception that log-mean concentrations will decline proportionally to time and that log-variance will remain constant. For a particular combination of sample size and frequency, statistical power was simulated by selecting a sample from the assumed distribution for each monitoring time step and the null hypothesis of no recovery was tested for the synthetic data based on the test statistic:

$$T = \frac{k - 0}{se(k)}$$

Which is approximately distributed as a Student's T random variable. In the power simulation process, the null hypothesis was rejected when T was smaller than a Student's T with n-2 degrees of freedom, where n is the number of samples included in the analysis. Power was estimated by repeating this procedure 1000 times and counting the proportion of times out of 1000 that the null hypothesis was rejected. Statistical power was plotted against sample size and the resulting analysis was used to determine number of samples per year and frequency of monitoring time steps providing 80% power to detect a specified rate of natural recovery. Results are summarized in Table 1 for both 5% and 10% recovery rates to allow resource managers to make appropriate value judgements with respect to the final number of samples.

For detecting a 5% decay rate in PCB concentrations in smallmouth bass tissues with 80% power over 40 samples would be required. To detect a 10% recovery rate with similar 80% power approximately 12 individual fish would be necessary. Given the uncertainty in the actual decay rate that can be anticipated it is recommended that a compromise of approximately 20 fish be sampled in each spatially defined area wherein a temporal trend is to be estimated for fish.

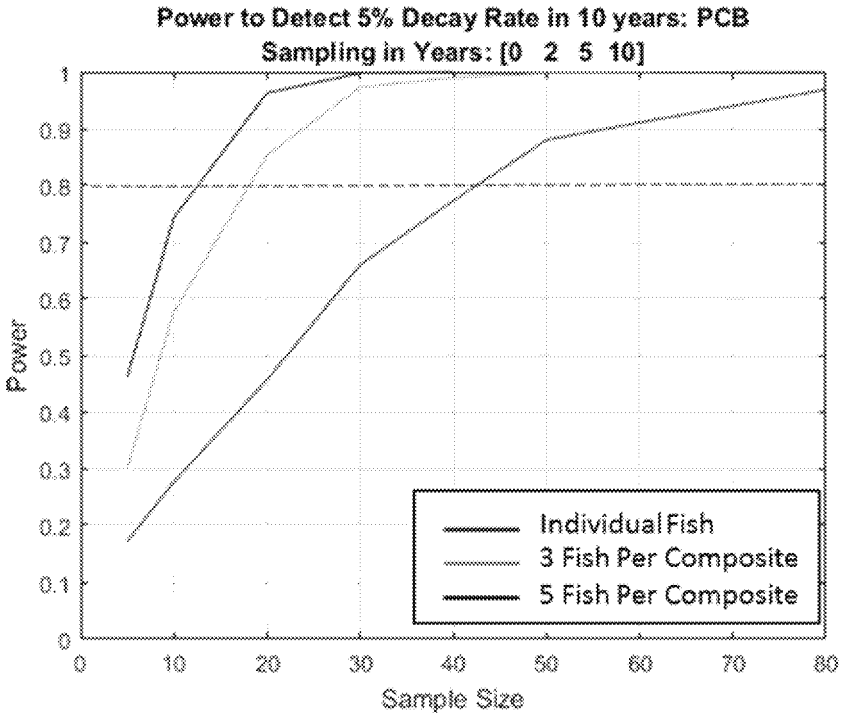


*Figure 5. Power to detect a 5% annualized recovery rate in **smallmouth bass** tissue by monitoring in year 0, 2, 5 and 10 for total PCBs.*

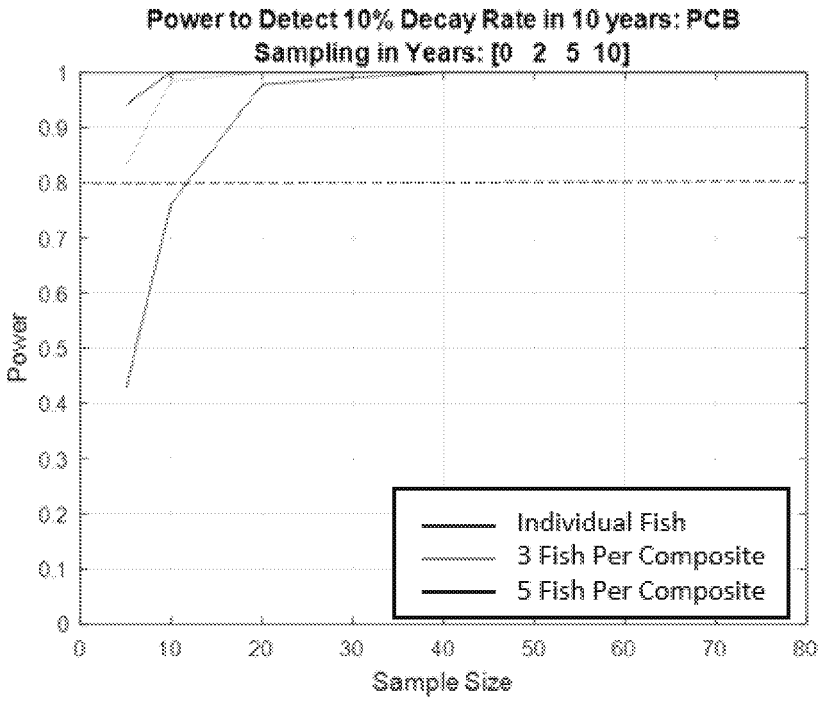


*Figure 6. Power to detect a 10% annualized recovery rate in **smallmouth bass** tissue by monitoring in year 0, 2, 5 and 10 for total PCBs.*

Figure 7 through Figure Figure 9 show power curves for detecting a 10% change in sediment COC levels which lead to approximate sample sizes of 65, 45 and 35 for PAH, DDx and PCB respectively. Although not included here, similar plots were also developed for 5% decay rates in sediment which resulted in sample sizes of more than 200, 160 and 130 for PAH, DDx and PCB respectively.

