EPA Comments added to Draft Work Plan (Text, Table 8 and Appendices) and Geosyntec 10/4/17 Memorandum

Prepared for:

United States Environmental Protection Agency, Region 10
1200 Sixth Avenue, Suite 900
Seattle, Washington 98101

and

Portland Harbor Pre-RD Group
Portland Oregon

DRAFT WORK PLAN

PORTLAND HARBOR PRE-REMEDIAL DESIGN INVESTIGATION STUDIES PORTLAND HARBOR SUPERFUND SITE

Prepared by:

Geosyntec consultants
engineers | scientists | innovators
520 Pike Street, Suite 1375
Seattle, Washington 98101

Project Number: PNG0767

22 September 2017

Confidential – For Purposes of Settlement Negotiations Only

TABLE OF CONTENTS

Commented [A1]: The Pre-RD Group Draft Work Plan (Work Plan) includes findings and conclusions from independent investigations that were not conducted under EPA oversight. EPA does not agree with any statement or position taken in the Work Plan that is in any way inconsistent or contrary to the ROD and particularly any response found in the Responsiveness Summary.

LIST OF TABLES

LIST OF FIGURES

Commented [A2]: General comments on the Work Plan tables are provided below. Additional comments are included on Table 8.

Commented [A3]: Tables 2 and 3. EPA would prefer that ROD Tables 17 and 21 be used rather than re-creating these tables to avoid transcription errors, omitted values, and maintain consistency with the ROD in the event of any subsequent errata.

Commented [A4]: These tables should be updated to incorporate EPA's comments on the work plan text.

Commented [A5]: The information provided in this table is insufficient for documenting the rationale for each new core location. The sampling and analysis plan to be prepared should include graphical presentations (e.g., plan view figures, cross-sections, 3D visualizations) of the existing core data and the proposed sediment core locations to illustrate how the proposed core locations were selected.

Commented [A6]: This table should be updated to incorporate EPA's comments on the work plan text. Additional comments on Table 8 are provided on the table.

Commented [A7]: See EPA attachment for EPA Sampling Plan Figure 2 "Surface Sediment Transect Locations for Baseline and Long-Term Monitoring". This figure shows location of 424 unbiased baseline monitoring surface sediment sample locations selected by EPA.

LIST OF APPENDICES

Commented [A8]: Several of the data collection efforts summarized in Appendix A were independent investigations that were not conducted under EPA- approved work plans. Inclusion of the findings and conclusions drawn from these studies, as stated in the work plan and Appendix A, should not be construed as an endorsement or acceptance by EPA. EPA does not agree with any statement or position taken in the Work Plan that is in any way inconsistent or contrary to the ROD and particularly any response found in the Responsiveness Summary.

Commented [A9]: See EPA attachment for Appendix A of EPA Sampling Plan "Statistical Basis for Sampling Approach"

LIST OF ACRONYMS AND ABBREVIATIONS

μm	micrometer
95UCL	95% upper confidence limit
AECOM	AECOM Technical Services
Alt F Mod	Alternative F Modified
ARARs	applicable or relevant and appropriate requirements
ASAOC	Administrative Settlement and Agreement Order on Consent
ASTM	American Society of Testing Materials
BAZ	biologically active zone
BERA	baseline ecological risk assessment
BHHRA	baseline human health risk assessment
bml	below mudline
BMPs	best management practices
CDM Smith	CDM Smith, Inc.
cfs	cubic feet per second
COCs	contaminants of concern
COPCs	contaminants of potential concern
CRD	Columbia River Datum
CSM	Conceptual Site Model
cy	cubic yards
D/F	dioxins/furans
D/U Reach	the Downtown Reach and the Upstream Reach
DDx	sum of dichlorodiphenyltrichloroethane and its derivatives
DQMP	Data Quality Monitoring Plan
DQOs	data quality objectives
DUOs	data use objectives
EIM	Environmental Information Management
ENR	enhanced natural recovery
EPA	United States Environmental Protection Agency
ERDC	Engineer Research and Development Center

ESD	Explanation of Significant Differences
FC	Field Coordinator
FWM	food web model
Geosyntec	Geosyntec Consultants, Inc.
Germano	Germano and Associates
GSI	GSI Water Solutions, Inc.
Hart Crowser	Hart Crowser, Inc.
HSP	Health and Safety Plan
ICs	Institutional Controls
L	liter
LSS	Legacy Site Services
LWG	Lower Willamette Group
MDL	method detection limit
MNA	monitored natural attenuation
MNR	monitored natural recovery
NELAP	National Environmental Laboratory Accreditation Program
NPL	National Priorities List
NWIS	National Weather Information System
ODEQ	Oregon Department of Environmental Quality
ORP	oxidation-reduction potential
OSHA	Occupational Safety and Health Administration
PAHs	polycyclic aromatic hydrocarbons
PCBs	polychlorinated biphenyls
PCI	Participation and Common Interest
PDI	pre-remedial design investigation
PES	polyethersulfone
PHSS	Portland Harbor Superfund Site
PII	Personal Identification Information
Pre-RD	Pre-Remedial Design
PRP	potentially responsible party

PTW	Principal Threat Waste
PVC	polyvinyl chloride
QA	quality assurance
QAPPs	quality assurance project plans
QC	quality control
RALs	Remedial Action Levels
RAOs	Remedial Action Objectives
RI/FS	Remedial Investigation/Feasibility Study
RM	river mile
ROD	Record of Decision
SAP	Sampling and Analysis Plan
Site	Portland Harbor Superfund Site
SMA	sediment management area
SMB	smallmouth bass
SOW	Statement of Work
SWACs	surface weighted average sediment concentrations
TOC	total organic carbon
TSS	total suspended solids
TZW	transition zone water
USACE	United States Army Corps of Engineers
USGS	United States Geological Survey

1. INTRODUCTION

The Record of Decision (ROD) described a post-ROD sampling effort for the Portland Harbor Superfund Site (Site or PHSS) to delineate and better refine the sediment management area (SMA) footprints, refine the Conceptual Site Model (CSM), determine baseline conditions, and support remedial design (United States Environmental Protection Agency [EPA] 2017).

The pre-remedial design investigation (PDI) study is intended to meet many of the ROD objectives and to support the allocation process. Table 1 lists the data that will be collected to satisfy these objectives. This PDI scope of work focuses on Site-wide studies that will more thoroughly define the remedial actions to support the 30% design, including refining the surface delineation of the SMAs, reevaluating technology assignments throughout the Site, evaluating the horizontal and vertical extent of the dredging and capping areas; and refining the extent of active versus passive remedial actions at the Site. It also re-baselines the Site and achieves many of the baseline dataset objectives outlined in the ROD. The Work Plan does not include SMA-specific design-level sampling, nor source control evaluations, which could be conducted during future remedial design. The data collected as part of this scope of work are not intended to provide final conclusions for the Site. Additional data collection as a part of separate scopes of work will be needed to support future remedial design efforts.

Commented [A10]: EPA's objectives differ from many of those identified by the Pre-RD Group in this work plan. Some of Pre-RD objectives are not ROD required or consistent with EPA's objectives and planned evaluations.

Commented [A11]: Technology assignments can be evaluated during design phase but not with limited data to be collected under the PDI.

This Work Plan, prepared by Geosyntec Consultants, Inc. (Geosyntec) and AECOM Technical Services (AECOM), is a focused and foundational step in what will likely be a multi-phase effort to bring current the collection of data over the past 15 years. It provides an overview of studies that will be prepared for pre-remedial design investigation at the PHSS located in Portland, Oregon. The Work described in this Work Plan will be conducted by a group of industrial parties called the Pre-Remedial Design Investigation (Pre-RD) Group. This Work Plan was prepared in general accordance with the Superfund Remedial Design and Remedial Action Guidance document (EPA 1985).

EPA and the Pre-RD Group recognize that the data gathered in PDI is not the complete data set for remedy design/implementation and that EPA's review of the data reports or data analysis may include an assessment as to whether the data relied on is sufficient to support the certain PDI evaluations, refinements, recalculations and updates. EPA reserves the right to review all submittals prepared under the Work Plan.

1.1 Site Description

The Site extends from river mile (RM) 1.9 near the mouth of the Willamette River to RM 11.8 (Figure 1). The Willamette River is a dynamic waterbody that originates within Oregon in the Cascade Mountain Range and flows approximately 187 miles north to its confluence with the Columbia River. Its average flow rate is 33,000 cubic feet per second (cfs), with high season rates of 200,000 cfs or higher (EPA 2016a).

The Site includes a water-dependent, highly industrialized area, which contains a multitude of facilities and both private and municipal outfalls. Land use along the Lower Willamette River in the Portland Harbor includes marine terminals, manufacturing and other commercial and municipal operations, and public facilities, parks, and open spaces (EPA 2016a). The Downtown Reach, which includes the urbanized area of downtown Portland, is defined by EPA as extending from RM 11.8 to RM 16.6. EPA defines the Upriver~~stream~~ Reach extending from RM 16.6 to RM 28.4. For purposes of the PDI, the Work Plan is focusing on RM 11.8 to RM 20 for data collection to assess incoming contaminant loads to the Site. ~~Collectively, the Downtown Reach and Upstream Reach are referred to as the Upriver Area for purposes of this PDI.~~

Commented [A12]: To be consistent with the ROD, terminology throughout the work plan should be revised to use "Upriver Reach" (not Upstream Reach) to describe the portion of the Willamette River from RM 16.6 to RM 28.4. Samples from the Upriver Reach should be obtained from the entire reach and not be limited to RM 16.6 to RM 20.

Commented [A13]: Data from Downtown Reach and Upriver Reach to be evaluated separately so no need to introduce "Upriver Area" which is inconsistent with the ROD.

The shorelines along most of the Portland Harbor area have been developed for industrial, commercial, and municipal operations; the Portland Harbor area serves as a major shipping route for containerized and bulk cargo. In addition, the Portland Harbor area has historically received, and currently receives, discharges from industrial and municipal sources including point- and non-point sources that discharge to the Lower Willamette River. Common shoreline features within the harbor include constructed bulkheads, piers, wharves, buildings extending over the water, and steeply-sloped banks armored with riprap or other fill materials. Site background and other site characteristics are described in detail in the Final Remedial Investigation (RI) Report (EPA 2016a).

On 1 December 2000, the Site was listed on the National Priorities List (NPL) by EPA mainly due to concerns about contamination in the sediments and the potential risks to human health and the environment from consuming fish. The most widespread contaminants found at the Site include, but are not limited to, polychlorinated biphenyls (PCBs), polycyclic aromatic hydrocarbons (PAHs), dichlorodiphenyltrichloroethane and its derivatives (DDx), and dioxins/furans (D/F). A remedial investigation and feasibility study (RI/FS) was initiated in 2001 by a small subset of potentially responsible parties

(PRPs) known as the Lower Willamette Group (LWG), and completed by EPA in 20167 (EPA 2016a and 2016b7).

In June 2017, the Pre-RD Group developed and offered a pre-remedial design scope of work focused on defining current conditions for specific media and refining delineation of active remedial areas to EPA. Over the next two months, EPA and the Pre-RD Group held several scoping meetings to negotiate study objectives, data collection activities, and data interpretation; the scope was expanded to address several baseline study elements desired by EPA. In September 2017, the EPA entered into an Administrative Settlement and Agreement Order on Consent (ASAOC) with the Pre-RD Group to conduct the agreed upon Work at the Site. This Work Plan supports the Statement of Work (SOW) which is an attachment to the ASAOC, and describes the specific field investigation activities, data analyses, schedule, and deliverables for the PDI. The Work Plan is included as an attachment to the SOW.

1.2 Remedy of Record

The remedy selected in the ROD (EPA 2017), called Alternative F Modified (Alt F Mod), identified 394 acres of engineered remediation with a combination of remedial technologies (Figure 2). The remedy includes 365.4 acres of capping and dredging contaminated sediment above Remedial Action Levels (RALs) and 28.2 acres of enhanced natural recovery (ENR) within the Site. The RALs are listed in ROD Table 21 2. Alt F Mod addresses all areas where contaminant concentrations exceed the cleanup levels (see ROD Table 173) through a combination of dredging, capping, ENR, monitored natural recovery (MNR), and Institutional Controls (ICs). The ROD indicates that EPA expects 215.2 acres of sediment will be dredged to varying depths and 140.1 acres will be capped, or partially dredged and capped. Additionally, 23,305 lineal feet of riverbank are assumed to be excavated and covered with either an augmented reactive cap or an engineered cap using beach mix or vegetation after excavating. Under Alt F Mod, approximately 3,017,000 cubic yards (cy) of contaminated sediment and 123,000 cy of soil were estimated by EPA to be removed and transported to off-site disposal facilities. About 1,774 acres are designated for MNR (EPA 2017).

The SMAs represent areas which EPA considered to have contaminant concentrations in surface sediment where natural recovery is not occurring or is not likely to be effective in reducing concentrations of contaminants of concern (COCs) within a reasonable time frame (EPA 2017). Additionally, EPA used the presence of Principal Threat Waste

(PTW) as defined in its FS (EPA 2017), and used in-situ treatment areas for PTW to delineate SMAs.

The ROD states that the in-river construction duration for Alt F Mod will be approximately 13 years at a pre-engineering estimated cost of $1.7 billion (non-discounted). The remedy will likely change somewhat during the remedial design and be adapted during the multi-year construction process. Changes to the remedy will be documented using a technical memorandum in the Administrative Record, an Explanation of Significant Differences (ESD), or ROD amendment (EPA 2017).

The remedial actions identified in the ROD address nine narrative remedial action objectives (RAOs) that EPA developed for the Site for environmental media of interest and exposure pathways, including exposure routes and receptors. The ROD defined numeric, concentration-based cleanup levels to achieve these RAOs for each exposure route (and tissue targets for seafood consumption RAOs). The cleanup levels considered conservative risk assessments, applicable or relevant and appropriate requirements (ARARs-based), and background concentrations (background-based). Achieving the RAOs relies on the remedy's ability to meet cleanup levels or tissue targets. Fish tissue targets will be used to update fish advisories, assess whether the remedy will achieve RAOs, make adjustments to best management practices (BMPs); their uses will be further defined in the monitoring plans. ROD Table 173 presents the COCs for the Site and respective cleanup levels by media ~~as presented in the ROD~~. Site-specific cleanup levels were developed for each RAO for the following media: sediment (including beaches) and riverbank soil, surface water, and groundwater (EPA 2017).

1.3 Pre-Remedial Design Data Use Objectives

Commented [A14]: EPA's objectives differ from many of those identified by the Pre-RD Group in this work plan. Some of Pre-RD objectives are not ROD required or consistent with EPA's objectives and planned evaluations.

The Pre-RD Group proposes to conduct a comprehensive 2018 synoptic sampling program of surface sediment, select sediment cores, fish tissue, surface water, background porewater, and bathymetry/fish tracking studies. These investigation activities are focused on achieving the following goals:

1. Implement investigation baseline sampling to update existing site-wide data;
2. Gather data to be used as part of a long-term trend analysis;
3. Define more clearly the remedial actions that will be performed at the Site, including further delineation of SMAs, reevaluation of technology

assignments throughout the Site, refinement of the horizontal and vertical extent of the dredging and capping areas and the scope and extent of active remedial actions at the Site;

4. Collect data to facilitate completion of the third-party allocation amongst PRPs;
5. Collect additional data regarding upstream conditions and contaminant loading into the Site; and
6. Update the baseline human health risk and refine understanding of the food web model (FWM) using new 2018 data.

Table 1 lists the data that will be collected to achieve these goals. The Pre-RD 2017/2018 data will be used to determine current SWACs, human health risks, and background concentrations associated with both upriver and Site conditions. This sampling program provides a technically sound and robust balance of: (i) near-term initiation of field work; (ii) prioritizing field data collection of studies that provide informative updates to the Site baseline; and (iii) supporting the Participation and Common Interest (PCI) Allocation Team's need to reduce remedy uncertainty.

This Work is further supported by the ROD and the goal of considering new data. As stated in Section 2.7.3 of the ROD Responsiveness Summary (EPA 2017), "*EPA agrees with the importance of considering new data during decision making and that decisions should have built in flexibility to accommodate an updated understanding of site conditions. However, it is important to have a representative data set that establishes 'baseline conditions' prior to initiating a response action.*" And "*EPA expects remedial footprints to be refined based on data collected during remedial design.*" Also, "*Pre-design sampling will be used to ensure that the natural recovery is factored into the design and implementation of the sediment remedy and post construction monitoring will be used to evaluate natural recovery following remedy implementation.*"

Commented [A15]: Technology assignments can be evaluated during design phase but not with limited data to be collected under the PDI.

Commented [A16]: Achieving this goal is not the intent of the PDI. Concentrations that were estimated using the FWM should not be recalculated.

Commented [A17]: Not relevant for this phase of sampling.

2. SITE BACKGROUND

This section provides a summary of Site physical conditions, risks above protective levels, COCs, and Site investigations completed after the data cut-off for completion of the RI/FS.

2.1 Physical Conditions

The lower Reach of the Willamette River extending from RM 0 to approximately RM 26.5 is a shallow segment that is tidally influenced with river flow reversals occurring during low-flow periods as far upstream as RM 15. The portion of the river where the federal navigation channel is maintained at -40 feet Columbia River Datum (CRD) defines Portland Harbor and extends upstream from the Columbia River (RM 0) to the Broadway Bridge (RM 11.7) (EPA 2017). The Willamette River channel, from the Broadway Bridge (RM 11.6) to the mouth (RM 0), varies in width from 600 to 1,900 feet.

The high tide can influence Willamette River levels by up to 3 feet in Portland Harbor when the river is at a low stage. Tidal fluctuations during low river stage can result in short-term flow reversals (i.e., upstream flow) in late summer to early fall. Low water typically occurs during the regional dry season from August to November. The winter (November to March) river stage is relatively high, but variable, due to short-term changes in precipitation levels in the Willamette basin. Finally, a distinct and persistent period of relative high water occurs from late May through June when the Willamette River flow into the Columbia River are slowed during the spring freshet by the high-water stage in the Columbia River (EPA 2016a).

Factors controlling river flow dynamics, sediment deposition and erosion, and riverbed character appear to be the river cross-sectional area, thalweg location, and navigation channel width. The upstream boundary of the Site to Willamette Falls is narrower, more confined by bedrock outcrops, and faster flowing than the Portland Harbor Reach. The river widens as it enters the Site and becomes increasingly depositional, most notably in the western portion of the river, until RM 7. From approximately RM 5 to RM 7, the river and navigation channel narrow; this Reach is dominated by higher energy environments with little deposition. From RM 5 to approximately RM 2, the river widens again and becomes depositional, particularly in the eastern portion.

Long-term net sedimentation rates based on time-series bathymetric surveys show patterns of general shoaling in wider reaches. Wide areas of deposition occur in the channel and along channel margins in the broader sections of the river (RM 1.5 to 3 [eastern margin], RM 4 to 5, and RM 7 to 10). These areas are known to be long-term sediment accumulation areas based on historical dredging records. Shoaling is the dominant change observed, with 26% of the riverbed surveyed showing net accretion (January 2002 to January 2009) exceeding 1 foot (30 centimeters), whereas net erosion exceeding 1 foot is noted in only 5% of the riverbed overall.

Downstream of the Site, the river narrows as it turns and converges with the Columbia River. The Multnomah Channel exits at RM 3, reducing direct discharge to the Columbia River. From 1973 through 2007, average annual mean flow in the Willamette River was approximately 33,800 cfs at the Morrison Bridge (near RM 12.8) (http://waterdata.usgs.gov/or/nwis/sw) (EPA 2016a).

2.2 Summary of Site Contaminants and Risks

The baseline human health risk assessment (BHHRA, Kennedy/Jenks 2013a) and the baseline ecological risk assessment (BERA, Windward 2013) concluded that contamination within the Site poses potential unacceptable risks to human health and the environment from numerous contaminants of potential concern (COPCs) in surface water, groundwater, sediment, and fish tissue. The RI/FS reduced the list of COPCs to COCs, as presented in Table 17 of the ROD.

As stated in Section 10.1 of the ROD, "The COCs used to define the SMA boundaries encompassed most of the spatial extent of contaminants posing the majority of the risks as identified in the baseline risk assessments. However, since it is difficult to design a range of alternatives for multiple COCs that have different distributions in various media throughout the Site, the FS alternatives were developed using COCs that were the most widespread and posed the greatest risk, called focused COCs." "*These focused COCs, were developed by evaluating colocation of all COCs, their toxicity, and significance in the risk assessments, as well as other factors outlined in the RI.*"

The focused COCs are:

- PCBs;
- DDx;

- total PAHs; and
- Dioxins/Furans.

The remedial footprint of the focused COCs encompasses the majority of the COCs at the Site (EPA 2017). To establish 2017/2018 baseline conditions at the Site, this Work will develop a representative data set by including the full list of ~~sediment~~ media-specific COCs presented in ROD Table 17~~3~~ for surface sediment, surface water, and fish tissue for the initial round of sampling. The data will also be evaluated for purpose of potentially reducing the list of COCs for future monitoring events.

The environmental media contaminated by Site-related contaminants include surface sediment (0 to 30 centimeters depth below mudline [bml]), subsurface sediment (>30 bml), suspended sediment, surface water, groundwater, biota, and riverbanks. The surface sediment sample interval (0 to 30 centimeters) is the point of compliance for the RAOs and cleanup levels, as it represents the biologically active zone (BAZ) and the active mixing zone depth, which is the portion of the sediment column that has the potential to be disturbed or transported under typical conditions (EPA 2017).

Several locations within the Site have relatively high surface sediment concentrations of more than one contaminant. Overall, the patterns of contaminant distribution are as follows:

- Nearshore areas have greater sediment contaminant concentrations than sediments offshore and in the navigation channel;
- Subsurface sediments have greater organic contaminant concentrations than surface sediments;
- Some contaminants, such as DDx and PAHs, have higher concentrations and are more commonly found in the downstream portion of the Site;
- Sediment grain size and concentrations of certain metals are correlated; and
- Multiple contaminants are co-occurring, that is they are co-located with other COCs with respect to horizontal and vertical distribution in the river/sediments (EPA 2016a).

2.3 Summary of Data Collected Since the RI/FS

From 2008 to 2016, eight environmental studies relevant to this Work Plan have been conducted since the RI/FS data were collected. Environmental media included surface sediment grabs, subsurface sediment cores, and smallmouth bass (SMB) fish tissue samples for various COCs. Several studies focused on mainly PCBs. The eight studies are summarized in Appendix A and include the following:

Commented [A18]: Several of the data collection efforts listed below were independent investigations that were not conducted under EPA- approved work plans. Inclusion of the findings and conclusions drawn from these studies, as stated in the work plan and Appendix A, should not be construed as an endorsement or acceptance by EPA.

- Downtown Portland Sediment Characterization Phase I and II (GSI Water Solutions, Inc. [GSI] and Hart Crowser, Inc. [Hart Crowser] 2010);
- Smallmouth Bass Tissue Sampling (GSI 2011);
- Smallmouth Bass Tissue Study (Kennedy/Jenks 2013b);
- Sediment Profile Imaging (Germano and Associates [Germano] 2014);
- Final Supplemental RI/FS Study, River Mile 11 East (GSI 2014);
- Sediment Sampling Data Report (Kleinfelder 2015);
- Concentrations and Character of PAH in Sediments in Area of River Miles 5 to 6, (NewFields 2016); and
- Sediment Sampling Data Report, Swan Island Lagoon (Geosyntec 2016).