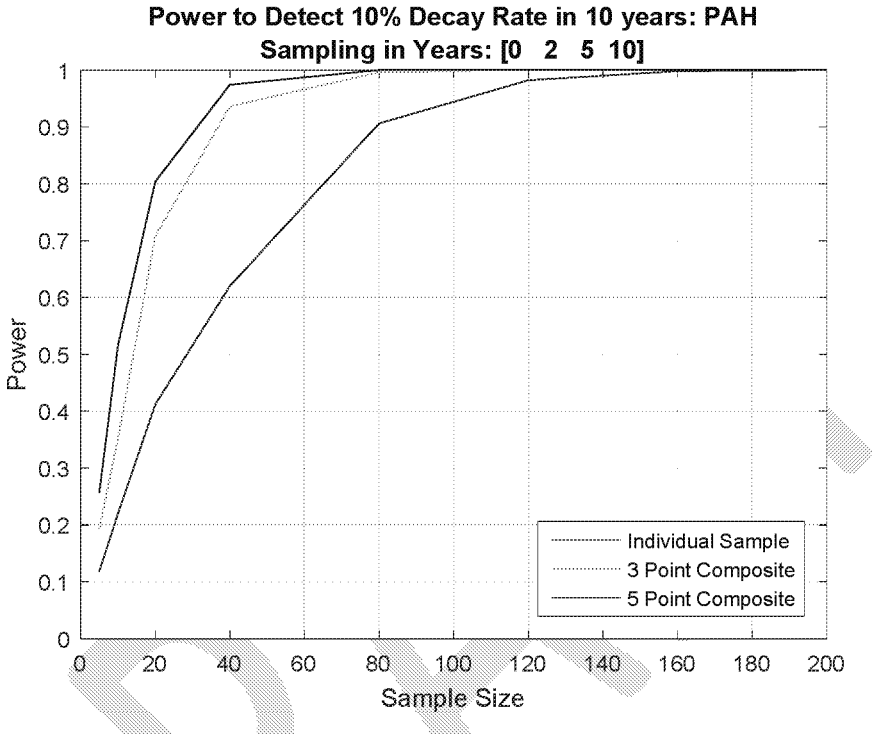


*Figure 7. Power to detect a 10% annualized recovery rate in **sediment** by monitoring in year 0, 2, 5 and 10 for total PAH.*

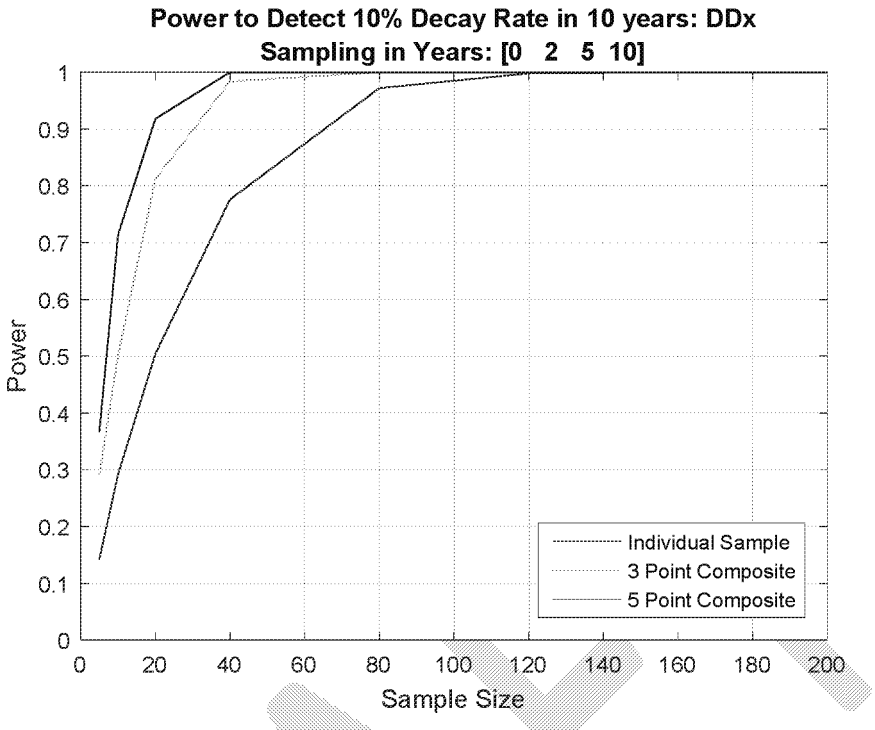


*Figure 8. Power to detect a 10% annualized recovery rate in **sediment** by monitoring in year 0, 2, 5 and 10 for total DDx.*

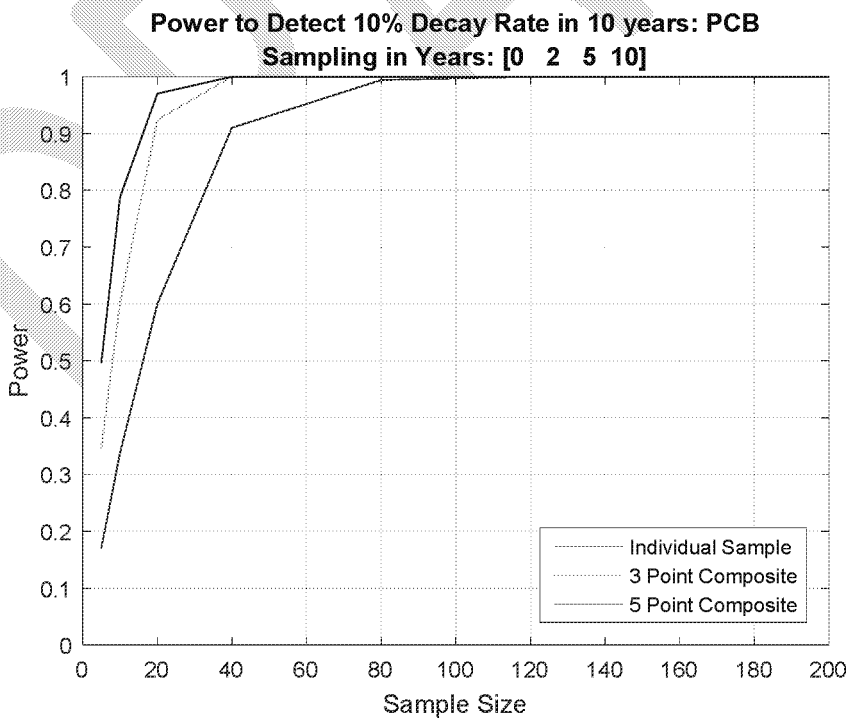


*Figure 9. Power to detect a 10% annualized recovery rate in **sediment** by monitoring in year 0, 2, 5 and 10 for total PCBs.*

Equivalence Evaluation

Attainment of RAOs occurs when sediment cleanup levels as specified in the ROD are attained. For PCBs this cleanup level is 9 ppb. For the purposes of evaluating remedy effectiveness and evaluating whether the remedy is functioning as intended, during the five-year review, site data will also be compared with fish and sediment data from areas upstream of the site using an equivalence analysis (McDonald and Erickson, 1994). This type of evaluation assesses whether the site data are statistically equivalent to upstream data representing background which may differ from 9 ppb in the future.

After site remediation and natural recovery, site and upstream concentrations are expected to converge and be similar (equivalent). This convergence will be evaluated by estimating the ratio of site geometric mean to upstream geometric mean. Geometric means are used in this situation because ratios of geometric means can be more precisely estimated than ratios of arithmetic means. Equivalence will be established when the upper 90% confidence limit for this ratio is no more than 1.5.

The number of samples necessary for equivalence analysis was developed using the simulation approach described above followed by identifying the number of samples necessary to achieve an upper confidence limit less than 1.5 based on the distribution of background COC data.

Rationale for Equivalence Test

Post remedial monitoring is intended to provide measures of remedial effectiveness and to provide site managers data suitable to determine when a site has met remedial goals. Ideally a remedial goal is achieved when data "demonstrate" that site conditions are equal or equivalent to conditions that would have been present absent the release. Prior to establishment of a need for remedial action the site is presumed to be un-contaminated or equivalent to reference condition and data are used to reject the classical null hypothesis of no difference between site and reference conditions. In this context the classical framework for hypothesis testing is appropriate—assumption of no difference unless and until data demonstrate otherwise with 95% confidence. This approach leads to development of tolerance limits based on background data which are generally used as points of compliance to identify areas that are or are not contaminated.

Conversely, when a release has been documented and there is known risk and a need for remediation, the null hypothesis of equivalence is no longer appropriate. The appropriate null hypothesis represents the current understanding of the site which is that contamination and risk have been established and one must assume the site to remain contaminated until data demonstrate with a high degree of confidence that the site is no longer contaminated. This suggests demonstration of remedial success by rejecting the reverse-null hypothesis:

$$H_o: \mu_{site} \geq \mu_{reference}$$

in favor of the alternative:

$$H_a: \mu_{site} < \mu_{reference}.$$

But in order to reject this reverse null it is necessary for the site mean to be lower than the reference condition, which is generally unachievable in practice. It would be unfair to require responsible parties to clean up to levels lower than background conditions.

Tests of bioequivalence (McDonald and Erickson, 1994) acknowledge this situation and provide a workable alternative that maintains the proper assumption of contaminated until data prove otherwise, without requiring cleanup to concentrations that are below background conditions. This is achieved by inserting a coefficient of equivalence ($R_0 > 1.0$) and testing the null hypothesis:

$$H_o: \mu_{site} \geq R_0 \times \mu_{reference}$$

against the alternative:

$$H_a: \mu_{site} < R_0 \times \mu_{reference}$$

where the equivalence coefficient is chosen to represent a scientifically meaningful value.

For example, at Portland Harbor the cleanup value for PCBs is 9 ppb and one could insure that the true site mean is less than 18 ppb by selecting R_0 to be 2.0. Site and reference means would be declared equivalent when an upper confidence limit for the sample ratio $\hat{R} = \frac{\hat{\mu}_{site}}{\hat{\mu}_{reference}}$ is less than 2.0.

Tests for equivalence are intuitive when expressed as confidence intervals for the ratio of site to reference means or when data are right skewed geometric means. Equivalence is established by rejecting the null hypothesis that the true Ratio is greater than a specified scientifically meaningful value which is equivalent to comparing the upper confidence limit to the specified value. When the UCL is less than the specified ratio (e.g. defining equivalence) equivalence is established.

This equivalence testing approach has several advantages relative to the test of classical null hypothesis:

1) For the traditional test of hypothesis where equal means (i.e. R=1.0) is the null hypothesis, remedial success is concluded when the null hypothesis is not rejected. Although common, this is an inappropriate conclusion as failure to reject the null simply means that the data and study design were inadequate to reject the null, which may or may not be true. Declaring remedy effectiveness in such a situation would be inappropriate because the true ratio could be substantially higher than the sample estimate as shown in the top panel of *Figure 10*. In this situation, the confidence interval for R captures 1.0 indicating that the classical null hypothesis would not be rejected (i.e. concluding successful remediation) when in fact the true ratio could be more than a factor of 2 greater than the reference area mean.
2) In the top panel of *Figure 10* the UCL exceeds the specified value of 2.0 which also leads to failure to reject the null hypothesis of non-bioequivalence and site managers would not change assumption that the site remains more contaminated than the reference site (i.e. site not proven to be equivalent to background). This is appropriate because the assumption that the site is contaminated should only change when statistically reliable evidence falsifies the assumed null condition.
3) The lower panel in *Figure 10* illustrates the desired outcome where the sample ratio and its' upper confidence limit is lower than the specified limit of 2.0 which leads to the rejection of the null hypothesis of non-equivalence in favor of the alternative hypothesis that site and reference data are "equivalent".