Conclusions provided in recent~~Recent~~ surface sediment sampling studies conducted by responsible party groups~~results compared to RI/FS data~~ indicate that newly deposited sediments are covering and/or mixing with the older surface sediments and that natural recovery is occurring in many areas of the Site (Geosyntec 2016; Germano 2014; Henderson 2015). The Oregon Department of Environmental Quality (ODEQ) Downtown Reach investigation found that COCs were much lower than those found in the Site and ODEQ believes the Downtown Reach is not a significant ongoing upstream source (ODEQ 2011).

Analysis of SMB tissue sampling results conducted by responsible parties found that the mean 2012 tissue concentrations were lower than the mean concentrations of the combined 2002 and 2007 SMB data that were used in the RI/FS and statistical comparisons of the two data sets on a Study Area-wide scale. Total PCB congener concentrations in whole body SMB tissue show a statistically significant decrease from the 2002 and 2007 data (Kennedy/Jenks 2013b, Legacy Site Services [LSS] 2015). The

2012 SMB data support that natural recovery is occurring on a system-wide scale. ~~Among these studies, data that have been properly validated will be incorporated into the project database.~~Data obtained from independent investigations will be approved by EPA prior to incorporating these data into the project database.

3. SCOPE OF WORK TASKS

Each task and field study included in the Work is briefly summarized below and in Table 3. Project goals for each component of the Work are provided in Table 4. Many of the field data will serve multiple data use objectives (DUOs), as shown in Table 5. Table 6 lists the Work studies including media, sample counts, and analyses. Figures 4, 5, and 6 show the approximate sampling locations for sediment, tissue, and surface water, respectively. The surface sediment sample plan shown in Figure 4 reflects an earlier draft of the sampling plan[1]. The final sample design for surface sediment will be depicted in the Sampling and Analysis Plan (SAP) for sediment sampling and follow the approach described in Appendix B.

3.1 Task 1: Quality Assurance Project Plans (QAPP, SAP, DQMP)

Details regarding these sampling efforts will be further refined in quality assurance project plans (QAPPs) to be prepared following this work plan. These documents will include a QAPP, SAP, and Data Quality Monitoring Plan (DQMP), and will be prepared in accordance with EPA standards and previously-approved RI documents for the Site. A health and safety plan (HASP) will also be prepared. These PDI project plans will be focused and targeted plans, or addendums; they will appropriately reference the RI plans as source documents and then describe and document any changes relevant to the PDI scope of work.

The QAPP will address all sample collection activities as well as sample analysis and data handling regarding the Work. The QAPP will be developed in accordance with *EPA Requirements for Quality Assurance Project Plans, QA/R-5, EPA/240/B-01/003* (March 2001, reissued May 2006); *Guidance for Quality Assurance Project Plans.*, QA/G-5, EPA/240/R 02/009 (December 2002); and *Uniform Federal Policy for Quality Assurance Project Plans*, Parts 1 3, EPA/505/B-04/900A though 900C (March 2005).

---

[1] The earlier plan (as shown in Figure 4) indicated 345 unbiased surface sediment samples. The revised plan will consist of 428 unbiased samples, placed randomly according to the procedures detailed in Appendix B. The final arrangement of the 428 unbiased sediment samples and additional 212 surface sediment samples for SMA delineation will be provided in the SAP.

A FSP, incorporated as a subsection of the QAPP, will provide objectives and minimum sampling requirements. The FSP will include guidelines for sediment, surface water, porewater, and SMB tissue sampling.

The DQMP will include analytical laboratory Quality Assurance Plans and internal data validation procedures, along with standard quality assurance/quality control (QA/QC) procedures. All chemical analysis will be performed by a National Environmental Laboratory Accreditation Program (NELAP) accredited lab. Analytical laboratories will conduct QA/QC as detailed by their respective laboratory quality control procedures and manuals. Standard method and operating procedures, calibration, internal QC, and preventative maintenance are examples of QA/QC processes to improve accuracy and precision. All laboratory QC analysis results will be reported with the final data report. Failure of any QC samples to meet QC criteria will be noted and the data which corresponds to these samples will be adequately qualified in the final report. Records of QA/QC will be maintained for review as needed. Field QA/QC procedure will include the collection of field duplicate samples which will be analyzed for the same set of physical and chemical analyses, along with equipment blank and field blank samples as appropriate. It will be the responsibility of the analytical laboratories to provide accurate results in electronic and hard copy formats, along with Level III Data Validation packages consistent with laboratory Quality Assurance Plans. Data provided by the laboratory will undergo data validation by a third party. Data validation is analyte- and sample-specific, and extends the evaluation of data beyond method, procedural, or contractual compliance to determine the analytical quality of a data set. Data will be validated and qualified as outlined in project specific QAPPs.

The HASP will describe all activities to be performed to protect onsite personnel and area residents from physical, chemical, and all other hazards posed by the Work. The HSP will be developed in accordance with EPA's Emergency Responder Health and Safety and Occupational Safety and Health Administration (OSHA) requirements under 29 C.F.R. §§ 1910 and 1926.

A few notable details relevant to these Project Plans include:

- Surface sediment will be collected from the 0- to 30-centimeter interval, which is the point of compliance throughout the Site and incorporates the BAZ and the active mixing zone depth;

- Chemical analyses for surface sediment, surface water, and fish tissue will include the full list of COCs for each media [excluding PAHs in tissue] as presented in Table 6;
- Chemical analyses for subsurface sediment will include the focused COCs (PCBs, PAHs, D/F, and DDx) which have corresponding RALs, and the additional contaminants on ROD Table 21 which have PTW thresholds;
- A DUO for surface sediment and tissue is to establish baseline Site-wide, rolling river mile (one-side), -and SDU segment-wide 95% upper confidence limit on the mean (95UCL) concentrations of theor SWACs for the media-specificfocused COCs;
- The sampling program of synoptically-collected surface sediment, SMB tissue, and surface water data collected from the Site is a substantial baseline study effort (further described in Section 3);
- Sediment and SMB tissue data collected from the Upriver Area will be evenly distributed between the Downtown Reach and the Upstream Reach; and sampling locations will target areas of the sediment bed likely to influence downstream contaminant concentrationsfine-grained sediment; the PDI objective is to evaluate background concentrations of COCs coming into the Site; and establish upstream sediment bed concentrations and SMB tissue concentrations in the Downtown and Upriver Reaches for comparison to site concentrations using an equivalency analysis;
- The home range of SMB will be evaluated over a year-long study in collaboration with the United States Army Corps of Engineers (USACE);
- The Site-wide bathymetry survey is intended to refresh and update the surface bed elevations to current conditions and fill-in no coverage areas (especially nearshore) to support the 30% design; and
- Background concentrations of naturally-occurring metals in porewater will be evaluated for arsenic and manganese; sampling locations will be developed in collaboration with EPA.

Commented [A19]: Evaluation of background concentrations is beyond the scope of this phase of work. Primary purpose for data collection in Downtown and Upriver Reaches is for comparison to site concentrations using an equivalency analysis.

3.2 Task 2: Sampling and Analysis

The DUOs, sampling design, and analytical methods for the PDI are discussed below. The PDI includes the following tasks involving several multi-media sampling and analytical testing activities:

- Site-wide bathymetry survey;
- Surface sediment sampling;
- Fish tissue sampling;
- Surface water sampling;
- Sediment coring;
- Fish tracking study;
- Camera survey;
- Downtown/Upstream Reaches (Upriver) background study;
- Background porewater study; and
- Reporting

Commented [A20]: Evaluation of background concentrations is beyond the scope of this phase of work. Primary purpose for data collection in Downtown and Upriver Reaches is for comparison to site concentrations using an equivalency analysis.

3.2.1 Bathymetry Survey

A bank-to-bank bathymetry survey throughout the Site will document current bed elevations relative to the remedial technology assignment requirements (per the ROD decision tree) and to assess changes in elevation/sedimentation over the past 15 years and to evaluate mudline elevations. Multibeam sonar will be used to collect high-resolution data with up to 100% coverage of the riverbed. The Site-wide multibeam bathymetry survey will be supplemented with lead-line measurements along the shoreline banks and difficult-to-access areas for better coverage than provided by multi-beam alone.

This survey will produce an up-to-date bathymetric data set with a high level of detail and accuracy. The multibeam bathymetric data will be used to create a digital terrain model of the riverbed morphology from which hill-shade images will be generated. Results may also serve as a line of evidence relevant for the evaluation of riverbed slope conditions, natural recovery, and bed stability (e.g., erosional versus depositional areas).

The most recent bathymetry survey was performed in 2002. The new bathymetry data will also be used to help identify target areas for surface sediment sampling, refine the elevation clearances for dredging and capping, and adjust the estimated dredge volumes (to reduce uncertainty for allocation associated with the extent of the active remedial footprint and remedial technologies assigned to them). The anticipated schedule for the bathymetry survey is the end of 2017 (December 2017).

3.2.2 Surface Sediment Sampling

Many of the surface sediment data for the Site are over 10 years old. The goals of new surface sediment data collection are to: re-baseline the river bed to establish current conditions and SWACs, refine the active remedial SMA footprints, and evaluate natural recovery changes since 2004. Because Portland Harbor is part of a dynamic river system, current concentrations for all COCs are expected to be different than the data set used in the RI. The synoptic surface sediment sampling, fish tissue, and surface water samples (discussed below) data will provide an empirical ~~and statistically valid~~ dataset for ~~re-baselining the river and~~ evaluating the CSM.

~~Nine~~ River Segments

To be consistent with and support the decision framework in the ROD, data should be evaluated on a Sitewide, rolling RM 1-side, and SDU basis for a total of up to 21 segments (10 river mile segments with east and west side, plus Swan Island Lagoon).~~Previous analyses (Wolf 2015a; Wolf 2015b; and Toll et al. 2015) found that the river is spatially and chemically unique and can be properly stratified into five river segments including the Swan Island Lagoon, each about 2 to 3 miles long. The Pre-RD Group will divide the Site into five segments spanning the length of the Site for evaluation of surface concentrations and SWACs, based on physical features, river flow dynamics, contaminant distributions, and fish home ranges (Figure 3). Four of the five segments will be further divided down the center of the navigation channel into two segments each, east and west, thereby forming eight segments. A ninth segment is between RM 8 and RM 9 at Swan Island Lagoon. The nine segments from upstream to downstream (as shown in Figure 3) are: Segment 1 E&W (RM 11.8 to 9), Segment 2 E&W (RM 9 to 7.5), Segment 3 E&W (RM 7.5 to 5), Segment 4 E&W (RM 5 to 1.9), and Segment 5 (Swan Island Lagoon).~~

Formatted: Highlight

Commented [A21]: The primary purpose is to collect an unbiased baseline data set to enable comparison to future unbiased sediment sampling data collected under long-term monitoring to evaluate natural recovery.

A comparison of 2018 surface sediment data from reoccupied 2004 sample locations with previously collected 2004 samples is statistically invalid because 108 of 428 surface sample locations are biased samples. A valid statistical approach for evaluating MNR for the Site must be based on a comparison of unbiased surface sediment data (i.e., 2018 baseline data and future sediment sampling data collected under long-term monitoring) to evaluate natural recovery.

Commented [A22]: The work plan and Table 8 states that the surface sediment data will be used to evaluate natural recovery changes since 2004. Although a comparison of 2018 surface sediment data reoccupying 2004 sample locations (biased data) with previously collected 2004 samples may indicate natural recovery trends, it is statistically invalid. A valid statistical approach for evaluating MNR for the Site must be based on a comparison of unbiased surface sediment data (i.e., 2018 baseline data and future sediment sampling data collected under long-term monitoring) to evaluate natural recovery.

Commented [A23]: The previous studies cited (Wolf 2015a; Wolf 2015b; and Toll et al. 2015) were conducted by responsible parties independent of EPA oversight.

The nine segments defined in this section are inconsistent with the ROD and segments should follow the revised text.

If the parties are unwilling to conduct the analysis as per the ROD, EPA will be conducting this to meet requirements of the ROD and move cleanup forward.

A key use of the new data (along with the fish tracking results and determination of SMB home ranges) will be to confirm the representativeness of these segment delineations, then estimate SWACs Site-wide and other spatial scales.

GeoStatistics

Several stratified/random/equal allocation methods of statistical analysis were used to estimate the appropriate sample size within the Site needed to satisfy the DUOs described above. A summary of the geostatistical analysis, approach, and findings is summarized in Appendix B. As detailed in Appendix B, the sample count was determined by considering the number of surface sediment samples needed in each segment to maintain or improve upon the level of variability in the SWACs generated using the 2004 data, and, in most areas and assessment segments, enable the design to statistically detect differences (α = 0.05) between 2004 SWACs and current SWAC estimates with an approximate 80% level of statistical power. Based on this analysis, an estimated 640 discrete surface sediment samples are needed to yield a statistically-robust new data set for calculating SWACs, with 428 unbiased locations within the Site and 212 additional samples specifically located for accurate SMA delineation. This new dataset will replace older RI/FS data for the purposes of refining the SMA footprints and technology selections described in the ROD.

Sampling Methods

Surface grab samples will be collected with a hydraulic power grab sampler from 640 locations (Figure 4). The surface sediment sample locations presented in Figure 4 reflects an earlier draft of the sampling plan, are shown in this document for example purposes only, and do not reflect final placement and numbers of the final design. The final sample design for surface sediment sampling will be depicted in the SAP and follow the approach detailed in Appendix B. The final sampling effort will include approximately 108 stations co-located with previously analyzed stations in the RI/FS. The hydraulic power grab sampler will collect sediment from the upper 0 to 30 centimeters of sediment at three sampling points at each sample location (without adjusting vessel location), and homogenized into a three-point composite sample for chemical analysis of the full list of sediment COCs, total organic carbon (TOC), and grain size. The three-point composite sample will be collected within a relatively small footprint around the anchored sampling vessel. For example, grab #1 will be deployed, accepted, and processed on the deck of the vessel. The sampling vessel's overhead winch may pivot 5 to 10 feet from the original

Commented [A24]: The most useful objectives for the fish acoustic tracking study are to establish the home range of the species used in long-term biological monitoring to inform the degree to which that species can be associated with remedial outcomes.

The proposed plan is not sufficient for recalculating SWACs because not all of the 428 surface sediment samples are unbiased.

Commented [A25]: The sampling methods described in this section, Table 1 and Appendix B indicate that of 428 surface sediment samples identified as "unbiased", 108 surface samples will reoccupy sample locations collected in 2004. Although the cells for these samples were randomly selected, because these 108 locations will be moved to 2004 sampling locations, 108 of the 428 samples are biased.

Figure 2 from EPA's sampling plan shows EPA's 424 unbiased sample locations located along transects advanced every 0.2 mile from RM 1.9 to 11.8. This figure is included in EPA's attachment to this RLSO document. 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.

Although EPA proposed the transect approach, EPA will consider other sampling approaches as long as all locations are truly unbiased.

Commented [A26]: The 108 surface samples that will reoccupy sample locations collected in 2004 are biased samples.

sample location, and the process will be repeated until there is an equal volume of sediment from the three grabs. The volume will be homogenized until uniform in color and texture, then processed (described in QAPP/FSP).

For consistency purposes, surface sediment grab samples will be collected using the RI data collection protocols. The anticipated schedule for the surface sediment sampling is the first quarter of 2018.

3.2.3 Fish Tissue Sampling

The primary objectives of the fish tissue sampling study include collecting the data needed to:

- Characterize current levels of ~~selected~~ fish tissue COCs in resident fish tissue (SMB) on a Site-wide basis and smaller spatial scale (pending results of the fish tracking study);
- Characterize upriver background concentrations in resident fish tissue (SMB);
- Update statistically-based evaluations of PCB differences and changes in fish tissue; and
- Evaluate the bioaccumulation model used to relate sediment and tissue concentrations (input new SWACs, fish tissue, and surface water concentrations into the model).

The study includes collection of synoptic SMB data to re-baseline resident fish tissue concentrations in the river, evaluate MNR changes, refine the understanding of the FWM, and update human health fish consumption risks. The scope includes collection of 95 whole body discrete (non-composited) samples from the Site, plus 20 from the Downtown Reach and 20 from the Upriver~~stream~~ Reach (D/U Reach). While 95 SMB samples within the Site and 40 SMB samples from the D/U Reach will be targeted, the number collected will be to the extent sufficient numbers of~~r~~ fish are present. The overall sample design is consistent with the approved 2012 SMB sample design.[2] The sample design targets 20 to

Commented [A27]: The primary purpose of fish tissue sampling is to obtain baseline data for comparison to future data to evaluate progress toward achieving the target tissue levels specified in ROD Table 17

Commented [A28]: Revision of the FWM is not an objective of the PDI.

Commented [A29]: These evaluations are not the intent of the PDI. MNR effectiveness will be evaluated as part of long-term monitoring. Concentrations that were estimated using the FWM should not be recalculated.

[2] The design is also consistent with the 2011 SMB study performed by EPA, the State of Oregon, and the City of Portland (GSI 2011). The analytical laboratory contracted by EPA incorrectly prepared 75% of the

30 samples in each of the 4 segments (described previously), including 5 samples in Swan Island Lagoon (Figure 5). A statistical analysis of the 2012 SMB data indicates that replicating the 2012 program sample size will allow detections of statistically significant ($p<0.05$) concentration differences for PCBs in SMB. A summary of the statistical power analysis performed for fish tissue sample size is provided in Appendix C. Within the D/U Reach, the 40 samples will be collected from locations throughout the Reach. Consistent with the 2012 sampling, SMB that are 225 to 355 millimeters in total length (approximately 9 to 14 inches) will be targeted.

All fish tissue samples will be analyzed for lipids and the COCs presented in Table 6 (with the exception of PAHs). Samples will be analyzed as individual whole-body specimens, and fillet concentrations will be estimated using the SMB whole body to fillet ratios presented in the final Feasibility Study (EPA 2016b), as supported by analysis of the Round 3B SMB tissue data (see Appendix D). Collection methods will include hook and line with electroshock back-up if needed. The anticipated schedule for the fish tissue sampling is late summer of 2018, consistent with previous tissue sampling events.

3.2.4 Surface Water Sampling

The objective of surface water sampling is to re-baseline river conditions with synoptic data (sediment, fish tissue, surface water), evaluate surface water current conditions and changes, and provide 2017/2018 data to update current risks. Surface water will be collected from seven transect locations over three sampling rounds. Figure 6 presents the locations of the transects, located at approximately:

- RM 1.8 at the downstream boundary (within the Site, Segment 4);
- Downstream boundary in Multnomah Channel; entrance to channel is near RM 3;
- RM 4;
- RM 7;
- RM 8.8;

---

samples as skin-off fillets, discarding the remainder of the carcass instead of processing the whole fish. Thus, tissue chemistry results from the 2011 sampling effort are limited.

- RM 11.8, just upstream of Site, near Upstream boundary; and
- RM 16.2, further upstream, near the Downtown Reach/Upstream boundary.

These locations will effectively characterize the four segments of the Site RM 1.9 to 5 (Segment 4); RM 5 to 7.5 (Segment 3); RM 7.5 to 9 (Segment 2); and RM 9 to 11.8 (Segment 1).

The purpose of sampling is to characterize the flow and quality of surface water passing through the river's cross section at each location. These locations were targeted to provide spatial coverage and analyze physical changes in the river dynamics.

One composited sample will be collected per transect (similar to the RI/FS data use approach). The sample will be vertically-composited and horizontally-composited along the transect. Composite samples will be collected by sampling equal volumes from three locations (east shore, navigation channel, and west shore) and at three depths per location – upper (three feet below water surface), near bottom (three feet above sediment surface) and middle (equal distance between upper and bottom). The objective of the composite sample design is equal volume across the cross-sectional area of the segment. The target volume will be collected from nine discrete subsample locations across the transect. However, if the nearshore areas have shallow water depths (i.e., less than 10 feet), fewer subsamples may be collected. Volumes will be adjusted such that equal volume of surface water is collected from the east shore, navigation channel, and west shore.

Surface water sampling will be conducted over three events targeting different months and flow conditions to capture seasonal variability of surface water conditions. For PDI investigations, sampling will target three seasonal events: (i) August, during summer low flow conditions; (ii) January/February, during winter high flow conditions; and (iii) November/December, targeted for storm flood-influenced conditions and consistency with previous sampling events. For the PDI effort, the time of year for sampling will be the primary factor (coverage throughout the year) and the water level/river flow will be a secondary consideration factor for selecting when to sample.

During the RI, two sets of sampling events occurred: one during Round 2A and one during Round 3A (EPA 2016a). Each set of sampling events targeted a low-flow, high-flow, and stormwater-influenced condition. Low-flow events occurred throughout the year in November 2004, March 2005, July 2005, and September 2006. For all low-flow events, the average flow was less than 19,400 cfs. High flow events occurred in January 2006 and January 2007, during which average flow was greater than 59,800 cfs. One stormwater influenced event occurred in November 2006, during which average flow was 23,000 cfs.

Flood conditions will be checked relative to the United States Geological Survey (USGS) Real Time National Weather Information System (NWIS) Database for the Morrison Bridge station 14211720. Consistent with the RI, high river flow events will be characterized as >50,000 to 100,000 cfs and low flow events will be characterized as within or less than historic average flows (15,400-24,700 cfs from 1998 to 2003). Surface water average monthly discharge, velocity, height, and rainfall are shown in Figure 7 for 2010 through the most recent data available in 2016.

Surface water data will be collected using a high-volume pumping system connected to a XAD-2 resin filter and column (hydrophobic polyaromatic resin) to collect hydrophobic organic compounds for analysis by ultra-low detection limit analytical methods (consistent with RI/FS approaches and methods). Surface water will be pumped through a Teflon lined polypropylene tubing, 140-micrometer (μm) stainless-steel pre-filter, 0.5-μm glass fiber filter cartridges, and XAD-2 resin beads packed inside stainless-steel canisters. A target volume of 300 liters (L) of water will be pumped through the system from three discrete vertically composited locations per transect for a single composite sample per transect as described above. XAD sampling is expected to result in method detection limits (MDLs) for DDx, Chlordanes, PCBs, and PAH that are at or below ARARs, and MDLs for Aldrin and D/F that are equivalent to those achieved in the RI.

Surface water samples will also be collected using a peristaltic pump for the analysis of semi-volatile organic compounds (SVOCs), metals, and non-chemical parameters (e.g. total suspended solids [TSS]). Surface water samples collected via peristaltic pump will be collected in accordance with RI procedures. In brief, surface water will be drawn through pre-cleaned acid washed Teflon tubing, following purging of at least 5 times the required sample volume. Surface water from each vertically-integrated location will be combined in a laboratory provided container. Once all locations on a single transect have been sampled, surface water will be homogenized, filtered as needed and allocated to individual laboratory sample containers for analysis. Peristaltic pump samples are expected to reach MDLs at or below ARARs for all metals with the exception to arsenic. Select organics measured via peristaltic pump are expected to achieve MDLs similar to those in the RI; however, some remain above ARARs. SVOCs will not be analyzed via XAD sampling due to analytical interferences and to remain consistent with RI methods.

However, lower MDLs than those listed in Table 6 may be achievable in coordination with the selected lab.

Each sample will be analyzed for suspended solids in the water column (TSS), particle size distribution, and chemical testing for the COCs presented in Table 6. XAD analytes (PCB congeners, DDx, PAHs,[3] and D/F) will be reported as dissolved and particulate phase results. Filtered peristaltic pump samples will be analyzed for dissolved metals and dissolved organic carbon; unfiltered samples will be analyzed for total metals, select organics (see Table 6), tributyltin, and conventionals. Field parameters will include temperature, pH, dissolved oxygen, oxidation-reduction potential (ORP), flow rate velocity, and conductivity. Field parameters will be measured real-time in situ at each location using a YSI Multiprobe Water Quality Meter or equivalent. Water quality meters will be calibrated per manufacturer's specifications each day prior to sampling.

The anticipated schedule for the surface water sampling is winter 2017 through summer 2018. A total of 15 water samples (7 total, 7 dissolved, and 1 quality assurance/quality control [QA/QC]) will be analyzed per event, for a total of 45 samples.

3.2.5 Subsurface Sediment Sampling

Subsurface sediment (core) sampling will be conducted in targeted areas within or along the boundaries of SMAs that have limited data coverage to refine the active footprint boundaries of the Alt F Mod SMA footprints. The goal of this study is to refine the horizontal and vertical extent of contamination at concentrations greater than the RALs at depth for the purpose of supporting the 30% design, to confirm the CSM, and to refine the dredge volumes for 30% design cost estimation.

A total of 90 core locations are planned based on visual distribution of subsurface contamination, using 250- to 300-foot distance as a general guide to the next nearest coring location. In some cases, stations will be reoccupied to determine the vertical extent of contamination where previous cores did not "tag bottom", and in other cases, a new core will be collected in an active footprint area where none previously existed.

---

[3] Surface water PAH samples may be collected by peristaltic pump method pending additional review of previous data.

Core locations, rationale, target depths, and analytes are provided in Table 7a. Table 7b presents the rationale for the core placement location and target depth. Target depths were based on the vertical extent of contamination observed in surrounding cores, anticipated depth of native material, or an additional 2- to 4-foot sample depth if previous cores did not reach the bottom of contamination ("tag bottom"). Cores will be advanced using a vibracore, impact core, diver push core, or similar device from a floating platform with an experienced subcontractor and field collection team. The QAPP/FSP will provide more details.

Cores will be visually logged using American Society of Testing Materials (ASTM) and RI procedures (e.g., correcting for compaction), then subsampled into 2-foot increments unless stratigraphy indicates otherwise. Planned coring locations (Figure 4) may be adjusted after the ~~SMA footprints have been revised based on~~ 2018 surface sediment sampling results are reviewed. Subsurface sediment samples will be analyzed for focused COCs, TOC, and grain size as outlined in Table 7a. Deeper intervals will be archived frozen pending the chemical results (> RALs) of selected intervals.

Commented [A30]: To reduce the need for re-coring locations to establish the vertical extent of contamination, EPA recommends that cores be drilled to adequate depth and samples not analyzed are archived until the laboratory results are received. Obtaining laboratory data in 1-foot increments near and below the base of contamination will enable dredge depths to be more accurately defined. A 2-foot sample core will allow to positively identify contamination above RALs, but a greater core interval resolution (1 foot or less) will be needed for remedial design to determine where active remediation areas have reached levels below RALs.

Commented [A31]: The surface and subsurface sediment samples collected during the PDI will not be sufficient to define SMAs—but can assist in developing the SMA locations to be further defined through additional sampling as part of remedial design.

Geotechnical characterization of subsurface sediment will include index testing (e.g., moisture content, grain size, and TOC) and relevant field parameters (field torvane test as a measure of shear strength).

3.2.6 Fish Acoustic Tracking Study

An acoustic fish tracking study is planned to capture fine-scale temporal and spatial movement of SMB at the Site, pending pilot study results. A pilot study conducted in June 2017 involved deployment of an array of acoustic receivers from two vendors in two different types of environments within the river system (quiescent and active). Willamette Cove was selected as the more quiescent location, and RM 11.5 East as the more active location (Figures 8 and 9). The acoustic receivers were mounted on the bottom of the river and deployed for one week (13 through 19 June 2017). The pilot test included mobile and stationary testing of acoustic tags to evaluate the range of reception and position accuracy of both vendor's systems.

Preliminary analysis of pilot study results supports the technical feasibility of a Site-wide study that will provide data on SMB movement in the Lower Willamette River. A properly designed array of acoustic receivers can provide fine-scale and presence/absence data that can be used to understand SMB movement in the Study Area. The full-scale study will be conducted over a one-year period to capture seasonal home range patterns. Using a more refined acoustic telemetry approach than the historical (2000-2003) radiotracking study (Freisen 2005), the results will re-evaluate where and to what extent SMB stay within the 1-mile exposure areas assumed in the risk assessments and FWM.

The results will be used to: (i) inform the fish tissue sampling plan scheduled for late summer 2018; (ii) refine the SWAC segments used to evaluate changes in surface sediment concentrations; (iii) refine understanding of the FWM and reduce uncertainty about remedy effectiveness for fish tissue recovery; and (iv) help inform a future ICs plan. The anticipated schedule for the full-scale fish tracking study is fourth quarter 2017 through 2018. The work will be performed in collaboration with Dr. Karl Gustavson, EPA Office of Superfund Remediation and Technology Innovation (formerly USACE) and experienced staff from USACE Engineer Research and Development Center (ERDC).

Commented [A32]: The fish acoustic tracking study results should not be used to refine the SWAC segments used to evaluate changes in surface sediment concentrations, and refine understanding of the FWM. The most useful objectives for the fish acoustic tracking study are to establish the home range of the species used in long-term biological monitoring to inform the degree to which that species can be associated with remedial outcomes.

3.2.7 Camera Survey of Anglers

A camera survey of anglers is planned to collect data on the location and frequency of people fishing along the river. The results will provide an empirical line of evidence on frequency and duration of angler trips over a year-long period that can be used to support the development of ICs (e.g., targeted locations and seasons for messaging). During the fish tracking pilot study, wildlife cameras were field-tested to assess their suitability and effectiveness for monitoring activity at the river. Four ZenNutt High Definition wildlife cameras were installed to provide coverage of the two offshore areas where fish tracking equipment was deployed (two cameras per location) (Figures 8 and 9). The motion-activated cameras were deployed for the one-week duration of the pilot study, and captured images with clear definition, color, and contrast. The results of the camera pilot study indicate that this type of field camera would be suitable for a larger program in the Study Area.

Commented [A33]: EPA maintains that the proposed camera survey information will not provide empirical evidence with any certainty to be useable for the intended data end-use objectives suggested in this section and presented in Table 8. In this regard, it would be reasonable to anticipate that surveillance cameras that are perceived as part of a government (or industry) study on the capture of contaminated fish from a specific area would deter fishing near surveillance cameras.

The camera survey program will consist of photographic documentation of human activity using a network of cameras in the Site. The scope includes installation of 12 stationary cameras at select locations that are known or suspected to be used by anglers or are popular points of access to the river based on prior studies and angler knowledge. Cameras will be pre-set to take photos at 30-minute intervals during daylight hours. This time interval is anticipated to be less than the average amount of time it would take for an angling activity to begin and end within the field of view.

The survey will be conducted over a one-year period to account for seasonal variation in use. Photographs will be uploaded monthly to digital photographic software for visual review and tagging with descriptors (e.g., date, location, time, number of individuals), and stored in a project database. Photographs will be reviewed and facial features and other personal information (e.g., license plates, boater registration) will be redacted to protect privacy. Camera data will be processed and provided to EPA with all Personal Identification Information (PII) removed.

The anticipated schedule for the camera survey is fourth quarter 2017 through 2018.

3.2.8 Downtown/Upriver~~stream~~ Reach ~~(Upriver)~~ Study

The Downtown Reach is immediately south (upstream) of the Site between RM 11.8 and RM 16.6 (as defined by EPA). According to EPA, it is bounded between the Site (RM 1.9 to RM 11.8) and the Upriver~~stream~~ Reach (RM 16.6 to RM 28.4) and located in the heart of the downtown Portland urban center. It has historically had a higher level of contamination than the Upriver~~stream~~ Reach and is in immediate proximity to the Site (EPA 2017); remedial actions have been completed in this area during the last decade. The Upriver~~stream~~ Reach was selected as the reference area for evaluating background sediment concentrations in the RI. This area extends from the upstream end of Ross Island Lagoon to approximately 2.5 miles above the Willamette Falls, which was considered generally representative of upstream sediment loading to Portland Harbor. Early (2006) memoranda describe the background Reach from RM 15.5 (upper end of Ross Island) to RM 26 (Willamette Falls), although this was revised to RM 15.3 to RM 28.4 in the Draft RI (in 2009). The upstream extension was to capture the EPA West Linn and Blue Heron Sediment Investigation data from 2007, located upstream of Willamette Falls. The lower boundary changed from RM 15.5 to RM 15.3 due to additional samples collected downstream of RM 15.5.

This component of the Work focuses on the Downtown Reach (RM 11.8 to RM 16.6) and Upriver~~stream~~ Reach (RM 16.6 to RM 28.4) to characterize incoming contaminant ~~background~~ loadings to the Site. ~~For purposes of this PDI, it is collectively referred to as the "Upriver Area" extending from RM 11.8 to RM 20 and encompasses the Downtown Reach and part of the Upstream Reach.~~ Figure 10a presents the distribution of fine-grained sediment based on historic samples. Sampling of this area includes surface water, surface sediment, sediment traps, and fish tissue sampling in the D/U Reach (Figure 10b), and samples will be collected assuming sufficient fined-grained sediment and fish

Commented [A34]: To be consistent with the ROD, the work plan should be revised to use "Upriver Reach" (not Upstream Reach) to describe the portion of the Willamette River from RM 16.6 to RM 28.4. Samples from the Upriver Reach should be obtained from the entire reach and not be limited to RM 16.6 to RM 20.

Commented [A35]: The work plan states that upstream sampling will not be conducted above RM 20. The boundaries for sampling should remain consistent with the ROD where the Downtown Reach includes RM 11.8 to 16.6, and Upriver Reach includes RM 16.7 to 28.4.

availability. Half of the targeted surface sediment samples will be collected from the Downtown Reach and the other half will be collected within the Upstream Reach. Data collected during the Work will be used to: (i) better characterize the concentration of COCs immediately upstream of the Site; (ii) better characterize the concentrations of COCs entering the Site in surface water and suspended sediments to assess potential recontamination post-remedy.; (iii) refine background concentrations of COCs in surface sediments reflective of an urban background; and (iv) reset achievable remedy targets/ actions.

Commented [A36]: Primary purpose for upstream sediment sampling is to collect samples over time to develop "equivalency"--determining incoming load or sediment quality that would dictate what cleanup with natural recovery could achieve. Multiple sampling rounds of the Downtown Reach and Upriver Reach will be required to evaluate equivalency in order to assess background concentrations and achievable remedy targets.

Surface Sediment

An additional 60 surface sediment samples will be collected from the D/U Reach, with locations targeting fine grain sediments to characterize the mobile sediments likely to be deposited throughout the Site. While a total of 60 surface sediment samples from the D/U Reach will be targeted, the number collected will be to the extent reasonably or technically practicable, based on sufficient fine-grained sediment presence. A desktop study and reconnaissance survey will be conducted in the D/U Reach to identify areas with fine grain sediments prior to sampling. The desktop study will research previous sediment study available grain size data and bathymetry data to select target areas. A reconnaissance survey will be performed to further identify target areas and ground truth results from the desktop study. Figure 10 shows 30 random, locations in the Downtown Reach and 30 locations in the Upstream Reach (locations to be confirmed pending grain size evaluation) with about 25% of the dataset as re-occupied stations (n = 5 and n = 5 from Downtown and Upstream, respectively). Surface sediment samples will be collected as described in Section 2.3.2 and analyzed for the full list of COCs. All validated and acceptable data will be considered in data evaluation (i.e., the topic of potential outliers associated with system errors will be handled by the Peer Review Panel if not resolved by the project team) to fully characterize potential upstream sources. Grain size and organic carbon content will be considered when comparing samples from the Downtown and Upstream Reaches to Site concentrations.

Commented [A37]: Sample locations in the Downtown Reach and Upriver Reach should be unbiased and not reoccupy previous sample locations.

Fish Tissue

An additional 40 SMB samples will be collected from the D/U Reach. Fish tissue samples will be collected as described in Section 2.3.3 and will include whole-body analysis of the COCs presented in Table 6 (with the exception of PAHs). Fish tissue sample locations are presented on Figure 10b.

Surface Water

As noted in Section 2.3.4, two upstream transects for surface water sampling will be included – one in the Downtown Reach at RM 11.8 and one in the Upstream Reach at RM 16.2. Surface water samples will be collected as described in Section 2.3.4 and analyzed for the surface water COCs presented in Table 6. Analysis of total and dissolved analytes via XAD and filtered/unfiltered peristaltic pump samples will match Site surface water sampling as described Section 3.2.4.

Sediment Traps

Commented [A38]: The Work Plan includes sediment traps upstream of the Site, with the objective of determining incoming sediment particle concentrations, but no corresponding sediment traps on the downstream boundary of the Site. Obtaining data on sediment leaving the site would provide an indicator of remedy effectiveness and natural recovery. If sediment trap data is collected upstream of the Site, corresponding sediment traps are needed downstream of the Site.

Sediment traps will be deployed to provide a line of evidence on incoming sediment load to the Site that targets fine-grained, more mobile suspended sediment, and higher-TOC material that is more likely to move downstream and be deposited at the Site. Consistent with methods in the RI, sediment traps will consist of four glass tubes approximately 10 centimeters in diameter and 55 centimeters long.[4] Tubes will be placed inside protective polyvinyl chloride (PVC) sleeves, which will be attached together and secured to a rebar post driven into the sediment floor by divers. The diver will affix the sediment trap assembly to the rebar so that the open tops of the cylinders are 3 feet above the mudline elevation. Two sediment traps will be deployed along each transect (total of four traps). Settling particulate material will be collected in the glass tubes of the sediment trap. For recovery, a diver will cover the tops of each glass tube with foil, detach the trap assembly from the rebar, and the trap will be raised to the surface with the vessel's winch. The glass tubes will be removed from the assembly, kept upright, and allowed to resettle, if needed. The thickness of accumulated sediment will be measured at multiple points around each tube to account for sloping of sediment within the tube. Overlaying water will then be siphoned or pumped off, sediments collected in a stainless-steel mixing container, homogenized until uniform color and consistency is achieved, and placed in the appropriate laboratory provided sample jars. Sediments will be analyzed for the full list of sediment COCs, TOC, and grain size (Table 6). Sediment traps will be sampled in coordination with the surface water sampling program (three events over one year, coordinated with the surface water sampling program).

[4] Note: the top of the trap is oriented parallel to mudline and perpendicular to the direction of flow; need to check with EPA, as their approach as described in the 6 June 2017 scoping plan may be different.

3.2.9 Background Porewater Sampling

Background metals concentrations in porewater were not defined during the RI, and the focus of a background porewater characterization would be naturally-occurring metals. Background metals porewater concentrations should be developed and cleanup levels adjusted accordingly. Metals, especially arsenic and manganese, are present in relatively high concentrations in volcanic rocks, which are the primary source of Willamette River sediment. Porewater concentrations above ROD cleanup levels may occur in the transitional zone water (TZW) near the mudline, as a result of the geochemistry which favors dissolution of these metals from the mineral components of the sediment.

This component of the Work is intended to place dialysis equilibrium passive porewater samplers (referred to as peepers) in the sediment bed in areas that are representative of background metals in porewater (during periods of low redox, target July/August). Peepers include a glass or polyethylene vial covered with a 0.45-µm polyethersulfone (PES) membrane (see photograph in Figure 11). The interior of a peeper vial consists of rows of chambers that are filled with distilled deionized water prior to deployment. During deployment, the deionized water approaches diffusive equilibrium with the porewater, over a 2- to 4-week period; the peepers are then retrieved.

Commented [A39]: It is recommended that the passive porewater samplers be evaluated for achievement of equilibrium by placement of chemical tracers in each probe—otherwise equilibrium is assumed rather than demonstrated.

Porewater peepers will be deployed in triplicate (for three-point composite samples) at eight locations in upstream areas, or other relevant areas from within the Site. Ideally, these stations would be co-located with surface sediment stations. Locations for porewater sampling will be selected to be representative of redox conditions and variation in source. In general, these areas will include thicker sediment zones, areas downgradient of wetlands or buried lakes, and will consider Columbia and Willamette River provenance. Sample locations will be pre-screened to ensure sediment concentrations are similar to background and redox potential is low. Two potential locations have been identified – adjacent to Port of Portland Terminal 5 at approximately RM 1.8 and adjacent to Miller Creek at the mouth of the Multnomah Channel.

Porewater peeper samples will be deployed from a vessel using a push pole deployment device, and will be deployed with a marker and weighted retrieval line. Porewater samples will be retrieved following two to four weeks of deployment, and porewater will be analyzed for freely-dissolved arsenic and manganese. Porewater results from passive samplers could be compared to laboratory-derived porewater samples from the upstream bulk sediment surface grab locations. A total of 8 samples (3 subsamples will be composited into 1 sample per location) will be collected from a one-time event during low flow conditions.

3.3 Task 3: Data Evaluation

Data collected as part of the Work will be summarized and analyzed to meet several DUOs (see Table 5). Table 8 outlines the data evaluation and interpretation plan. Following completion of field work and chemical analyses, data analyses will be completed and a PDI Evaluation Report will be submitted to EPA. The PDI Evaluation Report will include the following elements:

- Summary of the investigations performed;
- Summary of investigation results and identification of existing conditions;
- Summary of validated data (i.e., tables and graphics);
- Data validation reports (Tier II) and laboratory data reports;
- Photographs documenting the work;
- Angler survey information processed to eliminate all PII;
- Evaluation of current sediment/biota conditions along with background loading to refine active remedy and monitored natural attenuation (MNA) areas;
- Use of bathymetry data to refine the elevation requirements of the active remedy footprint, especially in the intermediate and shallow areas;
- Refinement of the CSM and understanding of current conditions;
- Refinement of sediment recovery curves based on empirical data changes (and confirm RALs);
- Re-calculation of Site-wide and segment-wide surface sediment SWACs using new data; may also consider other spatial scales;
- Evaluation of fish tracking results to refine the extent and segmenting of the river (for calculation of SWACs) and assess fish home ranges used in the FWM;

Commented [A40]: Many of the work elements described below represent revisiting the RI/FS and will result in significant delay in remedy implementation.

Commented [A41]: Many of the evaluation statements in this section are beyond the scope of this phase of work and not consistent with or support the ROD's data needs or objectives.

Without using the data quality objective process to fulfill a specific data need of the ROD's decision framework, the outcomes of these evaluations will be ancillary to and not directly influence remedial decision making.

Commented [A42]: The ROD states that sediment data will be used to delineate remedy areas; not fish trends and "background loading".

Commented [A43]: RAL curves presented in the ROD should not be revised.

Commented [A44]: The proposed plan is not sufficient for recalculating SWACs because not all of the 428 surface sediment samples are unbiased.

Commented [A45]: Fish tracking results should not be used to refine the extent and segmenting of the river. The primary objective for the fish acoustic tracking study is to establish the home range of the species used in long-term biological monitoring to inform the degree to which that species can be associated with remedial outcomes.

- Update the surface sediment, fish tissue, and surface water data based on the validated 2017/2018 information obtained during this investigation;
- Assessment of new bathymetry for bed stability and fish/sediment data for monitored natural recovery potential;
- Evaluation of current (2017/2018) upstream background concentrations;
- Update the active remedial footprint by running the new data through the ROD decision tree for assigning remedial technologies;
- Evaluation of the new ROD COC data for purpose of reducing the list of COCs for future monitoring rounds;
- Use of new data to refine our understanding of the FWM and reduce uncertainty about remedy effectiveness for fish tissue recovery and update calculations of baseline fish consumption risks; and
- Support and advance PCI allocation.

Technology assignments will be identified based on sampling data in all areas of the river, as indicated by the decision tree described in the ROD (2017 ROD Figure 28, Appendix I). The decision tree provides detail regarding how design data will influence design and construction and future maintenance dredging. The decision tree allows caps to be used in dredge areas if RALs are not achieved or if PTW remains. This is based on area-specific analysis (EPA 2017). The ROD decision tree describes four compliance regions (ROD Figure 28):

- Navigation Channel and Future Maintenance Dredge area;
- Intermediate Region (outside the navigation channel to -2 feet CRD);
- Shallow Region (-2 feet CRD to shore); and
- Riverbank Region (top of bank down to the river).

The riverbank areas are currently being evaluated under ODEQ-led investigations.

One important component of re-baselining the Site is to evaluate the extent of natural recovery processes as measured by changes in concentrations since the RI. As stated in the ROD (responsiveness summary), "*EPA concurs that natural recovery is occurring within Portland Harbor and that it should be utilized in sediment remedies, as evidenced*

Commented [A46]: Primary purpose for upstream sediment sampling is to collect samples over time to develop "equivalency"--determining incoming load or sediment quality that would dictate what cleanup with natural recovery could achieve. PDI will not obtain sufficient data to evaluate current background concentrations.

Commented [A47]: Detailed sampling results obtained for remedial design will be needed to update active remedial footprints. The surface and subsurface sediment samples collected during the PDI will not be sufficient to define SMAs—but can assist in developing the SMA locations to be further defined through additional sampling as part of remedial design.

Commented [A48]: These evaluations are not the intent of the PDI. Concentrations that were estimated using the FWM should not be recalculated. Fish tissue data can be used to inform fish advisories but not calculate fish consumption risks.

Commented [A49]: The primary purposes for this AOC/SOW are to collect data for Pre-RD SMA delineation, which will assist with allocation and to collect baseline data to establish existing site conditions for long-term monitoring. Notably, technology assignments within SMAs are based on other information components (e.g. erosion/deposition, slope, propeller wash) that are not entirely informed by Pre-RD SMA delineation and baseline sampling and more appropriate for the design phase that is outside this AOC/SOW.

by the fact that MNR represents the response action assigned to between 64 and 90 percent of the total area of the Site for all alternatives carried through the detailed analysis in the June 2016 feasibility study. However, the rate of natural recovery is expected to vary by location. Pre-design sampling will be used to ensure that the natural recovery is factored into the design…".

The Work sampling program will be statistically robust to support calculation of Site-wide SWACs and assess spatial patterns without reliance on older data. Figure 12 presents a summary of the PDI field sampling tasks.

3.4 Task 4: Data Compilation

The purpose of this task is to identify, review, compile, and summarize Site and upstream data collected since the RI/FS that are relevant to the Work. This task includes compilation of data collected after 2008, including data collected as part of the Work. These data will be incorporated into the project database. A summary of investigations from 2008 to 2017 are included as Appendix A.

Commented [A50]: Only EPA approved data will be incorporated into the project database.

EPA and the Pre-RD Group recognize that the data gathered in PDI is not the complete data set for remedy design/implementation and that EPA's review of the data reports or data analysis may include an assessment as to whether the data relied on is sufficient to support the certain PDI evaluations, refinements, recalculations and updates. EPA reserves the right to review all submittals prepared under the Work Plan.

The data collected since the RI/FS that has been approved by EPA will be compiled and uploaded into the project database and may include the following:

- Site data – sediment, porewater, fish (SMB) tissue, and bank soil data collected from 2008 to 2015;
- Downtown/Upstream data – sediment and tissue data collected from 2008 to 2015; and
- PDI data – sediment, fish tissue, surface water, and porewater data collected as part of this study.

Available data will be acquired from LWG, ODEQ's Environmental Information Management (EIM) database, and participating parties. Site data (i.e., sediment, tissue, surface water, and porewater data) will undergo a data quality review to determine if they meet data quality objectives (DQOs) consistent with those developed for the RI/FS using Superfund guidance. If so, the data will be summarized, compiled in the project database, and determined acceptable for all uses. If data do not meet DQOs, they will be summarized, compiled in the Site database, and flagged for conditional use. For example, data from the EIM database did not meet DQOs because QC backup was not available. Data (including surface and subsurface sediment and porewater data) collected at locations that were subsequently dredged or remediated will also be excluded from the compilation as these no longer represent current conditions.