4) Equivalence tests provide direct and quantifiable evidence that two populations are similar with a level of confidence and an a priori definition of similarity (i.e. equivalence).
5) The plot in the lower panel of *Figure 10* also illustrates that the classical null hypothesis can result in data that are "too" precise. Note that the lower limit fails to capture 1.0 indicating that in this case the null hypothesis of "equality" would be rejected and one would conclude that the site had not been adequately remediated when the site mean was in fact just 4.5ppb greater than the 9 ppb cleanup level—a statistically significant difference lacking in biological meaning.

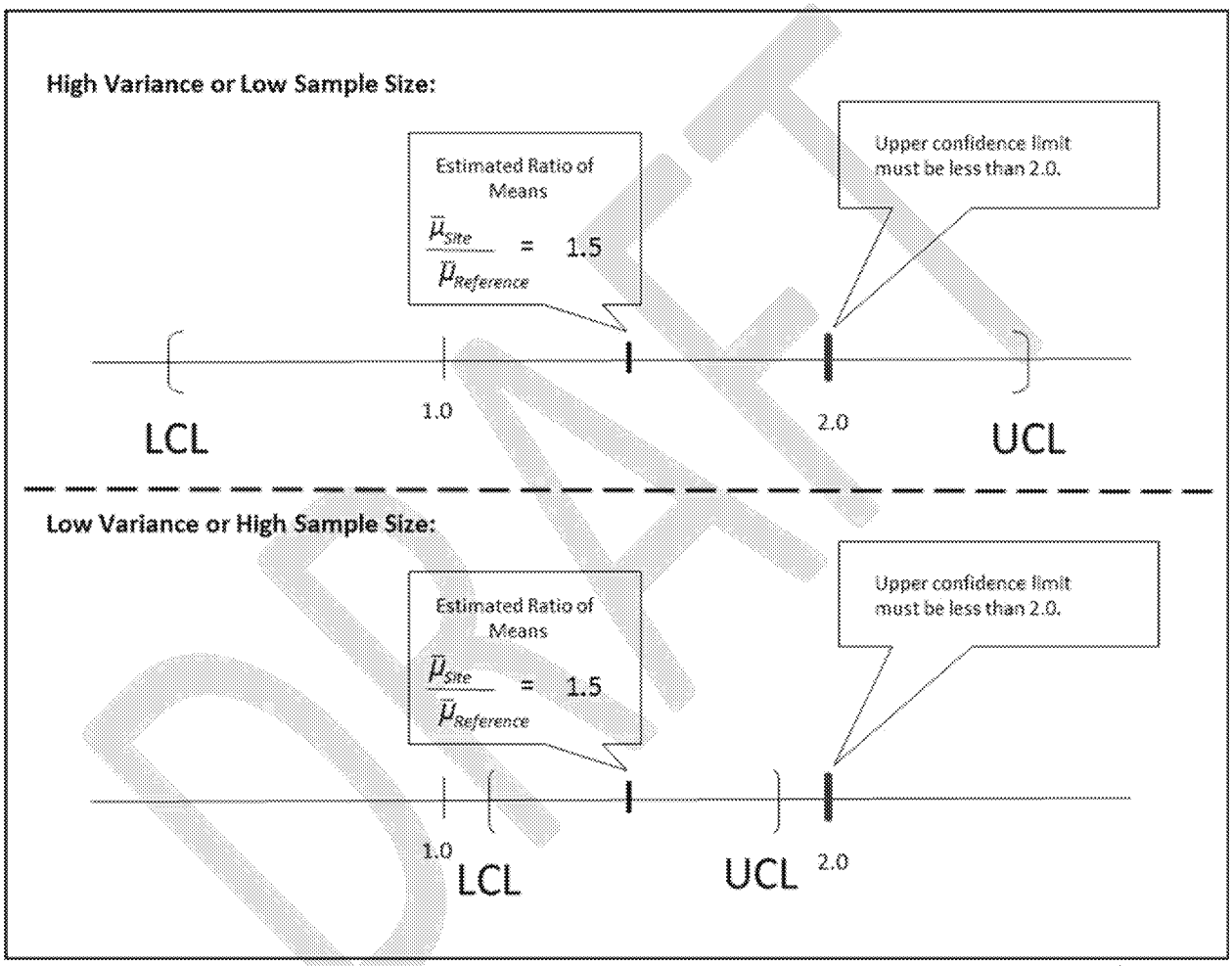


Figure 10. Comparison of reverse and classical null hypotheses with high variance and/or small sample size (top panel) and with low variance and/or large sample size (bottom panel)

Selection of Equivalence Coefficient and Sample Size

For equivalence testing it is necessary to identify a balance between a rigorous equivalence coefficient (R_0) and the number of samples necessary to reliably assure that when site and reference data are equivalent, (i.e. when $\hat{R} = 1.0$) there is high power to correctly reject the null hypothesis of non-equivalence. Using the simulation approach described above, the upper 95% UCLs from 1000 samples of

sizes 10 through 200 were plotted and compared with a trial equivalence coefficient of 1.5 and it can be seen that 100% of simulated sample sets had upper confidence limits below the 1.5 when 50 samples were selected in each of the reference and site data sets (N=100 total). Assuming a geometric mean of approximately 9 ppb in the reference area, bioequivalence would be established when the UCL is less than 13.5 ppb and importantly the sample geometric mean would, by necessity, be very close to 9 ppb as desired. Although at first blush it may sound like the agencies are defining the goal to be 50% higher than the cleanup levels, one must recognize that if the actual ratio is 1.5 the chances of rejecting the null hypothesis is just 5%, virtually impossible. Demonstration of equivalence requires the site conditions to be very close to background levels, high quality low variance data must be carefully collected, and the number of samples must be adequate to insure very narrow confidence limits on the ratio. Taken together the bioequivalence approach incentivizes a high quality and statistically rigorous study design.