Commented [A51]: Only EPA approved data will be incorporated into the project database.

3.5 Task 5: Reporting

Reporting and deliverable are discussed in Section 5.

4. WORK MANAGEMENT STRATEGY

The following information generally describes the duties, responsibilities of personnel and firms involved in the Work; project organization; reporting relationships; lines of communication; and management authorities.

4.1 Roles and Responsibilities

4.1.1 EPA

EPA is the lead agency overseeing the Work. EPA has the authority to review and approve this PDI work plan, supporting FSP and QAPP documents, and reporting deliverables. EPA will be assisted in the review of technical documents by oversight contractor CDM Smith, Inc. (CDM Smith). In addition, Karl Gustavson, from EPA headquarters and EPA Office of Superfund Remediation and Technology Innovation and Contaminated Sediments Technical Advisory Group, will continue to provide regulatory and technical support throughout the project. A peer review process will be followed per the ASAOC.

4.1.2 Participating Parties

Those participating in the PDI studies being performed by Pre-RD Group and its contractors will be determined at a later date. Once the participating parties are determined, they collectively will be responsible for implementing the studies.

4.1.3 Selected Contractor

Geosyntec is coordinating activities including management of all subcontractors, field sampling, analysis, and reporting scoping tasks in preparation of this Work Plan. The contractor to lead the field sampling will be determined at a later date.

The Project Manager will be responsible for overall project coordination and providing oversight on planning and coordination, work plans, all project deliverables, and performance of the administrative tasks needed to ensure timely and successful completion of the project. Geosyntec will also be responsible for coordinating with Pre-RD Group and EPA on schedule, deliverables, and other administrative details.

The Field Coordinator (FC) will be responsible for managing field activities and general field QA/QC oversight. The selected FC will ensure that appropriate protocols for sample collection, preservation, and holding times are observed and oversee delivery of environmental samples to the designated laboratory for chemical analyses. Deviations from this QAPP/FSP will be reported to the Project Manager for consultation. Significant deviations from the QAPP/FSP will be further reported to representatives of the Pre-RD Group and EPA.

The lead subcontractor will oversee data management to ensure that analytical data are incorporated into the PDI database with appropriate qualifiers following acceptance of the data validation. QA/QC of the database entries will ensure accuracy for use in the PDI study. The testing laboratories (TBD) and field contractors (TBD) all play supporting roles.

4.1.4 Peer Review Panel

To be determined.

4.2 <u>Communication Strategy</u>

To be determined.

5. SCHEDULE AND DELIVERABLES

5.1 Schedule

The goal is to complete the Work by June 2019. Figure 13 presents an example project schedule through 2019 (to be updated). The field schedule for the Work includes time for development of QAPP and other project plans in 2017 and completion of field investigation activities by the end of 2018. The PDI scope of work is planned for completion by June 2019 and the draft PDI Evaluation Report is targeted for delivery to EPA by June 2019. An updated project schedule will be provided to EPA when revisions are made, and EPA will be given a minimum of 2 weeks' notice prior to the start of each field activity.

5.2 Deliverables

Laboratories will provide all data for field investigations in electronic format and QA/QC reports, including a narrative of the standard QA/QC protocols. Data validation of laboratory results will be performed by the lead contractor. Following data validation, all data, supplementary information, and validator qualifiers will be compiled into an SQL Server database for the project. Data summary files will be provided to EPA as they become available after data validation and database management. Deliverables include:

- FSP, QAPP, and DQMP describing how the work will be conducted;
- HASP describing worker safety for hazards posed by the Work;
- Monthly Progress Reports;
- Pre-RD Remedial Footprint Report; and
- PDI Evaluation Report.

Commented [A52]: The surface and subsurface sediment samples collected during the PDI will not be sufficient to define SMA boundaries, but can assist in developing the SMA locations to be further defined through additional sampling as part of remedial design.

Deliverables for the PDI Evaluation Report will include data summary tables, data graphics such as box-and-whisker plots, maps depicting the spatial distribution of sediment chemistry for selected analytical parameters, a comparison of Site conditions to the active Alt F Mod remedial footprint, analysis of differences and changes, and new SMA boundary maps.

Commented [A53]: See above comment.

This document comprises the total work scope agreed upon by the Pre-RD Group and EPA.

6. REFERENCES

EPA, 1985. Superfund Remedial Design and Remedial Action Guidance. United States Environmental Protection Agency, Office of Emergency and Remedial Response, Hazardous Site Control Division. Washington DC. February.

EPA, 2016a. Portland Harbor RI/FS, Final Remedial Investigation Report, Portland Oregon. United States Environmental Protection Agency Region 10, Seattle, Washington. 8 February.

EPA, 2016b. Portland Harbor RI/FS, Final Feasibility Study, Portland Oregon. United States Environmental Protection Agency Region 10, Seattle, Washington. June.

EPA, 2017. Record of Decision Portland Harbor Superfund Site, Portland Oregon. United States Environmental Protection Agency Region 10, Seattle, Washington. January.

Friesen, T.A., 2005. Biology, Behavior, and Resources of Resident and Anadromous Fish in the Willamette River, Final Report of Research, 2000-2004. Oregon Department of Fish and Wildlife. Contracted by the City of Portland. March.

Geosyntec, 2016. Sediment Sampling Data Report, Swan Island Lagoon, Portland, Oregon. Prepared for The Marine Group, LLC and BAE Systems San Diego Ship Repair, Inc. 12 August.

Germano and Associates (Germano), 2014. Characterization of the Lower Willamette River with Sediment Profile Imaging, Changes in Space and Time. June.

GSI Water Solutions, Inc. (GSI), and Hart Crowser, Inc., 2010. Field and Data Report, Downtown Portland Sediment Characterization Phase II, Willamette River, Portland, Oregon. Prepared for Oregon Department of Environmental Quality. June.

GSI, 2011. Sampling and Analysis Plan, Portland Harbor 2011 Baseline Smallmouth Bass Tissue Study, Willamette River, Portland, Oregon. Prepared for EPA, United States Army Corps of Engineers, and City of Portland. September.

GSI, 2014. Supplemental Remedial Investigation/Feasibility Study Field Sampling and Data Report, River Mile 11 East, Portland, Oregon. Prepared for RM11E Group. July.

Henderson, J., 2015. Statistical Trend Evaluation of Site-wide 2014 PCB Aroclor Data for Portland Harbor Superfund Site, Presentation. Prepared by TruGround Environmental, LLC. 3 Feb.

Kennedy/Jenks, 2013a. Portland Harbor RI/FS, Final, Appendix F Baseline Human Health Risk Assessment, Final. Prepared for Lower Willamette Group for submittal to EPA Region 10. 28 March.

Kennedy/Jenks, 2013b. Portland Harbor RI/FS, 2012 Smallmouth Bass Tissue Study, Data Report. Draft. Prepared for Lower Willamette Group. 13 March.

Kleinfelder, 2015. Sediment Sampling Data Report, Portland Harbor, Portland, Oregon. Prepared for de maximis Inc. 1 June.

Legacy Site Services (LSS), 2015. An Assessment of the Coupled Sediment Recovery and Dynamic Food Web Model: Predicting the Concentrations of Total PCBs in Lower Willamette Fish Tissue Based on 2002 to 2012 Sampling Data. Legacy Site Services LLC, White Salmon, Washington. August.

NewFields, 2016. Concentrations and Character of PAH in Sediments in the Proposed Remedial Alternatives Area of the Portland Harbor Superfund Site, River Miles 5-6 2015 Investigation. Prepared for ExxonMobil. 29 March.

Oregon Department of Environmental Quality (ODEQ), 2011. Portland – Willamette River Sediment Evaluation Downtown Reach – Phase II Follow-up Summary. 4 November.

Toll, J, J. Devary, and F. Wolf, 2015. Temporal Data Provide Confirmation of Portland Harbor Bioaccumulation Model Projections and Evidence of Natural Recovery. Presentation at 36th Annual Meeting, SETAC North America, Salt Lake City, Utah. 5 November.

Windward, 2013. Final Baseline Ecological Risk Assessment. Prepared on behalf of Lower Willamette Group. Appendix G of the Portland Harbor RI/FS. 16 December.

Wolf, F., 2015a. An Assessment of the Coupled Sediment Recovery and Dynamic Food Web Model: Predicting the Concentrations of Total PCBs in Lower Willamette Fish Tissue Based on 2002 to 2012 Sampling Data. Technical Report, available on ResearchGate.net. Presentation at Battelle Seventh International Sediments Conference in New Orleans, January. Prepared by Legacy Site Services LLC. August.

Wolf, F., 2015b. Natural Recovery in the Lower Willamette: What Do We Know About the Lines of Evidence? Presentation to LWG and State of Oregon. Prepared by Frederick Wolf, RETIA, USA. Portland, Oregon. 19 August.

Table 8. Data Interpretation and Analysis Plan
Portland Harbor Pre-Remedial Design Investigation Work Plan
Portland, OR

Data Utilization	Description
Current Conditions and Sediment SWACs for baseline dataset	Generate summary tables and maps with the new 2018 data and re-baseline the river for sediment, SMB fish tissue, and surface water for focused and selected other COCs. Generate sediment SWACs on a site-wide and other spatial scales using Theissen Polygons for the 640-sample plan, and the statistical mean for both the unbiased 428 and 640 sample plans.
Concentration Changes Over Time	Provide plots(s) of the tissue data over time. Provide surface water data with simple data plots of baseline and previous applicable surface water data. Compare the new 2018 data to the 2004 dataset for sediment, tissue (2002, 2007 and 2012), and surface water media and empirically evaluate changes since the RI: (1) The new sediment SWAC and arithmetic mean of unbiased dataset will be compared to the 2004 data at site-wide and segment-wide scales to look for differences in the last 14 years. (2) Datasets will also be compared to the Downtown/Upriverstream Area (are the two populations different, how different, and is site recontamination expected?). (3) 10887 surface sediment locations will be re-occupied from 2004, and these two populations will be compared for changes. *Note: The data permits statistical comparisons at the site-wide, segment-wide, or RM-scales.*
Alt F Mod Active SMA Footprint (1)	Confirm the active remedial footprint using new data collected during the PDI. SMA delineation activity to evaluate these data, bathymetry, and applicable historical data will be run through the ROD decision tree to support allocation. Run the new 2018 surface sediment data and bathymetry data through the ROD decision tree to refine the active remedy footprint. The 2018 core data will be combined with RI/FS subsurface coring data to update the conceptual site model understanding of subsurface contamination, collectively these data will also run through the ROD decision tree to refine the active remedy footprint and dredge volume estimate. Changes in sediment elevation will be a modifying factor for volume estimates to inform the 30% design estimates.
Downtown/Upriverstream Baseline	Report data separately to EPA by reach (Downtown / Upriverstream Reach). Evaluate current upriver conditions by SWAC and summary distribution statistics for other media, for future long-term comparative analysis with Site. Generate tables, maps, and summary statistics for all new 2018 data (sediment, SMB tissue, surface water, and sediment traps) as 2018 baseline conditions. These data may be compared to new site data, and qualitatively compared to older downtown/upstream data to evaluate changes and provide a first look at what may be achievable at the site for focused COCs.
Background Porewater	Provide porewater data to EPA, provide sufficient data to derive porewater background for metals using the passive porewater samples from the upriver reach or other appropriate background areas (see EPA March 13, 2017 framework).
Fish Tracking and Camera Survey Results	Provide fish tracking data to be presented as a tabular spreadsheet deliverable that includes location, tag IDs, and time stamps for each of the tagged fish. Data processing for data report limited to tabular and graphical outputs showing locations/tag IDs of fish movement (e.g. heat maps). Evaluate the fish tracking results to evaluate the home range of SMB. Maps, home range estimates, and summary tables will be generated. *Note: results may inform the fish tissue sampling program and the appropriate scale for calculating baseline conditions with respect to fish; refine understanding of the FWM and reduce uncertainty about remedy effectiveness for fish tissue recovery and update calculations of baseline fish consumption risks.* Camera data will be processed and provided with all Personal Identification Information (PII) removed.
Data Design for Long-term Monitoring	The PDI Evaluation Report will include an evaluation of the 2017/2018 data for purpose of reducing the list of COCs for future monitoring rounds.

Notes:
(1) Pre-Design Core data collection will have limitations for characterizing final SMA footprint delineation. Accordingly, any final decision on the SMA footprint will be pending full remedial design and confirmation sampling results obtained during remedy implementation.

Abbreviations:
COCs - chemicals of concern; RI/FS - remedial investigation/feasibility study; RAO - remedial action objective; ROD - Record of Decision; SWAC - surface weighted average concentrations; SMB - small mouth bass

Commented [Author1]: The 108 sample locations that reoccupy 2004 sample locations are biased samples.

Commented [Author2]: Comparison of 2018 unbiased data to existing biased data is not statistically valid

Commented [Author3]: As stated in Note 1, PDI sediment data suitable for SMA delineation for allocation purposes. Design-level data will be needed to confirm active remedy footprint.

Commented [Author4]: The fish acoustic tracking study results should not be used refine understanding of the FWM. The most useful objectives for the fish acoustic tracking study are to establish the home range of the species used in long-term biological monitoring to inform the degree to which that species can be associated with remedial outcomes.

DRAFT

APPENDIX A

SUMMARY OF DATA COLLECTED SINCE THE REMEDIAL INVESTIGATION/FEASIBILITY STUDY

Pre-Remedial Design Work Plan
Portland Harbor Superfund Site

SUMMARY OF DATA COLLECTED SINCE THE RI/FS

Commented [Author1]: Several of the data collection efforts summarized in Appendix A were independent investigations that were not conducted under EPA-approved work plans. Inclusion of the findings and conclusions drawn from these studies, as stated in the work plan and Appendix A, should not be construed as an endorsement or acceptance by EPA. EPA does not agree with any statement or position taken in the Work Plan that is in any way inconsistent or contrary to the ROD and particularly any response found in the Responsiveness Summary.

This appendix summarizes the eight environmental studies that were conducted after the remedial investigation (RI) was completed in 2008. The studies were conducted between 2008 and 2016 and included collection of surface and subsurface sediments, smallmouth bass (SMB) tissue samples, and sediment profile imaging (SPI).

Field and Data Report, Downtown Portland Sediment Characterization Phase I and II, GSI and Hart Crowser, Inc. 2008 and June 2010

Phase I of the Downtown Portland Sediment Characterization (DPSC) was initiated by Oregon Department of Environmental Quality (ODEQ) in 2008 to assess the presence of environmental contaminants within the downtown Reach (River Mile [RM] 12 to RM 16). Between May and June 2008, 81 grab samples and 36 core samples were collected and analyzed for polychlorinated biphenyls (PCB) Aroclors, butyltins, dioxins/furans (D/F), metals, pesticides, polycyclic aromatic hydrocarbons (PAHs), semi-volatile organic compounds (SVOCs), and total petroleum hydrocarbons (TPHs). Surface sediment samples were collected by power-grab (with a maximum penetration capability of 22 centimeters) (GSI Water Solutions, Inc. [GSI] and Hart Crowser 2010).

ODEQ conducted a preliminary screening level evaluation of this data to identify areas where additional sampling was warranted to confirm the detection of elevated contaminants of concern (COCs), identify sources if possible, and delineate areas needing remediation. Concentrations of COCs in sediment were compared to screening values developed for the Portland Harbor Joint Source Control Strategy. Based on the relative proportion of samples exceeding screening values, ODEQ identified nine Focus Areas for additional investigation and source identification efforts (GSI and Hart Crowser 2010).

In Phase I, PCB Aroclors were detected in 69% of 100 surface and 70% of 20 subsurface sediment samples, with concentrations ranging from <1 microgram per kilogram (µg/kg) to 4,200 µg/kg and <1 µg/kg to 610 µg/kg, respectively. Total sum of dichlorodiphenyltrichloroethane and its derivatives (DDx) was detected in 88% of 100 surface and 75% of 20 subsurface samples, with concentrations ranging from <0.047 µg/kg to 144 µg/kg and <0.13 µg/kg to 300.5 µg/kg, respectively. Total PAHs were detected in 99% of 100 surface and 100% of 20 subsurface samples, with concentrations ranging from <0.28 µg/kg to 40,310 µg/kg and 72 µg/kg to 7,802 µg/kg, respectively.

Total D/F were detected in 93% of 58 surface and 14 subsurface samples, with concentrations ranging from <4.14 nanograms per kilogram (ng/kg) to 15,400 ng/kg and <2.88 ng/kg to 4,594 ng/kg, respectively (GSI and Hart Crowser 2010).

Phase II of the DPSC was conducted in 2010 to better understand the nature and extent of potential COCs within nine Focus Areas and the TriMet Supplemental Sampling Area. ODEQ identified the following Focus Areas: RM 12.1E, 12.4W, 12.5E, 12.9W, 13.1E, 13.3E, 13.5E, 14.1W, and 15.1E. Along with analysis of archived Phase I samples in these Focus Areas, an additional 27 grab samples and 9 core samples were collected between February and March 2010. Surface grabs were collected via Van Veen sampler, pneumatic power-grab sampler, diver-assisted grab samples, and, due to low water levels, dry-land sampling methods for one location. The average grab sample recovery depth was 11 centimeters below mudline (bml). Core samples were collected by vibracore with an average recovery depth of 4.8 feet bml. Surface grab samples and cores were analyzed for a focused set of target parameters (with a few exceptions). The "Partial Analyte Group" included PCB Aroclors, total organic carbon (TOC) and total solids; TriMet samples included grain size, metals, PAHs, pesticides, and TPHs (GSI and Hart Crowser 2010).

In Phase II, PCB Aroclors were detected in 79% of 38 surface and 89% of 9 subsurface sediment samples, with concentrations ranging from <1.3 µg/kg to 520 µg/kg and <1.3 µg/kg to 147 µg/kg, respectively. Total DDx was detected in 98% of 40 surface and 100% of 10 subsurface samples, with concentrations ranging from <0.05 µg/kg to 73 µg/kg and <0.14 µg/kg to 73 µg/kg, respectively. Total PAHs were detected in 100% of 29 surface and 10 subsurface samples, with concentrations ranging from 4.0 µg/kg to 32,030 µg/kg and 76 µg/kg to 5,680 µg/kg, respectively. Total D/F were detected in 100% of 16 surface and 4 subsurface samples each, with concentrations ranging from 7.7 ng/kg to 7,021 ng/kg and 112 ng/kg to 2,351 ng/kg, respectively (GSI and Hart Crowser 2010).

Upon review of the Phase II data, ODEQ identified four areas which warranted follow-up evaluations: RM 12.1E, RM 12.5E, RM 12.9W, and RM 15.1E. These areas were referred to ODEQ's Site Assessment Program to evaluate potential sources and the need for source control. As of 2011, ODEQ did not recommend additional in-river investigation in the Downtown Reach. ODEQ expects that concentrations will decline over time as sources are identified and addressed and natural recovery occurs. The Phase II investigation found that COCs were much lower than those found in the Site and ODEQ believes the Downtown Reach is not a significant ongoing upstream source (ODEQ 2011).

Smallmouth Bass Tissue Sampling, GSI, September 2011

In September 2011, United States Environmental Protection Agency (EPA) and the City of Portland performed SMB sampling throughout the Study Area to support the Remedial Investigation/Feasibility Study (RI/FS). The study design and methods are described in the 2011 sampling and analysis plan (SAP) (GSI 2011). The SAP identified collection of individual (non-composited) SMB fish from 136 locations between RM 1 and RM 16, with 4 samples from RM 1 to RM 1.9, four from Multnomah Channel, 123 from the Study Area (including 11 from Swan Island Lagoon), and five from RM 15E. Each sample was identified for analysis of the full suite of PCB congeners, SVOCs, PAHs, and organochlorine pesticides (GSI 2011). The analytical laboratory contracted by EPA incorrectly prepared 75% of the samples as skin-off fillets, discarding the remainder of the carcass instead of processing the whole fish. Thus, results from the 2011 sampling effort are limited. Of the 32 reconstituted whole body Site samples with total PCB data, the mean concentration was 530 µg/kg with a standard deviation of 868 µg/kg (Legacy Site Services [LSS] 2015).

Smallmouth Bass Tissue Study, Data Report, Kennedy/Jenks, March 2013

In late summer/early fall of 2012, the Lower Willamette Group (LWG) conducted fish tissue sampling and analysis under the oversight of EPA. The primary purpose of the sampling was to provide an additional line of evidence to support the monitored natural recovery (MNR) Site-wide evaluation presented in the draft FS (Kennedy/Jenks 2013a). A total of 83 discrete SMB samples were collected in the Study Area and 9 SMB samples were collected from RM 15 to RM 18. With one exception, 4 to 12 samples were collected per RM and in Swan Island Lagoon (n = 8); one sample was collected in RM 2. All fish were caught using conventional rods and reels, with the assistance of contract anglers from the Oregon Bass & Panfish Club and The Bass Federation of Oregon. All 92 samples were analyzed as whole-body individuals for lipids and PCB congeners.

In the Study Area, the concentrations of total PCBs in whole body SMB ranged from 0.092 milligrams per kilogram (mg/kg) to 6.47 mg/kg. The mean concentration of total PCBs in whole body SMB was 0.65 mg/kg, with a standard deviation of 1.19 mg/kg. Upriver, the concentrations of total PCBs in whole body SMB ranged from 0.051 mg/kg to 0.63 mg/kg. The mean concentration of total PCBs in whole body SMB was 0.23 mg/kg, with a standard deviation of 0.19 mg/kg. This data was not included in the FS.

In summary, the mean 2012 PCB SMB tissue concentrations were lower than the mean concentrations of the combined 2002 and 2007 SMB data that were used in the RI/FS and risk assessments on an RM and Study Area basis, except for RM 10 (Kennedy/Jenks 2013a). Based on statistical comparisons of the two data sets on a Study Area-wide scale, total PCB congener concentrations in whole body SMB tissue show a statistically significant ($p<0.05$) decrease from the 2002 and 2007 data (Kennedy/Jenks 2013b, LSS 2015). The 2012 SMB data support that natural recovery is occurring on a system-wide scale.

Formatted: Highlight

Commented [Author2]: This statement is based on a study performed without EPA oversight. EPA would need to review the data and associated documents to evaluated the validity of this statement.

Formatted: Highlight

<u>Characterization of the Lower Willamette River with Sediment Profile Imaging, Changes in Space and Time, Germano and Associates, June 2014</u>

The purpose of the study was to provide information on the physical and biological features of the surface sediments in the Lower Willamette River through specialized photography and compare to similar work performed in many of the same locations as work performed in 2001 by the LWG during the RI. The 2014 effort used the same people and protocol as the 2001 work, but was enhanced by updated technology (Germano and Associates 2014). Results of the SPI showed significant recovery in benthic infaunal successional stage compared to the 2001 RI results. These results support the fish tissue studies indicating natural recovery is occurring throughout the Study Area.

<u>Final Supplemental RI/FS Study Field Sampling and Data Report, River Mile 11 East, GSI, July 2014</u>

The River Mile 11 East Early Action Area is part of the Portland Harbor Superfund Site (PHSS) and was identified as a PCB "hot spot" which required accelerated remediation as per the EPA's settlement agreement in 2013. Surface sediment sampling was conducted in October 2013. The limited-access surface sediment samples were collected by divers using a hand-coring device during May 2014 (GSI 2014).

Nine surface (bank) soil samples were collected and analyzed for PCB Aroclors, hydrocarbons (diesel range and residual range hydrocarbons), PAHs, pesticides, metals, phthalates, SVOCs, TOC, total solids, grain size, and D/F. Surface soil samples were composited and represented the 0- to 1-foot depth (GSI 2014).