Figure 11 through Figure 13 show that the UCL for the ratio of geometric mean COCs in sediment can be expected to fall below the target 1.5 level for 50, 25 and 50 samples for PAH, DDx and PCB respectively.

Demonstration of equivalence for PCB in smallmouth bass tissue approximately 25 to 30 individual fish samples (Figure 14).

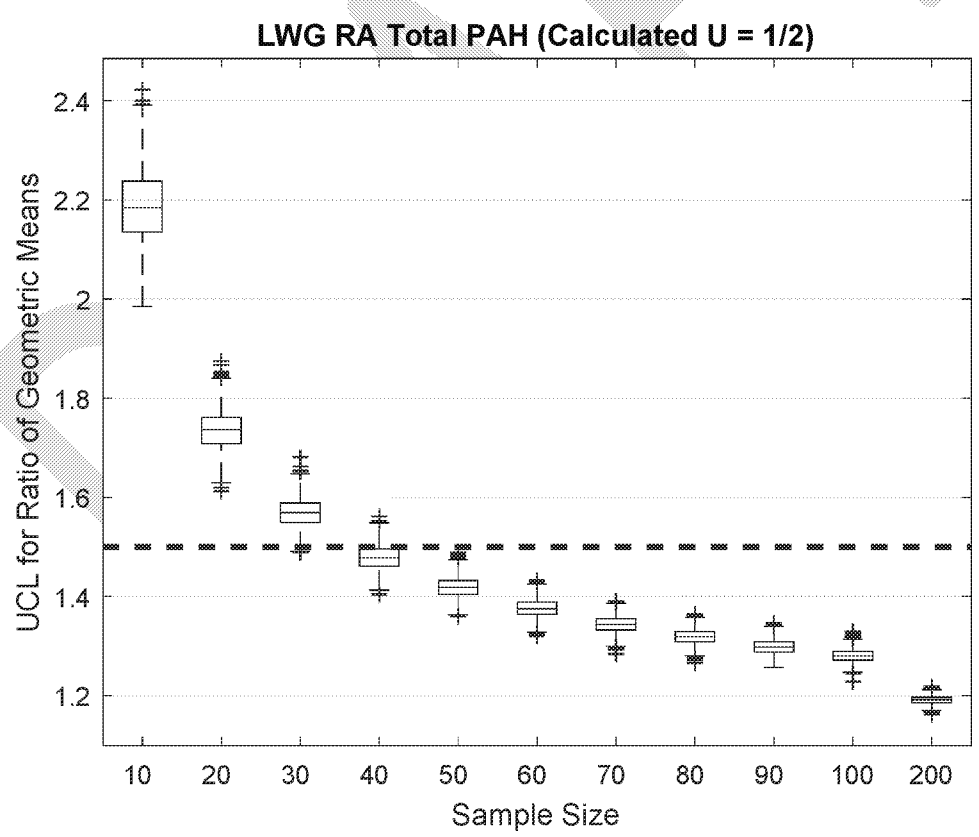


Figure 11. Upper 95% confidence limits for 1000 simulated samples when site and reference data are equivalent for Total PAH in sediment.

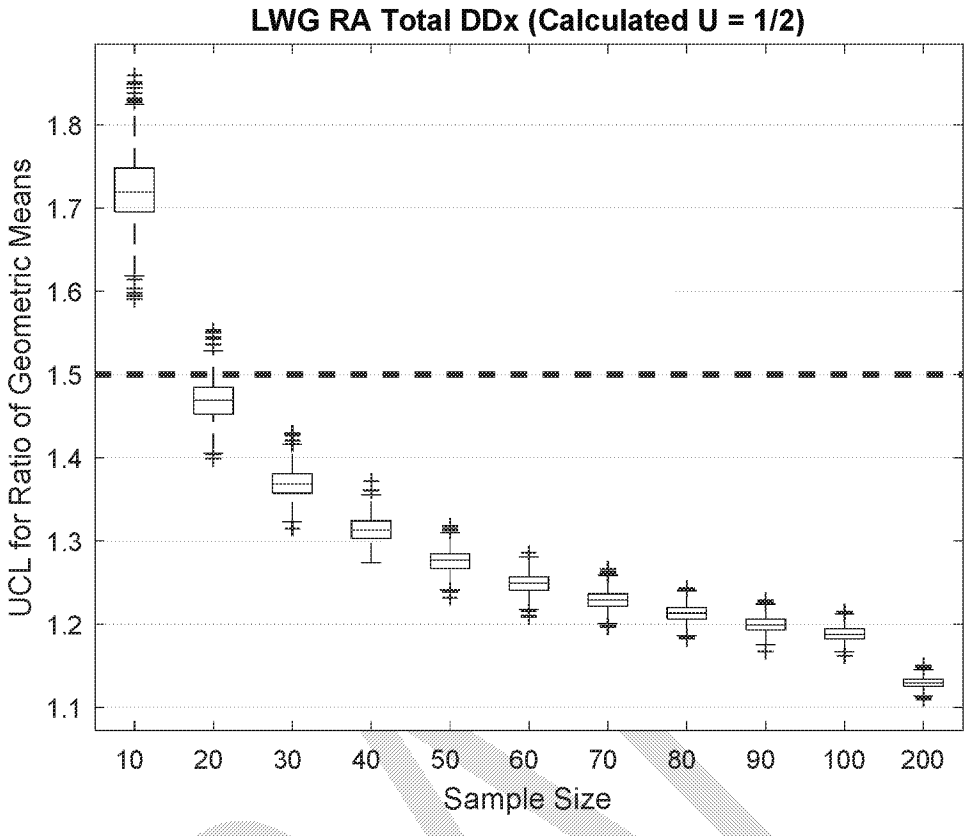


Figure 12. Upper 95% confidence limits for 1000 simulated samples when site and reference data are equivalent for Total DDx in sediment.

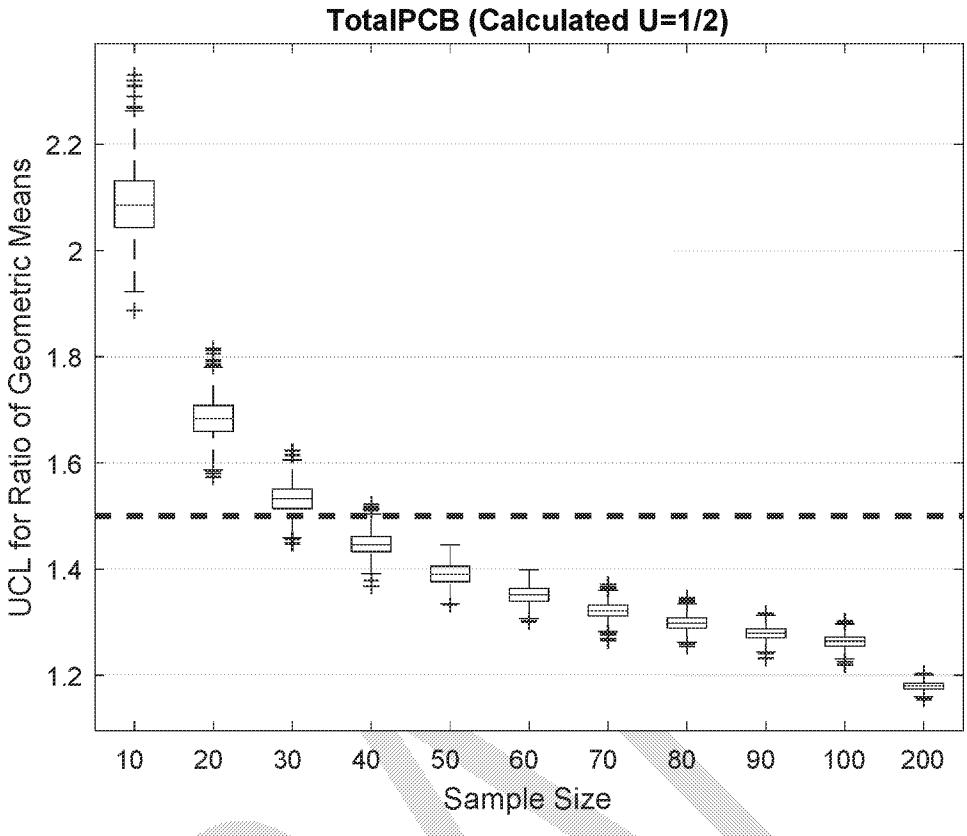


Figure 13. Upper 95% confidence limits for 1000 simulated samples when site and reference data are equivalent for Total PCBs in sediment.

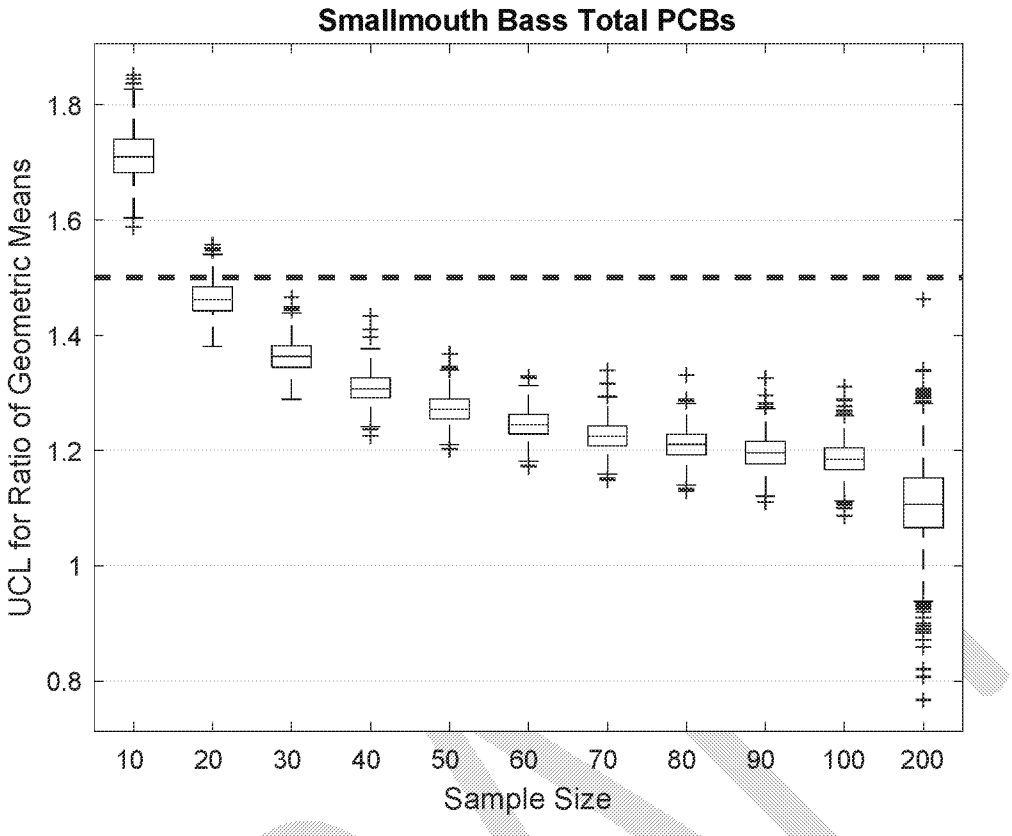


Figure 14. Upper 95% confidence limits for 1000 simulated samples when site and reference data are equivalent for Total PCBs in whole body smallmouth bass..