22 surface sediment samples were collected and analyzed for PCB Aroclors, TOC, total solids, and grain-size distribution. In addition, samples from six re-occupied stations were analyzed for organochlorine pesticides. Of the 22 samples, 12 were collected by a pneumatic power-grab sampler, and 10 samples were collected by divers. The target depth for surface sediment samples was 30 centimeters bml (with a minimum acceptable penetration of 20 centimeters) (GSI 2014).

Total PCB Aroclors were analyzed in 22 surface sediment samples with 100% detection frequency. Concentrations of total PCB Aroclors ranged from 2.5 µg/kg to 1,405 µg/kg, with a median concentration of 93.5 µg/kg. Total dichlorodiphenyltrichloroethane (DDT) was analyzed in six surface sediment samples with 100% detection frequency. Concentrations of total DDT ranged from 0.24 µg/kg to 9.5 µg/kg, with a median concentration of 2.05 µg/kg. Total D/F were analyzed in four samples with 100% detection frequency. Concentrations of total D/F ranged from 556 ng/kg to 2,160 ng/kg, with a median concentration of 1,357.5 ng/kg (GSI 2014).

Sediment Sampling Data Report, Portland Harbor, Kleinfelder, June 2015

The purpose of the 2014 sediment investigation was to: (i) assess the current concentrations of PCB Aroclors in surface sediments (0-30 centimeters) from RM 2 to RM 16.2; (ii) provide data to compare with prior results and with concentrations predicted by the sediment recovery food web model (FWM); and (iii) develop a dataset representative of current PCB concentrations to be used in developing future remedial actions. Samples were collected from November to December 2014 (Kleinfelder 2015).

Within the PHSS, 98 surface sediment samples were collected, and 27 surface sediment samples were collected within the Downtown Reach (RM 11.8 to RM 16). Samples were collected using a hydraulic power-grab sampler (maximum penetration of 30 centimeters bml) and analyzed for PCB Aroclors, TOC, and grain size (Kleinfelder 2015).

Total PCBs were detected in 113 of 125 (90%) surface sediment samples. In Site sediment samples, total PCBs ranged from <0.7 µg/kg to 5,180 µg/kg. One sample was reported at 7,420 µg/kg; however, due to analytical interference, this sample was flagged as non-detect. In the upstream area (RM 11.8 to 16.2), total PCBs ranged from <0.7 µg/kg to 61.1 µg/kg. TOC ranged from 820 mg/kg wet weight to 35,000 mg/kg wet weight. Grain size results for the upstream area showed a lower percentage of fines (silt and clay) compared to Site samples (Kleinfelder 2015). Results showed that PCBs in surface sediments were generally lower when compared to RI data co-located stations suggesting that natural recovery is occurring.

Concentrations and Character of PAH in Sediments in Area of River Miles 5 to 6, 2015 Investigation, NewFields, March 2016

Two sampling events were conducted during 2014 and 2015 to investigate the nature and extent of PAHs in sediments in an area between RM 5 and the St. Johns Bridge (RM 6). The potential for Principal Threat Waste (PTW) was also assessed for the various possible dredge horizon intervals. Sediment samples were analyzed for PCBs, D/F, DDx, and carcinogenic polycyclic aromatic hydrocarbons (cPAHs) (Benzo[a]pyrene Equivalents [BaP Eq]) (NewFields 2016).

The sediment bed depth intervals sampled in this study included: (i) Interval A, Surface, 0 to 1-foot bml; (ii) Interval B, Future Channel, -48 to -49 feet below Columbia River Datum (CRD); (iii) Interval C, Future Overdredge, -51 to -52 feet below CRD; and (iv) Interval D, Future Overdredge (plus cap buffer), -53 to -54 feet below CRD. 53 samples were taken at Sampling Interval A, 15 samples were taken at Interval B, 34 samples were taken at Interval C, and 40 samples were taken at Interval D (NewFields 2016).

The highest cPAH concentrations were detected in the proximity of the former Gasco property and were consistent with pyrogenic manufactured gas plant (MGP)-derived tar wastes. Very few sediment samples from this study detected petroleum-derived PAHs. The mean concentrations of other COCs (PCBs, DDx, selected chlorinated D/F isomers, BaP Eq) in this study did not exceed PTW classifications (NewFields 2016).

Total PCB Aroclors were detected in 26 of 31 (84%) sediment samples, with concentrations ranging from <0.02 µg/kg to 27.8 µg/kg. Total DDx were detected in 22 of 31 (71%) samples, with concentrations ranging from <0.036 µg/kg to 58.3 µg/kg. Total PAHs were detected in 100% of 150 samples, with concentrations ranging from 1.3 µg/kg to 1,376,830 µg/kg (reported as sum of 17 PAHs). Total D/F were detected in 100% of 31 samples, with concentrations ranging from 5.9 ng/kg to 5,291 ng/kg (NewFields 2016).

Sediment Sampling Data Report, Swan Island Lagoon, Geosyntec Consultants, August 2016

20 surface sediment (0-30 centimeters) samples were collected within Swan Island Lagoon (RM 8 to 9) during March 2016. Samples were analyzed for PCB Aroclors, TOC, and grain size. Sediment sample locations were co-located with previously sampled locations by the LWG for the RI/FS (1998-2007). The purpose of the study was to evaluate if natural recovery of sediments is occurring in Swan Island Lagoon by comparing the 2016 results to the older RI/FS results (Geosyntec 2016).

PCB Aroclors were detected in all 20 surface sediment samples, with concentrations ranging from 33.6 µg/kg to 996 µg/kg. 75% of these samples showed reduced PCB concentrations when compared with sample results collected over 10 years ago. These results also confirmed trends seen with PCB concentrations found in surface sediment samples collected by the 2015 Kleinfelder study. Recent sampling indicates that newly deposited sediments are covering and/or mixing with the older surface sediments both river-wide and in Swan Island Lagoon (Geosyntec 2016).

REFERENCES

Geosyntec Consultants, Inc. (Geosyntec), 2016. Sediment Sampling Data Report, Swan Island Lagoon, Portland, Oregon. Prepared for The Marine Group, LLC and BAE Systems San Diego Ship Repair, Inc. 12 August.

Germano and Associates, 2014. Characterization of the Lower Willamette River with Sediment Profile Imaging, Changes in Space and Time. June.

GSI Water Solutions, Inc. (GSI), and Hart Crowser, Inc., 2010. Field and Data Report, Downtown Portland Sediment Characterization Phase II, Willamette River, Portland, Oregon. Prepared for Oregon Department of Environmental Quality. June.

GSI, 2011. Sampling and Analysis Plan, Portland Harbor 2011 Baseline Smallmouth Bass Tissue Study, Willamette River, Portland, Oregon. Prepared for EPA, United States Army Corps of Engineers, and City of Portland. September.

GSI, 2014. Supplemental Remedial Investigation/Feasibility Study Field Sampling and Data Report, River Mile 11 East, Portland, Oregon. Prepared for RM11E Group. July.

Kennedy/Jenks, 2013a. Portland Harbor RI/FS, 2012 Smallmouth Bass Tissue Study, Data Report. Draft. Prepared for Lower Willamette Group. 13 March.

Kennedy/Jenks. 2013b. Statistical Comparison of Historical and 2012 Smallmouth Bass Data. Prepared for Lower Willamette Group. 6 March.

Kleinfelder, 2015. Sediment Sampling Data Report, Portland Harbor, Portland, Oregon. Prepared for de maximis Inc. 1 June.

Legacy Site Services (LSS), 2015. An Assessment of the Coupled Sediment Recovery and Dynamic Food Web Model: Predicting the Concentrations of Total PCBs in Lower Willamette Fish Tissue Based on 2002 to 2012 Sampling Data. Legacy Site Services LLC, White Salmon, Washington. August.

NewFields, 2016. Concentrations and Character of PAH in Sediments in the Proposed Remedial Alternatives Area of the Portland Harbor Superfund Site, River Miles 5-6 2015 Investigation. Prepared for ExxonMobil. 29 March.

ODEQ, 2011. Portland – Willamette River Sediment Evaluation Downtown Reach – Phase II Follow-up Summary. 4 November

DRAFT

APPENDIX B

APPROACH FOR SAMPLING OF SURFACE SEDIMENT

Pre-Remedial Design Work Plan
Portland Harbor Superfund Site

INTRODUCTION

This appendix details the proposed design and evaluation of the proposed 640-sample surface sediment sampling plan. It is organized into two sections, as shown below:

- Section B1: Sample Placement and Design; and
- Section B2: Sample Plan Statistical Evaluation.

The proposed plan consists of 428 unbiased samples and 212 biased samples located in the Sediment Management Areas (SMAs), for a total of 640 samples. The density of unbiased samples is approximately 20 samples per bisected River Mile segment (distributed in in 20 segments, plus a 20-sample segment in Swan Island Lagoon and 8 additional samples in inlet areas). Approximately 160 of the unbiased sample locations will fall within/immediately adjacent to the SMAs, enabling the delineation of the SMAs to take advantage of approximately 370 samples (212 SMA sample plus approximately 160 unbiased samples = approximately 370 samples).

This density of SMA sampling achieves the level of precision needed to delineate SMAs within an approximate 30% level of uncertainty, sufficient for the needs of the Pre-RD group and an initial dataset to inform on additional SMA delineation sampling events needed for higher levels of engineering design. Additionally, the sampling plan is sufficient to meet the needs of the Pre-RD group in evaluating natural recovery since 2004 and developing statistically robust Spatially Weighted Average Concentrations (SWACs) on Site-wide and other spatial scales. The proposed 428 unbiased samples also meets the unbiased sample count necessary to meet EPA's data objectives, in addition to the Pre-RD group's objectives.

Section B1: Sample Placement and Design

Placement of the 640 surface sediment samples will be achieved through the following steps, and a conceptual figure showing the distribution of the samples is shown in Figure 1:

1. A sample grid consisting of 428 cells will be devised for the Site (River Miles 1.9 to 11.8). This grid will be used to place the unbiased (randomly placed) samples. There will be three types of cells:

Commented [Author3]: As presented in this appendix, a substantial portion of the baseline program centers around the objective of testing for temporal changes in COC concentrations by comparing 2018 samples collected in the vicinity of selected 2004 sampling locations. Because, smallmouth bass provide the exposure pathway for human health risks that are being managed, estimated natural recovery rates must accurately reflect broad river-mile or site-wide scale changes in sediment COC concentrations so that recovery rates can be extrapolated to expected rates in biota. While the EPA agrees there is a need to understand natural recovery rates, comparison of 2018 samples with 2004 samples which were located using a biased sampling procedure will preclude reliable estimation of the natural recovery rate in sediment. Because the 2004 sampling locations were selected using a biased subjective sampling design, there is no way to extrapolate the 2004 data to create an unbiased estimate of conditions in 2004. Pre-RD group has proposed Thiessen polygon weighting as a means to calculate estimates, but this method cannot be relied upon for unbiased estimation. The Pre-RD group is either unaware or has not been forthcoming with this situation. The documents as they stand may mislead the reader with regard to the utility of Thiessen Polygon and other weighting schemes for integrating biased and unbiased samples. Lacking this unbiased characterization in 2004, comparisons with similarly biased locations in 2018 do not provide a valid basis to draw conclusion about recovery rates over site-wide or river-mile scales. Because of this limitation the EPA maintains the position that corrupting the unbiased sampling design for baseline data is not warranted and wastes resources which could be otherwise allocated to support statistically rigorous evaluations of sediment analytical data.

Commented [Author4]: The simulation study supporting SMA delineation sampling density fails to recognize that the key components to uncertainty in the SMA footprint are the strength of spatial correlation and the magnitude of short scale heterogeneity (nugget effect).

Commented [Author5]: The document fails to address the data needs for equivalence testing which EPA has proposed as an approach to determine when remedial goals have been met. As a result, EPA assumes that the Pre-RD group plans to consider the cleanup level of 9 ug/kg PCBs and other similar background based values for other CoCs as the fixed values which must be attained on sight irrespective of COC concentrations in the downtown reach or upstream of the downtown reach. If this was an oversight by the Pre-RD group, EPA recommends that the Pre-RD group increase sample sizes to those identified by EPA's Appendix A based on equivalence testing. The equivalence DQO is the driver of sample size for most of the program recommended by EPA. Sample sizes proposed in the Pre-RD work plan and appendices are in adequate to demonstrate equivalence.

a. **Shoal cells:** Shoal cells will be placed parallel along the center channel flow path (thalweg) of the center of the navigation channel in the areas of the river outside of the navigational channel (shoals). One set of shoal cells will be placed to the east of the navigational channel, and one set of shoal cells will be placed to the west side of the navigational channel. The length of each shoal cell will be approximately 0.067 miles such that approximately 30 shoal cells (15 on east shoal, 15 on west shoal) will be present per River Mile. The width of each shoal cell will be the distance between the navigational channel boundary and the river bank. 150 shoal cells will be placed on the east side of the river and 150 will be placed on the west side of the river for a total of 300 shoal cells.

b. **Navigational cells:** Navigational cells will be placed parallel along the thalweg of the center of the navigation channel in the areas of the river within the navigational channel. One set of navigational cells will be placed to the east of the navigational channel thalweg centerline, and one set of navigational cells will be placed to the west side. The length of each navigational cell will be approximately 0.2 miles such that approximately 10 navigational cells (5 on eastern half of the channel, 5 on western half of the channel) will be present per River Mile. The width of each navigational cell will be the distance between the navigational channel boundary and the navigational channel thalweg centerline. 50 navigational cells will be placed on the eastern half of the navigational channel, and 50 navigational cells will be placed on the eastern half of the navigational channel.

c. **Inlet cells:** Five areas of the Site (e.g., Swan Island Lagoon) do not conform well to the shoal grid placement, so these areas will be identified as inlets. In these areas, inlet cells will be placed such that the length of each cell is approximately 0.067 miles and the width is the bank-to-bank width of the inlet. Thus, these areas will be sampled at the same density as the shoal areas that border the navigational channel. 8 inlet cells will be placed in the smaller inlets, and 20 cells will be placed within Swan Island Lagoon.

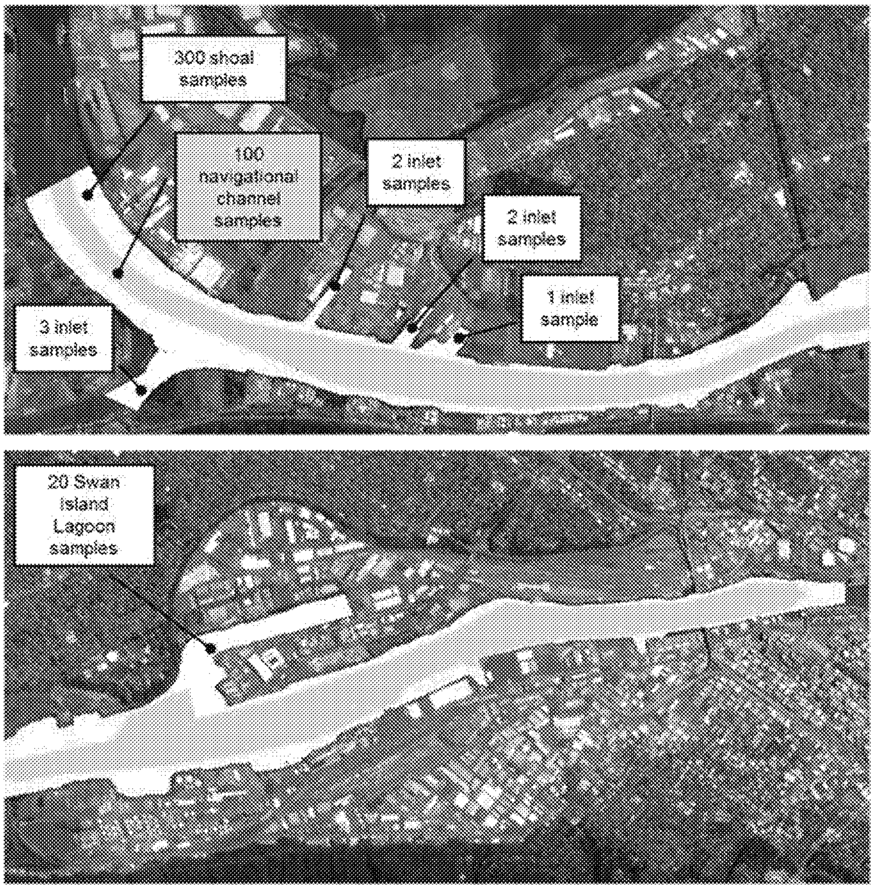


Figure 1. Conceptual distribution of the 428 unbiased samples among the shoal (light blue), navigational channel (purple), and inlet/Swan Island Lagoon (pale yellow) areas.

The 428-cells will be designed such that:

- 3 times the number of samples will be placed in the shoals relative to the navigational channel. The addition of inlet samples (which are in inlet sections of the river shoals) will slightly increase the number of shoal samples relative to navigational channels, such that 23% of the final sample count will represent navigational channel samples.
- The sample design will be evenly balanced between east and west. Due the presence of inlets, approximately: 225 (53%) of samples will be on the eastern side and 203 (47%) of will be on the western side.
- Approximately 40 randomized samples will be present per each River Mile (20 on east and 20 on west). This targets 15 samples on each shoal and 5 samples on each side of the navigational channel.
- Robust sample sizes will be achieved in four segments (Segment 1 = River Mile 9 to 11.9, Segment 2 = River Mile 7.5 to 9, Segment 3 = River Mile 5 to 7.5, and Segment 4 = River Mile 1.9 to 5), as well as the four segments bisected into east and west portions, and Swan Island Lagoon (these segments are shown in Figure 3 of the Workplan).

2. In each bisected River Mile (starting with 1.9 to 2.9 on the East side), 6 cells containing at least 1 surface sediment sample station sampled in 2004 will be randomly selected for reoccupation. In 11E and 11W, no cells can be selected for reoccupation due to lack of 2004 sample locations. This will result in approximately 108 cells selected for reoccupation.
3. In each cell, a location will be randomly determined to receive the sample location. The sample location will be selected via a random number generator to determine the x- and y-coordinates. The randomization process will exclude areas that were within 20 feet of the shoreline to avoid placing samples in the intertidal zones or in areas that could not be reached with the sample vessel. Random sample locations will be visually inspected (aerial imagery) to ensure that the sample will not be placed in an area that could not be reached with the sample vessel (e.g., underneath piers or other fixed infrastructure), with a re-randomization of the location if needed.

Commented [Author6]: The document refers to 428 unbiased locations whereas there are only 320 unbiased because 108 of the 428 locations will be moved to 2004 sampling locations, and so are biased. To meet the DQOs, the 108 sample locations must be truly unbiased.

Biased samples should clearly tracked so that these data can be properly identified in databases as biased data attempting to relocate on 2004 sampling locations.

Commented [Author7]: These 108 sample locations are biased because they will be moved to biased 2004 sampling locations.

4. If the cell will be randomly selected for reoccupation (step 2 above), a location will be selected at random from the cell and then shifted to reoccupy the nearest 2004 sample location.

5. 212 additional samples will be added to enable 2-dimensinoal delineation of the Alternative F Mod active Sediment Management Area (SMA) footprint. These samples ("SMA Samples") are not unbiased, randomly placed samples. They will be placed in two successive steps.

   a. The first step in SMA sample placement will assume a surface sediment sample will be placed at each of the 60 proposed sediment core locations (the 30 proposed bank core locations are not included).

   b. The second step in SMA sample placement will consist of selecting 152 sample locations within and immediately adjacent to the SMAs such that a sediment sample would be present approximately every 300 feet within the SMAs.

The 212 SMA samples specifically placed within and immediately adjacent to the SMAs plus the unbiased samples placed randomly within the SMAs (approximately 166 unbiased samples) will result in approximately 378 samples total within and immediately adjacent to the SMAs. Additionally, many unbiased sample locations near the SMAs will also be used to aid in delineation.

In summary, the 640 surface sample count will have the following distributions:

- 428 unbiased samples and 212 biased SMA samples (640 total);
- 108 re-occupied stations are expected to be evenly distributed across the range of PCB concentrations, as compared to the full 2004 dataset (Figure 2);
- Approximately 20 samples per bisected River Mile (distributed in in 20 segments plus Swan Island Lagoon), achieving EPA's goal of approximately 20 samples per rolling River Mile can be achieved in the 20 segments, plus Swan Island Lagoon;
- Among the unbiased samples, 100 will be located in the navigation channel and 328 will be located in the shoaled areas (approximate 25/75 distribution); and

- The unbiased samples will be distributed among the 9 segments as shown in Table 1 (Figure 3 of the workplan).

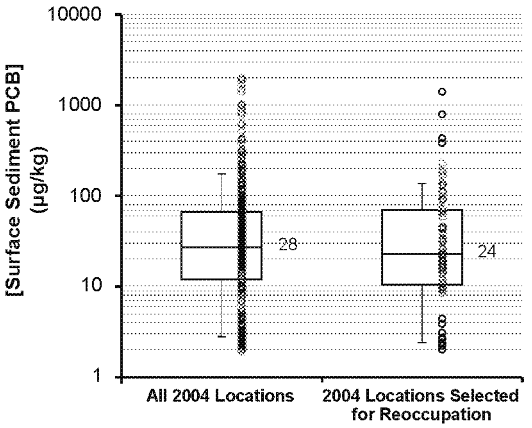


Figure 2. Concentrations of PCBs in surface sediments in all 2004 surface sediment sample locations compared to the 87 locations selected for re-occupation in an earlier draft proposed plan. This figure will be updated in the Sampling and Analysis Plan (SAP) following the placement of the approximate 108 sample reoccupations. The boxes represent the interquartile ranges of the data, whiskers represent the 10th and 90th percentiles, the empty circles represent individual data points, and the labels indicate the geometric mean of concentrations.

Table 1. Expected distribution of the 428 unbiased samples within each of the 9 segments.

Commented [Author8]: Of the 428 samples, 108 sample locations are biased because they will be moved to biased 2004 sampling locations.

Segment	Number of Unbiased Samples
1W	58
1E	59
2W	29
2E	30
Swan Island Lagoon	20
3W	50
3E	48
4W	63
4E	71
Total	428

Section B2: Sample Plan Statistical Evaluation

The ability of the sample plan to meet various data objectives in a statistically robust manner was evaluated using several statistical approaches, as detailed below. It should be noted that these approaches are conservative and tend to overestimate sample size requirements, as they are based on the 2004 PCB data. Recently-collected data have suggested concentrations of PCBs in surface sediment have decreased significantly since 2004 such that many areas are converging on lower, ambient levels of PCBs. This convergence would result in lower variability in hypothetical 2018 data (and future data), which would greatly improve the precision and power of data and lower necessary sample size requirements for the data evaluations discussed below.

Additionally, the below statistical evaluations are considered initial efforts and may be augmented by additional or alternate evaluations (as needed) if any revisions to the sampling plan are considered or other aspects of the data objectives are emphasized.

The Plan is Sufficient for SMA Delineation

Commented [Author9]: The data will not be sufficient to define SMAs—but can assist in developing the SMA locations to be further defined during remedial design through additional sampling.

The proposed sample placement density of the 378 samples within/adjacent to the SMAs is approximately 1 sample per 70,000 ft^2 and a spacing of approximately 290 to 300 ft between each sample. Via a Monte Carlo simulation using a hypothetical 20-sample evaluation area and an approximated Site-wide 2004 PCB sediment data distribution, we calculated the average area that would be indicated by the number of samples that would exceed the PCB RAL of 75 µg/kg and indicate designation for inclusion within an SMA. The Monte Carlo simulation calculated the sum of the individual 70,000-ft^2 areas 1000 times to simulate potential results for the final total SMA area sizes. 80% of the total SMA area estimates (area that would be indicated above the PCB RAL) were within ± 25% to 38% (average of 31%) of the median estimate. This approximate level of ±30% precision was considered satisfactory for the initial SMA delineation.

The Plan is Sufficient for Spatially Weighted Average Concentration (SWAC) Determination

Commented [Author10]: The proposed plan is not sufficient for recalculating SWACs because not all of the 428 surface sediment samples are unbiased.

Spatially Weighted Average Concentrations (SWACs) for PCBs in the proposed plan were calculated using a simplified approach in which equal polygon areas were assumed for each proposed sample point and a random concentration of PCBs were assigned from approximated 2004 PCB sediment data distributions for each of the 8 segments. SWACs were calculated for each of the 8 segments shown in Table 1, assuming approximately 40 samples were present in each segment, except for segment 2W, which is expected to receive fewer samples due to its shoreline configuration and River Mile length. Actual sample sizes for the final sample plan are expected to be higher (Table 1), and this analysis will be updated in the SAP. Using Monte Carlo simulation, the PCB SWAC and 95% Confidence Intervals (95% CIs) around the SWAC were calculated 1000 times for each the 8 segments. Swan Island Lagoon, within 2E, was also assessed in an additional evaluation (segment 9) and it was noted that the proposed sample size and placement would be sufficiently robust to meet SWAC data objectives if Swan Island Lagoon is evaluated as its own segment.

The projected SWACs are shown in Figure 3, and are compared to the SWACs obtained from the 2004 data. On average, the 95% CIs of the projected SWACs[1] (that would be obtained using the sample plan) are 10% more precise than the 95% CIs calculated using the actual 2004 data and spatial placement of 2004 samples. Thus, the sampling plan will be able to generate SWACs that are as precise or more precise than the SWACs generated with the historical data used as the basis for the remedial investigation, feasibility study,

1 The median of the 1000 SWAC 95%CIs (for each segment) was evaluated.

and record of decision at the scale of evaluation (8 segments plus Swan Island Lagoon, as shown in Figure 3 of the work plan).

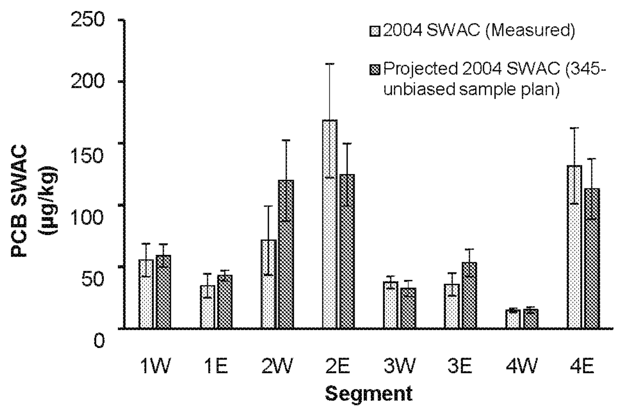


Figure 3. SWACs calculated using the 2004 sample locations and results (grey bars) and projected using a Monte Carlo simulation of the sample sizes in each segment, as placed in the sample plan (green bars). Note: this figure was generated using an earlier draft proposed plan that included fewer samples. This figure will be updated in the SAP, and overall statistical robustness of the plan will increase due to the additional samples.

It should also be noted that this projected SWAC analysis was performed assuming that the samples would not be composited. The 3-point composting approach is likely to reduce variability by approximately 50% on a Site-wide basis (see below), so the precision of the projected SWACs (shown in Figure 3) is likely to be reduced significantly such that precision of the SWACs in all segments (including 4W and 3W) should be more precise from those calculated in 2004.

Commented [Author11]: Figure 3 is provided in efforts to argue that the locations to be re-occupied are not biased relative to the 2004 sampling locations, however, this figure merely shows that the 2018 data locations will reproduce the bias of the 2004 data. These distributions may not be representative of the actual distribution of COCs in general because of the unknown bias (most likely high bias) in the 2004 data which were located based on a subjective sampling design. As proposed, approximately 25% of 2018 samples will be biased toward 2004 locations causing an unknown bias in future analyses dependent on these sample locations. This bias will carry forward into the LTM program as it is anticipated that monitoring will be conducted by re-occupying the same locations through time. All future analyses will need to be provided with the caveat that parameter estimates are biased and the direction and magnitude of this bias is unknown.

This will bias future comparisons to long term monitoring data and limit the scope of statistical inference to the sample locations themselves. In other words inference to the unknown means and recovery rates (*i.e.* the subject of the studies) will be limited and estimates will be subject to unknown biases. These limitations are unnecessary because there is little or no advantage to locating 25% of the sample at the 2004 locations. The Pre-RD group has provided no evidence that any of their planned comparisons will have greater accuracy or precision as a result of re-occupying previously sampled locations.

The Plan is Sufficient for Comparing Differences in Concentrations Since 2004

A variety of data evaluation approaches will likely be used to evaluate differences in concentrations of PCBs in surface sediment between 2004 and 2018 (or going forward by EPA), including regression and multi-variate analyses and consideration of various other data, including water depth, grain size, organic carbon, and other sample location-specific factors that influence natural recovery processes. Data will also be modeled and evaluated spatially, including statistical comparisons of SWACs estimated in 2004 and 2018.

A simple analysis would include a comparison of 2004 and 2018 data at the stations proposed for reoccupation. A Site-wide power analysis indicated that approximately 90 reoccupations of the 2004 samples, if paired and compared on a log-transformed basis (n = 90), will be sufficient to detect an approximate 40% or more reduction in PCB concentrations (power of 0.8, α of 0.05). This is likely conservative (i.e., lower levels of reduction will likely to be detectable) given the analysis was based on single samples (not composited samples, as identified in the sampling) plan, and the fact that 2018 are likely to be of lower variability than 2004 data. The approximate 40% reduction from 2004 to 2018 (14 years) is less than that hypothesized by EPA in the natural recovery trend analysis (i.e., assumption of 5% to 10% per year, which would equate to a 50% or greater reduction in concentrations over the 2004 to 2018 time period). Additionally, as proposed, approximately 108 samples will be used for the reoccupation, further increasing power (analysis will be updated in the SAP).

Assuming a reduction in variability as a result of the 3-point compositing approach proposed in the workplan[2] the 87 reoccupations should be able to detect a 25% (or greater) difference at a Site-wide scale. At the segment level, this approach should enable approximate 30 to 40% differences (or more) to be detected. The statistical power at this level should be further augmented by additional variables and approaches (as described above) or sensitivity analyses that include portions (or the entirety) of data from samples collected in the SMAs. In particular, spatial weighing of the data (using SWAC or equivalent geostatistical analyses) is likely to greatly improve the power of the statistical

Commented [Author12]: This term is inaccurate. There are no geostatisical analyses in the report.

2 Variability of the data should decrease from an approximate 40% coefficient of variation (CV) to a 20% CV, based on Monte Carlo simulation.

comparisons, enabling in lower levels of differences between 2004 and 2018 SWACs to be detected at the 8-segment scale.

The Plan is Sufficient for PCB Natural Recovery Power Analysis

Commented [Author13]: The analysis supporting numbers for sediment is based on precision of site wide SWAC and has no real power analysis associated with it. See EPA Appendix A for sample size design using power analysis (provided in EPA attachment).

~~As noted above, the proposed plan will achieve approximately 20 samples per bisected River Mile, and each of these samples is represented by a 3-point composite. EPA's statistical power analysis (Figure 9 of Appendix A of the June 6 draft plan[3], shown in Figure 4 below) indicates that 20 samples bisected River Mile segment would likely be sufficient to achieve statistical power to evaluate their stated goal of evaluating natural recovery trends for PCBs (the basis of the 20 unbiased samples per bisected River Mile segment sample size determination). The proposed sampling approach is projected at a level of statistical power of 0.9 power for PCBs according to EPA calculations[4].~~

~~3 EPA. 2017. Portland Harbor Superfund Site Sampling Plan for Pre-Remedial Design, Baseline, and Long-Term Monitoring. Revised Working Draft. June 6.~~

~~4 This power analysis was based on PCBs; the same analysis based on DDx and PAHs indicates higher sample sizes are needed, but these chemicals are located in hotspots at the Site such that the power analysis likely overestimates the sample sizes needed.~~

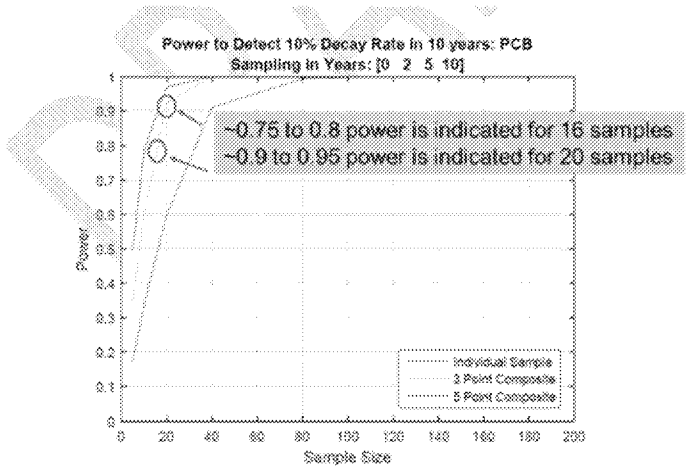


Figure 9. Power to detect a 10% annualized recovery rate in sediment by monitoring in year 0, 2, 5 and 10 for total PCBs.

~~**Figure 4.** Reproduction of Figure 9 in Appendix A of EPA's draft June 6, 2017 sampling plan. The green ovals, arrows, and text box have been added (see discussion).~~

Summary of Plan Evaluation

In summary, the above evaluations indicate the proposed plan is sufficiently robust for addressing the Pre-RD group data needs:

Commented [Author14]: See EPA's above comments

- SMA areas will be delineated at an approximate ± 30% level of uncertainty;
- SWACs for each of the segments will be as precise or more precise than previously-calculated SWACs; and
- Differences in sediment concentrations (including SWAC and other data expressions) between 2004 and 2018 will be statistically detectable ($\alpha = 0.05$) an approximate 0.8 level of statistical power, assuming concentrations have decreased approximately 25-40% or more since 2004.

In addition, the proposed sampling density achieves a statistical power of 0.9 based on one of the key power analyses used by EPA to support their approximate sample size requirements.

DRAFT

APPENDIX C

POWER ANALYSIS FOR FISH TISSUE SAMPLE SIZE
PREPARED BY AECOM

Pre-Remedial Design Work Plan
Portland Harbor Superfund Site

POWER ANALYSIS FOR FISH TISSUE SAMPLE SIZE, PREPARED BY AECOM

A statistical analysis was performed using the 2012 smallmouth bass (SMB) tissue data set for the Portland Harbor Superfund Site (Kennedy/Jenks Consultants, 2013) to estimate sample sizes for re-baselining fish tissue chemistry. Prior analyses of 2002, 2007, and 2012 data for polychlorinated biphenyls (PCBs) in SMB indicate fish tissue concentrations have declined over time. The primary focus of this analysis was to estimate the number of SMB samples needed to determine with statistical significance whether PCB concentrations have continued to decline since 2012. Comparison of historical and new means and 95% confidence intervals, as well as trend analysis, provide robust and statistically-valid approaches for evaluating the data to evaluate natural recovery, and are likely to be used following data collection in 2018.

An important aspect of monitoring design is power analysis, which was performed to estimate the difference that can be detected between the 2012 and re-baselining fish tissue data sets. This was illustrated by calculating the effect size (d) or the standardized difference between two means that could be detected (Cohen, 1988); and the minimum detectable difference (MDD) as a function of sample size and variance. The MDD is the difference between two means that must exist to detect a statistically significant effect/difference. The MDD is a commonly employed technique to indicate the potential significant difference at a given sample size in fish tissue monitoring (e.g., United States Environmental Protection Agency [EPA], 2000).

METHODS AND RESULTS

Commented [Author15]: The fish numbers are based on a flawed analysis that develops numbers of samples to estimate 30% - 40% change in log-base 10 concentrations. This is misleading because a 30% change in log-10 numbers averaging 2.56 ug/kg (e.g. site wide mean SMB) is equivalent to an approximately 83% change in actual concentrations {1-10^(2.56*0.7)/10^2.56}. We do not expect to see 83% changes for decades so the analyses are wholly inappropriate for the time frames at which we will want to identify changes in concentration and remedial effectiveness.

Two scenarios were evaluated: (i) assuming sample size equivalent to the 2012 smallmouth bass data set (n=83 Study Area, n= 9 upriver background); and (ii) assuming sample size equivalent to the targets set forth in the 2012 work plan (n=95 Study Area, n=10 upriver background) (Windward Environmental, 2012). The sample size calculations were performed considering two spatial scales: (i) site-wide; and (ii) four 2 to 3-mile segments.[1] Both spatial scales are relevant to evaluation of SMB in the risk assessment and bioaccumulation modeling.[2] Sample size estimation

[1] The segmentation is based on prior analysis of the available data (Wolf, 2015).

[2] A spatial scale of 2 to 3 mile segments may better reflect the home range of SMB which ranges from 0.5 to 5.5 miles (Scott and Crossman, 1998). Based on a radio-tracking study of predator species in the Lower Willamette River, the median total distance traveled (upstream and downstream) by SMB was 4.3 kilometers (km) (2.7 miles), with 25th and 75th percentiles of 0.8 km and 8.0 km (Friesen, 2005).

to support statistical analysis of trend/natural recovery on an individual river mile (RM) basis is not a data use objective; large sample sizes would be needed to detect trends with adequate power. The power analysis was performed using the following assumptions:

- Confidence = 95% (alpha [α] = 0.05);
- Power = 80% (1-ß = 0.80); and
- PCB data lognormally distributed.

The assumption that PCB concentrations in the 2012 whole body SMB data set are lognormally distributed is based on distribution testing performed using @Risk v.6 (Palisades, 2012). Table 1 presents summary statistics for the 2012 SMB data set. PCB fish tissue concentrations were log (base 10) transformed for the power analyses. The assumptions of 95% confidence ($\alpha = 0.05$) and power of 80% are used by convention to support statistically significant results (e.g., EPA, 2000). However, statistically robust changes may still be concluded using less stringent hypothesis test statistics (e.g., lower confidence or power).

The calculation of effect size (Cohen's d) was performed using the pwr package (Champely, 2017) in the R statistical computing (R Core Team, 2015). The calculation of the minimum detectable difference (MDD) as a function of the sample size and variance in the 2012 data was performed using the following formula and expressed as a percentage of the 2012 mean[3] (Harcum and Dressing, 2015):

MDD = √[(4σ2 (Z(1-(α/2)) + Z(1-β)) 2) ÷ N]

where:

N = total sample size (number of samples in 2012 and new baseline)
σ = standard deviation (assumed equal to 2012 sample populations)
Z(1-(α/2)) = 1.96
Z(1-ß) = 0.84
MDD = minimum detectable difference between 2012 and new baseline means

[3] While log transforming the data results in the power analysis being on the population geometric mean, results using the techniques described here are considered to be adequate practical approximations for the purpose of this analysis.

Commented [Author16]: The minimum detectable difference for smaller areas with less fish samples (i.e. N=20 or so) would be as high as 63.5% in log-10 scale which is equivalent to 97% reduction in actual concentration (ug/kg) scale. In other words, the design, according to this calculations, would not be able to detect changes for decades. **Yet the footnote on this page states:** ***"While log transforming the data results in the power analysis being on the population geometric mean, results using the techniques described here are considered to be adequate practical approximations for the purpose of this analysis."***

This statement acknowledges that the analysis is incorrect for original scale data, but is incorrect in asserting that it is applicable to the geometric mean. The analysis and reported percentages are applicable to log-10 scales. The corresponding changes in geometric mean are what is described above. The Minimum Detectable Differences for geomean are very large and will not be expected for decades.

The effect size analysis indicates that replicating the 2012 sample sizes (actual or target sample sizes) will allow for moderate differences between the means to be detected on a site-wide basis (Cohen's d of ~0.4). For the four segments, the 2012 sample sizes will allow for large differences between the means to be detected (Cohen's d of ~0.8).[4]

Results of the MDD analysis are summarized in Table 2 for the full Site (RM 2-12), each of the four segments, and the upriver background area. The MDD analysis was also performed excluding Segment 1 (RM 9-12), which has the highest mean and variance; eight of the ten highest 2012 PCB SMB samples were from Segment 1, which includes the RM11 area. Table 2 includes MDD results for combined Segments 2, 3, and 4 (RM 2-9).

As shown in Table 2, sample sizes consistent with the 2012 program (actual or target) result in a MDD of about 30% on a site-wide basis. The MDD improves slightly (about 1%) when the 2012 target sample size is used (increase of 12 samples site-wide). When the area with high variance is removed (Segment 1), the MDD is about 23% for the remainder of the Site.

On a river segment basis, sample sizes of 20 to 23 result in MDDs of about 28% to 40% in Segments 2, 3, and 4. In Segment 1 (RM 9-12) where variance is highest, sample sizes of 22 to 28 result in a MDD of about 60%. For upriver background, a sample size of 9 to 10 results in an MDD of about 60%. For upriver background, the means for the 2002 (n=6 composite samples) and 2012 (n=9 discrete samples) SMB data sets are similar (170 micrograms per kilogram [µg/kg] and 230 µg/kg, respectively).

Based on prior trend analysis indicating a decline of about 4% per year (Nielsen, 2015), the maximum change that could be expected in the new baseline data would be approximately 24% in 2018 from the 2012 site-wide mean assuming a linear response (declines may become asymptotic over time as conditions reach equilibrium). Based on the MDD values calculated in this analysis, sample sizes consistent with the 2012 program should be sufficient to detect a change of this magnitude on a site-wide basis (excluding Segment 1/RM11 area) with a high degree of statistical significance.

---

[4] The Cohen's d statistic is a standardized measure of the size of the effect that can be observed between two means, with smaller values indicating smaller differences can be observed. Per Cohen (1977): 0.2 = small effect; 0.5 = moderate effect; and 0.8 = large effect.

REFERENCES

Champely, S., 2017. pwr: Basic Functions for Power Analysis. R package version 1.2-1. http://CRAN.R-project.org/package=pwr

Cohen, J., 1988. Statistical power analysis for the behavioral sciences (2nd ed.). Hillsdale, New Jersey: Lawrence Erlbaum.

Friesen, T.A., 2005. Biology, Behavior, and Resources of Resident and Anadromous Fish in the Willamette River, Final Report of Research, 2000-2004. Oregon Department of Fish and Wildlife. Contracted by the City of Portland. March.

Harcum, J.B., and S. A. Dressing, 2015. Technical Memorandum #3: Minimum Detectable Change and Power Analysis. Developed for U.S. Environmental Protection Agency by Tetra Tech, Inc., Fairfax, Virginia. October.

Kennedy/Jenks Consultants, 2013. Portland Harbor RI/FS, 2012 Smallmouth Bass Tissue Study, Data Report. Draft. Prepared for Lower Willamette Group. 13 March.

Nielsen, D., 2015. Temporal Regression Analysis of PCBs in Smallmouth Bass. Memorandum to Frederick Wolf, RETIA, USA. Draft. 2 July. Attachment C of Wolf (2015).

Palisades, 2012. @Risk. Risk Analysis and Simulation, Add-in for Microsoft Excel. Version 6, July 2012. Palisade Corporation, Ithaca, New York.

R Core Team, 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL. https://www.R-project.org/

Scott, W., and E. Crossman, 1998. Freshwater Fishes of Canada. Galt House Publications.

United States Environmental Protection Agency (EPA), 2000. Guidance for Assessing Chemical Contaminant Data for Use in Fish Advisories Volume 1 Fish Sampling and Analysis Third Edition. Office of Water. EPA 823-B-00-007. November.

Windward Environmental, 2012. Portland Harbor RI/FS: 2012 Modifications to the Field Sampling Plan for Bass Tissue. Prepared for Lower Willamette Group. 15 August.