Summary

Sample size results for each combination of media chemical and statistical analysis of interest are summarized in Table A-1 below. It is anticipated that site managers will utilize this table to develop final numbers of samples as the basis for the baseline and long term monitoring programs. In the simplest way one would identify the maximum number needed across all fish tissue evaluations which would lead to approximately 45 fish per group. Similarly, because all sediment concentrations are based on the same samples one could identify the maximum number per each statistical evaluation across chemicals (row) followed by identifying the maximum within this row-wise calculation which would lead to over 200 sediment samples within each analysis group.

The power analysis developed in this appendix represents only the mathematical and statistical aspects of study design. Final implementation of these results in developing a sampling plan must also incorporate other non-statistical factors such as practical concerns related to implementation, cost and the relative importance of each combination of media and chemicals. The results here provide a guide to final selection of sampling numbers and frequency by site managers.

It should also be noted that the results summarized here are approximate, based on simulations derived from site data that may be imperfect representatives of future conditions. In general, one can expect

future data distributions to recover in absolute concentration, but also it can be anticipated that with declining mean values will also come declining variability. Generally, analyses provided in this report can be considered somewhat pessimistic, so balancing perfection with practicality is fully acceptable and even expected when planning a monitoring program. With this in mind, on the order of 20 to 45 fish and 65 to 130 may be considered reasonable target sample sizes for groups of interest. These include a total of 8 groups as currently envisioned; upstream and downtown reaches and 6 groups within the site—left right and mid channel subdivided into approximately 3 mile subsections. Additional power and precision could be achieved with potentially increased power by considering compositing or by increasing the size of analysis groups--equivalently defining fewer groups.

Table A-1. Sample sizes adequate to support baseline and long term monitoring objectives in fish and sediment at the Lower Willamette River Superfund Site.

	Smallmouth Bass	Sediment		
Hypothesis (Effect Size)	PCBs	PAH	DDx	PCBs
Pre-Post Comparison (20% Reduction at 80% Power)	10	30	>200	60
Temporal Trend (5% Decay Rate at 80% Power)	45	>200	160	130
Temporal Trend (10% Decay Rate at 80% Power)	15	65	45	35
Equivalence (Defined by ratio of 1.5 and confidence level of 90%)	25	50	25	50

References

McDonald, L.L. and W.P. Erickson, 1994. Testing for bioequivalence in field studies: Has a disturbed site been adequately reclaimed? In *Statistics in Ecology and Environmental Monitoring,* ed. D.J. Fletcher and B.F.J. Manly, 183-197. Dunedin, New Zealand: University of Otago Press.

Hollander, M., and D. A. Wolfe. 1999. Nonparametric Statistical Methods. Hoboken, NJ: John Wiley & Sons, Inc.

Attachment 1

Direction of the Null Hypothesis for Demonstrating Remedial Success

Attachment A1

Direction of the Null Hypothesis for Demonstrating Remedial Success

Objective: The objective of post remediation monitoring is to collect data adequate to demonstrate to regulatory agencies and the public with a strong level of confidence that selected remedial actions have been successful. In this context, the default position of the public and the regulatory agencies is to assume that the site remains un-remediated until new post remedial data demonstrate reliably that site conditions have changed in response to natural process and implementation of the remedial action. Success and completion of the remedy is demonstrated when site conditions have met stipulated cleanup goals and/or equivalence with reference conditions expected to prevail at the site over the long term using a test of bioequivalence.

Null Hypothesis for Traditional Research Questions: Traditional statistics and hypothesis testing is generally geared toward evaluations comparing treatments with controls, where the treatment is assumed to be ineffective until experimental data demonstrate convincingly that it is superior to the control group. In this context, the null hypothesis is defined such that control and treatment means are equal

$$H_0: \mu_{treatment} = \mu_{control}$$

and rejection of this null hypothesis at 5% level of significance demonstrates with 95% confidence that the treatment differs from the control.

Using the traditional test of equal means for demonstrating remedial effectiveness would suggest concluding the site had been effectively remediated when the null hypothesis was accepted. This approach is inappropriate because statistical tests cannot prove the null hypothesis. When the null hypothesis is not rejected, it is impossible to distinguish the following two conditions;

1) The null hypothesis is false, but data were insufficient to distinguish treatment and control groups relative to the variation in the measurements; or
2) The null hypothesis is indeed true.

To rectify this situation, some experimenters have imposed statistical power conditions which would preclude the un-detected difference in means from exceeding some magnitude at a selected probability level. However, to ensure such a constraint, the regulatory agency is required to effectively impose sample size requirements on the responsible parties, which may be difficult many responsible parties and stakeholders must all come to agreement.

Bioequivalence Test: Test of bioequivalence effectively reverse the null hypothesis from one of equality of means to one of inequality, assuming that the site remains contaminated until data can be used to nullify this default position.

Why is this an appropriate position for the agencies to take?

After responsible parties have implemented the remedy, they are essentially asking the agencies to agree that the overall process of risk assessment, remedial investigation and feasibility studies and remedial design and implementation have culminated in an effective remedy. The agencies are

effectively being asked to endorse the outcome with a clean bill of health. It is then incumbent on the agencies to minimize the risk that the remedy may have been unsuccessful. In this context, it is correct for the agencies to assume the remedy was unsuccessful until data demonstrate otherwise. Rejection of the null hypothesis of inequality in favor of the alternative of equivalence at a specified level of significance (equivalently high confidence) provides the agencies and the public with a demonstration that the responsible parties' claim of success are in fact reliable. Unfortunately, simply reversing the null hypothesis is not enough to develop a fair and appropriate decision approach, because rejecting the null hypothesis of equality vs the alternative that the site has lower concentrations than reference would be unfair in that cleanup levels would necessarily be lower than natural or anthropogenic background conditions. A "reasonably" attainable and scientifically meaningful level of similarity is required to define equivalence. Tests for equivalence then test the null hypothesis that site concentrations are greater than reference conditions by a pre-set and agreed to amount defined as "equivalent". Rejection of the null hypothesis of non-equivalence at the 5% level of significance provides the agencies with assurance that the site mean is less than the specified threshold with 95% level of confidence. Setting the equivalence threshold at a biologically meaningful difference insures that the remediation will be declared successful only when there is strong evidence that the true site mean concentration is within biologically meaningful margin of error of the reference condition—the site is statistically equivalent to reference.