Wolf, F., 2015. "An Assessment of the Coupled Sediment Recovery and Dynamic Food Web Model: Predicting the Concentrations of Total PCBs in Lower Willamette Fish Tissue Based on 2002 to 2012 Sampling Data". Technical Report, available on Researchgate.net. Presentation at Battelle Seventh International Sediments Conference in New Orleans, January 2015. Prepared by Legacy Site Services LLC. August, 2015

Table 1
Summary Statistics for 2012 Smallmouth Bass Data Set
Portland Harbor Superfund Site
Pre-Remedial Design Baseline Study

Domain	River Miles	2012 SMB Data (a)					
		n	Minimum (ug/kg)	Maximum (ug/kg)	Mean (ug/kg)	SD (ug/kg)	CV
Site-wide	2-12	83	92.4	6470	648.7	1185.9	1.8
Segment 1	9-12	22	134	6470	1531.6	2057.7	1.3
Segment 2	7.5-9 & SIL	23	118	1060	331.3	203.7	0.6
Segment 3	5-7.5	19	92.4	440	225.1	85.4	0.4
Segment 4	2-5	19	148	1280	434.2	299.8	0.7
Upriver Background	15-18	9	50.7	634	234.0	187.5	0.8

Notes
(a) PCB wet weight data from 2012 Smallmouth Bass Tissue Study (whole body). Kennedy/Jenks 2013.
n = sample size
CV = coefficient of variation
SD = standard deviation
SMB = smallmouth bass

Table 2
Summary of Power Analysis for Smallmouth Bass Sample Size
Portland Harbor Superfund Site
Pre-Remedial Design Baseline Study

Domain	River Miles	2012 SMB Data (a) n	Mean log10 PCB (ug/kg)	SD log10PCB (ug/kg)	Geometric Mean (ug/kg)	Sample Size Equals 2012 Actual Catch: MDD log PCB (ug/kg)	MDD (%)	Sample Size Equals 2012 Targets: n	MDD log PCB (ug/kg)	MDD (%)	Increase in MDD due to attaining targets
Site-wide	2-12	83	2.56	0.38	364	0.17	31.6%	95	0.16	30.7%	0.9%
Segments 2, 3 & 4 (b)	2-9	61	2.45	0.23	282	0.12	23.9%	67	0.12	23.4%	0.5%
Segment 1	9-12	22	2.87	0.52	745	0.44	63.5%	28	0.41	61.2%	2.4%
Segment 2	7.5-9 & SIL	23	2.46	0.22	290	0.18	33.8%	23	0.18	33.8%	0.0%
Segment 3	5-7.5	19	2.32	0.16	211	0.14	28.3%	21	0.14	27.6%	0.6%
Segment 4	2-5	19	2.56	0.27	359	0.24	42.8%	23	0.23	41.2%	1.6%
Upriver Background	15-18	9	2.26	0.33	180	0.44	63.8%	10	0.43	62.8%	1.0%

Notes
(a) PCB data from 2012 Smallmouth Bass Tissue Study. Kennedy/Jenks 2013.
(b) Site-wide domain after exclusion of Segment 1 (highest 2012 SMB PCB levels were observed in RM11 area).
Analyses performed using alpha of 0.05 and power of 0.80.
n = sample size
MDD = Minimum Detectable Difference
SD = standard deviation of the log base 10 transformed mean
SMB = smallmouth bass

DRAFT

APPENDIX D

CALCULATION OF WHOLE BODY – FILLET RATIOS FOR FOCUSED CHEMICALS OF CONCERN IN SMALLMOUTH BASS TISSUE PREPARED BY AECOM

Pre-Remedial Design Work Plan
Portland Harbor Superfund Site

Appendix D
Calculation of Whole Body - Fillet Ratios for Focused Chemicals of Concern in Smallmouth Bass Tissue

Location	Fraction (a)		DDx (b)				Dioxins_Furans (b)				PCB_Congeners (b)			
	Body w/o fillet	Fillet	Body w/o fillet µg/kg	Fillet µg/kg	Whole Body (c) µg/kg	Ratio WB:F	Body w/o fillet pg/g	Fillet pg/g	Whole Body (c) µg/kg	Ratio WB:F	Body w/o fillet pg/g	Fillet pg/g	Whole Body (c) µg/kg	Ratio WB:F
SB010E	0.71	0.28	92.02	7.87	67.95	8.63	4.78	0.55	3.57	6.45	857040	83375	635686	7.62
SB010W	0.71	0.29	254.22	15.55	184.21	11.85	13.50	1.17	9.88	8.48	1117067	122842	825177	6.72
SB011E	0.76	0.23	43.11	6.41	34.29	5.35	9.66	1.12	7.61	6.81	8162565	1481605	6554652	4.42
SB011W	0.75	0.24	86.47	7.36	66.60	9.05	10.37	1.25	8.07	6.46	675258	83874	526546	6.28
SB02E	0.68	0.31	127.54	14.91	91.78	6.16	8.19	0.87	5.86	6.76	1997918	199566	1427353	7.15
SB03E	0.72	0.28	129.37	13.75	96.38	7.01	5.60	0.66	4.19	6.39	371940	40694	277425	6.82
SB03W	0.73	0.26	135.96	14.37	103.62	7.21	7.33	0.91	5.62	6.21	264052	26961	201006	7.46
SB04E	0.72	0.27	123.33	14.85	93.24	6.28	10.30	1.37	7.82	5.70	1855376	241418	1407655	5.83
SB04W	0.73	0.27	172.39	18.07	129.92	7.19	6.93	1.15	5.34	4.62	379264	38773	285558	7.36
SB05W	0.76	0.24	184.29	17.03	143.62	8.43	7.75	0.89	6.08	6.85	345978	34276	270171	7.88
SB06E	0.74	0.26	116.97	11.38	89.25	7.84	7.03	0.78	5.39	6.90	872325	82290	664932	8.08
SB06W	0.75	0.24	279.18	20.47	215.42	10.52	16.54	1.67	12.87	7.70	352513	28952	272752	9.42
SB07E	0.72	0.27	153.26	12.34	114.06	9.24	34.49	3.71	25.92	6.98	2687788	210859	1998729	9.48
SB07W	0.77	0.23	1843.30	180.58	1461.08	8.09	430.22	57.27	344.43	6.01	678402	65752	537571	8.18
SB08E	0.75	0.25	81.59	9.46	63.19	6.68	12.12	1.56	9.43	6.05	369306	43015	286074	6.65
SB08W	0.66	0.33	593.47	48.25	408.18	8.46	43.44	4.21	30.10	7.16	651255	51441	447423	8.70
SB09E	0.65	0.34	100.62	10.99	69.18	6.29	8.13	0.97	5.62	5.78	496991	53205	341323	6.42
SB09W	0.71	0.28	238.42	20.06	175.81	8.76	10.83	0.93	7.99	8.63	1318052	104333	970079	9.30
Average	**0.72**	**0.27**			**Average Ratio**	**7.95**			**Average Ratio**	**6.66**			**Average Ratio**	**7.43**
					Standard Deviation	**1.64**			**Standard Deviation**	**0.96**			**Standard Deviation**	**1.33**
					Table B3-3 of Final FS (EPA 2016)	**7.17**			**Table B3-3 of Final FS (EPA 2016)**	**6.13**			**Table B3-3 of Final FS (EPA 2016)**	**8.02**

Notes:
(a) Fractions calculated using fillet and body without fillet weights in Table 3-3 of Round 3B Fish and Invertebrate Tissue and Collocated Sediment Field Sampling Report (Integral and Windward 2008). Average of individuals in composite sample.
(b) Smallmouth bass tissue chemistry data from SCRA database (sum of detected congeners for dioxins_furans and PCBs).
(c) Whole body concentration calculated as sum of (fillet concentration X fillet mass fraction) + (body w/o fillet concentration X body w/o fillet mass fraction).
Abbreviations:
w/o - without; DDx - sum of dichlorodiphenyltrichloroethane and its derivatives; µg/kg - microgram per kilogram; pg/g - picogram per gram; WB - whole body; F - fillet; FS - Feasibility Study

MEMORANDUM

Date: October 4, 2017

To: The Portland Harbor Superfund Site Pre-RD Group

From: Jason Conder and Anne Fitzpatrick, Geosyntec

Re: **Sample Placement Approach (Grids/Cells vs. Transects) for the Unbiased Surface Sediment Sample Plan**

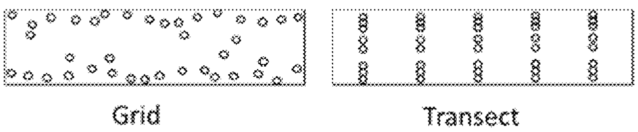
Geosyntec believes (as previously discussed with EPA, below) that placement of the unbiased surface sediment sampling locations using a grid-cell pattern, instead of transects proposed by EPA in their June 2017 scoping plan, will provide better spatial coverage in this dynamic river system.

> **Commented [SEC1]:** EPA has no issues with the grid approach and considers the method valid as long as the points are randomly located within each grid box and are truly unbiased.

Advantages of Grid Cell Placement:

- The grid cell approach will provide a more spatially-balanced dataset for SWAC calculation since the distances between the sample points are more equidistant.
- The transect system compresses the sample location to particular points on the river and leaves large gaps (~0.2 miles) between transects. These un-sampled gaps result in higher spatial uncertainty, and result in fewer rolling River Mile averages that can be calculated. Example: 40-sample layouts for a River Mile (river flow right to left):

> **Commented [SEC2]:** This comment is incorrect as it ignores EPA's sample scheme for subsequent sampling events after the baseline sampling is performed. Section 5.4.1 of the EPA Sampling Plan states: "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."



- Transects are not randomly placed since they follow a simple line across the river. The placement of sample locations in the grid cells is random. Given the variability in physical processes of this dynamic river system, the transect approach could miss certain depositional/erosional environments.
- A spatially-weighted evaluation of surface sediment provides a way to optimize the dataset for future monitoring designs;
- ~~Use of transects would not facilitate re-occupation of locations sampled in 2004 for analysis of recovery potential;~~
- ~~A grid approach is being used/considered at other large, complex Region 10 Superfund sediment sites including PSNS Bremerton and the Lower Duwamish Waterway.~~

> **Commented [SEC3]:** EPA does not agree with comparing the baseline data set with re-occupied (biased) 2004 sample locations. This evaluation will not provide a meaningful recovery evaluation.

> **Commented [SEC4]:** EPA does not agree that these sites mentioned are as large and complex as Portland Harbor. They are not comparable to Portland Harbor. The PSNS Bremerton site is small and not within a river system and the Lower Duwamish remedy is less than a quarter (19%) of the size of the Portland Harbor remedy.

EPA Discussion/Agreement on Grid Cell Approach:

- 8-9-17 Meeting: EPA and Pre-RD group verbally agreed to evaluate sample placement approach for 345 unbiased samples.

> **Commented [SEC5]:** This number of 345 unbiased samples was not verbally agreed to by EPA. This number was the Pre-RD group's original sample placement approach. At this meeting EPA agreed to work with the Pre-RD Group on the sample placement approach to refine and bring into better alignment with EPA's 424 unbiased, spatially distributed, sample location needs. This resulted in the Pre-RD Group's proposed Grid Cell Placement approach presented during the 8-17-17 meeting.

- 8-17-17 Meeting: The Pre-RD group presented the grid cell placement approach to EPA. EPA verbally agreed to evaluate the approach further and requested the grid cell GIS files. Geosyntec provided the grid cell GIS files to EPA on 8/18/17.
- 8-22-17 Meeting: During web meeting between Pre-RD group and EPA technical staff (including S. Sheldrake, J. Kern, and K. Gustavson), EPA verbally agreed with the grid cell approach and requested that samples be placed randomly in each grid cell instead of at the center of each grid cell. EPA requested additional unbiased samples (420 instead of 345). Since that meeting, the Pre-RD Group has agreed to increase the number of unbiased samples to 428 as reflected in the latest draft of the Work Plan.

Commented [SEC6]: The fundamental issue EPA has with the latest Draft of the Work Plan is that the Pre-RD Group has selected 108 of the 428 unbiased samples for re-occupying 2004 sample locations. These re-occupied locations are biased and therefore should not be included in the count of 428 unbiased samples.

Last revised by JMC and AGF on 10/4/2017; saved in Geosyntec Seattle server in projects: Portland Harbor/ technical/AOC scope/ surface sed and geostats

EPA ATTACHMENT

Figure 2 from EPA Sampling Plan

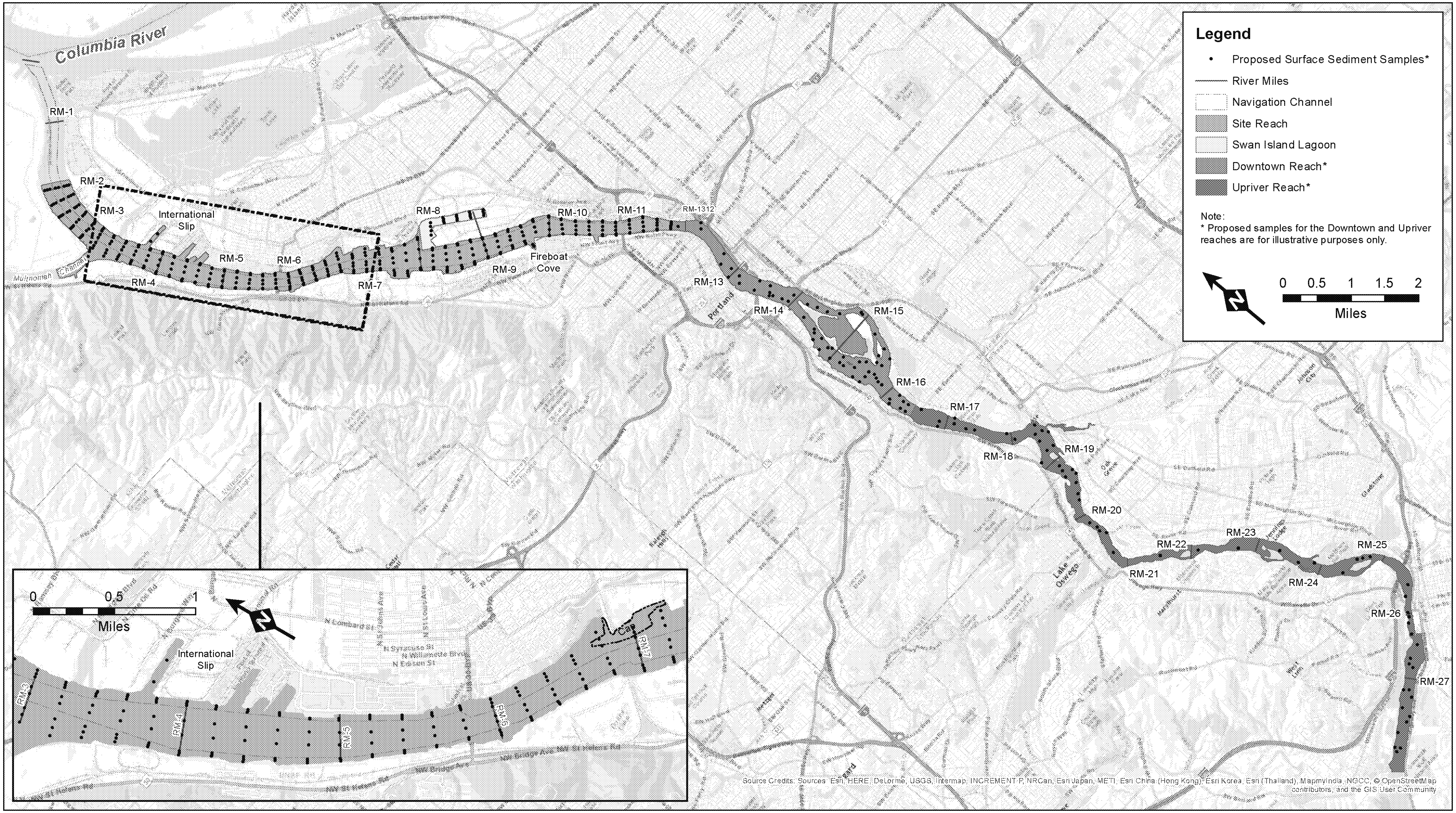


Figure 2. Surface Sediment Transect Locations for Baseline and Long-Term Monitoring

Portland Harbor Superfund Site

Date: 5/31/2017

Appendix A of EPA Sampling Plan

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.

<u>Simulation of Power and Precision</u>

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.

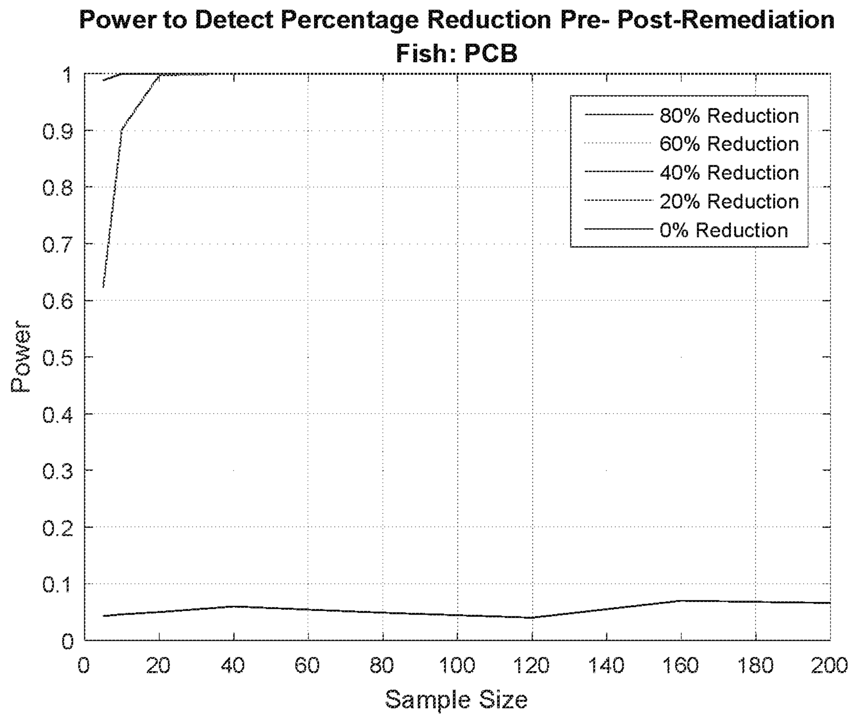


Figure 1. Power to detect a percentage reduction in concentration after completion of the remedy for total PCBs in whole body smallmouth bass tissue.

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.

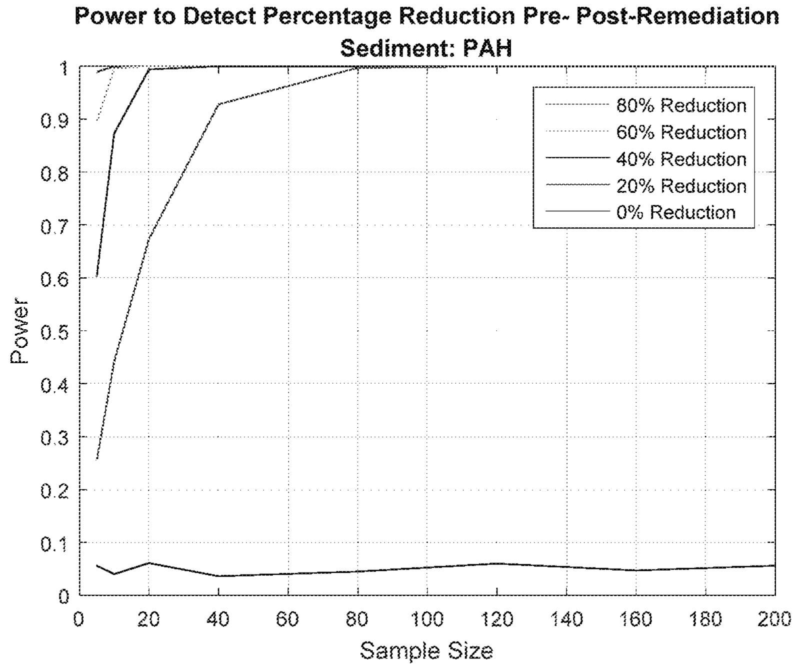


Figure 2. Power to detect a percentage reduction in concentration after completion of the remedy for total PAHs in sediment.

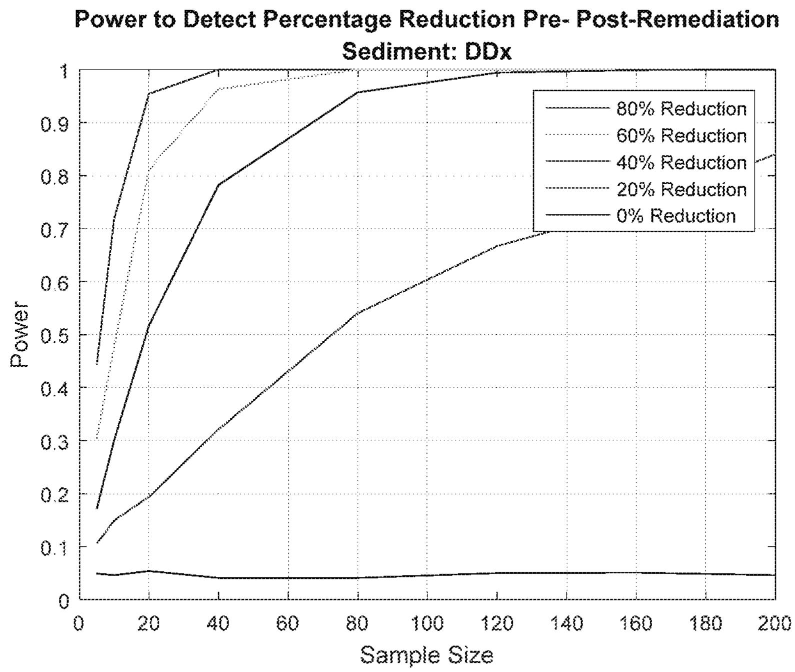


Figure 3. Power to detect a percentage reduction in concentration after completion of the remedy for total DDx in sediment.

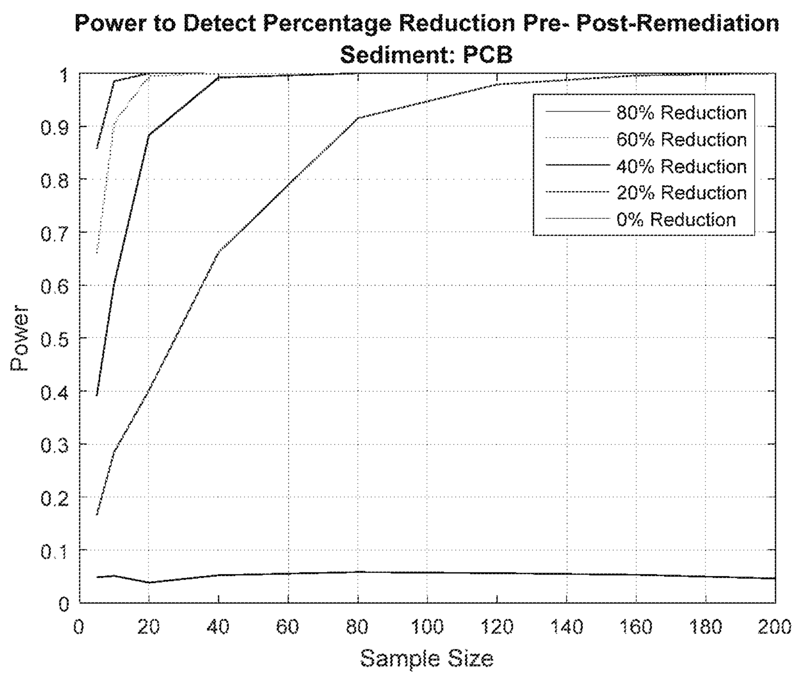


Figure 4. Power to detect a percentage reduction in concentration after completion of the remedy for total PCBs in sediment.

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.

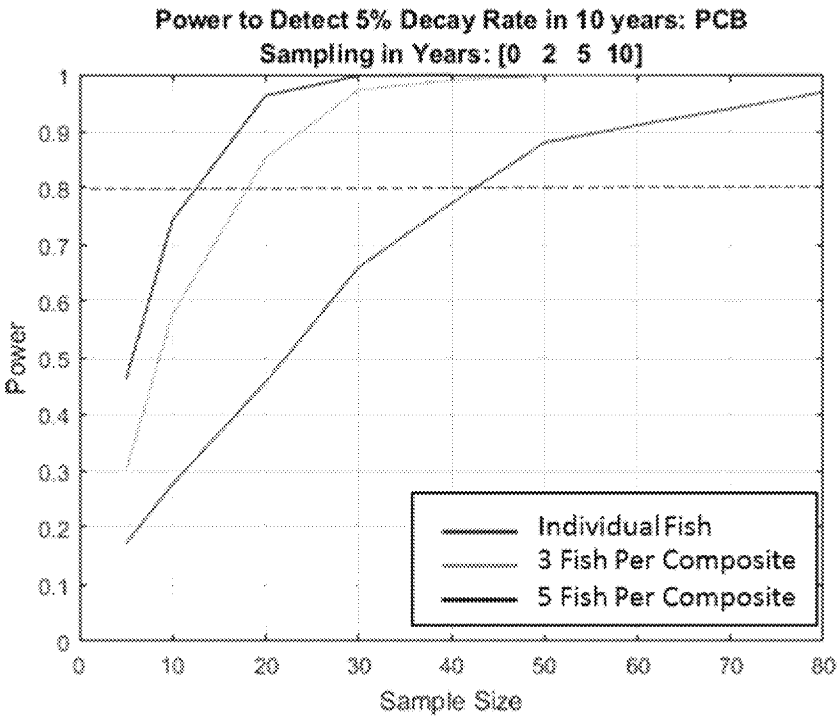


*Figure 5. Power to detect a 5% annualized recovery rate in **smallmouth bass** tissue by monitoring in year 0, 2, 5 and 10 for total PCBs.*

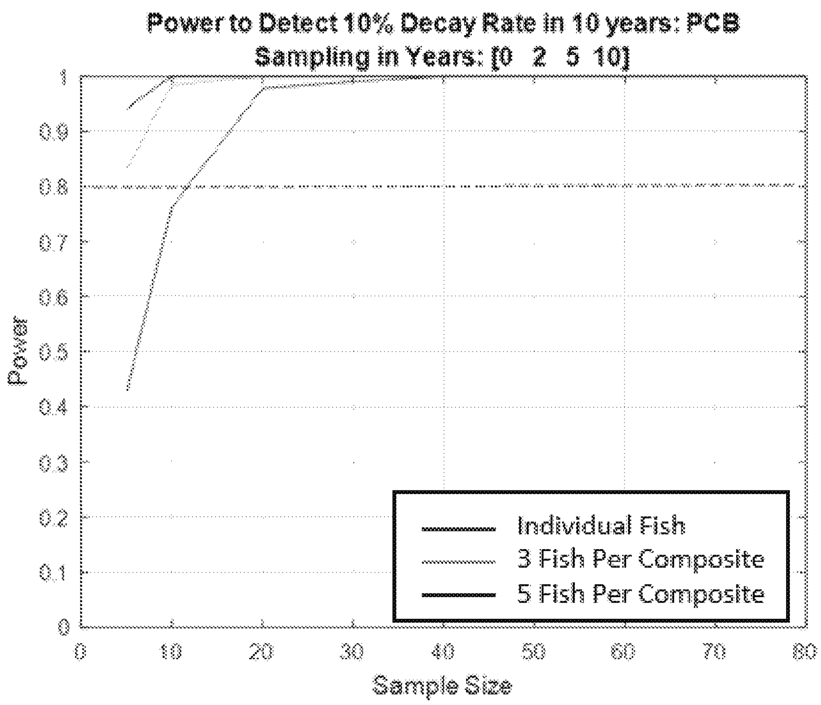


*Figure 6. Power to detect a 10% annualized recovery rate in **smallmouth bass** tissue by monitoring in year 0, 2, 5 and 10 for total PCBs.*

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.