Comparison of Approaches

Four red horizontal lines are plotted in Figure 1 representing confidence intervals for 4 hypothetical situations that may result from monitoring data. The solid vertical line at 0.0 represents the null hypothesis of no difference and the traditional null hypothesis would be accepted when intervals include this line representing no difference. The dashed vertical line represents a hypothetical value deemed to represent equivalence, and the hypothesis of non-equivalence is rejected when the upper confidence limit is less than this value. The remedy is declared successful when the null hypothesis is accepted under the traditional framework, whereas rejection of the null hypothesis demonstrates success under the equivalence framework.

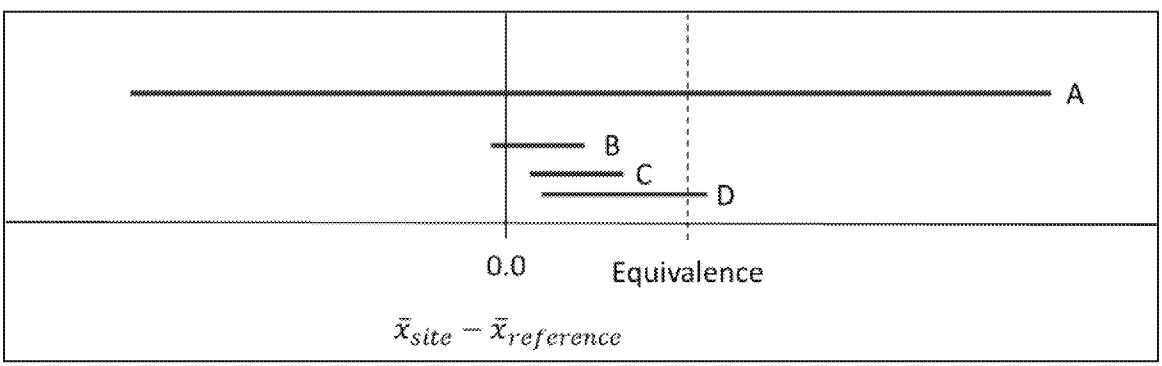


Figure 1. Hypothetical confidence intervals (red lines) for the difference in control and reference concentrations and comparison of traditional and equivalence tests under four different situations (A through D).

Situation A is the primary reason for which the traditional null hypothesis is understood to be inappropriate. In this case a wide confidence interval does not precisely bound the true mean, which

could be considerably higher than the reference mean, yet the remedy would be declared successful. Wide intervals are indicative of small sample sizes, high variance and/or poor laboratory practices. Under the traditional approach there is incentive for responsible parties to negotiate for small sample sizes to improve chances for the wide interval most likely result in accepting the traditional null hypothesis—declaring success. Conversely because the wide interval also captures the equivalence threshold, the remedy would not be declared successful at this point in time under the equivalence testing framework. This could lead to any number of response actions, including continue monitoring in the future, or collecting more data contemporaneously to refine the imprecise confidence interval to better understand the situation.

Situation B shows the case where both frameworks would result in the same outcome, that the site was not statistically different from background condition and that any differences that may exist are less than pre-specified definition of equivalence—the remedy would be declared successful at this time.

Situation C is an ironic condition where under the classical null framework, too much precision can be a problem in that the mean for the site is statistically greater than that for the reference area, but the difference is biologically indifferent. In this case the remedy would be declared successful under bioequivalence and unsuccessful under the traditional no-difference null framework. This situation illustrates that the traditional null framework incentivizes small sample sizes, and high variance to insure against an unlucky outcome due to a too-narrow confidence interval. The test of equivalence avoids this situation entirely in that narrow intervals are required to demonstrate success and these are the result of a well implemented powerful study. High quality science is incentivized by correctly stating the null hypothesis with a pre-defined biologically meaningful threshold for equivalence.

Situation D may be the most common outcome early in a monitoring program where the data suggest that the true mean may be within the equivalence region, but the interval is too wide to demonstrate equivalence statistically. In this situation it may be to the responsible parties advantage to simply collect more data in efforts to narrow the confidence interval in hopes of achieving statistical evidence of equivalence. This opens the door for an adaptive monitoring approach where smaller sample sizes could be considered early in the monitoring program until evidence suggests that site conditions may be approaching equivalence, at which time it would be up to responsible parties to increase the power of the design with more sampling effort to develop statistically significant evidence of success.

Summary

The discussion above provides general rationale for selection of the direction of the null hypothesis for demonstrating similarity to reference condition. The EPA is recommending the equivalence testing approach because it eliminates the need for inappropriate interpretation of nonsignificant statistical tests, stimulates discussion of what will be considered a biologically meaningful difference before sampling and analysis, and because the method incentivizes sound high quality science with powerful statistical testing methods.

Within the monitoring program, other analyses will be conducted that are unrelated to reference area comparisons. These will take on more familiar statistical frameworks and are not described in detail here. It is expected that for all media first order decay functions will be fit to samples to estimate post-remedial recovery rates. It is also expected that for each medium, baseline and post-remedial data will be compared to demonstrate the short-term effects of the remedy, immediately after completion.

Statistical procedures for evaluating the effectiveness of the remedy will take on the traditional null hypothesis framework because the questions are naturally and appropriately framed with a null hypothesis of no change (equality of metrics) with rejection of the traditional null signifying remedial effectiveness.