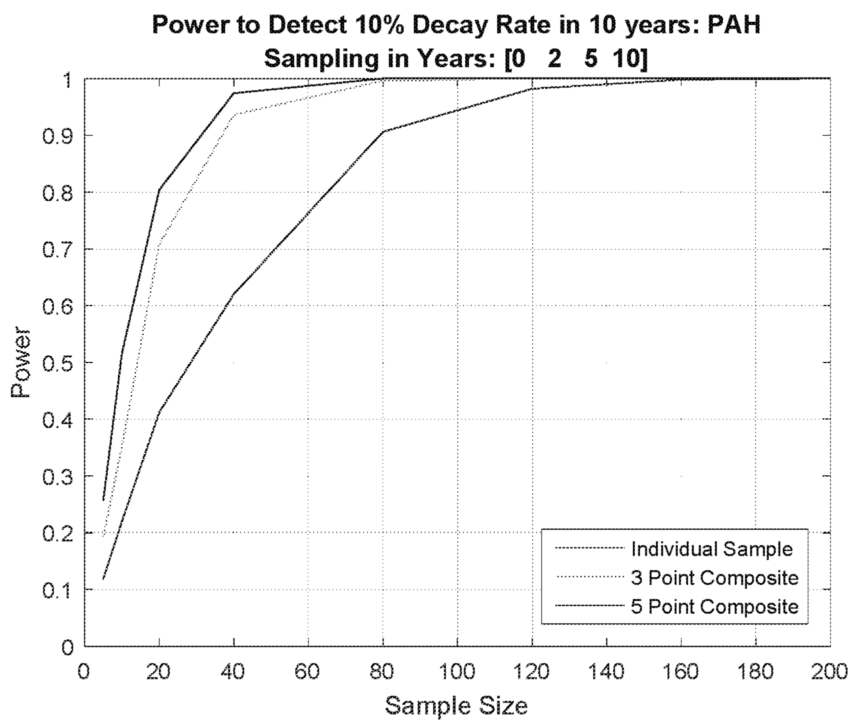


*Figure 7. Power to detect a 10% annualized recovery rate in **sediment** by monitoring in year 0, 2, 5 and 10 for total PAH.*

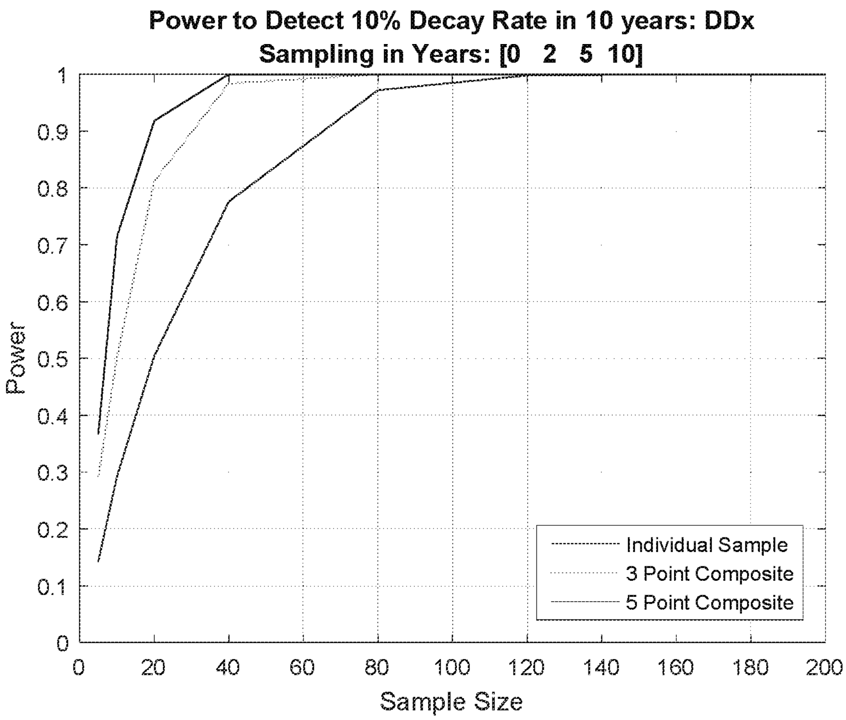


*Figure 8. Power to detect a 10% annualized recovery rate in **sediment** by monitoring in year 0, 2, 5 and 10 for total DDx.*

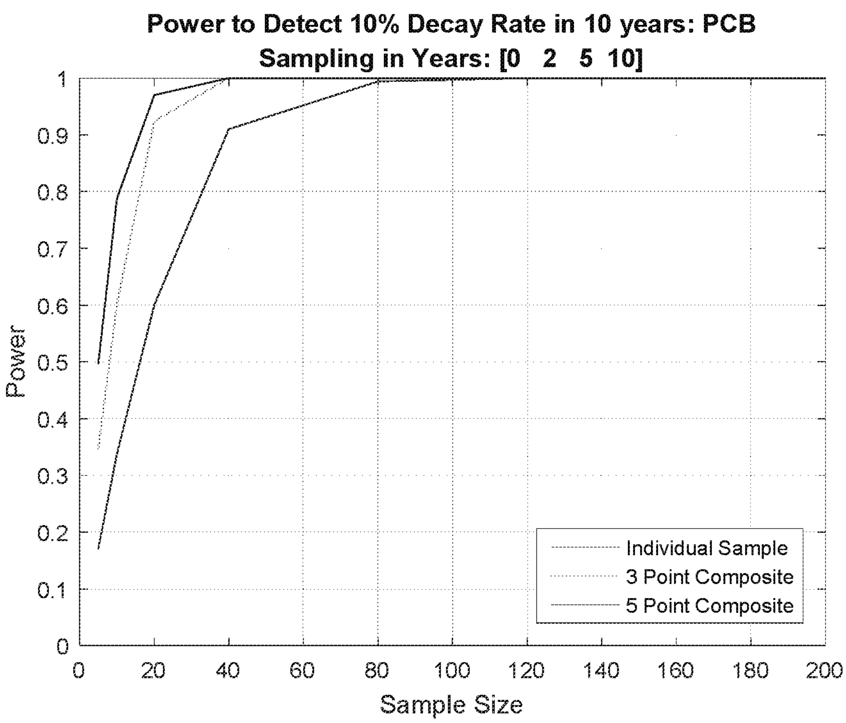


*Figure 9. Power to detect a 10% annualized recovery rate in **sediment** by monitoring in year 0, 2, 5 and 10 for total PCBs.*

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.

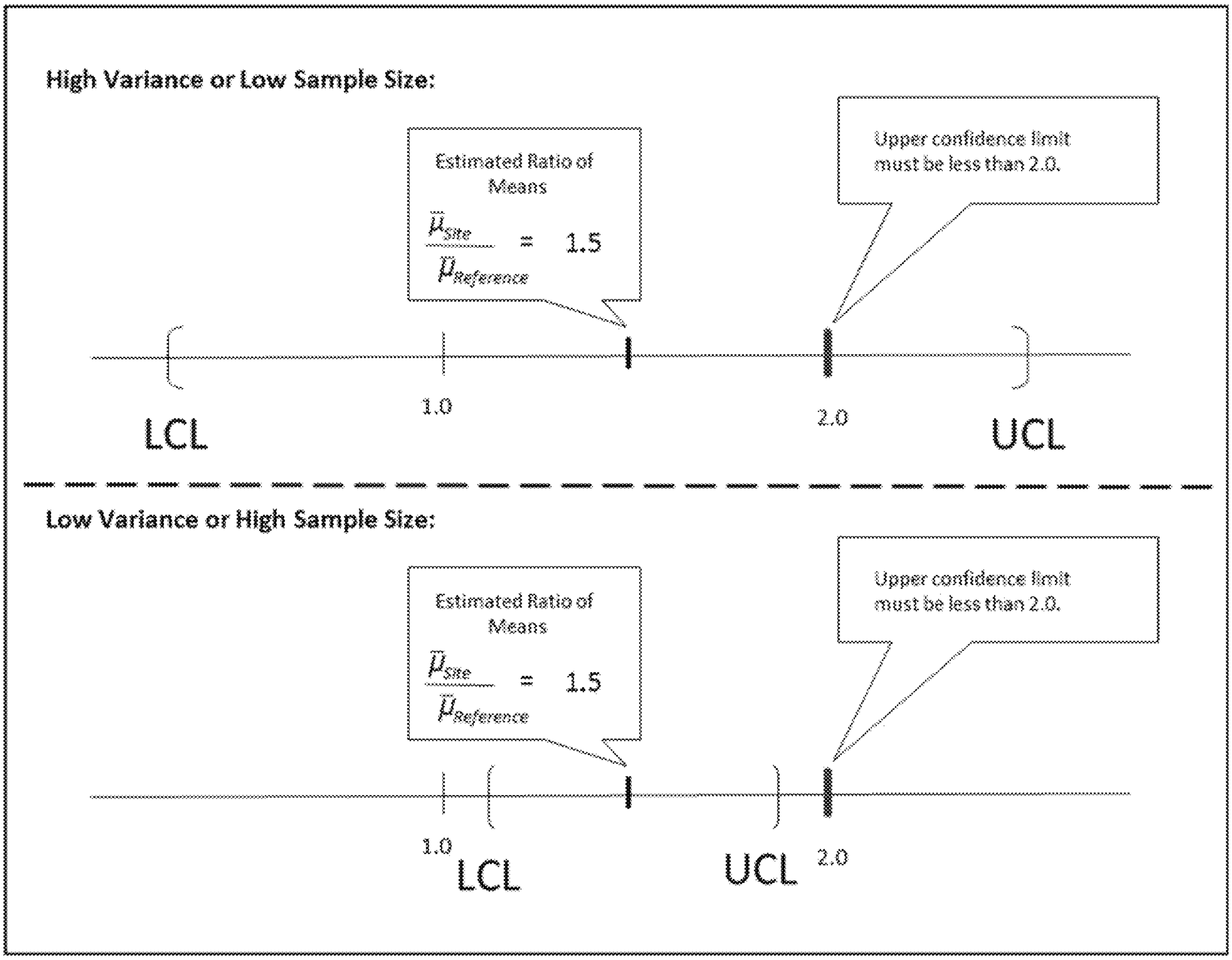


Figure 10. Comparison of reverse and classical null hypotheses with high variance and/or small sample size (top panel) and with low variance and/or large sample size (bottom panel)

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).

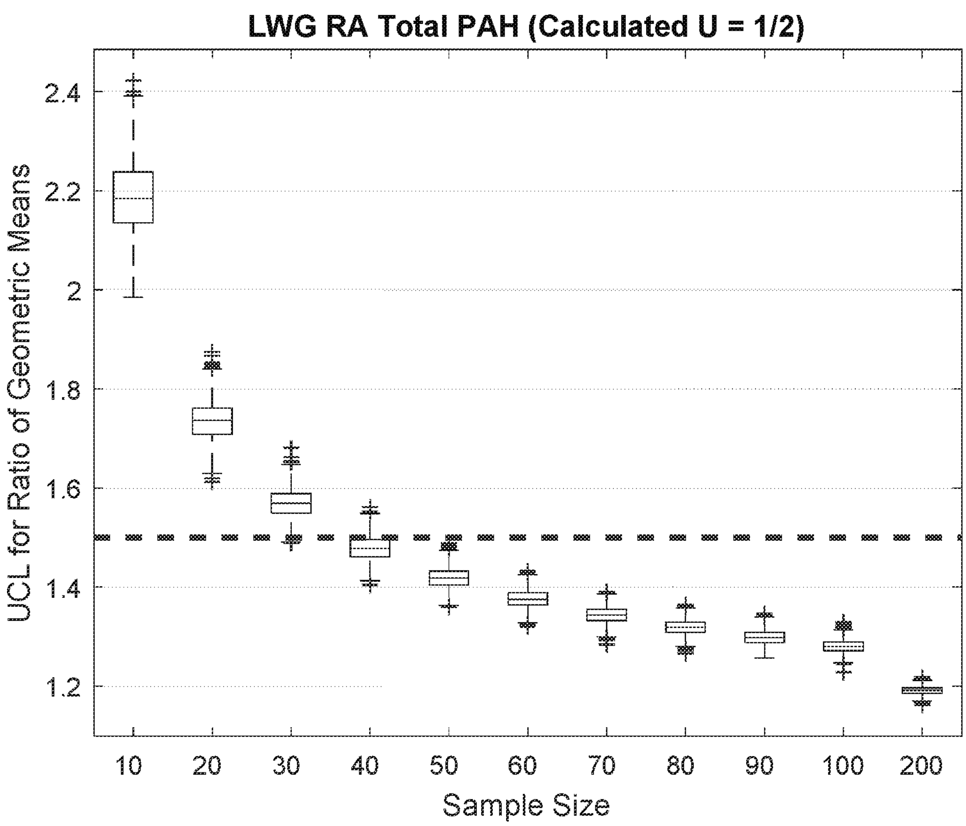


Figure 11. Upper 95% confidence limits for 1000 simulated samples when site and reference data are equivalent for Total PAH in sediment.

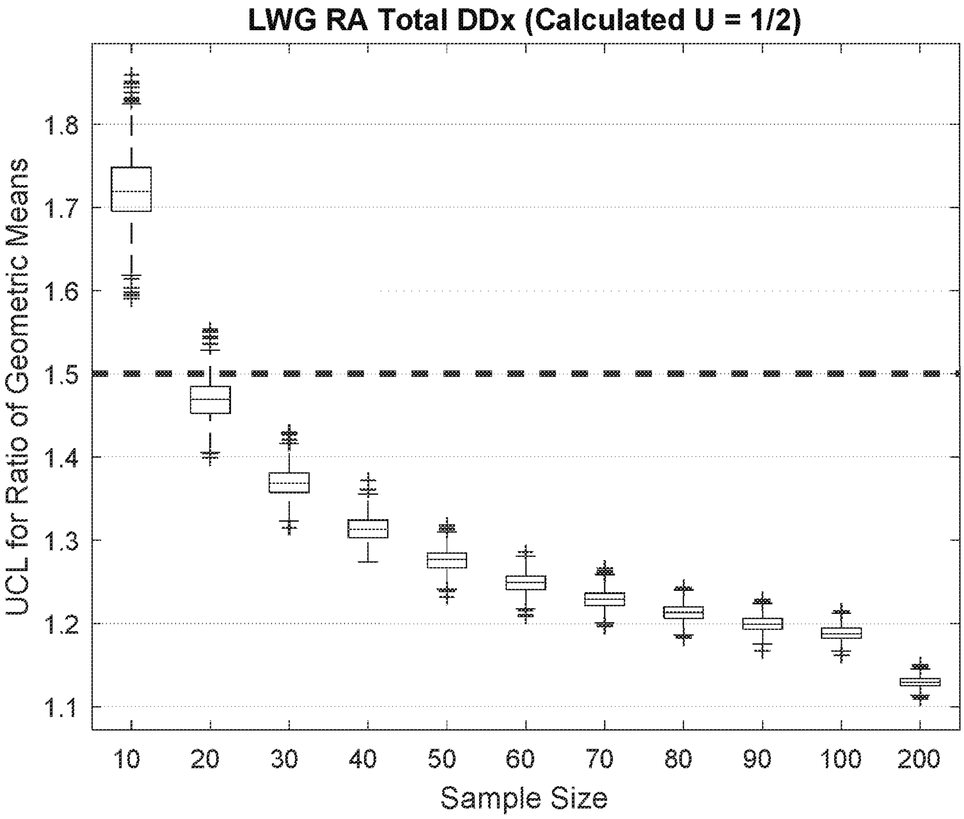


Figure 12. Upper 95% confidence limits for 1000 simulated samples when site and reference data are equivalent for Total DDx in sediment.

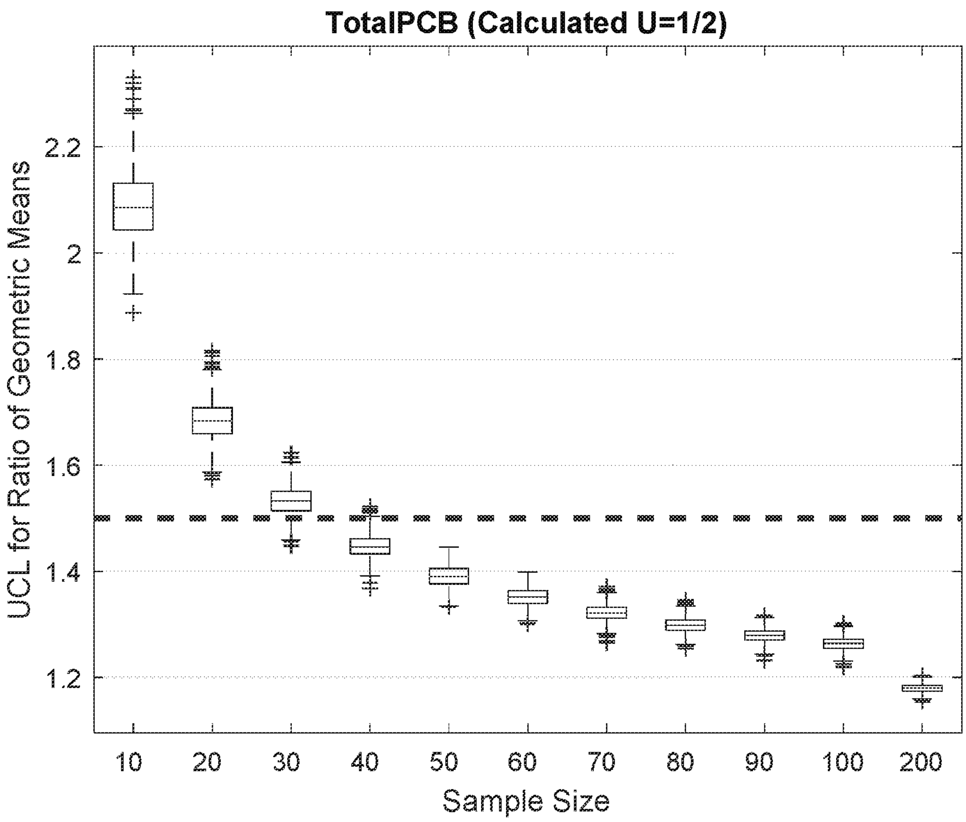


Figure 13. Upper 95% confidence limits for 1000 simulated samples when site and reference data are equivalent for Total PCBs in sediment.

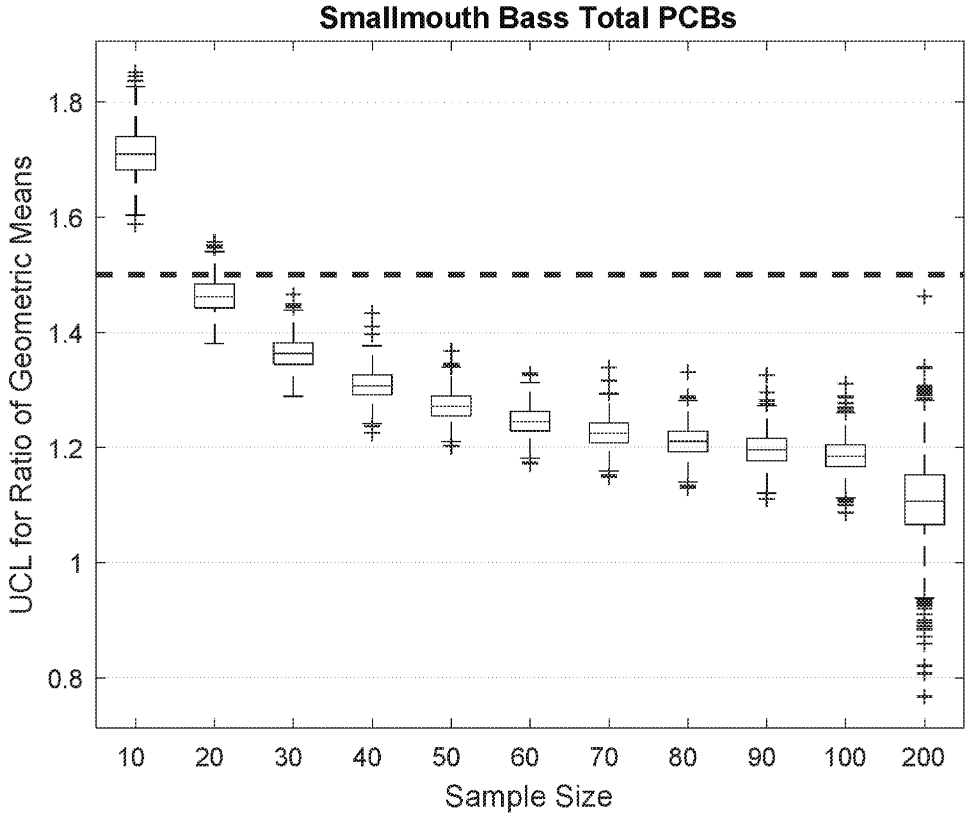


Figure 14. Upper 95% confidence limits for 1000 simulated samples when site and reference data are equivalent for Total PCBs in whole body smallmouth bass..

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.

Hypothesis (Effect Size)	Smallmouth Bass	Sediment		
	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.

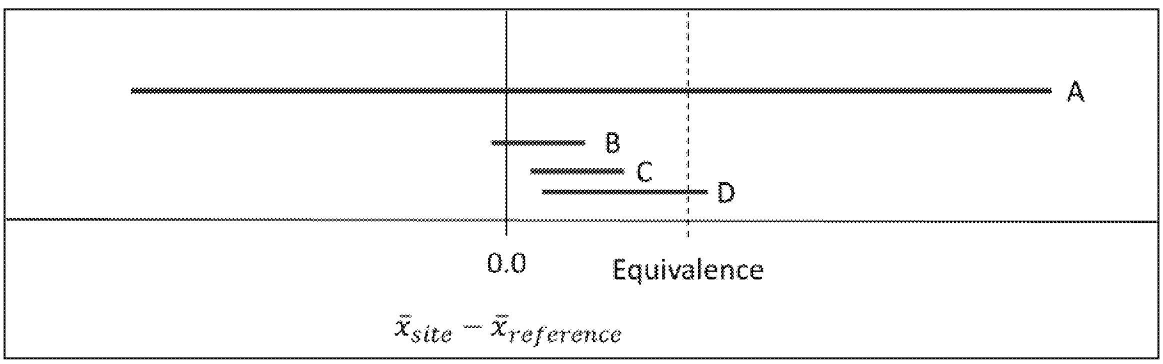


Figure 1. Hypothetical confidence intervals (red lines) for the difference in control and reference concentrations and comparison of traditional and equivalence tests under four different situations (A through D).

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.